2017 IX CONGRESSO NAZIONALE SISMEC

La statistica a supporto della salute:
dalla prevenzione alla continuità delle cure

ATTI

sismec
Società Italiana di Statistica Medica ed Epidemiologia Clinica

2017 IX CONGRESSO NAZIONALE SISMEC

La statistica a supporto della salute: dalla prevenzione alla continuità delle cure

LA STATISTICA A SUPPORTO DELLA SALUTE: DALLA PREVENZIONE ALLA CONTINUITÀ DELLE CURE

ATTI DEL IX CONGRESSO NAZIONALE SISMEC

Segreteria di rediazione: Anna Bossi, Elena Spada, WebMarketingMedia
Edizione: Marzo 2018
A cura di: Consiglio Direttivo SISMEC

ISBN 978-88-943456-0-5

Si ingraziano Bristol-Myers Squibb e Novartis per il contributo non condizionato alla realizzazione degli Atti

Bristol-Myers Squibb

NOVARTIS

sismec
Società Italiana di Statistica Medica ed Epidemiologia Clinica

Comitato Scientifico
Elia Biganzoli (*Università degli Studi di Milano*)
Flavia Carle (*Università Politecnica delle Marche*)
Giovanni Corrao (*Università degli Studi di Milano-Bicocca*)
Ivan Cortinovis (*Università degli Studi di Milano*)
Adriano Decarli (*Università degli Studi di Milano*)
Monica Ferraroni (*Università degli Studi di Milano*)
Ciro Gallo (*Università degli Studi della Campania "Luigi Vanvitelli"*)
Carlo La Vecchia (*Università degli Studi di Milano*)
Umberto Genovese (*Università degli Studi di Milano*)
Francesco Masedu (*Università degli Studi dell'Aquila*)
Rocco Micciolo (*Università di Trento*)
Silvano Milani (*Università degli Studi di Milano*)
Antonella Piga (*Università degli Studi di Milano*)
Patrizio Pasqualetti (*AFaR-Associazione Fatebenefratelli per la Ricerca - Roma*)
Paolo Trerotoli (*Università degli Studi di Bari*)
Maria Grazia Valsecchi (*Università degli Studi di Milano-Bicocca*)
Anna Zolin (*Università degli Studi di Milano*)

Consiglio Direttivo
Franco Cavallo (Presidente)
Anna Bossi (Presidente eletto)
Simona Villani (Segretario)
Rosaria Gesuita (Tesoriere)
Giulia Barbati
Lucia Simoni
Simone Accordini
Paolo Chiodini

Comitato Organizzatore
Anna Bossi
Ivan Cortinovis
Valeria Edefonti
Elena Spada
Anna Zolin

Segreteria Organizzativa
WebMarketingMedia
webmarketingmedia.it
LECCO - MONZA - BERGAMO - BARI – VARSAVIA

Gli articoli pubblicati su prestigiose riviste scientifiche (per un pubblico di esperti) e le notizie riportate dai *media* (per un pubblico più vasto) ci informano, quasi giornalmente, sulle conseguenze per la salute, dell'invecchiamento, di errati stili di vita, della ricomparsa di malattie che si ritenevano eradicate, delle strategie di prevenzione e dell'uso di nuovi farmaci.
Spesso però le informazioni riportate sono contrastanti, di non facile interpretazione o provengono da analisi non corrette, da studi mal programmati o male interpretati.
Infatti, le tecnologie dell'informazione rendono disponibili, facilmente e in tempo reale, i risultati della ricerca biomedica e la "medicina basata sulle evidenze" (EBM) insegna come utilizzare appropriatamente questi risultati nella pratica corrente. Tuttavia, come riportato dal Prof. Vettore (Presidente emerito della Società Italiana di Pedagogia Medica) sul Bollettino d'Informazione sui Farmaci dell'AIFA - … *non bisogna farsi prendere da "deliri di onnipotenza": i risultati scientifici non sono sempre corretti, e soprattutto non sono mai certi, completi e definitivi, ma solo probabili e provvisori; e poi nella clinica anche la migliore delle conoscenze scientifiche va filtrata attraverso l'esperienza del medico e le aspettative del paziente, perché l'EBM – come diceva Sackett – non è un "libro di ricette da cucina"*.
Prendere decisioni razionali, e corrette, in ambito sanitario richiede quindi non solo competenza clinica, ma anche capacità di usare metodi adeguati a valutare l'efficacia degli interventi che si attuano per la salute dei singoli e della collettività.
Inoltre, perché tutti capiscano la razionalità dei risultati delle ricerche scientifiche e li accolgano, è di importanza cruciale che i mezzi di comunicazione (specializzati, o di massa) forniscano informazioni corrette e comprensibili sugli effetti che gli stili di vita, le condizioni ambientali, nonché le strategie di prevenzione, diagnosi, cura e riabilitazione hanno sulla salute di ciascuno.
In questo ambito, il metodo statistico riveste un ruolo indispensabile sia nell'assicurare l'attendibilità delle statistiche sanitarie correnti, sia nel pianificare studi osservazionali e sperimentazioni cliniche eticamente accettabili, sia infine nell'interpretare i risultati e disseminarli in modo chiaro e non distorto.
Il IX Congresso Nazionale SISMEC intende contribuire a consolidare il rapporto di collaborazione tra ricercatori clinici e statistici per migliorare la salute di tutti.
Con questo intento, l'argomento principale della prima giornata sarà relativo all'impatto dello stile di vita sulla salute e quello della seconda, alle nuove strategie terapeutiche e alla valutazione del danno e dell'accertamento della colpa nel caso di "errori" nella cura. Infine, nella terza giornata, sarà discusso come poter monitorare la continuità delle cure quando sono coinvolti molteplici interlocutori e differenti modalità assistenziali.
Con l'augurio che la partecipazione attiva dei Soci della Società e dei ricercatori e professionisti della salute, interessati ad approfondire e discutere l'applicazione del metodo biostatistico ed epidemiologico per arricchire le conoscenze e promuovere la salute, possa far germogliare e crescere alberi robusti…

Continua a piantare i tuoi semi, perché non saprai mai quali cresceranno – forse lo faranno tutti

Albert Einstein

EPIDEMIOLOGIA GENERALE E CLINICA

CARDIOVASCULAR DISEASE RISK ESTIMATION IN THE WORKING POPULATION: DISCRIMINATION ABILITY OF LIFESTYLE RISK FACTORS AND JOB-RELATED CONDITIONS

Veronesi Giovanni[1], Gianfagna Francesco[1,2], Borchini Rossana[3], Grassi Guido[4], Iacoviello Licia[1,2], Cesana Giancarlo[4], Tayoun Patrick[5], Ferrario Marco Mario[1,3]

1. *Research Centre in Epidemiology and Preventive Medicine, Department of Medicine and Surgery, University of Insubria, Varese, Italy.*
2. *IRCCS Neuromed, Pozzilli, Italy*
3. *Occupational Medicine Unit, Varese Hospital, Varese, Italy*
4. *Department of Medicine, University of Milano-Bicocca, Monza, Italy. 5. School of Medicine, University of Insubria, Varese, Italy*

Introduction

Lifestyle and job-related (LS&JR) conditions are recognized risk factors for cardiovascular disease (CVD), but their prognostic utility remains to be established in prospective studies. We investigated the discrimination ability at 10 years of LS&JR risk factors in a Northern Italian working male population.

Methods

N=2532 men, 35-64 years, free of CVD and employed at the time of recruitment (1989-1996) in either the MONICA-Brianza and PAMELA (3 population-based surveys) or the SEMM (1 factory-based survey) studies, were available for the analyses. At baseline, the following LS&JR conditions were ascertained: measured height and weight; self-reported smoking (current vs. non-current) and alcohol intake (drinks/day; less than 1, 1-3, 4-5 and 6 or more); job strain (high vs. non-high; Job Content Questionnaire); physical activity (PA) at work (low, intermediate and intense, according to sample tertiles) and doing sport (minutes/week of moderate or intense activity) from the Baecke questionnaire. Workers were followed-up to the first occurrence of coronary event, acute revascularizations, or ischemic stroke, fatal and non-fatal. A 10-year risk estimation model was developed using LS&JR risk factors satisfying the Akaike Information Criterion for the selection of candidate predictors, and contrasted to a standard risk score including blood lipids, blood pressure, smoking and diabetes. Model discrimination was estimated by the Area Under the ROC-Curve (AUC), in the overall sample and among workers at “low” risk and therefore not qualifying for preventive actions according to European guidelines [***1***].

Results

During 14 years of median follow-up, we observed n=162 events (10-year risk: 4.3%). Body mass index was not associated with the endpoint. The following risk factors met the AI Criterion and entered into the LS&JR model: smoking (Hazard Ratio=2.49, p<.0001); alcohol intake (less than 1 drink/day: HR=1.52, 95%CI 1.03-2.23; 6+ drinks/day: HR=1.81, 1.11-2.95; 3df p=0.07); job strain (HR=1.39, p=0.06); combined sport and occupational PA (5df p=0.02), as the HRs for sport PA changed between workers at low (HR=0.42) and intense (HR=1.55) occupational PA (interaction test p=0.001). The model was well-calibrated (Gronnesby-Borgan chi-square statistic 7.7, p=0.6) and its discrimination ability (AUC=0.750, bootstrapped 95% CI: 0.702-0.780) did not differ from the standard model (AUC=0.749) in the overall sample. The AUC for the LS&JR model was 0.743 among “low” risk workers (1832, with 91 events). Of these, 38% could have been selected for preventive action based on their estimated LS&JR risk; 1 every 16 was a CVD case.

Conclusions

In our working male population, lifestyle and job-related conditions had the same discriminant ability than clinical and biological risk factors in identifying future cardiovascular events, and they may improve stratification of the overwhelming majority of workers classified at low risk by standard scores. A LS&JR risk score may increase feasibility and lower costs of CVD screening at the workplace.

References

[**1**] Piepoli MF, Hoes AW, Agewall S, et al. European Guidelines on cardiovascular disease prevention in clinical practice: The Sixth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice. Eur Heart J 2016;37:2315-81.

SEDENTARINESS AND EDUCATION CONTRIBUTE SIGNIFICANTLY TO SOCIOECONOMIC INEQUALITIES IN NON-COMMUNICABLE DISEASES

Matranga Domenica[1], Bono Filippa[2]

1. *Dipartimento di Scienze per la Promozione della Salute e Materno-Infantile "G. D'Alessandro", Università degli Studi di Palermo*
2. *Dipartimento di Scienze Economiche, Aziendali e Statistiche, Università degli Studi di Palermo*

Introduction

In Europe, the main non-communicable diseases (NCDs), including diabetes, cardiovascular diseases, cancer, chronic respiratory diseases and mental disorders, all together account for an estimated 86% of the deaths and 77% of the disease burden. Of the six WHO regions, Europe is the most affected by NCDs [***1***]. The detection and control of physiological and behavioral risk factors (BRFs) remain the essential preventive strategy to counteract not only the average population's exposure to the main NCDs, but also socioeconomic inequalities, which are related to chronic diseases.
The scope of this work is to investigate socioeconomic inequalities among the European elderly in NCDs and BRFs for NCDs, namely tobacco consumption, obesity, unhealthy nutrition and physical inactivity, between 2004 and 2015.

Methods

Data are drawn from the Survey of Health, Ageing and Retirement in Europe, which is a panel database of microdata on health, socioeconomic status and social and family networks of people aged 50 years and over, covering most of the European Union [***2***]. From waves 1 and 6, release 6, information has been obtained about ten European countries (Austria, Germania, Svezia, Spagna, Italia, Francia, Danimarca, Grecia, Svizzera e Belgio), for a total of 25016 people for year 2004 and 43916 people for year 2015. Socioeconomic inequalities are measured by means of Wastgaff's concentration index and people have been ranked form poorest to richest according to both income and wealth (C) [***3-5***]. The number of NCDs is predicted through negative binomial regression model, with socioeconomic, physical and behavioral covariates. The predicted number of NCDs is used to estimate the concentration curve and to find the contributions (CO) of determinants to socioeconomic inequalities in NCDs. In order to estimate change over time in socioeconomic inequalities in NCDs, the Oaxaca decomposition is used to discriminate how much of this variation is due to change in elasticity and how much is due to changes in inequality of determinants.

Results

Among European elderly people, the number of chronic diseases is significantly associated to all SES determinants and BRF's both in 2004 and 2015. The inequality in the number of NCDs disfavor the poorer in both years, but the effect is decreasing from 2004 (C=-0.191) to 2015 (C=-0.161). This inequality can be mostly attributed to sedentariness and education in both years, even if the role of these determinants is exchanged between 2004 ($CO_{education}$=-0.021, $CO_{sedentariness}$=-0.014) and 2015 ($CO_{education}$=-0.013, CO_{bmi}=-0.016). In 2015, inequalities in all determinants disfavor the poorer in both years. Among SES determinants, the most concentrated in both years are education ($C_{education}$=0.055 in 2004 and $C_{education}$ =0.053 in 2015) and marital status ($C_{marital}$ =-0.055 in 2004 and $C_{marital}$ =-0.045 in 2015). Among BRFs, the most concentrated is sedentariness in both years ($C_{sedentariness}$ =-0.221 in 2004 and $C_{sedentariness}$ =-0.211 in 2015) (***Figure 1***).

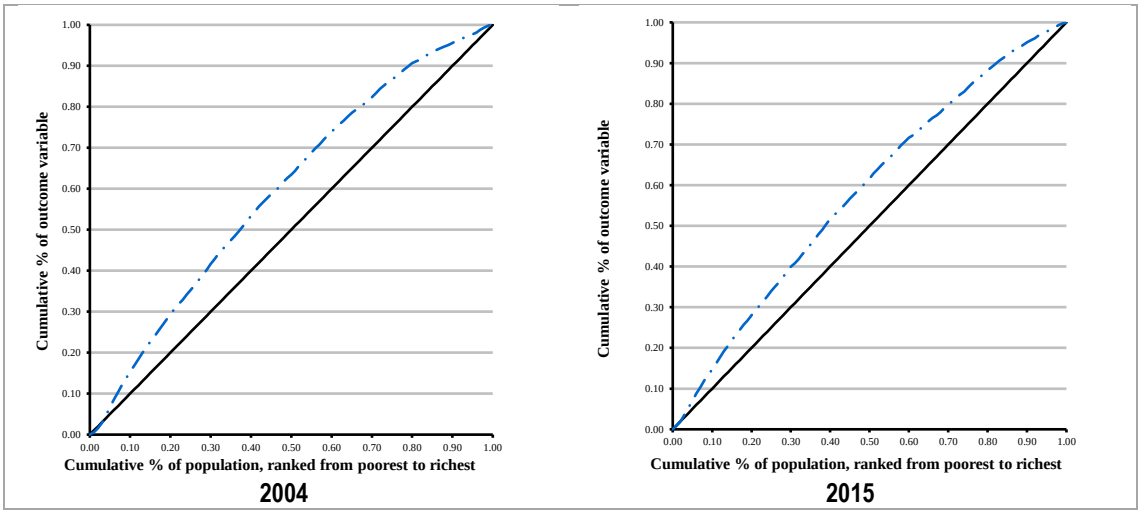

Figure 1. Concentration curve for the number of NCDs among European elderly, years 2004 and 2015

Conclusions

The effect of sedentariness in our study has been found increased in the study period and is the most unequally distributed to the disadvantage of poor people. Other studies addressing the association between socioeconomic status (SES) and sedentariness have shown that poor living conditions and primary education are associated with sedentariness in old age. Among SES determinants, we have found education as the main determinant of both the average number of NCDs and of socioeconomic inequalities in NCDs. The inverse relationship between higher education and the number of NCDs can be explained because educated people have major knowledge about risk behaviors, preventive care and medical treatments and have major access to health care services and use them more efficiently. Wealth more than income has been shown as an important socioeconomic determinant of both NCDs and inequalities in NCDs.
Sedentariness, tobacco consumption and unhealthy nutrition not only contribute to determine the burden of NCDs for European countries but they have been found among the most important determinants of socioeconomic inequalities of NCDs. These are exacerbated by education and wealth. Effective actions to reduce NCDs inequalities include programs to enhance education and economic development and healthy lifestyle promotion.

References

[1] World Health Organization. Non communicable diseases country profiles. Available from: http://www.who.int/nmh/publications/ncd-profiles-2014/en/ [Accessed March 14, 2017].

[2] Börsch-Supan A, Brandt M, Hunkler C, et al. Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE). Int Journal Epidemiol 2013;42:992-1001.

[3] Wagstaff A, Paci P, van Doorslaer E. On the measurement of inequalities in health. Soc Sci Med 1991;33:545-57.

[4] Wagstaff A. Socioeconomic inequalities in child mortality: comparisons across nine developing countries. Bulletin of the World Health Organization 2000;78(1):19-29.

[5] Kakwani N, Wagstaff A, Van Doorlsaer E. Socioeconomic inequalities in health: Measurement, computation and statistical inference. Journal of Econometrics 1997;77(1):87-104.

ANALISI DEL PROFILO CLINICO E SOCIO-ECONOMICO DEGLI ADOLESCENTI CON DISTURBI DEL COMPORTAMETNO ALIMENTARE E DIABETE DI TIPO 1

Skrami Edlira[1], Carle Flavia[1], Ferrito Lucia[2], Iannilli Antonio[2], Cherubini Valentino[2], Gesuita Rosaria[1]

1. *Centro di Epidemiologia, Biostatistica ed Informatica medica, Università Politecnica delle Marche, Ancona, Italia*
2. *Division of Paediatric Diabetes, Women's and Children's Health, AOU Ancona, Salesi Hospital, Ancona, Italy*

Introduzione

I disordini del comportamento alimentare appartengono alle complesse condizioni legate alla nutrizione, molto spesso tra loro interconnesse e associate a bassi livelli di autostima, che assumono forme cliniche patologiche: il denominatore comune è una eccessiva preoccupazione per il proprio peso e forma fisica, che determina una radicale alterazione delle abitudini alimentary [***1***]. È una condizione considerata tradizionalmente tipica dell'adolescenza, ma di recente si è riscontrata una insorgenza sempre più precoce [***2***], con casi di bambine di 8-9 anni di età [***3***]. Dato il progressivo incremento riscontrato negli ultimi decenni, il Ministero della Salute nel 2010, nel Piano Nazionale della Prevenzione, ha affermato che "la diffusione dei Disturbi Alimentari ha una rapidità ed una rilevanza sconcertante: non si ha alcun altro esempio di malattia psichiatrica con una simile propagazione e con le caratteristiche di una vera e propria epidemia sociale".
Gli adolescenti affetti da diabete di tipo 1 (DT1) sono a maggior rischio di sviluppare disordini del comportamento alimentare (DCA) rispetto ai loro pari non diabetic [***4, 5***]: Gli studi a riguardo riportano una elevata variabilità nella prevalenza di questa condizione, con stime comprese fra il 10% e il 49% [***6, 7***], e un recente studio longitudinale ha evidenziato un rischio di sviluppare DCA pari a 79% [***8***]. Gli adolescenti con DT1 tendono a omettere intenzionalmente la terapia insulinica come strumento di controllo del peso corporeo: la presenza di entrambe le condizioni assume, quindi, elevata rilevanza clinica, perché associata ad un cattivo controllo glicemico e ad alti tassi di complicanze e di mortalità.
Dato che la coesistenza di diabete di tipo 1 e DCA ha una prognosi sfavorevole, è di fondamentale importanza individuare precocemente i soggetti a rischio attraverso opportuni strumenti di screening.
Studiare l'associazione fra le caratteristiche cliniche, metaboliche e socio-economiche e i disordini del comportamento alimentare negli adolescenti con DT1, utilizzando e validando il questionario Revised Diabetes Eating Problem Survey come strumento di screening.

Metodi

È stato condotto uno studio trasversale di popolazione nella Regione Marche, fra novembre 2015 e maggio 2016; tutti gli adolescenti di età compresa fra 11 e 20 anni con DT1, appartenenti al Registro regionale del Diabete di tipo 1, sono stati invitati a partecipare in occasione della periodica visita medica programmata. Le variabili cliniche (altezza, peso, BMI-z score per età e genere, numero di episodi di ipoglicemie severe nei precedenti tre mesi, numero di ore dedicate all'attività fisica per settimana, tipo di terapia insulinica, dose totale giornaliera di insulina, numero di somministrazioni di insulina per giorno, numero di somministrazioni di insulina omesse, presenza di co-morbosità) e di laboratorio (HbA1c, trigliceridi, HDL, LDL) sono regolarmente registrate sulla cartella clinica informatizzata.
Durante la visita medica sono state raccolte le informazioni sociodemografiche, quali il livello di educazione (scuola primaria e/o secondaria di primo grado, scuola secondaria di secondo grado, laurea) e il tipo di occupazione (riclassificando i 9 livelli proposti dall'ISTAT, in alto, se manageriale, dirigenziale e professionisti, e in basso, altrimenti) dei genitori ed i partecipanti sono stati invitati a compilare la versione italiana del questionario Revised Diabetes Eating Problem Survey (DEPS-R) [***9***].

Il profilo di istruzione della famiglia è stato categorizzato in *basso*, se entrambi i genitori non erano in possesso di un diploma di scuola secondaria di secondo grado, *medio* se uno dei due possedeva un diploma di scuola secondaria di secondo grado, *alto*, se almeno uno dei genitori risultava laureato. Il profilo di occupazione della famiglia è stato categorizzato in *basso*, se entrambi i genitori avevano un tipo di occupazione basso, *alto*, se almeno uno dei genitori aveva un tipo alto di occupazione.
Il DEPS-R è un questionario specifico per adolescenti con DT1, composto da 16 item, ciascuno dei quali strutturato attraverso una scala Likert a 6 livelli (0 = mai, 1 = raramente, 2 = qualche volta, 3 = spesso, 4 = di solito, 5 = sempre). Il punteggio finale si ottiene effettuando la somma dei singoli item e, quindi, assume valori fra 0 (assenza di disordini) e 80. Il questionario DEPS-R è stato tradotto in italiano, seguendo la procedura di adattamento linguistico e culturale, ed è stata analizzata la sua validità attraverso la valutazione della:

- coerenza interna, calcolando l'α di Cronbach generale e 16 indici item-specifici, ottenuti escludendo un item per volta;
- struttura, applicando l'analisi fattoriale per evidenziare la presenza di eventuali sotto-domini;
- validità discriminante, confrontando i punteggi del questionario dei soggetti con il miglior controllo metabolico (HbA1c<7) con quelli dei soggetti caratterizzati dal peggior controllo glicemico (HbA1c>8.)

La prevalenza di DCA è stata stimata considerando un punteggio di DEPS-R ≥20, e calcolata per genere, classi di età, tipo di terapia insulinica, profilo familiare di istruzione e di occupazione, con i relativi intervalli di confidenza al 95% (IC95%).
L'analisi delle Corrispondenze Multiple (MCA) è stata utilizzata per valutare l'associazione fra i livelli di DEPS-R, il controllo metabolico, (HbA1c <7% e ≥7%), le caratteristiche socio-demografiche personali e della famiglia, e le caratteristiche cliniche e specifiche della patologia dei pazienti.
L'associazione tra la presenza di DCA (variabile risposta) e le caratteristiche cliniche, demografiche e socio-economiche (variabili esplicative) è stata valutata attraverso l'analisi della regressione logistica. Il Likelihood Ratio (LR) test and il test di Hosmer-Lemeshow sono stati utilizzati per individuare il modello più parsimonioso e per valutare la bontà di adattamento del modello ai dati, rispettivamente.
Le analisi sono state effettuate utilizzando il pacchetto R, versione 3.2.4 e la significatività statistica è stata valutata utilizzando un livello di probabilità del 5%.

Risultati

Cento-sessantatré adolescenti (48.5% maschi), su un totale di 219 pazienti presenti nel Registro di patologia, hanno partecipato allo studio (tasso di risposta 74.4%), con una età mediana di 15.4 anni e una durata mediana della malattia pari a 6.4 anni.
La versione italiana del DEPS-R è risultata caratterizzata da un buon livello di consistenza interna (α=0.81, $IC_{95\%}$ 0.76-0.85) e di validità discriminante, riscontrando punteggi significativamente minori nei soggetti con un ottimo controllo metabolico rispetto ai soggetti con un controllo metabolico non ottimale. L'analisi fattoriale ha identificato 3 fattori con autovalori >1, che complessivamente spiegavano il 33% della varianza totale, probabilmente dovuta alla debole correlazione fra gli item; non si è evidenziata, quindi, la presenza di sotto-domini.
In totale 56 adolescenti risultarono avere un punteggio DEPS-R≥20, con una prevalenza pari a 34,4% ($IC_{95\%}$ 27-42), 27% ($IC_{95\%}$ 17-38) nei maschi e 42% ($IC_{95\%}$ 31-53) nelle femmine. La prevalenza risultava significativamente più elevata nei pazienti sovrappeso/obesi (65.7%, $IC_{95\%}$ 47.8–80.9) che nei soggetti normopeso (25.8%, $IC_{95\%}$ 18.5–34.3); non sono emerse differenze statisticamente significative in relazione al genere, età, tipo di terapia insulinica, profilo familiare di istruzione e di occupazione.
MCA ha identificato 2 dimensioni che complessivamente spiegavano il 64% dell'inerzia totale. In particolare ha permesso di evidenziare il profilo clinico dei soggetti positivi al DEPS-R, caratterizzato dalla presenza di obesità, una scarsa attitudine all'attività fisica (meno di due ore per settimana), un basso livello socio-

economico (profilo familiare di educazione e di occupazione basso), un controllo metabolico non ottimale (HbA1c>7), una tendenza ad omettere la terapia insulinica.
La ***tabella 1*** mostra i risultati della regressione logistica: la probabilità di avere un punteggio di DEPS-R ≥20 aumentava del 63% per ogni aumento unitario di HbA1c, di 36% per ogni aumento unitario di omissione di insulina per settimana e diminuiva del 20% per ogni ora aggiuntiva di attività fisica per settimana; inoltre, gli adolescenti sovrappeso o obesi erano sei volte a maggior rischio di avere un punteggio di DEPS-R ≥20 rispetto ai soggetti normopeso.
rispetto ai loro comportamenti alimentari per prevenire l'insorgenza di disturbi di natura patologica.

Tabella 1. Fattori associati ad un punteggio DEPS-R≥20 (analisi della regressione logistica)

Variabili indipendenti	OR	IC95%	p
HbA1c (%)	1.63	1.05; 2.58	0.031
Insulin Delivery System (CSII vs MDI)	0.72	0.27; 1.88	0.513
Dose di Insulina(U/Kg/die)	2.91	0.60; 14.52	0.185
N° di omissioni di insulina per settimana	1.36	1.04; 1.81	0.027
BMI: Sovrappeso/Obeso vs Normopeso	6.32	2.36; 18.13	<0.001
Percentili di Trigliceridi (>75mo vs <=75mo)	0.99	0.33; 2.83	0.983
Percentili di LDL (>75mo vs <=75mo)	0.96	0.32; 2.78	0.939
Attività fisica (ore per settimana)	0.80	0.69; 0.91	0.001
Profilo familiare di istruzione (medio vs basso)	2.01	0.72; 6.00	0.192
Profilo familiare di istruzione (alto vs basso)	1.75	0.50; 6.40	0.384

Hosmer and Lemeshow test: χ^2_8 =2.82, p=0.945
LR test: χ^2_{10}=52.7, p<0.001

Conclusioni

Il presente lavoro è il primo studio di popolazione condotto in Italia sui disordini del comportamento alimentare negli adolescenti affetti da diabete di tipo 1. Più di un terzo degli adolescenti è risultato avere un punteggio di DEPS-R superiore al valore soglia e la prevalenza di DCA è risultata elevata sia nelle femmine che nei maschi. Lo studio, inoltre, ha permesso di evidenziare quali fattori tra le caratteristiche socio-economiche, antropometriche e cliniche, possono essere utili nell'identificazione dei soggetti vulnerabili alla DCA: eccesso di peso, attitudine a omettere la terapia insulinica, scarsa attitudine all'attività fisica, basso profilo familiare di istruzione ed occupazione dovrebbero essere considerati un campanello d'allarme della presenza di disordini del comportamento alimentare negli adolescenti e pre-adolescenti affetti da diabete di tipo 1. Infine, lo studio, effettuando l'analisi della validità della versione italiana del DEPS-R, ha fornito uno strumento semplice e veloce da utilizzare nella pratica clinica al fine di monitorare periodicamente i pazienti in età adolescenziale e pre-adolescenziale

Bibliografia

[1] Dalle Grave R. Eating disorders: progress and challanges. Eur J Int Med. 2011;22:153-60.
[2] Favaro A, Caregaro L, Tenconi E, et al. Time trends in age at onset of anorexia nervosa and bulimia nervosa. J Clin Psychiatry 2009;70:1715-21.
[3] Dalla Ragione L. I disturbi del comportamento alimentare: un'epidemia della modernità. In: Presidenza del Consiglio dei Ministri, Dipartimento della Gioventù. Il Coraggio di guardare: Prospettive e incontri per la prevenzione dei disturbi del comportamento alimentare, Eye03, Roma 2012;19-34.

[**4**] Ackard DM, Vik N, Neumark-Sztainer D, et al. Disordered eating and body dissatisfaction in adolescents with type 1 diabetes and a population-based comparison sample: comparative prevalence and clinical implications. Pediatr Diabetes 2008;9:312-19.

[**5**] Young V, Eiser C, Johnson B, et al. Eating problems in adolescents with type 1 diabetes: a systematic review with meta-analysis. Diabet Med 2013;30:189-98.

[**6**] Wisting L, Frøisland DH, Skrivarhaug T, et al. Disturbed eating behavior and omission of insulin in adolescents receiving intensified insulin treatment: a nationwide population-based study. Diabetes Care 2013;36(11):3382-7.

[**7**] Colton PA, Olmsted MP, Daneman D, et al. Five-year prevalence and persistence of disturbed eating behavior and eating disorders in girls with type 1 diabetes. Diabetes Care 2007;30:2861-62.

[**8**] Colton PA, Olmsted MP, Daneman D, et al. Eating Disorders in Girls and Women With Type 1 Diabetes: A Longitudinal Study of Prevalence, Onset, Remission, and Recurrence. Diabetes Care 2015;38(7):1212-7.

[**9**] Markowitz JT, Butler DA, Volkening LK, et al. Brief screening tool for disordered eating in diabetes: internal consistency and external validity in a contemporary sample of pediatric patients with type 1 diabetes. Diabetes Care 2010;33 (3):495-500.

L'OSPEDALIZZAZIONE PER USO DI SOSTANZE NELLA POPOLAZIONE AQUILANA COLPITA DAL SISMA DEL 6 APRILE 2009: RISULTATI PRELIMINARI

Cofini Vincenza, Vittorini Pierpaolo, Necozione Stefano, Fabiani Leila

Dipartimento Medicina clinica, sanità pubblica, scienze della vita e dell'ambiente e Scuola di Specializzazione Igiene e Medicina Preventiva - Università degli Studi dell'Aquila.

Introduzione

Le evidenze scientifiche indicano che le conseguenze a lungo termine di un disastro naturale, possono provocare cambiamenti importanti in campo sociale economico e sanitario. Dopo il terremoto dell'Aquila sono stati condotti diversi studi sulla salute della popolazione, soprattutto con riferimento a problematiche di natura mentale. In particolare, la prevalenza di consumo di alcol a rischio è risultata elevata [***1***], anche se i dati di letteratura appaiono controversi rispetto all'aumento o decremento dell'abuso di sostanze dopo un disastro naturale [***2***]. Non è stato ancora indagato se l'andamento dei ricoveri per uso di sostanze, abbia subito nel tempo dei cambiamenti importanti, anche in considerazione del segnalato decremento a livello nazionale nello stesso periodo di riferimento [***3***]. L'obiettivo dello studio è quello di valutare il trend dei ricoveri riferiti a disordini per uso di sostanze registrati nel periodo 2009-2015 nella popolazione aquilana colpita dal sisma del 6 aprile 2009, confrontando la popolazione esposta e quella non esposta.

Metodi

Attraverso un progetto di linkage tra i flussi informativi delle SDO abruzzesi ed i flussi informativi del Comune dell'Aquila [***4***], sono stati selezionati i ricoveri effettuati per uso di sostanze in Abruzzo, attraverso l'approccio delle diagnosi principali delle schede di dimissione ospedaliera (SDO), secondo la versione italiana della "International Classification of Diseases 9th revision Clinical Modification" (ICD-9-CM) [***5, 6***] Le analisi hanno riguardato i ricoveri effettuati in Abruzzo dai residenti con età maggiore di 15 anni. La popolazione colpita dal sisma è quella residente nei 42 comuni del cosiddetto cratere sismico della provincia dell'Aquila. Sono stati calcolati i tassi di ospedalizzazione con riferimento alla popolazione residente ISTAT [***7***]. La standardizzazione dei tassi si basa sul metodo diretto attraverso il software STATA 14.

Risultati

Sono stati analizzati 16351 ricoveri. Il tasso di ospedalizzazione è risultato pari a 23 ricoveri per 10000 abitanti nella popolazione esposta al sisma ($IC_{95\%}$ 21.56-23.73) mentre è risultato pari a 17x10000 ($IC_{95\%}$ 17.05-17.61) in quella non esposta. Nella popolazione non esposta l'ospedalizzazione decresce nel tempo, nella popolazione colpita dal sisma, dopo circa 2 anni dal sisma, l'ospedalizzazione subisce un forte aumento, con un tasso del 44*10000 nel 2013, decresce nel 2014 ma aumenta nel 2015 (***Figura 1***).
I tassi standardizzati per genere (anni 2011-2015) indicano differenze significative tra la popolazione femminile esposta al sisma rispetto a quella non esposta (28x10000, $IC_{95\%}$ 25.54-29.47 vs 15x10000, $IC_{95\%}$ 14.77-15.63), mentre non si rilevano differenze per la popolazione maschile. La standardizzazione per età è in corso.

Conclusioni

L'analisi ha evidenziato un andamento diverso dell'ospedalizzazione nella popolazione aquilana colpita dal sisma rispetto alla popolazione abruzzese, soprattutto nelle donne. La standardizzazione per età potrebbe indicare eventuali classi a rischio, il fenomeno andrebbe monitorato con specifica attenzione per una pronta risposta assistenziale oltre che per un'adeguata azione preventiva.

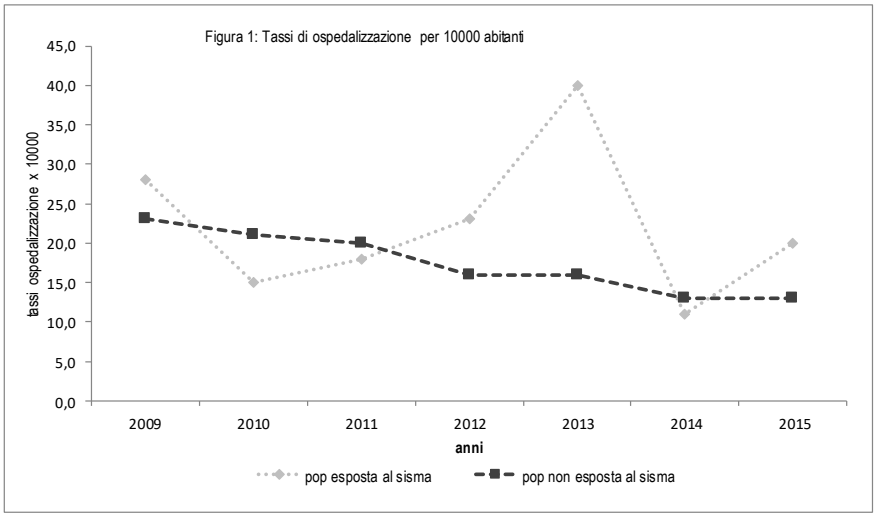

Figura 1. Tassi di ospedalizzazione (per 10000 abitanti) tra il 2009 e il 2015 nella popolazione esposta e non esposta al sisma del 6 aprile 2009

Bibliografia

[1] Minardi V, Gigantesco A, Mancini C, Quarchioni E, D'Argenio P, Cofini V. Behavioural risk factors in L'Aquila (Central Italy) 3-5 years after the 2009 earthquake. Epidemiol Prev 2016;40 (2):34-41.

[2] Pollice R, Bianchini V, Roncone R, Casacchia M. Marked increase in substance use among young people after L'Aquila earthquake. Eur Child Adolesc Psychiatry 2011;20(8):429-30.

[3] Burgio A, Grippo F, Pappagallo M, Crialesi R. Hospitalization for drug-related disorders in Italy: trends and comorbidity. Epidemiology Biostatistics and Public Health 2015;12(1):Suppl 1.

[4] Vittorini P, Angelone AM, Cofini V, et al. (2017) A Case Study on the Integration of Heterogeneous Data Sources in Public Health. In: Rojas I., Ortuño F. (eds) Bioinformatics and Biomedical Engineering. IWBBIO 2017. Lecture Notes in Computer Science, vol 10209. Springer, Cham.

[5] Beck CA, Southern DA, Saitz R, Knudtson ML, Ghali WA. Alcohol and Drug Use Disorders among Patients with Myocardial Infarction: Associations with Disparities in Care and Mortality. PLOS ONE 2013;8(9):e66551.

[6] Ministero del lavoro, della salute e delle politiche sociali. Classificazione delle malattie, dei traumatismi, degli interventi chirurgici e delle procedure diagnostiche e terapeutiche versione italiana della ICD-9-CM "International Classification of Diseases 9th revision Clinical Modification" 2007.

[7] ISTAT. Tassi di ospedalizzazione con riferimento alla popolazione residente. Disponibile in: http://dati.istat.it/

Acknowledgements

Lavoro realizzato con il contributo dell'Agenzia Sanitaria Regionale d'Abruzzo.

FALLS AMONG AGED IMPATIENTS: A STEPPED-WEDGE TRIAL FOR THE EVALUATION OF EFFICACY OF A CARE-BUNDLE INTERVENTION

Di Gennaro Gianfranco[1], Chamitava Liliya[1], Cazzoletti Lucia[1], Mosci Daniela[2], Zanolin Maria Elisabetta[1]

1. *Università degli Studi di Verona, Dipartimento di Diagnostica e Sanità Pubblica*
2. *Azienda Ospedaliero-Universitaria Policlinico S. Orsola Malpighi*

Introduction

Accidental falls are major adverse events in the hospitalization of the elderly. This type of event is a considerable burden on both the quality of life of patients and in terms of healthcare sustainability.
Falls are often the result of the interaction between intrinsic factors linked to the clinical and anamnestic elements of individuals and the environment, due to the physical characteristics of the context (flooring, lighting, furnishings, aids for mobilization etc.). Literature has shown that prevention of falls in the hospital is possible through the adoption of multimodal strategies, which address these different factors. Considering that a universal set of care bundles has not been yet identified, further research is needed [***1, 2***].
In this study, an intervention called Care Bundle, consisting in a series of operations including patient education, improvement of environmental safety, and continuous patient monitoring, is compared with the current clinical organization.
This study aims to assess whether the application of a set of inter-professional interventions (Care Bundle) reduces the incidence of accidental falls in elderly hospitalized patients.

Methods

The study involved 10 hospital units (geriatrics, internal medicine, rehabilitation) of Bologna University Hospital and of Bologna USL. These 10 units (clusters) were randomised in a Stepped Wedge design. Stepped Wedge design consists in a mono-directional crossover-cluster trial, in which clusters switch from control to intervention at different time-points. The benefits of using this kind of design are, in general, the requirement of smaller sample size, with consequent economic, organizational and ethical benefits. Conversely, this type of study requires more advanced methods in terms of statistical analysis [***3***].
The primary endpoint of the study is the count of falls for patient, assessed by the current Incident Reporting Systems used by the hospital units. Incidence rates were calculated considering the total number of falls and the total days of hospitalization.
In addition, autonomy and *a priori* risk of fall were calculated using Barthel and Conley scales [***4, 5***].
For the power calculation, the current rate of falls, obtained from estimates of 2012, of about 4x1000 days of hospitalization was used. Considering a reduction to 2x1000 days of hospitalization, and expecting to observe about 158,000 days of hospitalization, the study will have a 99% power with a significance level of 0.05. This power estimate was obtained using the formula proposed by Woertman et al (2013) with an interclass correlation coefficient of 0.10 [***6***].
Statistical analysis was performed as follow: incidence rates of falls in both the control and treated periods were calculated considering the patient-days of exposure. An overall crude relative risk was calculated with its 95% confidence interval. Moreover, the Care Bundle protocol proposes to develop a Poissonian-linked GEE (Generalized Estimating Equation) model to obtain a more precise population averaged Relative Risk estimation. Care Bundle effect is going to be corrected for age, sex, timing of cluster cross-over and variables concerning clinical severity of patients. In case Poissonian assumptions fail (for example in the case of data over-dispersion), a Negative-Binomial model will be used. The strength of the results obtained will be assessed by the conduction, as sensitivity analysis, of linear mixed models.

Results

The ten participating hospital units were under investigation in the period April 2015- December 2016. Participating units, their randomization in the Stepped Wedge Design and the crude incidence rates are reported in ***figure 1***.

The total number of falls during the overall treatment periods was 200 (exposure days: 72190), while 181 falls were reported during the 56786 days of unexposed hospitalization.

A preliminary crude relative risk estimation was performed: the value of 0.87 (95% CI: 0.71-1.07) indicates a protective effect, though statistical significance was not reached. Complete analysis plan conduction will reveal a more precise estimation of the relative risk of fall.

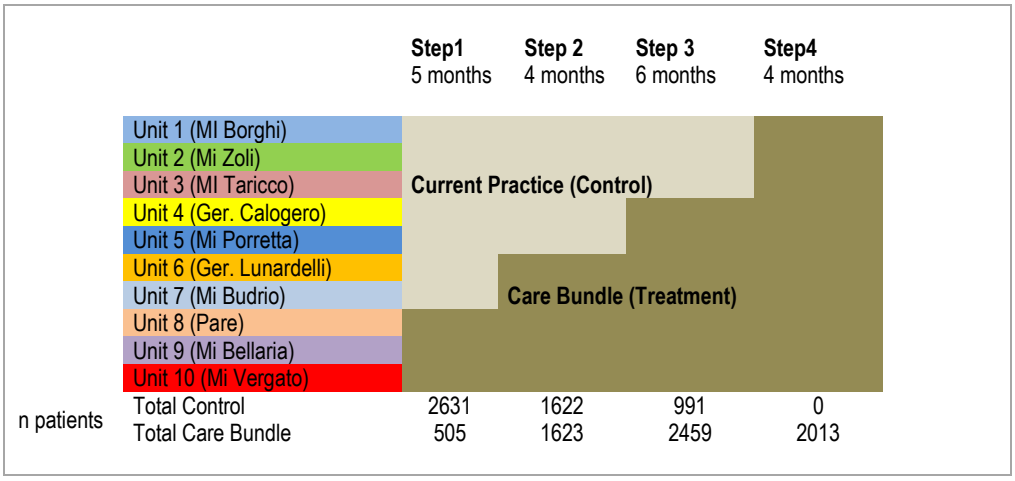

Figure 1. Stepped Wedge Design and Data Summary. 10 hospital units were randomized between control and Care-Bundle arms. Units were followed along 4 periods (Steps) and falls were reported.

Conclusions

A 13% reduction in falls due to the Care Bundle implementation was observed, even if it did not reach statistical significance. An analysis in deep, accounting for confounders, will be performed in a short time.

References

[1] Miake-Lye IM, Hempel S, Ganz DA, Shekelle PG, et al. Inpatient fall prevention programs as patient safety strategy: a systematic review. Ann Intern Med 2013;158:390-6.

[2] Cameron ID, Gillespie LD, Robertson MC, et al. Interventions for preventing falls in older people in nursing care facilities and hospitals. Cochrane Database of Systematic Reviews 2012, Issue 12. Art. No.: CD005465.

[3] Barker D, McElduff P, D'Este C, Campbell MJ,, et al. Stepped wedge cluster randomised trials: a review of the statistical methodology used and available. BMC Med Res Methodol 2016;16:69-88

[4] Mahoney FI, Barthel DW. Functional Evaluation: the Barthel Index. Md State Med J 1965;14:61-5.

[5] Conley D, Schultz AA, Selvin R. The challenge of predicting patients at risk for falling: development of the Conley Scale. Medsurg Nurs 1999;8(6):348-54.

[6] Woertman W, de Hoop E, Moerbeek M, Zuidema SU, Gerritsen DL, Teerenstra S. Stepped wedge designs could reduce the required sample size in cluster randomized trials. J Clin Epidemiol. 2013;66(7):752-8.

CIGARETTE SMOKING AND HEART RATE VARIABILITY: RESULTS FROM THE CHRIS STUDY

Murgia Federico[1], Melotti Roberto[1], Foco Luisa[1], Gögele Martin[1], Motta Benedetta[1], Meraviglia Viviana[1], Merati Giampiero[2,3], Schmidt Georg[4], Rossini Alessandra[1], Pramstaller Peter P[1,5], Pattaro Cristian[1]

1 *Institute for Biomedicine, Eurac Research, Affiliated Institute of the University of Lübeck, Bolzano, Italy*
2. *Dipartimento di Scienze Biomediche per la Salute, Università degli Studi di Milano.*
3. *Centro di Medicina dello Sport, Fondazione Don C. Gnocchi, Milano.*
4 *Medizinische Klinik und Poliklinik, Klinikum rechts der Isar, Technische Universität München, Munich, Germany.*
5 *Department of Neurology, General Central Hospital, 39100 Bolzano, Italy*

Introduction

Heart rate variability (HRV) represents the temporal variation between consecutive heartbeats (R-R interval), reflecting modulation of the autonomous nervous system (ANS) and rhythmicity of the sinoatrial node. Reduced HRV predicts increased mortality in patients with myocardial infarction. In the general population, HRV is associated with heart failure, diabetes and core cardiovascular risk factors, notably cigarette smoking. Several studies have assessed the relationship between cigarette smoking and HRV. However, multiple facets of smoking behaviour present conflicting results.
The aim of our study was to assess the specific effects of smoking status, history, and intensity on HRV in a large population-based rural population.

Methods

We analysed data from 4774 18-to-93 year old participants (55% female) of the Cooperative Health Research in South Tyrol (CHRIS) study. Participants underwent a resting 20-minute electrocardiogram (ECG) in supine position. The standard deviation (SD) of the R-R interval (SDNN) and the root mean square of successive R-R interval differences (RMSSD), measured the global and the parasympathetic dimensions of the ANS, respectively. Interviewers' administered questionnaires adapted from the Cooperative Health Research in the Augsburg Region and the European Community Respiratory Health Survey were harmonized to derive measures of tobacco smoking status (never, past, and current smoking), smoking history among ever smokers (cumulative packyears) and smoking intensity among current smokers (grams/day). Linear mixed models explored the relationship between log-transformed HRV indexes and functions of sex, age and smoking constructs, including interactions, to allow for subjects relatedness, non-linearity and moderating effects. Fractional polynomial (FP) transformations for spike-at-zero variables specifically investigated the role of smoking constructs.

Results

Overall, there were 52.2%, 29.7%, and 18.1% never, past, and current smokers, respectively. SDNN and RMSSD had means of 55.7 (SD=24.2) and 40.1 (SD=22.2) ms, respectively. Ever smokers had mean packyears of 13.6 (SD=12.8) and current smokers had mean smoking intensity of 12.3 (SD=7.5) grams/day. Past, but not current smokers, had higher HRV than never smokers: +0.02 (95%CI: 0.00, 0.05) log(SDNN) and +0.05 (95%CI: 0.02, 0.08) log(RMSSD), in base models. FP models highlighted a negative linear trend of packyears among ever smokers, which was only apparent among current smokers when including interaction between smoking status and packyears. For each increase of 5 units of packyears among current smokers, log(SDNN) was -0.02 (95%CI: -0.033,-0.006) and log(RMSSD) was -0.02 (95%CI: -0.041, -0.004). Further analysis returned a linear reduction of -0.09 (95%CI: -0.12, -0.05) in log(SDNN) and of -0.08 (95%CI: -0.13, -0.03) in log(RMSSD), for each increase of 10 tobacco grams/day. The observed associations were

robust to, and independent of, body mass index, hypertension, diabetes, cardiovascular events, and physical activity.

Conclusions

Current smoking intensity but not cumulative smoking per se is associated with decreased HRV, supporting an acute effect of smoking on the ANS, independent of common cardiometabolic risk factors. Survivor bias may have hidden the detrimental effects of smoking history on HRV. Our comprehensive outline of the smoking-HRV association, based on a large general population sample with wide age range, may help identify better narrative for prevention campaigns about smoking and cardiovascular risk.

DOSE-RESPONSE IN THE CAUSATION OF MESOTHELIOMA FROM ENVIRONMENTAL EXPOSURE: A SYSTEMATIC REVIEW

Magnani Corrado [1], Ferrante Daniela[1], Mirabelli Dario[2]

1. *Unit of Medical Statistics and Epidemiology, Department of Translational Medicine, University of Eastern Piedmont, Novara, and CPO-Piemonte, Novara, Italy.*
2. *Unit of Cancer Epidemiology, CPO Piemonte and University of Turin, Turin.*

Conflict of interest: CM and DM reported that they served as expert witness for the public prosecutor in court trials on asbestos related diseases.

Introduction

The relation between risk of mesothelioma and asbestos exposure has been studied in greater detail at the levels of intensity corresponding to occupational exposure [***1***] while no systematic comparison have been conducted using the information from studies reporting on the risk at lower intensity, as those corresponding to environmental exposures. Non-occupational exposure, including environmental and domestic sources, accounted for 8.3% of cases of pleural malignant mesothelioma (PMM) in a report using data from the Italian National Registry of Mesothelioma – ReNaM [***2***]. Non- occupational cases were younger than those with occupational exposure. Women represented 51% of all cases with environmental exposure, but 84% of those with familial exposure. Subsequent reports from ReNaM confirmed these figures, albeit on more limited statistical analyses [***3***]. The potential relevance of non-occupational exposure to asbestos had been addressed since long [***4***] but limited data existed on the quantification of exposure.

Methods

We investigated the relation between cumulative asbestos exposure and pleural malignant mesothelioma (PMM) in areas with environmental asbestos exposure from human activities and asbestos material in place, using our studies and a literature review.

In Casale Monferrato we conducted a population-based case-control study on PMM, including PMM cases diagnosed between January 2001 and June 2006 (200 PMM and 348 controls). Based on individual, quantitative assessment of exposures, among subjects never occupationally exposed we observed an exposure-response relationship consistent with that caused by occupational exposures. More details were provided in the report by Ferrante et al [***5***].

A systematic literature search was conducted using PubMed, searching for 'Asbestos' and 'Mesothelioma'. The list of titles obtained (4747 items) was examined; abstracts and full papers were examined on the basis of the selection from the titles. The selected items were compared with the selection obtained adding 'Environmental' in the search criteria and with the list of references included in papers and previous reviews. The selected items had to include a quantitative assessment of exposure to asbestos fibres and to analyze the specific contribution of non-occupational exposure [***5-13***].

Results

Casale Monferrato (North West Italy) presents high PMM incidence caused by asbestos contamination at work and in the general environment from the asbestos cement Eternit plant that operated until 1986. ORs were 3.8 (95%CI 1.3 to 11.1) for cumulative exposure from ≥0.1 to <1 f/ml-year, 14.8 (5.7 to 38.6) for ≥1–<10 f/ml-y and 23.3 (95% CI 2.9 to 186.9) for >10 f/ml-y (reference: background level of asbestos exposure).

ORs of about 2, statistically significant, were observed for domestic exposure and for living in houses near buildings with large asbestos cement parts.
Similar trends had been observed in previous studies that explored the exposure-response relationship in the low dose range [***6-8***]. These results are consistent with those observed in the area of Wittenoom, where the exposure intensity, albeit of environmental origin, was higher [***9***]. Kurumatani et al found an exposure-response relationship of PMM with distance from a point source, but could not provide a quantitative estimation of intensity or dose [***10***]. Even more limited information for the purpose of the present analyses could be provided by other studies that did not quantitatively estimate exposures [***11***] or focused on domestic [***12, 13***] or occupational exposures [***14***].

Conclusions

PMM risk increased with cumulative asbestos exposure in analyses limited to subjects non-occupationally exposed and in the environmental exposure range. These results suggest that a definite risk is associated with common sources of environmental exposure and are relevant for the evaluation of residual risk after the cessation of asbestos industrial uses.

References

[**1**] Hodgson JT, Darnton A. The quantitative risks of mesothelioma and lung cancer in relation to asbestos exposure. Ann Occup Hyg. 2000;44:565-601.
[**2**] Mirabelli D, Cavone D, Merler E, et al. Non-occupational exposure to asbestos and malignant mesothelioma in the Italian National Registry of Mesotheliomas. Occup Environ Med 2010; 67:792-4
[**3**] Registro Nazionale Mesoteliomi (ReNaM) Quinto Rapporto Edizioni INAIL Roma 2016. Available from: https://www.inail.it/cs/internet/comunicazione/pubblicazioni/catalogo-generale/il-registro-nazionale-dei-mesoteliomi-v-rapporto.html [Accessed June 15, 2017]
[**4**] Gardner MJ, Saracci R. Effects on health of non-occupational exposure to airborne mineral fibres. IARC Sci Publ 1989;90:375–97.
[**5**] Ferrante D, Mirabelli D, Tunesi S, et al. Pleural mesothelioma and occupational and non-occupational asbestos exposure: a case-control study with quantitative risk assessment. Occup Environ Med 2016;73:147-53.
[**6**] Iwatsubo Y, Pairon JC, Boutin C, et al. Pleural mesothelioma: dose-response relation at low levels of asbestos exposure in a French population-based case-control study. Am J Epidemiol 1998;148:133-42.
[**7**] Lacourt A, Gramond C, Rolland P, et al. Occupational and non-occupational attributable risk of asbestos exposure for malignant pleural mesothelioma. Thorax 2014;69:532-9.
[**8**] Rödelsperger K, Jöckel KH, Pohlabeln H, et al. Asbestos and man-made vitreous fibers as risk factors for diffuse malignant mesothelioma: results from a German hospital-based case-control study. Am J Ind Med 2001;39:262-75.
[**9**] Hansen J, de Klerk NH, Musk AW, Hobbs MS. Environmental exposure to crocidolite and mesothelioma: exposure-response relationships. Am J Respir Crit Care Med 1998;157:69-75
[**10**] Kurumatani N, Kumagai S. Mapping the risk of mesothelioma due to neighbourhood asbestos exposure. Am J Respir Crit Care Med 2008;178:624-9.
[**11**] Magnani C, Agudo A, Gonzales CA, et al. Multicentric study on malignant pleural mesothelioma and non-occupational exposure to asbestos. Br J Cancer 2000;83:104-11.
[**12**] Howel D, Arblaster L, Swinburne L, et al. Routes of asbestos exposure and the development of mesothelioma in an English region. Occup Environ Med 1997;54:403-9.

[**13**] Howel D, Gibbs A, Arblaster L, et al. Mineral fibre analysis and routes of exposure to asbestos in the development of mesothelioma in an English region. Occup Environ Med 1999;56:51-8.
[**14**] Rake C, Gilham C, Hatch J, et al. Occupational, domestic and environmental mesothelioma risks in the British population: a case-control study. Br J Cancer 2009;100:1175-83.

VALUTAZIONE DEGLI EFFETTI A LUNGO TERMINE DELL'ESPOSIZIONE DOMESTICA AD AMIANTO IN UNA COORTE DI MOGLI DI LAVORATORI EX-ESPOSTI

Ferrante Daniela[1], Mirabelli Dario[2], Magnani Corrado[1]

1. *Unità di Statistica Medica ed Epidemiologia dei Tumori, Università del Piemonte Orientale, Alessandria,Novara,Vercelli e CPO Piemonte*
2. *Università degli Studi di Torino, Servizio Epidemiologia dei Tumori, CPO Piemonte, Torino*

Introduzione

Lo studio riguarda la valutazione della mortalità generale e causa-specifica e l'incidenza di mesotelioma maligno della pleura (MM) nella coorte di mogli dei lavoratori dello stabilimento "Eternit" di Casale Monferrato, uno dei maggiori produttori di manufatti in cemento-amianto in Italia attivo dal 1907 al 1986, interessate da esposizione indiretta ad amianto derivante dal contatto e dal lavaggio delle tute da lavoro dei mariti.

Lo studio pubblicato da Ferrante et al. nel 2007 [***1***], con aggiornamento del follow-up al 2003, ha rilevato un aumento della mortalità per tumori dell'apparato respiratorio, soprattutto dovuto all'aumento della mortalità per tumore pleurico (21 oss vs 1,2 attesi; p<0,01), oltre ad un aumento dell'incidenza di mesotelioma maligno della pleura (11 oss vs 0,44 attesi; p<0,01). Si è evidenziata una flessione nell'eccesso di mortalità per tumore polmonare e pleurico per periodi superiori a 40 anni di latenza (tempo dalla prima esposizione al termine del follow-up) mostrata anche in altri studi. Il presente studio è il più importante studio di coorte di familiari dell'amianto e unico per numerosità e durata di follow-up.

Lo studio si propone di valutare gli effetti a lungo termine di esposizioni di tipo domestico ad amianto nella coorte delle mogli dei lavoratori dello stabilimento "Eternit" di Casale Monferrato. In particolare si intende valutare la mortalità e l'incidenza di mesotelioma maligno della pleura in soggetti caratterizzati da esposizioni differenti da quella occupazionale.

L'estensione del follow-up di mortalità al 2014 permette di studiare lo stato in vita delle donne dopo lungo periodo di latenza e di valutare l'esperienza di mortalità per le principali patologie asbesto correlate e altre patologie respiratorie non tumorali il cui eccesso è stato rilevato nell'ambito dello studio di coorte degli esposti professionali ad amianto [***2***].

Metodi

La coorte delle mogli è costituita da 2410 donne, di cui 2017 incluse nel follow-up (sono state escluse le donne con esposizione professionale ad amianto).

L'esposizione è stata definita in base alla sovrapposizione del periodo lavorativo del marito e del matrimonio. Sono state considerate "esposte" le donne sposate ad un dipendente Eternit durante il suo periodo di lavoro in azienda. Sono state considerate "non esposte" le donne sposate ad un dipendente Eternit dopo la cessazione del periodo lavorativo del marito, o con data di fine del matrimonio precedente al primo periodo lavorativo del marito. Sono stati calcolati i rapporti di mortalità standardizzati (SMR) per il periodo 1965-2014 utilizzando come tassi di riferimento i tassi di mortalità della Regione Piemonte considerando il periodo di esposizione domestica e la latenza.

I casi incidenti di MM sono stati identificati tramite il Registro dei Mesoteliomi Maligni (RMM) del Piemonte [***3***], che costituisce il Centro Operativo Regionale del Registro Nazionale dei Mesoteliomi (ReNaM). I tassi di incidenza sono stati standardizzati con metodo indiretto utilizzando come riferimento i tassi della popolazione piemontese.

Risultati

La coorte di donne "esposte" ha incluso 1779 donne. Al 31.12.2014 risultavano: 743 donne vive (41,8%), 1020 decedute (57,3%), 14 emigrate all'estero (0,8%), 2 perse al follow-up (0,1%). Durante il periodo 1965-2014 sono state osservate 62.542 persone anno. È risultato statisticamente significativo l'incremento di mortalità per tutti i tumori maligni (SMR=1.22; p<0.01), tumori dell'apparato respiratorio (SMR=2.9; p<0.01) e per tumore della pleura (SMR=17.3; p<0.01). È stato osservato un aumento di mortalità per tumore peritoneale (4 casi osservati vs 2.1 attesi), tumore dell'ovaio (16 oss vs 12.1 att) che non raggiunge la significatività statistica. Per quanto riguarda il tumore della pleura si è osservato un aumento dell'SMR dopo 20 anni di latenza e per durata di esposizione.

Nel periodo 1990-2012, considerando solo i casi "certi" secondo classificazione ReNaM, si sono osservati 18 casi di mesotelioma pleurico (attesi 1.02; Rapporto Standardizzato di Incidenza (SIR): 17.7 IC95%: 10.5–28.0). Un incremento dei SIRs si è evidenziato in particolare dopo i 30 anni di latenza.

Conclusioni

*T*ale studio permette di migliorare le conoscenze riguardanti gli effetti dell'esposizione domestica ad amianto studiandone le caratteristiche con maggiore dettaglio e valutandone gli effetti a lungo termine. Tale studio è utile per monitorare lo stato di salute della popolazione ex-esposta in soggetti con periodi di latenza superiori ai 40 anni.

Bibliografia

[**1**] Ferrante D, Bertolotti M, Todesco A, Mirabelli D, Terracini B, Magnani C. Cancer mortality and incidence of mesothelioma in a cohort of wives of asbestos workers in Casale Monferrato, Italy. Environ Health Perspect 2007;115:1401-5.

[**2**] Magnani C, Ferrante D, Barone-Adesi F, et al. Cancer risk after cessation of asbestos exposure: a cohort study of Italian asbestos cement workers. Occup Environ Med 2008;65:164-70.

[**3**] Mirabelli D, Stura A, Gangemi M, Bertolotti M, Maule MM, Magnani C. Incidence of malignant mesothelioma in Piedmont, 1990-2001. Epidemiol Prev 2007;31:132-8.

KNOWLEDGE AND USE OF E-CIGARETTE AMONG NURSING STUDENTS: RESULTS FROM A CROSS-SECTIONAL SURVEY IN NORTH-EASTERN ITALY

Finocchio Eliana[1], Canzan Federica[2], Moretti Francesca[3], Vincenzi Silvia[2], Vesentini Roberta[1], Adda Veronica[1], Brugnolli Anna[4], Tchepnou Kouaya Alex[2], Marognolli Oliva[2], Poli Albino[3], Verlato Giuseppe[1]

1. *Unit of Epidemiology & Medical Statistics, Department of Diagnostics and Public Health, University of Verona, Italy*
2. *School of Nursing, University of Verona, Italy – [3]Unit of Hygiene, Department of Diagnostics and Public Health, University of Verona, Italy – –[4]Centre of Higher Education for Health Sciences, Trento, Italy*

Introduction

Data on e-cigarettes use among the Italian population are sparse [***1, 2***], in particular among health professional students, who could play a central role in promoting healthy habits and counselling their patients about smoking cessation. Moreover, conflicting reports exist on the safety of e-cigarettes and on their possible role in promoting smoking cessation [***3***].
The present study aims to i) investigate the diffusion of e-cigarette among nursing students in North-Eastern Italy and ii) explore its association with tobacco smoking.

Methods

In September 2015, 2020 nursing students of the University of Verona were anonymously administered a questionnaire focused on e-cigarette use and tobacco smoking habits. School of nursing comprises 5 centres in North-Eastern Italy, three (Verona, Vicenza, Legnago) located in the plane (Po Valley) and two (Trento, Bolzano) located inside the Alps. The influence of tobacco smoking (never, past, occasional, regular smoker) on e-cigarette ever use (yes/no) was investigated by a multivariable logistic model, controlling for centre, sex, university class, family history of smoking habits, smoking habits among current housemates. Standard errors were adjusted for intra-centre correlation.

Results

Among the 2020 students attending the courses, 1463 (72.4%) answered and returned the questionnaire. Most responders were female (1108/1438, 77.1%) and mean (SD) age was 23.2 (4.2) years.
Nearly all students (1379/1456=94.7%) had heard about e-cigarettes. About one third (442/1460 = 30.3%, 95% CI 27.9-32.7%) had ever used e-cigarettes but only 2.1% (31/1452, 95% CI 1.5-3.0%) in the last month. Very few (21/1015=2.1%) among those who had never used e-cigarettes, were willing to try them. Prevalence values were much higher for tobacco smoking: 40.9% (591/1444) of responders reported to be current tobacco smokers, either occasional or regular, and only 10.1% (146/1444) past smokers. A strong association was observed between tobacco smoking and e-cigarette use: ever use and current use of e-cigarettes were respectively reported by 57.2% and 4.4% of current tobacco smokers and by 12.0% and 0.6% of the other students ($p<0.001$).The association between e-cigarette ever use and tobacco smoking was confirmed in the multivariable logistic model, where the risk of e-cigarette ever use significantly increased with incremental use of tobacco smoking.
Seventeen of 26 dual users (students declaring to currently use both electronic and tobacco cigarettes) provided the motivations to use e-cigarette: tobacco cessation, decrease of tobacco consumption and reduction of harmful health effects were reported respectively by 11 (65%), 3 (18%) and 3 (18%) nursing students of current users. Of note, only three students reported to have completely stopped smoking thanks to e-cigarette.

Conclusions

Use of e-cigarettes seem to be rather rare among Italian nursing students, and mainly restricted to current smokers. E-cigarette does not seem useful to promote smoking cessation. An alarming high prevalence of tobacco smoking was detected, suggesting the need for school-based prevention programs [***4***].

References

[**1**] Gallus S, Lugo A, Pacifici R, et al. E-Cigarette awareness, use, and harm perceptions in Italy: a national representative survey. Nicotine & tobacco research 2014;12:1541-48.

[**2**] Caponnetto P, Campagna D, Papale G, et al. The emerging phenomenon of electronic cigarettes. Expert Rev Respir Med 2012;6(1):63-74.

[**3**] Polosa R, Caponnetto P, Morjaria JB, et al. Effect of an electronic nicotine delivery device (e-Cigarette) on smoking reduction and cessation: a prospective 6-month pilot study. BMC Public Health 2011;11:786-98.

[**4**] Charrier L, Berchialla P, Galeone D, et al. Smoking habits among Italian adolescents: what has changed in the last decade? Biomed Research International 2014: Article Number: 287139

EFFECTS OF GRAND-MATERNAL AND PATERNAL SMOKING ON LUNG FUNCTION IN OFFSPRING

Accordini Simone[1]; Johannessen Ane[2]; Calciano Lucia[1]; Jögi Rain[3]; Martínez-Moratalla Rovira Jesús[4]; Benediktsdóttir Bryndis[5]; Jacobsen Bertelsen Randi[6]; Bråbäck Lennart[7]; Dharmage Shyamali[8]; Gomez Real Francisco[6]; Holm Mathias[9]; Malinovschi Andrei[10]; Marcon Alessandro[1]; Portas Laura[1]; Sánchez-Ramos José Luis[11]; Jarvis Deborah[12]; Schlünssen Vivi[13]; Svanes Cecilie[2]

1. *Unit of Epidemiology and Medical Statistics, Department of Diagnostics and Public Health, University of Verona, Verona, Italy;*
2. *Centre for International Health, Department of Global Public Health and Primary Care, University of Bergen, Bergen, Norway;*
3. *Lung Clinic, Tartu University Hospital, Tartu, Estonia;*
4. *Servicio de Neumología del Complejo Hospitalario Universitario de Albacete (CHUA), Servicio de Salud de Castilla-La Mancha (SESCAM), Albacete, Spain;*
5. *Faculty of Medicine, University of Iceland, Reykjavik, Iceland;*
6. *Department of Clinical Science, University of Bergen, Bergen, Norway;*
7. *Division of Occupational and Environmental Medicine, Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden;*
8. *Allergy and Lung Health Unit, School of Population and Global Health, University of Melbourne, Melbourne, Australia;*
9. *Department of Occupational and Environmental Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden;*
10. *Department of Medical Sciences: Clinical Physiology, Uppsala University, Uppsala, Sweden;*
11. *Department of Nursing, University of Huelva, Huelva, Spain;*
12. *Faculty of Medicine, National Heart & Lung Institute, Imperial College, London, UK;*
13. *Department of Public Health, Aarhus University, Aarhus, Denmark*

Introduction

Maternal exposures during pregnancy are important determinants of children's health, but little is known regarding their multi-generation effects. In addition, little is known regarding paternal pre-conception risk factors and offspring's health.

This analysis is aimed at assessing the effects of grandmothers' smoking in pregnancy and fathers' smoking in puberty (<15 years) on offspring's lung function within the paternal line.

Methods

A three-generation analysis was carried out on 190 fathers who had participated in the "European Community Respiratory Health Survey" (ECRHS; 1991-2013) and on their 276 offspring who had participated in the "Respiratory Health In Northern Europe, Spain and Australia" study (RHINESSA; 2012-2016), from Estonia, Norway and Spain. This analysis is part of the "Ageing Lungs in European Cohorts" study (ALEC; *www.alecstudy.org*), which has received funding from the European Union's Horizon 2020 research and innovation programme (grant agreement No. 633212).

The fathers had valid lung function measurements (pre-bronchodilator FEV_1 and FVC according the American Thoracic Society criterion for reproducibility) and provided complete information on their own smoking history, as well as information on their parents' (grandparents) smoking and respiratory diseases (asthma, chronic obstructive pulmonary disease, emphysema, chronic bronchitis). Their offspring had valid lung function measurements and provided complete information on their own smoking.

Two-level mediation models were developed within the paternal line. These models included: offspring's FEV_1 (or FVC) as the normally distributed outcome; father's FEV_1/FVC as the normally distributed mediator;

grandparents' respiratory diseases, grandmother's smoking history, father's age and smoking history as potential predictors of the mediator; grandmother's, father's and offspring's smoking history, offspring's gender, age and height as potential predictors of the outcome. All models had random intercept terms at level 2 (father), cluster-robust standard errors and a "2→2→1" configuration.

Results

Grandparents' respiratory diseases were associated with a reduced FEV_1/FVC in fathers (-2.8%; 95%CI: -4.8 to -0.8%) and, in turn, a lower FEV_1/FVC in fathers was associated with a reduced FEV_1 in their offspring (1-percent decrease: -1.0 cl; 95%CI: -1.6 to -0.4 cl). Grandmothers' smoking during pregnancy decreased FEV_1/FVC in their sons (-2.1%; 95%CI: -3.9 to -0.3%) and had a direct effect on FEV_1 in their grand-offspring (-6.0 cl; 95%CI: -10.9 to -1.1cl). Fathers' smoking in puberty reduced both FEV_1 (-6.7 cl; 95%CI: -9.4 to -3.9 cl) and FVC (-5.6 cl; 95%CI: -9.9 to -1.2 cl) in their offspring.

Conclusions

These results suggest that grand-maternal and paternal smoking may affect lung function across three generations.

DIETARY FLAVONOIDS AND CHRONIC RESPIRATORY DISEASES IN ADULTS

Mattioli Veronica[1], Cazzoletti Lucia[1], Zanolin Maria Elisabetta[1], Bono Roberto[2], Cerveri Isa[3], Ferrari Marcello[4], Pirina Pietro[5], Garcia-Larsen Vanessa[6]

1. *Unit of Epidemiology and Medical Statistics, Department of Diagnostics and Public Health, University of Verona - Verona (Italy)*
2. *Department of Public Health and Pediatrics, University of Turin - Turin (Italy)*
3. *Division of Respiratory Diseases, Istituto di Ricovero e Cura a Carattere Scientifico "San Matteo" Hospital Foundation, University of Pavia - Pavia (Italy)*
4. *Unit of Respiratory Diseases, Department of Internal Medicine, University of Verona, Verona (Italy)*
5. *Institute of Respiratory Diseases, University of Sassari - Sassari (Italy)*
6. *Johns Hopkins Bloomberg School of Public Health – Baltimore (USA)*

Introduction

Flavonoids are bio-compounds widely found in fruits, vegetables, herbs and red wine. Due to their antioxidant and anti-inflammatory effect in the airways, flavonoids have been suggested to reduce the severity or prevent the risk of lung diseases. Experimental studies in animals have found specific flavonoids to attenuate airway hyper-responsiveness and inflammation, but there is scant epidemiological evidence on the possible modulating role of flavonoids on lung function or chronic respiratory diseases such as asthma or COPD.

Within the Genes Environment Interaction in Respiratory Diseases (GEIRD), a population-based multi-centric multi-case control study in adults, we investigated the association between several chronic respiratory diseases and intake of flavonoids, fruits and vegetables.

Methods

The GEIRD project is a multi-case control study involving seven Italian centers. [***1***] In the first stage of the study, new random samples or preexisting randomly sampled cohorts (the Italian Study on Asthma in Young Adults (ISAYA) [2] and the Italian arm of the European Community Respiratory Health Survey (ECRHS-Italy) [***3***]) from the general population (20–84 years of age, male/female = 1/1) were mailed a screening questionnaire on respiratory symptoms. In the second stage, all the subjects reporting symptoms suggestive of chronic bronchitis (CB), chronic obstructive pulmonary disease (COPD) or asthma, and random samples of subjects reporting symptoms of rhinitis and without symptoms were invited to clinics.

During the clinical visit, each subject underwent a clinical interview, lung function tests (forced spirometry according to the American Thoracic Society reproducibility criteria [***4***]), the reversibility test, bronchial hyper-responsiveness to methacholine, skin prick test (SPT) to common allergens.

Information on usual dietary intake was collected using the Italian version of the validated European Investigation into Cancer and Nutrition (EPIC) food frequency questionnaire (FFQ). [***5***] The NAF software (Nutritional Analysis of Food Frequency Questionnaires, National Cancer Institute, Milan, Italy) was used to transform information about food composition into daily intake of food items, energy, macro- and micronutrients. [***6***] Intakes of flavonoids were derived using as primary sources, a constructed dataset based on values from the updated and expanded US Department of Agriculture (USDA) flavonoid content of foods and the proanthocyanidin databases and from the European database EuroFIR eBASIS (http://www.eurofir.org). Values for the individual flavonoid compounds were assigned to each of the foods listed in the FFQ, and if values for specific foods were not available, we imputed from similar foods if appropriate. [***7***]

Intakes of individual compounds were calculated as the sum of the consumption frequency of each food multiplied by the content of the specific flavonoid for the specified portion size. We derived intakes of the 6 main subclasses commonly consumed in the US diet, specifically flavanones (eriodictyol, hesperetin, and naringenin), anthocyanins (cyanidin, delphinidin, malvidin, pelargonidin, petunidin, and peonidin), flavan-3-ols (catechins and epicatachins), flavonoid polymers (including proanthocyanidins, theaflavins, and thearubigins), flavonols (quercetin, kaempferol, myricetin, and isohamnetin), and flavones (luteolin and apigenin). Total flavonoid intakes were derived by the addition of component subclasses (flavanones, anthocyanins, flavan-3-ols, polymers, flavonols, and flavones).
Overall, out of the 2189 subjects participating in the clinical stage of the survey in the four centres, 994 (45%) completed the EPIC FFQ.
Subjects with implausibly high and low total energy intake and subjects with incomplete FFQ (less than 80% of the 434 questions) were excluded from the analyses. Therefore, the final sample for this analysis was comprised of 869 subjects with clinical and nutritional information. These 869 subjects were hierarchically classified as cases, according to the following outcomes: asthma (CA; n=142); past asthma (PA; n=77); chronic bronchitis (CB; n=45); allergic rhinitis (AR; n=142); non-allergic rhinitis (NAR; n=125); and 338 subjects were grouped as controls.
The following dietary exposures were analysed: total flavonoids, flavanones, anthocyanins, flavan-3-ols, flavonols, flavones and polymers. Also, total intake of fruit, vegetables, citrus and berries were analysed. All exposures were categorized into quartiles on the basis of the distribution of the exposure in control subjects.
Analyses were adjusted for the potential confounders study sample/cohort (ISAYA, ECRHS-Italy, new random sample), centre (Verona, Pavia, Torino and Sassari), gender, age, body mass index (BMI), education, smoking habit, alcohol intake (g/day) and total energy intake (kcal/day).
To identify the associations between our primary exposures of interest and case-control status, several multinomial logistic regression models were fitted to the data, using a 6-level dependent variable (CA, PA, CB, AR, NAR, control).
Multivariable associations of exposures with case-control status were expressed by relative risk ratios (RRRs; using control as the reference category) and their 95%CIs. The p-value for trend was the resulting p-value for the associated model coefficient.

Results

There were no significant differences between groups in terms of age at the clinical visit, gender, BMI, physical activity and educational level. Smoking and drinking habits, and total alcohol intake were significantly different between the six case groups (p-values 0.035, 0.030 and 0.030 respectively): in particular, CB subjects showed the highest proportion of current smokers, current drinkers and of subjects drinking more than 15g/day of alcohol. Daily intake of total and specific flavonoids was not significantly different among groups.
Both unadjusted and adjusted analysis show that total fruit intake (g) was associated with a decreased risk of having CB (***Table 1***).
Intake of anthocyanins was associated with reduced risk of CA (RRR=0.68 95%CI: 0.39;1.18; RRR=0.55 95%CI: 0.30;0.98; RRR=0.61 95%CI: 0.32;1.16 for the 2nd, 3rd, 4th vs 1st quartile).
Flavonols were associated with a decreased risk of AR (RRR=0.53 95%CI: 0.30;0.94; RRR=0.60 95%CI: 0.34;1.06; RRR=0.62 95%CI:0.34;1.13 for the 2nd, 3rd, 4th vs 1st quartile).
Intake of flavanones was associated with reduced risk of NAR (adjusted RRR=0.25; 95% CI 0.11, 0.56 for the 4th vs. 1st quartile of intake). The reduction in the risk of NAR was confirmed when flavanones were considered as a continuous exposure (adjusted RRR=0.71; 95% CI 0.54, 0.92).
Vegetable intake was not associated with any of the outcomes studied.

Table 1. Unadjusted and adjusted RRR (and 95%CI) of being a case of CB (n=45) rather than a control, according to quartiles of fruit, vegetables, citrus and berries intake (based on the distribution of controls) in the GEIRD study.[#]

	Quartile of intake				
CB	1	2	3	4	P (Trend)
Fruits (unadjusted)	1.00	**0.32 (0.13;0.79)**	**0.36 (0.15;0.86)**	**0.37 (0.16;0.87)**	**0.013**
Fruits (adjusted)	1.00	**0.33 (0.13;0.85)**	**0.34 (0.13;0.84)**	**0.33 (0.12;0.86)**	**0.013**
Vegetables (unadjusted)	1.00	1.01 (0.36;2.82)	2.25 (0.93;5.45)	1.39 (0.53;3.63)	0.216
Vegetables (adjusted)	1.00	1.01 (0.35;2.96)	2.38 (0.92;6.14)	1.20 (0.42;3.42)	0.369
Citrus (unadjusted)	1.00	0.72 (0.31;1.67)	0.61 (0.25;1.46)	0.68 (0.29;1.61)	0.320
Citrus (adjusted)	1.00	0.65 (0.27;1.59)	0.64 (0.25;1.63)	0.57 (0.23;1.42)	0.240
Berries (unadjusted)	1.00	0.62 (0.27;1.46)	0.63 (0.27;1.47)	0.57 (0.24;1.36)	0.207
Berries (adjusted)	1.00	0.60 (0.25;1.46)	0.62 (0.25;1.51)	0.57 (0.23;1.48)	0.264

[#] Significant results are in bold.

Conclusions

Fruit consumption might have a beneficial effect against risk of CB. Flavanones seems to reduce the risk of having NAR. Flavonoids found in fruits might reduce the risk of CB and NAR in adults.
There are some evidences of an association of anthocyanins with a reduced risk of CA and of flavonols with a decreased risk of AR.

Reference

[1] de Marco R, Accordini S, Antonicelli L, et al. GEIRD Study Group. The Gene-Environment Interactions in Respiratory Diseases (GEIRD) Project. Int Arch Allergy Immunol 2010;152:255-63.
[2] de Marco R, Poli A, Ferrari M, et. al. ISAYA study group. Italian Study on Asthma in Young Adults. The impact of climate and traffic-related NO2 on the prevalence of asthma and allergic rhinitis in Italy. Clin Exp Allergy 2002;32(10):1405-12.
[3] de Marco R, Verlato G, Zanolin E, et al. Nonresponse bias in EC Respiratory Health Survey in Italy. Eur Respir J. 1994;7:2139-45.
[4] American Thoracic Society. Standardization of spirometry, 1994 update. Am J Respir Crit Care Med 1995;152:1107-36.
[5] Pisani P, Faggiano F, Krogh V, et al. Relative validity and reproducibility of a food frequency dietary questionnaire for use in the Italian EPIC centres. Int J Epidemiol. 1997;26:S152-S160.
[6] Pala V, Sieri S, Palli D, et al. Diet in the Italian EPIC cohorts: presentation of data and methodological issues. Tumori 2003;89(6):594-607.
[7] Cassidy A, O'Reilly ÉJ, Kay C, et al. Habitual intake of flavonoid subclasses and incident hypertension in adults. Am J Clin Nutr 2011;93(2):338-47.

ELECTROCARDIOGRAPHIC CHANGES IN CHRONIC HEPATITIS C PATIENTS TREATED WITH THE NEW DIRECT ACTING ANTIVIRAL AGENTS

Pafundi Pia Clara[1], Parrella Antonio[2], Vitrone Martina[2], Rago Anna[3], Andini Robert[2], Ricozzi Carmen[2], Zampino Rosa[2], Russo Vincenzo[3], Durante-Mangoni Emanuele[2]

1. *Internal Medicine, University of Campania "Luigi Vanvitelli", Department of Clinical and Experimental Medicine c/o Monaldi Hospital, Naples*
2. *Internal Medicine, University of Campania "Luigi Vanvitelli" c/o Monaldi Hospital (Unit of Infectious and Transplant Medicine), Naples*
3. *Unit of Cardiology, Monaldi Hospital, Naples*

Introduction

The hepatitis C virus (HCV) polymerase inhibitor sofosbuvir (SOF) is a major current option for the treatment of chronic hepatitis C (CHC) [***1***], exhibiting a pan-genotypic activity and a high genetic barrier to resistance [***2***]. In combination with ribavirin and/or the nonstructural 5A protein or protease inhibitors, SOF-based regimens show high efficacy and good tolerability [***3-6***].
Several case reports have recently brought to attention its possible role in the generation of cardiac bradyarrhythmias [***7, 8***], fatal in one patient, in the absence or not of amiodarone coadministration [***9***]. A systematic assessment of the effects of SOF, as well as other HCV direct acting antivirals (DAAs), on cardiac rhythm is lacking.
Thus, aim of this study was a detailed analysis of electrocardiographic (ECG) changes occurring during SOF- and non-SOF-based treatment of CHC.

Methods

The study involved CHC patients treated in our unit with various DAA regimens, according to current national guidelines. All patients >18 years, with a fibrosis stage ≥ F3, detectable HCV RNA, treated with any interferon-free regimen between June and December 2015, were included in this study. All subjects underwent a standard 12-lead surface ECG recorded at a paper speed of 25 mm/s and gain of 10 mm/mV, in the supine position, before and during treatment.
All ECG parameter measurements (P wave, QRS, QT interval, JT interval, T^{apex}-T^{end} interval duration) were manually performed with the use of a computer software from each of the 12 ECG lead. Among the laboratory parameters measured at baseline, serum levels of bilirubin, alanine aminotransferases (ALT), aspartate aminotransferases (AST) and gammaglutamyltranspeptidase (GGT) were recorded and included in the analysis. Changes of variables over time were analyzed through the Friedman test was applied.

Results

The study included 39 Caucasian patients with CHC, 26 of whom received a SOF-based therapy and 13 an interferon free regimen not including SOF. No adverse cardiovascular events were observed in the study group. There were no cases of palpitations, symptomatic bradycardia or syncope, and none of the patients had to receive antiarrhythmic medications. At treatment start, as well as after 28 days of therapy, ECG parameters were comparable in the two groups.
No differences in heart rate were observed. In the SOF group, QTc duration significantly rose after one week (p=0.013) and returned to baseline values later during therapy. In contrast, QT dispersion dropped since week 1 and remained significantly reduced up to the end of the observation period (p=0.003). JT dispersion reduced up to week 2 and subsequently returned to baseline at week 4 (p=0.010) (***Table 1***).

As concomitant beta blockers were hypothesized to play a role in the induction of bradycardia and sinus node dysfunction during SOF-based treatment, we analysed ECG changes in 15 SOF-treated patients who were on a beta-blocker at the time of DAA treatment start. Interestingly QTc duration, QT dispersion and JT dispersion changes observed in the overall SOF group did not occur in the beta-blocker subgroup. In contrast, these ECG changes remained statistically significant in SOF-treated patients not on betablockers. Comparing trends over time of each parameter considered, we did not observe significant differences between the two treatment subgroups (***Table 1***).

Table 1. Changes in ECG parameters at four different time-points during antiviral treatment.

Therapy	**t_0, median (IQR)**	**t_7, median (IQR)**	**t_{14}, median (IQR)**	**t_{28}, median (IQR)**	**P**
Sofosbuvir (n = 26)					
P-wave duration, m/s	90 (84.9-102.9)	92.9 (82.2-98.1)	85.5 (76.9-91.1)	88.1 (80-98.5)	0.220
P-wave dispersion, m/s	43 (32.5-64.25)	51 (42.5-58.25)	41 (32.25-52.75)	42 (33-48)	0.895
PR duration, m/s	159.8 (148.8-182.6)	167.4 (151.8-175.6)	162.8 (150.6-12.6)	155.7 (145.4-174.1)	0.360
QRS duration, m/s	91.3 (84.6-105.2)	92.6 (88.7-101.5)	91.4 (82.4-99.2)	93.8 (83-110.8)	0.172
QT dispersion, m/s	75 (60-118)	63 (49-85)	60 (49.5-72)	63 (50.2-81)	**0.003**
QTc duration, m/s	424.3 (411.6-432)	431.2 (419.6-443)	424.8 (410.6-428.2)	422.7 (408.9-435.5)	**0.013**
QTc dispersion, m/s	85.6 (63.7-111.6)	67.2 (50.5-106.1)	72.8 (54.3-80.3)	72.9 (61.1-95.7)	0.581
JT dispersion, m/s	73 (61-96)	68 (53-88)	66 (46-75)	71 (46.8-79)	**0.010**
JTc dispersion, m/s	86.1 (64-101.2)	70.6 (58.3-111)	69.2 (57.6-83.6)	75.7 (51.2-88.1)	0.334
T^{apex}-T^{end} dispersion, m/s	54 (42-68)	49 (31-61)	51 (35-66)	48.5 (35.2-67.5)	0.489
No Sofosbuvir (n = 13)					
P-wave duration, m/s	95.6 (90-100.5)	90.2 (76.4-95.5)	86.8 (82-99.4)	91.1 (87.1-99.7)	0.652
P-wave dispersion, m/s	49 (35-57)	38.5 (23.5-54.7)	37.5 (33.2-63)	44 (30.5-50.5)	0.289
PR duration, m/s	151 (135.8-166.3)	156.7 (145.4-165.7)	155.9 (151.9-162.8)	155.8 (139-162.1)	0.509
QRS duration, m/s	94 (80.6-106.4)	90 (78.8-99.2)	90.1 (80.1-97.7)	97 (85.4-105)	0.115
QT dispersion, m/s	63 (45-94)	58.5 (57-86.2)	63.5 (50.5-77.7)	75.5 (52.2-88)	0.466
QTc duration, m/s	417.3 (411.9-430.1)	416.2 (410.4-420)	416.3 (404.8-421.3)	421.6 (415-427.5)	0.896
QTc dispersion, m/s	66.3 (48.3-98.4)	74.6 (66.7-82.2)	74.3 (63.1-78.3)	76.5 (64.1-93.6)	0.319
JT dispersion, m/s	66 (58-91)	68.5 (57.3-70.8)	56 (47.5-67.5)	77 (62.8-90.3)	0.440
JTc dispersion, m/s	66.9 (46.8-104.8)	81.9 (69.2-94.4)	70.8 (66.7-87.5)	86.6 (70.8-99.1)	0.717
T^{apex}-T^{end} dispersion, m/s	41 (24-69)	50.5 (38.2-65.2)	40.5 (36.7-52.2)	43 (40.2-49)	0.263

*p-value computed through Friedman two-way ANOVA by ranks for correlated samples

Conclusions

None of the current DAA regimens induces changes in cardiac electrical activity potentially associated with clinically significant arrhythmias. In particular, we did not observe significant changes in P-wave, PR duration or QRS duration, consistent with the absence of symptomatic bradycardia [***10***], sinus node dysfunction or atrio-ventricular block. Also others ECG parameters largely remained within the normal range, with some significant early changes in ventricular repolarization that were mostly associated with reduced risk of malignant arrhythmias. We believe our study provides physicians caring for CHC patients and prescribing DAAs useful data confirming the cardiovascular safety of these drugs.

References

[**1**] Pawlotsky JM, Aghemo A, Back D, et al. European Association for Study of Liver. EASL recommendations on treatment of hepatitis C 2015. J Hepatol 2015;63:199-236.

[**2**] Keating GM. Sofosbuvir: a review of its use in patients with chronic hepatitis C. Drugs 2014;74:1127-46.

[**3**] Lawitz E, Poordad FF, Pang PS, et al. Sofosbuvir and ledipasvir fixed-dose combination with and without ribavirin in treatment-naive and previously treated patients with genotype 1 hepatitis C virus infection (LONESTAR): an open-label, randomised, phase 2 trial. Lancet 2014;383:515–23.

[**4**] Lawitz E, Sulkowski MS, Ghalib R, et al. Simeprevir plus sofosbuvir, with or without ribavirin, to treat chronic infection with hepatitis C virus genotype 1 in non-responders to pegylated interferon and ribavirin and treatment-naive patients: the COSMOS randomised study. Lancet 2014;384:1756-65.

[**5**] Afdhal N, Zeuzem S, Kwo P, et al. Ledipasvir and sofosbuvir for untreated HCV genotype 1 infection. N Engl J Med 2014;370:1889-98.

[**6**] Afdhal N, Reddy KR, Nelson DR, et al. Ledipasvir and sofosbuvir for previously treated HCV genotype 1 infection. N Engl J Med 2014;370:1483-93.

[**7**] Food and Drug Administration. FDA drug safety communication: FDA warns of serious slowing of the heart rate when antiarrhythmic drug amiodarone is used with hepatitis C treatments containing sofosbuvir (Harvoni) or Sovaldi in combination with another direct acting antiviral drug. Available from: http://www.fda.gov/Drugs/DrugSafety/ucm439484.htm. [accessed March 24, 2015].

[**8**] Fontaine H, Lazarus A, Pol S, et al. Bradyarrhythmias associated with sofosbuvir treatment. N Engl J Med 2015;373:1886-8.

[**9**] Back DJ, Burger DM. Interaction between amiodarone and sofosbuvir-based treatment for hepatitis C virus infection: potential mechanisms and lessons to be learned. Gastroenterology 2015;149:1315-7.

[**10**] Renet S, Chaumais M-C, Antonini T, et al. Extreme bradycardia after first doses of sofosbuvir and daclatasvir in patients receiving amiodarone: 2 cases including a rechallenge. Gastroenterology 2015;149:1378–80.

DOES FIRST-LINE BEVACIZUMAB-CONTAINING TREATMENT OF METASTATIC COLORECTAL CANCER IMPROVE OVERALL SURVIVAL? THE GRETA OBSERVATIONAL COHORT STUDY

Franchi Matteo[1], Barni Sandro[2], Tagliabue Giovanna[3], Ricci Paolo[4], Mazzucco Walter[5], Tumino Rosario[6], Caputo Antonietta[7], De Ceglie Maria Carolina[7], Corrao Giovanni[1].

1. *Lab of Healthcare Research & Pharmacoepidemiology, Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy*
2. *Department of Oncology- ASST Bergamo Ovest, Italy*
3. *Cancer Registry Unit, Department of Preventive and Predictive Medicine, Fondazione IRCCS National Cancer Institute, Milan, Italy*
4. *Epidemiology Unit, Health Protection Agency Mantua&Cremona, NHS Italy*
5. *Department of Health Promotion and of Maternal and Childhood Sciences, University of Palermo, Palermo, Italy*
6. *Histopathology Department and Cancer Registry, Provincial Health Authority, ASP Ragusa, Italy*
7. *Roche SpA, Monza, Italy*

Introduction

The efficacy of first-line bevacizumab added to chemotherapy (CT) in patients with metastatic colorectal cancer (mCRC) was proved by randomized clinical trials (RCTs) [***1-3***]. Based on these results, bevacizumab plus standard regimens is currently recommended as first-line treatment for mCRC. Despite these findings, however, there is still a relative paucity of data on clinical outcomes associated with bevacizumab-containing therapy used in large, unselected, general clinical practice populations, including elderly patients and subjects with less favourable prognostic factors than those enrolled in RCTs.
The "Generating Real-world Evidence on Therapy of metastatic colorectal cancer with bevacizumab-Avastin" (GRETA) observational cohort study was designed for comparing overall survival (OS) of mCRC patients treated with first-line bevacizumab plus CT (B+CT), as compared to those treated with CT alone, in the real-world setting of Italian clinical practice. Secondary aims were to assess the baseline characteristics of patients assigned to the different therapeutic schemes (B+CT vs. CT alone) and treatment patterns in clinical practice and duration of treatment.

Methods

The cohort of incident mCRC patients was identified during the period 2010-2012 from five population-based cancer registries from Northern (Provinces of Varese, Mantova and Cremona) and Southern (Provinces of Palermo and Ragusa) Italy, covering on the whole a target population of more than 3 million people. Cases were linked to the Regional HCU databases of the five areas covered by the Cancer Registries, in order to obtain the entire pathway of health services provided by the National Health Service to each patient, including outpatient dispensations of high-cost drugs, diagnostic and intervention codes for admission to public or private hospitals and the outpatient services. Only patients who started a first-line treatment with either B+CT or CT alone, within 90 days from the date of diagnosis of mCRC, were included in the study cohort. A propensity score (PS) method was also applied in order to take into account residual confounding.

Results

A cohort of 1,118 incident mCRC cases was identified. After excluding subjects who did not meet the inclusion criteria, a final study cohort of 480 subjects was selected, of which 101 received first-line B+CT and 379 CT alone. As compared to patients using CT alone, those using B+CT were younger and more likely received a surgical intervention. The median OS was 22.5 and 14.6 months in patients treated with or without bevacizumab, respectively (p=0.011). The corresponding adjusted hazard ratio was 0.82 (95% CI

0.62-1.08). The HR resulting from the PS matched analysis was 0.86 (95% CI 0.56-1.33). The gain in median OS was 8.3 months among patients aged less than 70 years (p=0.087) and 4.3 months among those aged 70 years or older (p=0.221). Young ages at baseline (≤70 years) and experiencing surgery were associated with a significant reduction of the risk of death (***Table 1***).

Conclusions

The current study suggested a beneficial effect, even not statistically significant, of adding bevacizumab to CT in the real-world clinical practice of mCRC patients. A modest clinical benefit of using bevacizumab was observed also in patients aged more than 70 years, suggesting that the use of bevacizumab should not be limited only to young patients.

Table 1. Hazards ratios (HR) of death, and 95% confidence intervals (CI), estimated by a multivariable Cox proportional hazard model.

	No.	No. (%) of deaths	HR (95% CI) [a]
At baseline			
First-line treatment			
Chemotherapy alone	379	292 (77.0%)	1.00 (ref)
Bevacizumab + Chemotherapy	101	68 (67.3%)	0.82 (0.62 to 1.08)
Gender			
Men	267	199 (74.5%)	1.00 (ref)
Women	213	161 (75.6%)	1.01 (0.82 to 1.26)
Age (years)			
<50	51	38 (74.5%)	0.80 (0.55 to 1.15)
50 – 59	85	53 (62.4%)	0.56 (0.41 to 0.77)
60 – 69	133	96 (72.2%)	0.80 (0.62 to 1.04)
≥70	211	173 (82.0)	1.00 (ref)
p-trend			0.005
Year of mCRC diagnosis			
2010	183	143 (78.1%)	1.00 (ref)
2011	147	112 (76.2%)	0.94 (0.73 to 1.21)
2012	150	105 (70.0%)	0.87 (0.67 to 1.13)
p-trend			0.296
Charlson comorbidity index			
<=8	417	309 (74.1%)	1.00 (ref)
>8	63	51 (80.9%)	1.23 (0.90 to 1.68)
Surgery			
No	195	166 (85.1%)	1.00 (ref)
Yes	285	194 (68.1%)	0.49 (0.39 to 0.62)
During follow-up			
Radiotherapy [b]	57	52 (91.2)	1.05 (1.01 to 1.10)
Computerised tomography [b]	364	293 (80.5)	1.05 (0.98 to 1.13)
Magnetic resonance [b]	86	68 (79.1)	1.24 (1.07 to 1.43)
X-ray of digestive system [b]	30	25 (83.3)	2.22 (1.45 to 3.39)
Surgery [b]	127	87 (68.5)	0.62 (0.49 to 0.77)

References

[**1**] Hurwitz H, Fehrenbacher L, Novotny W, et al. Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. N Engl J Med 2004;350(23):2335-42.

[**2**] Cassidy J, Clarke S, Díaz-Rubio E. et al. Randomized phase III study of capecitabine plus oxaliplatin compared with fluorouracil/folinic acid plus oxaliplatin as first-line therapy for metastatic colorectal cancer. J Clin Oncol 2008;26(12):2006-12.

[**3**] Saltz LB, Clarke S, Díaz-Rubio E. et al. Bevacizumab in combination with oxaliplatin-based chemotherapy as first-line therapy in metastatic colorectal cancer: a randomized phase III study. J Clin Oncol 2008; 6(12):2013-9.

FOLLOW-UP AND PREDICTIVE MODEL FOR THE RESPONSE TO RHGH TREATMENT IN PAEDIATRIC PATIENTS BASED ON DATA FROM THE GH REGISTRY IN PIEDMONT

Migliaretti Giuseppe[1], Ditaranto Serena[1], Stura Ilaria[1], Cappello Nazario[1], Vannelli Silvia[2], Matarazzo Patrizia[3], Cavallo Franco[1]

1. *Department of Public Health and Pediatric Sciences, University of Turin*
2. *Regina Margherita Hospital (Città della Salute e della Scienza di Torino), Auxology unit*
3. *Regina Margherita Hospital (Città della Salute e della Scienza di Torino), Endocrinology unit*

Introduction

Growth follow up and final height of patients affected by Growth Hormone (GH) deficiency undergoing rhGH therapy have been rarely investigated up to now. This lack is probably due to the shortage of regionally and nationally structured databases. Regional projects for monitoring rhGH treatments in Italy started in 1999 and are now populated and robust enough to offer reliable information. In this paper, we analyze the available data and a mathematical method to predict the patient's final height validated on them.
We pursue a dual aim: to describe the growth of paediatric patients affected by growth failure who were treated with rhGH and to provide a mathematical model predicting the final height of each single patient based on his/her data of the first two years of follow-up.

Methods

We selected 252 patients from the GH Registry who were treated for GH deficit, with a diagnosis of Idiopathic Growth Hormone Deficiency (IGHD) or of Organic Congenital Hormone Deficiency (OGHD), between January 2000 and October 2015. We included subjects who had completed follow up and reached their final height. For all of them, initial and final height, initial pubertal stage and parental Target Height were available. All evaluations are given in terms of SDS score.
The results are based on classical statistical descriptive methods and each indicator is shown with its relative 95% Confidence Interval. The therapeutic efficacy is evaluated in terms of:
- difference between initial and final height
- difference between final height and parental Target Height
The predictive mathematical model, called RBF-PSO, is a mixed method based on approximation and optimization theory [***1, 2***].

Results

68,5% of patients obtained a height gain ≥ +0,5 SDS; note that the majority of them started the treatment in Tanner stage 1; conversely, the majority of patients who did not have a growth gain started treatment in Tanner stage 3. The results seem to highlight the benefits of an early treatment starting. On the contrary, considering as clinical outcome the final Target Height, 64,2% of patients showed a response to therapy with no correlation with Tanner stage at start.
To assess the overall performance of the predictive model, we evaluated a maximum deviation of 5 cm (average RMSE = 3,8), confirming the reliability of the prediction (***Figure 1***).

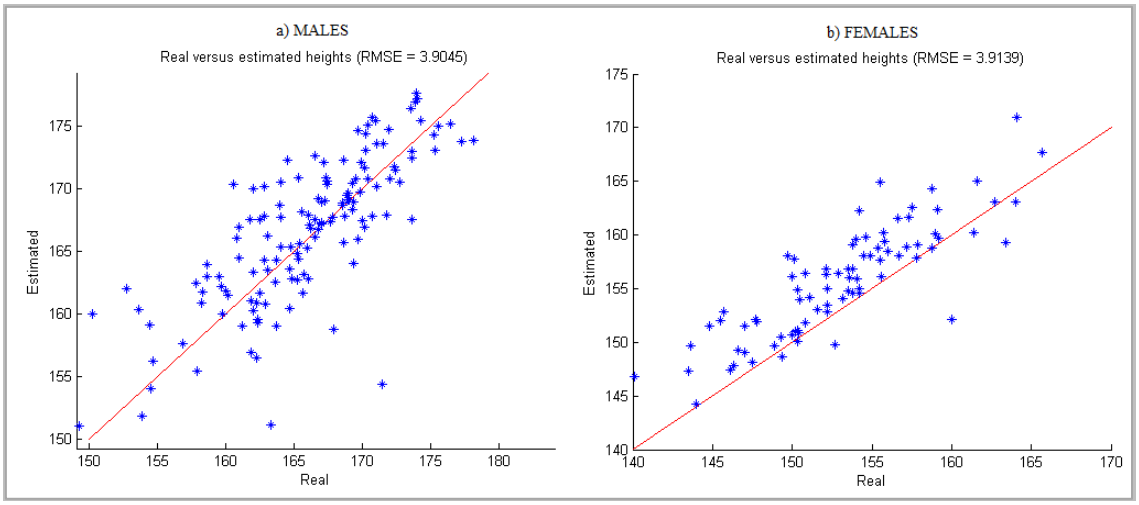

Figure 1. Real versus estimated final height of the considered a) males and b) females patients

Conclusions

These analyses show interesting food for thought, confirming the validity of rhGH treatment and suggesting the best period to start it. These suggestions will be further investigated in future analysis, based on a larger patients cohort, comprehensive of Organic Acquired GHD, that should be available in a short time from other Piedmont centres. In particular, a study on the effectiveness of treatments using also pre-therapy data is in progress.

References

[1] Gliozzi AS, Guiot C, Delsanto PP and Iordache DA. A novel approach to the analysis of human growth, Theor. Biol. Med. Model 2012;9:17-32.

[2] Perracchione E, Stura I. RBF kernel method and its applications to clinical data, Dolomites Res. Notes Approx 2016;9:13-5.

PROSTATE CANCER GROWTH AND RESPONSE TO TREATMENT: DATA AND MODEL REPOSITORIES

Stura Ilaria[1], Migliaretti Giuseppe[1], Cavallo Franco[1], Guiot Caterina[2]

1. *Department of Public Health and Pediatric Sciences, University of Turin, Torino,* ***Italy***
2. *Department of Neuroscience, University of Turin, Torino, Italy*

Introduction

Computational Horizon In Cancer (CHIC) is a European FP7 Project (Grant agreement n° 600841) completed in March 2017 aimed at developing tools, services and secure infrastructure that will support the clinical application of oncological research [***1***]. A double target were therefore selected: to create reliable, realistic and validated mathematical models (*in silico* medicine) to answer specific clinical questions and to create a computer infrastructure in order to connect them to patients' databases. CHIC specifically investigated four cancer types: nephroblastoma, glioblastoma, non-small cells lung and prostate cancer (PCa) (***Figure 1***).

Here we focus on the PCa scenario, with the specific aim of assessing the recurrence probability, their timing and the effectiveness of adjuvant therapies in radically prostatectomized patients. Their serum dosage of PSA is periodically monitored and biochemical recurrence occurs when the threshold value of 0.2 ng/ml is reached.

Methods

Several mathematical and statistical models were worked out in this study, and namely a nomogram to assess the risk of recurrence in 5 years, prediction models of free growth of the tumor [***2, 3***] and the response to adjuvant hormone therapy. All the models were inserted in the CHIC model repository and are available to the interested clinicians by the user-friendly interface called CRAF [***1***].

Results

The models were validated on a clinical dataset including some three thousand Italian patients and were appreciated as 'excellent' by the European Commission. They showed able to predict the risk of recurrence and the most probable treatment scenario on all the patients considered for the demonstration.

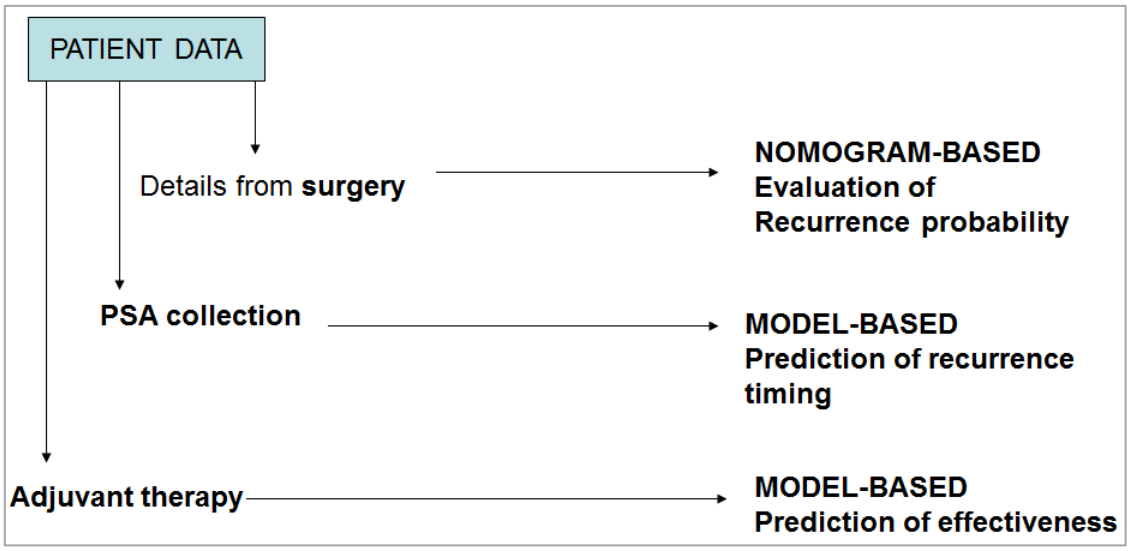

Figure 1. Infrastructure of the PCa study in the CHIC project

Conclusions

The CHIC approach can be hopefully extended to other cancer types and also to other non oncological diseases and the paradigmatic example of PCa can be used as inspiration for other similar study. Well-structured computer tools can be helpful in research and clinical practice, by providing validated in silico models, secure data storage and user-friendly interfaces.

References

[1] Christodoulou NA, Tousert NE, Georgiadi EC, Argyri KD, Misichroni FD, Stamatakos GS. A Modular Repository-based Infrastructure for Simulation Model Storage and Execution Support in the Context of In Silico Oncology and In Silico Medicine. Cancer Inform 2016;15:219-35.

[2] Perracchione E, Stura I. An rbf-pso based approach for modelling prostate cancer. In: AIP Conf. Proc. of ICNAAM 2015, volume 1738, 2016.

[3] Stura I, Gabriele D, Guiot C. A simple PSA-based computational approach predicts the timing of cancer relapse in prostatectomized patients Cancer Res 2016;76(17):4941-7.

INCREASED AIRWAY RESPONSIVENESS CAN PREDICT COPD: RESULTS FROM THE ALEC STUDY

Marcon Alessandro[1], Locatelli Francesca[1], Keidel Dirk[2,3], Beckmeyer-Borowko Anna[2,3], Cerveri Isa[4], Dharmage Shyamali[5], Fuertes Elaine[6,7,8], Garcia-Aymerich Judith[6,7,8], Heinrich Joachim[9], Imboden Medea[2,3], Janson Christer[10], Johannessen Ane[11], Leynaert Benedicte[12], Pascual Erquicia Silvia[13], Pesce Giancarlo[1], Schaffner Emmanuel[2,3], Svanes Cecilie[11,14], Urrutia Landa Isabel[13], Jarvis Deborah[15,16], Probst-Hensch Nicole[2,3], Accordini Simone[1]

1. *Unit of Epidemiology and Medical Statistics, Department of Diagnostics and Public Health, University of Verona, Italy;*
2. *Department Epidemiology and Public Health, Swiss Tropical and Public Health Institute, Basel, Switzerland;*
3. *University of Basel, Switzerland;*
4. *IRCCS San Matteo Hospital Foundation, University of Pavia, Italy;*
5. *Allergy and Lung Health Unit, School of Population and Global Health, University of Melbourne, Australia;*
6. *ISGlobal, CREAL, Barcelona, Spain;*
7. *Universitat Pompeu Fabra, Barcelona, Spain;*
8. *CIBERESP, Barcelona, Spain;*
9. *Institute and Outpatient Clinic for Occupational, Social and Environmental Medicine, University Hospital of Ludwig Maximilians University, Munich, Comprehensive Pneumology Centre Munich, German Centre for Lung Research, Germany;*
10. *Department of Medical Sciences: Respiratory, Allergy and Sleep research, Uppsala University, Sweden;*
11. *Centre for International Health, Department of Global Public Health and Primary Care, University of Bergen, Norway;*
12. *Inserm UMR 1152. Pathophysiology and Epidemiology of Respiratory Diseases; University Paris Diderot Paris 7, UMR 1152, F-75890, France;*
13. *Respiratory Department, Galdakao Hospital, OSI Barrualde-Galdakao, Biscay, Spain;*
14. *Department of Occupational Medicine, Haukeland University Hospital, Bergen, Norway;*
15. *Population Health & Occupational Disease, National Heart and Lung Institute, Imperial College London, UK;*
16. *MRC-PHE Centre for Environment and Health, Imperial College London, UK.*

Introduction

It has been debated but not yet established whether increased airway responsiveness can predict chronic obstructive pulmonary disease (COPD) [***1***]. Recognising this link may help in identifying subjects at risk.
To study prospectively whether airway responsiveness is associated with COPD risk.

Methods

Within the EU-funded ALEC study (Horizon 2020, GA #633212, https://www.alecstudy.org/), we pooled data from two multicentre cohort studies, ECRHS and SAPALDIA, which collected information at three time points using similar methods [***2, 3***]. Subjects aged 37 years in median (1^{st}–3^{rd} quartiles: 29–43) had airway responsiveness measured by methacholine challenge tests at the 1^{st} examination (1991–1994). We classified these subjects by quintiles of methacholine dose-response slope [1]. We excluded prevalent COPD cases at the 2^{nd} examination (1999–2003) in order to minimise reverse causation, and then analysed incidence rates of COPD (post-bronchodilator FEV_1/FVC < lower limit of normal [***4***]) over the follow-up time between the 2^{nd} and the 3^{rd} examination (2010–2014). Association estimates were adjusted for study, sex, education, FEV_1 predicted, age, BMI, asthma, active smoking, passive smoking and occupational exposures. We conducted stratified analyses by active smoking and asthma to explore effect modification.

Results

There were 119 new cases of COPD among 4817 subjects over a median follow-up of nine years. We found dose-response relationships between airway responsiveness and COPD risk. Crude incidence rates ranged between 0.7 and 5.0 per 1000/year for the groups with lowest and highest airway responsiveness respectively ($p<0.001$). Compared to the least responsive group, adjusted incidence rate ratios ranged from 1.40 (95%CI: 0.54–3.66) to 6.15 (95%CI: 2.88–13.13) for increasing airway responsiveness. Dose-response relationships were more evident among smokers and subjects without asthma compared to their counterparts.

Conclusions

Increased airway responsiveness is an independent risk factor for COPD even in subjects with no history of asthma. Young adults who smoke and have high airway responsiveness may particularly benefit from targeted smoking cessation advice. It is still unclear whether early treatment of COPD in patients with high airway responsiveness can slow down disease progression.

References

[1] Marcon A, Cerveri I, Wjst M, et al. Can an airway challenge test predict respiratory diseases? A population-based international study. J Allergy Clin Immunol 2014;133:104-10.

[2] Burney PG, Luczynska C, Chinn S, Jarvis D. The European Community Respiratory Health Survey. Eur Respir J 1994;7:954-60.

[3] Martin BW, Ackermann-Liebrich U, Leuenberger P, et al. SAPALDIA: methods and participation in the cross-sectional part of the Swiss Study on Air Pollution and Lung Diseases in Adults. Soz Praventivmed 1997;42:67-84.

[4] Quanjer PH, Stanojevic S, Cole TJ, et al. Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. Eur Respir J 2012;40:1324-43.

DETERMINANTI DELLA BUONA SALUTE IN UNA COLLETTIVITÀ DI ULTRANOVANTENNI DEL MUGELLO

Pasqualetti Patrizio[1], Strozza Cosmo[2], Simonelli Ilaria[1], Giordani Alessandro[1], Padua Luca[3], Egidi Viviana[2]

1. *Servizio di statistica medica in Information Technology Fondazione Fatebenefratelli per la ricerca e la formazione sanitaria e sociale, Roma*
2. *Dipartimento Scienze Statistiche Università Sapienza, Roma*
3. *Dipartimento di Geriatria, Neuroscienze, Ortopedia, Università Cattolica del Sacro Cuore, Roma e Fondazione Don Carlo Gnocchi, Milano*

Introduzione

La popolazione mondiale sta progressivamente invecchiando: questo fenomeno continuerà nei prossimi decenni con tempi e intensità differenti tra le aree del pianeta e tra paesi appartenenti alla stessa regione geografica e/o economica. Le principali ragioni di invecchiamento sono l'aumento dell'aspettativa di vita e dunque della sopravvivenza e il calo delle nascite negli ultimi decenni. La combinazione di questi due fenomeni ha fatto sì che il numero assoluto e la quota di persone anziane (>=65 anni) e dei grandi vecchi (>=80 anni) siano cresciuti negli ultimi anni e continuino a farlo. Ha quindi guadagnato interesse lo studio delle condizioni di salute della popolazione in età avanzata, in particolare quella dei grandi vecchi. Questi ultimi secondo le stime ONU, dovrebbero più che triplicare entro i prossimi 40-45 anni.
Il Mugello Study [***1***] è nato proprio allo scopo di valutare le condizioni di salute di un collettivo (504 casi) di ultranovantenni residenti nel "Mugello", nel nord della Toscana. Mediante un dettagliato questionario e una visita di uno specialista nel luogo di residenza degli anziani, è stato possibile collezionare molte informazioni. Sono stati rilevati diversi aspetti della salute: la salute percepita, la presenza di depressione, la capacità di svolgere delle attività della vita di tutti i giorni e il livello cognitivo.
Si è deciso di ricercare i possibili driver della buona salute in tre step.
In primo luogo, si è investigato sul ruolo rivestito dalle principali caratteristiche demografiche e socio-economiche (età, sesso, stato civile, istruzione, lavoro svolto in passato e condizione abitativa). Successivamente è stata condotta un'analisi per valutare se alcune caratteristiche, come la qualità del sonno, la presenza di malattie croniche gravi e la forza nella stretta di mano (nota in letteratura per essere una proxy della buona condizione fisica), potessero essere considerati dei predittori dello stato fisico e mentale. Tali relazioni sono state esplorate a parità delle caratteristiche demografiche e socio-economiche e solamente sugli individui in condizioni di salute non eccessivamente gravi. Allo stesso modo si è investigato sull'associazione tra i comportamenti a rischio per la salute - fumo, alcol, dieta, indice di massa corporea o attività fisica - e le diverse dimensioni della salute a parità delle caratteristiche demografiche, socio-economiche e degli aspetti maggiormente legati alla salute. L'analisi delle variabili riguardanti i comportamenti a rischio, trattandosi di uno studio trasversale, è stata necessariamente diversa rispetto alle altre caratteristiche analizzate, non potendo escludere che esse siano più frequentemente conseguenze piuttosto che cause dello stato di salute.
In ultimo, si è deciso di verificare la possibile esistenza di classi latenti dettate dalle condizioni di salute, cercando di capire quali siano le principali caratteristiche demografiche e socio-economiche dei gruppi eventualmente emersi.

Metodi

Preliminarmente all'analisi dei dati è stata condotta l'imputazione dei missing values. Tra le diverse tecniche (media e distribuzione non condizionata, media e distribuzione condizionata, hot deck e k-Nearest

Neighbour) si è deciso di utilizzare l'ultima citata. Questo tipo di imputazione si basa sulla donazione di un valore già osservato all'unità che presenta il dato mancante. Il valore imputato è il risultato di un'aggregazione dei valori dei k vicini considerati. La distanza per definire quali siano i "vicini più vicini" può essere calcolata per ogni tipo di variabile. Si riporta la formulazione generale:

$$d_{i,j} = \frac{\sum_{k=1}^{p} \omega_k \delta_{i,j,k}}{\sum_{k=1}^{p} \omega_k}$$

In cui ω_k sono i pesi eventualmente assegnati alle singole variabili e $\delta_{i,j,k}$ è il contributo della k-esima variabile alla distanza tra i e j. Il calcolo di quest'ultimo varia a seconda delle variabili considerate. Per le variabili continue il caso è il seguente, in cui $x_{i,k}$ rappresenta il valore dell'i-esima osservazione della k-esima variabile e r_k il range della k-esima variabile.

$$\delta_{i,j,k} = |x_{i,k} - x_{j,k}|/r_k$$

Per quanto riguarda le variabili binarie o nominali la formulazione è invece la seguente:

$$\delta_{i,j,k} = \begin{cases} 0, se\ x_{i,k} = x_{j,k} \\ 1, se\ x_{i,k} \neq x_{j,k} \end{cases}$$

Tra le varia funzioni di aggregazione conosciute, si sono utilizzate le più comuni quali rispettivamente la mediana, per l'imputazione dei valori continui, e il valore più frequente, per l'imputazione di quelli categoriali. Successivamente è stata condotta l'analisi descrittiva e bivariata sfruttando il test del chi quadrato e il t-test appropriatamente per tipo di variabile. Di seguito l'analisi multivariabile è stata condotta utilizzando regressioni di tipo lineare multipla, logistica binaria e logistica multinomiale in base al tipo di target in analisi. Infine è stata utilizzata la LCA (Latent Class Analysis), metodologia statistica volta a individuare un insieme di classi latenti, discrete ed esclusive, alle quali appartengono gli individui oggetto di studio, a partire dai valori che questi ultimi presentano in un insieme di variabili categoriali appositamente scelte per questa analisi.
Nella LCA tradizionale si stimano due insiemi di parametri, in primo luogo la probabilità di appartenenza alle classi per ognuna delle unità oggetto di studio e poi la probabilità di ottenere un determinato valore di un item condizionatamente alla classe di appartenenza. Essendo un modello di misura, l'LCA stima e rimuove la misura dell'errore dal vettore delle probabilità di appartenenza ad una classe latente.
Al fine di descrivere la metodologia statistica che caratterizza la LCA, si devono introdurre: γ, che rappresenta la probabilità di appartenere ad una classe latente e ρ ad indicare la probabilità di rispondere con un determinato item condizionatamente alla classe latente di appartenenza. Nel momento in cui si considerano delle covariate nel modello, si stima un insieme addizionale di parametri β che rappresentano i coefficienti della regressione logistica attraverso la quale si stimano le classi latenti. In questo caso particolare, ρ e β saranno stimati, mentre γ ricavato come funzione delle covariate sfruttando i coefficienti β.
Supponendo un modello a classi latenti con *C* classi da stimare a partire da m variabili categoriali, una covariata *x* e una variabile di raggruppamento *g*. Il vettore $Y_i = (Y_{i1}, \dots, Y_{iM})$ rappresenta le risposte di ogni individuo i alla M-esima domanda. Siano inoltre $c_i = 1, 2, \dots, C$ le classi latenti in questione e $I(y = k)$la funzione indicatrice che assume valore 1 se la risposta y equivale a k, 0 altrimenti. Supponendo inoltre che g_i rappresenta il valore del gruppo di appartenenza dell'i-esima unità, x_i il valore della covariata per l'unità i. Allora il modello a classi latenti può essere come segue:

$$P(Y = y|x_i, g) = \sum_{c=1}^{C} \gamma_{c|g}(x_i) \prod_{m=1}^{M} \prod_{k=1}^{r_m} \rho_{mk|cg}^{I(y_{im}=k)}$$

Dove $\gamma_{c|g}(x_i) = P(C_i = c|x_i, G_i = g)$ è una categoria di riferimento standard della regressione logistica multinomiale. Se ad esempio si considera una sola covariata x, il parametro γ si esprime come segue:

$$\gamma_{c|g}(x_i) = \mathrm{P}(C_i = c|x_i, G_i = g) = \frac{\exp\{\beta_{0c|g} + x_1\beta_{1c|g}\}}{1 + \sum_{j=1}^{C-1} \exp\{\beta_{0c|g} + x_1\beta_{1c|g}\}}$$

con $c = 1, 2, \dots, C - 1$ considerando la classe C come riferimento nella regressione logistica. A partire da questa equazione è possibile ricavare gli odds-ratio che un individuo appartenga ad una classe latente. La formulazione seguente si riferisce ad un modello con due classi latenti 1 e 2:

$$\log\left(\frac{\gamma_{1|1}(x_i)}{\gamma_{2|1}(x_i)}\right) = \beta_{01|1} + \beta_{11|1}x_i$$

L'esponenziale del coefficiente β è proprio l'odds-ratio che rappresenta l'incremento in odds di appartenenza ad una classe in corrispondenza di un incremento espresso in unità del valore della covariata.

Risultati

Gli ultranovantenni oggetto di studio si sono rivelati complessivamente in buono stato di salute, almeno coloro per i quali è stato possibile condurre tutte le prove programmate. Infatti, 110 dei 504 ultranovantenni selezionati sono risultati non testabili principalmente a causa delle loro pessime condizioni mentali. Se si considera il collettivo nel suo insieme, cioè comprensivo dei non testabili, gli uomini possono vantare migliori condizioni in ogni aspetto della salute considerato. Più di metà dell'intero collettivo e quasi i due terzi degli uomini ha dichiarato di essere in condizioni di salute dal buono all'eccellente. Inoltre, quasi metà del totale degli anziani intervistati è risultato non depresso: sono le donne principalmente ad esserlo, visto che solo una porzione inferiore al 40% non ha presentato questo disturbo. Stesso discorso vale per le capacità nello svolgere le attività della vita quotidiana: il 60% degli uomini dell'intero collettivo intervistato è in grado di compierle tutte in maniera autonoma, contro meno del 43% delle donne. Anche la condizione cognitiva ha mostrato questa netta distinzione: non presenta problematiche cognitive più del 60% degli uomini, mentre quasi due terzi delle donne ne risulta affetta per lo meno in forma lieve. Anche i punteggi ottenuti negli indicatori di salute fisica e mentale calcolati a partire da quanto dichiarato nel questionario SF-12 confermano quanto osservato per gli altri aspetti della salute: gli uomini ottengono mediamente tra i due e i 2 e 2,5 punti in più delle donne. È parso, dunque, interessante analizzare le determinanti della buona salute mostrata dagli anziani del Mugello, verificando tra l'altro se a parità delle altre caratteristiche permangono differenze di genere. In altri termini, si è inteso anche capire se a parità di altre caratteristiche (in particolare di età) permanga o meno lo svantaggio femminile osservato nei dati appena descritti che sono ovviamente al lordo dei confondenti.

Sesso ed età sono inserite in tutti i modelli riportati nei risultati: mentre il contributo del primo è completamente assorbito dalle altre covariate in analisi, l'età risulta spesso tra le maggiori determinanti delle peggiori condizioni di salute. Infatti, coloro che sono risultati non testabili, cioè ai quali non è stato possibile somministrare alcune parti del questionario principalmente per le loro pessime condizioni mentali, presentano un età media maggiore rispetto al resto del collettivo. Allo stesso tempo, avere anche un solo anno in più è risultato determinante per riscontrare peggiore salute sia dal punto di vista mentale che fisico. L'istruzione è probabilmente la variabile che più determina le diverse condizioni di salute, come risultato dalla rassegna bibliografica sull'argomento [***3***] [***4***]. Essere più istruiti porta a percepire e dichiarare una

migliore condizione di salute rispetto a chi ha frequentato meno anni di scuola. Inoltre, essere andati oltre la quinta elementare risulta essere un fattore di protezione per la salute mentale, anche se sembra associato anche a più frequenti condizioni di depressione. L'attività svolta durante la propria vita lavorativa è, invece, risultata una determinante meno incisiva. Si è comunque osservato che coloro i quali hanno svolto un lavoro d'ufficio o nel settore dei servizi sono in migliori condizioni di salute generali. Anche la condizione abitativa è stata inserita in questo gruppo di variabili, riscontrando la forte connessione tra cattiva salute e due specifiche situazioni: coesistenza con estranei (presumibilmente badanti e/o assistenti sanitari) o presenza in strutture di riposo. Non è possibile distinguere se un eventuale peggioramento della salute sia avvenuto prima o dopo il trasferimento in casa di riposo poiché lo studio è di tipo trasversale. In altri termini non è possibile dire se la necessità di ricevere cure più specifiche abbia determinato il ricovero o se, viceversa, sia stata la presenza nella struttura assistenziale a peggiorare la situazione (quantomeno psicologica) dell'anziano. Questo è quanto emerge però dai risultati dei modelli con i quali è stata analizzata la salute funzionale e mentale per l'intero collettivo e la salute percepita e la depressione per il sottoinsieme delle persone non testabili e in cattive condizioni di salute (cioè rispettivamente con pessima salute percepita e in stato di depressione). Non sembra determinare invece la presenza o assenza di depressione il luogo in cui si vive.

Il secondo gruppo di determinanti è composto dalle variabili maggiormente legate alle condizioni di salute. Delle più note in letteratura si è deciso di considerare: la qualità del sonno, la presenza di malattie croniche gravi e la forza nella stretta di mano. Anche in questo caso, la terminologia "fattore di rischio" o "fattore di protezione" va utilizzata con la massima cautela non potendo escludere inversioni del nesso causale. A sostegno dei risultati ottenuti vi sono tuttavia numerosi studi longitudinali, sia circa qualità del sonno e numero di patologie [***5***] che per la forza nella stretta di mano [***6***]. La qualità del sonno sembra determinare in senso positivo le condizioni di salute, sia fisica che mentale, degli ultranovantenni del Mugello. Chi dorme bene ha molta più probabilità di non essere depresso o di riuscire a svolgere le attività di vita quotidiana rispetto a chi dorme male. Ciò inoltre porta a dichiarare una migliore salute percepita ma non è significativo nel determinare le buone condizioni cognitive. Il numero di patologie croniche gravi di cui sono affetti gli anziani sembra invece determinare proprio le condizioni di salute cognitive, così come la capacità di svolgere le attività della vita quotidiana. Quanto emerso dalle analisi è un'ulteriore conferma dei riscontri empirici riportati in letteratura. Non è stato invece possibile confermare il ruolo della presenza o meno delle patologie croniche per la depressione o per la salute percepita. La forza nella stretta di mano è nota in letteratura per essere una proxy della buona condizione fisica. Anche nelle analisi condotte è risultata associata con tutte le variabili di salute considerate, eccezion fatta per l'indicatore sintetico mentale (MCS) e per la percezione della propria salute.

L'ultimo gruppo di determinanti considerate è quello delle variabili che illustrano i comportamenti a rischio per la salute. Mentre il fumo (attuale o pregresso) non è mai risultato una determinante delle condizioni degli ultranovantenni del Mugello, il minor consumo di alcol sembra portare a dichiarare una migliore condizione di salute. Per quanto riguarda l'aderenza alla dieta mediterranea, nel presente studio non vi sono state evidenze a riguardo. Probabilmente la collocazione geografica della collettività indagata, caratterizzata da ottime abitudini alimentari sia attualmente, sia nel passato, la rende omogenea rispetto a questa variabile. L'indice di massa corporea (BMI), invece, ha confermato quanto già riportato in letteratura. A parità delle altre condizioni, le persone sovrappeso si sentono meglio e stanno meglio (in questo secondo caso anche gli individui obesi) rispetto ai normopeso dal punto di vista mentale [***7***], allo stesso tempo è emerso che sono in condizioni peggiori per quanto riguarda la capacità di svolgere le attività della vita di tutti i giorni [***8***]. Infine, l'attività fisica è probabilmente il fattore risultato maggiormente associato alle buone condizioni di salute. Chiaramente, chi riesce a svolgere tali attività anche a età così avanzate è in buone condizioni di salute. Nella letteratura internazionale ci sono diversi studi che considerano questa variabile come una

determinante importante della salute, ma solo quando si dispone di indagini longitudinali è possibile verificarne il ruolo.
La ricerca di classi latenti è stata effettuata sia utilizzando i soli aspetti della salute misurati sull'intero collettivo, sia quando si sono aggiunte le informazioni rilevate esclusivamente su una parte degli intervistati (escludendo i cosiddetti soggetti non testabili perché non in grado di rispondere ad alcune domande),. Si sono ottenute quattro classi latenti chiaramente distinte e in grado di confermare molti dei risultati emersi nelle regressioni lineari e logistiche proposte. Infatti, in entrambe le analisi sono risultati due gruppi simili: uno di chi si dichiara ed è in buone condizioni complessive di salute (il terzo dell'analisi sull'intera collettività e il quarto dell'analisi sul gruppo dei soli testabili), che raccoglie il 32,7% degli anziani nella prima analisi e il 44,2% nella seconda, e un altro di chi si dichiara ed è in pessime condizioni di salute che rappresenta il 22,2% e il 29,4% rispettivamente dell'intero collettivo e dei soli testabili. Chi sta meglio ha un'età media più bassa e una minore proporzione di donne. Dall'analisi sull'intero collettivo di anziani si è ottenuto, inoltre, un interessante gruppo di anziani che non hanno dichiarato di essere in cattive condizioni pur essendolo. Si tratta di persone più anziane, con una prevalenza femminile, più spesso vedove e poco istruite che testimoniano l'atteggiamento più volte studiato in letteratura del relativo ottimismo con il quale all'avanzare dell'età e, spesso, al peggioramento delle condizioni di salute oggettive, i molto anziani valutano la propria salute [***9***]. Altro gruppo interessante è quello costituito da individui depressi sebbene in condizioni di salute relativamente buone. In questo caso la predominanza femminile risulta minore e l'età più giovane, si tratta però di persone che vivono da sole più spesso di quanto non succeda nell'intero collettivo.

Conclusioni

In base ai molteplici indicatori considerati risulta che gli ultranovantenni del Mugello sono prevalentemente in buona salute.
Le caratteristiche demografiche e socio-economiche, in particolare età, stato civile e istruzione risultano spesso determinanti delle condizioni di salute. Talvolta risulta rilevante anche l'associazione, in particolar modo della salute mentale e funzionale, con le variabili sanitarie: qualità del sonno e presenza di patologie croniche gravi. Inoltre tra di esse risalta l'associazione tra forza nella stretta di mano e salute mentale o depressione. Infine tra gli health behaviours e le condizioni di salute l'associazione si manifesta principalmente in termini di BMI e svolgimento di attività fisica.
La LCA ha sottolineato la multidimensionalità del concetto di salute, rivelando due gruppi di particolare interesse: uno in cattive condizioni di salute oggettiva ma che dichiara di star bene ed un altro in buone condizioni di salute oggettiva ma in stato depressivo.

Bibliografia

[**1**] Molino-Lova R, Sofi F, Pasquini G, et al. The Mugello study, a survey of nonagenarians living in Tuscany: design, methods and participants' general characteristics. Eur J Intern Med 2013; 24(8):745-9.
[**2**] Tigani X, Artemiadis A, Alexopoulos E, Chrousos G, Darviri C. Self-rated health in centenarians: A nation-wide cross-sectional Greek study. Arch Gerontol Geriatr 2012; 54(3):342-8,.
[**3**] Lantz P, Lynch J, House J, et al. Socioeconomic disparities in health change in a longitudinal study of US adults: the role of health-risk behaviors. Soc Sci Med 2001; 53(1):29-40.
[**4**] Dartigues J, Gagnon M, Letenneur L, et al. Principal lifetime occupation and cognitive impairment in a French elderly cohort (Paquid). Am J Epidemiol 1992;135(9):981-8.
[**5**] M. Cole e N. Dendukuri, «Risk factors for depression among elderly community subjects: a systematic review and metaanalysis,» Am J Psychiatry, 1147-1156, 2003.
[**6**] Silva N, Menezes TN. The association between cognition and handgrip strength among the elderly: an integrative review. Cien Saude Colet 2016;21(11):3611-36.

[**7**] T. Sobów, W. Fendler e R. Magierski, «Body mass index and mild cognitive impairment-to-dementia progression in 24 months: a prospective study,» Eur J Clin Nutr, 1216-1219, 2014.

[**8**] Freedman V, Martin L, Schoeni R, Cornman J. Decline in late-life disability: the role of early- and mid-life factors. Soc Sci Med 2008;66(7):1588-602.

[**9**] Golini N, Egidi V. The Latent Dimensions of Poor Self-Rated Health: How Chronic Diseases, Functional and Emotional Dimensions Interact Influencing Self-Rated Health in Italian Elderly. Social Indicators Research 2016;128(1):321-39.

A MEDIATION ANALYSIS OF THE IMPACT OF THEORY OF MIND ON SOCIAL INFORMATION PROCESSING IN CHILDREN WITH AUTISM SPECTRUM DISORDERS

Masedu Francesco[1], Mazza Monica[1], Mariano Melania[1], Peretti Sara[1], Pino Maria Chiara[1], Valenti Marco[1,2]

1 *Department of Applied Clinical Sciences and Biotechnology, University of L'Aquila, L'Aquila, Italy*
2 *Reference Regional Centre for Autism, Abruzzo Region, Health System, L'Aquila, Italy*

Introduction

Children with autism spectrum disorders (ASD) are characterized by a range of deficits in two domains: social communication and social interaction, and repetitive patterns of behavior [***1***]. In addition, children with ASD show deficit in social cognition abilities [***2***]. Social cognition includes the cognitive mechanisms based on the capacity to process the social world and emerges in early childhood through the development of "theory of mind" [***3***].

Most researchers have suggested that the ability to understand another person's cognitive or affective state plays an important role in the development and production of social behavior [***4***]. Several studies [***5***] have suggested that ToM is not a single process but involves two components: cognitive and affective. Specifically, the cognitive component is the ability to understand what others are thinking (i.e., their mental states); whereas affective ToM is the ability to make inferences regarding other people's emotions [***6***].

Ziv and Sorongon (2011) [***7***] proposed the Social Information Processing Interview (SIPI) to evaluate patterns of social information processing and social behavior of preschool children. The SIPI model suggests that the covert mental mechanisms mediating between an overt (social) stimulus and overt (social) response comprise the abilities to understand the thoughts of others and to select the appropriate social responses. According to this model, many mental steps take place before individuals enact a behavioral response to social cues, such as the encoding of social cues, interpretation of the cue, clarification of goals, generation of a behavioral response, response construction, response decision, and enactment of the behavioral response.

The ToM difficulties that characterize children with ASD have been found to be related to inadequate social behavior and poor social communication skills [***8***].

On this basis, the first goal of the present study was to evaluate the ToM sub-components, that is, beliefs, emotions, and intentions by using the Comic Strip Task [***9***] and social information processing abilities through the SIPI in children with ASD and to compare them to TD children.

In previous studies [***10***], the participants were preschoolers. There is evidence supporting a delay in the development of language in children with ASD [***11***]. However, a consistent relationship between language development and ToM has been shown. For this reason, our sample of children (both TD and with ASD) covered an age range from 5 to 13 years, because both tasks (CST and SIPI) require well-developed expressive and receptive language skills.

The second goal was to examine the role and effect of ToM on the social information processing patterns using a new analysis approach, that is, a mediation model. Our hypothesis was that a deficit of ToM could adversely affect the ability of social information processing and subsequently compromise the development of adequate social behavior. Therefore, we used the mediation analysis to examine the impact of different components of ToM on the development of social behavior, first in TD children and then in children with ASD. The mediation analysis is based on the assumption that a third variable (the mediator) can influence the relationship between the other two variables, that is, independent variable (IV, X) and dependent variable (DV, Y). The mediational effect, in which X leads to Y through M, is called the indirect effect, which represents the portion of the relationship between X and Y that is mediated by M [***12***]. In our study, we used

the ToM sub-components (intentions, beliefs, and emotions) as mediators and all components of the SIPI (encoding, interpretation, response construction, and response evaluation) both as independent and dependent variables to perform our mediation models.

Methods

One-hundred and seven children participated in our study: 52 children (42 male, 10 female) with ASD ranging in age from 5 to 13 years and 55 TD, age-matched children (37 male; 18 female) ranging in age from 5 to 12.25 years. No differences between groups (ASD and TD) emerged for chronologic age ($F_{1,102}$ = 3.03; p = 0.095) and mental age ($F_{1,102}$ = 1.85; p = 0.177). The TD children were selected from a nursery (for children of 5 years), a primary school (for children from 6 to 10 years), and a lower secondary school (for children from 11 to 13 years) located in Frosinone, central Italy.

Participants with ASD were inpatients of the Reference Regional Centre for Autism of L'Aquila, central Italy. According to the principles established by the Declaration of Helsinki, ethical approval was obtained by the Ethical Committee of the NHS Local Health Unit (Azienda Sanitaria Locale 1). The children's parents provided informed consent to participate in the study. Diagnosis of ASD was established by experienced clinicians using the Autism Diagnostic Observation Schedule, Second Edition [***13***], according to the new criteria of the DSM-5 (American Psychiatric Association 2013). Verbal mental age was assessed with the Test for Reception of Grammar, Version 2 [***14***].

Social Cognition Measures

Social Information Processing Interview [***7***]

The SIPI is an interview based on a storybook depicting a series of vignettes in which a protagonist is either rejected by two other peers or provoked by another peer.

According to Ziv and his collaborators [***15***] (2014), the scores correspond to four of the social information processing mental steps proposed by Crick and Dodge's (1994) model [***16***]: (1) encoding, (2) interpretation of cues, (3) response construction, and (4) response evaluation.

The encoding component evaluates the level of detail that the child recalls across the four stories.

The interpretation component evaluates the hostile attribution to others' behavior.

The response construction score is derived from the child's responses to the open-ended question: "What would you say or do if this happened to you?" For each story, the examiner encodes the response as competent or non-competent and assigns a score of 1 if the child's response is classified as competent and score of 0 if the answer is classified as non-competent.

The response evaluation items examine the way the child assesses the behavior of other people as right or wrong.

Comic Strip Task [***9***]

The CST is a novel 21-item measure developed to assess three aspects of ToM; namely, the understanding of others' beliefs, intentions, and emotional states, and comprises the three sub-components testing belief-understanding (beliefs), intention-understanding (intentions), and emotion-understanding (emotions), respectively.

There are 5 items in each sub-component, each comprising a 5-picture comic strip illustrating everyday social scenarios involving interpersonal interactions familiar to young children. Each sub-component has a maximum score of 5, with a total test score of 15 (i.e., higher scores correspond to better ToM).

Sivaratnam and collaborators assessed the internal consistency reliability, as measured by Cronbach's alpha: the beliefs sub-component demonstrated low and negative internal consistency (Cronbach's $\alpha = -0.04$); the emotions and intentions sub-components demonstrated moderate internal consistency (Cronbach's $\alpha = 0.69$ and $\alpha = 0.70$, respectively). In their study, Sivaratnam et al. [***10***] found a lower internal consistency; for this reason, we calculated the internal consistency of the overall CST and each sub-component (intentions 0.56, beliefs 0.66, emotions 0.23).

Data Analysis

ANOVA Analysis

One-way ANOVA was used to test differences between groups (ASD, TD) regarding demographic parameters and in the component measures of the SIPI (encoding, interpretation, response construction, and response evaluation) and the CST (emotions, beliefs, and intentions).

Correlations Analysis

Exploratory Pearson's correlations were computed to assess the relationships between the components of the SIPI (encoding, interpretation, response construction, and response evaluation) and CST (emotions, beliefs, and intentions), both in the ASD and TD group. On the basis of the correlation results, the mediation models were performed.

Regression Analysis

A linear regression analysis was performed in order to evaluate the relations between the dependent variable and one or more explanatory variables. In this study, the variables for the regression analysis were the components of the SIPI and the sub-components of CST.

Mediation analysis

The mediation model is important for understanding the mechanism through which the causal variable affects the outcome [***17***].

In the mediation process, the relationship between the independent variable (X) and the dependent variable (Y) is hypothesized to be an indirect effect that exists due to the influence of a third variable (M, the mediator). According to Baron and Kenny's suggestion, we followed a four-step approach, in which several linear regression analyses are conducted, and the significance of the coefficients is examined at each step (***Figure 1A, B***).

In this study, we found full mediation for each mediation model. Finally, we used the Sobel test [***18***] to explore the significance of a mediation effect. Particularly, in our study, the Sobel test was performed to evaluate whether the social information processing ability, measured by the SIPI, was mediated by the ToM abilities measured by the CTS in both the ASD and TD group. The indirect effect of X on Y through M can then be quantified as the product of a and b (i.e., ab). The total effect of X on Y is quantified with the unstandardized regression weight c (***Fig. 1a***). The Sobel test involves computing the ratio of ab to its estimated standard error (SE). The overall statistical significance of the model was set at the 0.001 level. The Statistical Package for the Social Sciences (SPSS) software (version 22; SPSS Inc., Chicago, IL, USA) was used.

Results

The ANOVA of the SIPI components showed that the ASD group had lower scores compared to the TD group with regard to social information processing competences. Specifically, children with ASD showed difficulties in all components of the SIPI: encoding ($F_{1,102} = 26.96$; $p = 0.0001$), interpretation ($F_{1,102} = 14.85$; $p = 0.0001$), response evaluation ($F_{1,102} = 78.80$; $p = 0.0001$), and response construction ($F_{1,102} = 86.80$; $p = 0.0001$), compared to the TD children.

Furthermore, the ASD group had lower scores in beliefs ($F_{1,102} = 44.84$; $p = 0.0001$) and emotions ($F_{1,102} = 13.02$; $p = 0.0001$), but there were no significant differences in the intentions ($F_{1,102} = 0.214$; $p = 0.645$) sub-component of the CST compared to the TD group.

Significant correlations were found in the TD group between the emotions CST sub-component (affective ToM) and three components of the SIPI (interpretation, response construction, and response evaluation). Specifically, we found a significant negative correlation between the emotions sub-component of the CST and the interpretation component of the SIPI ($r = -0.388$; $p = 0.002$).

A significant positive correlation was also found between the CST emotions sub-component and both the response construction ($r = 0.378$; $p = 0.002$) and response evaluation ($r = 0.251$; $p = 0.032$) components of

the SIPI. Furthermore, the three SIPI components (interpretation, response construction, and response evaluation) showed significant correlations with the beliefs sub-component of the CST (cognitive ToM).

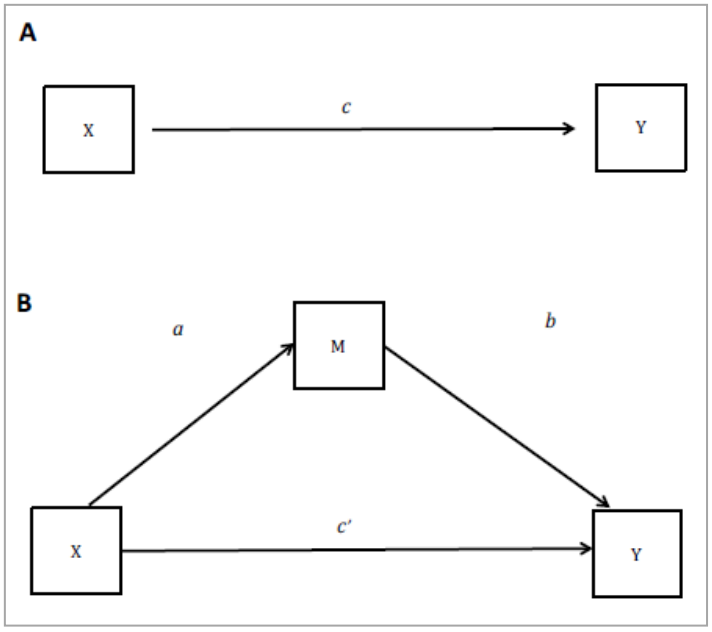

Figure 1. A four-step approach.

A significant positive correlation was found in the ASD group between the emotions sub-component of the CST (affective ToM) and the response construction component of the SIPI ($r = 0.270$; $p = 0.035$). Significant correlations were also found between the intentions sub-component of the CST (cognitive ToM) and the three SIPI components (interpretation, response construction, and response evaluation).
Specifically, the CST intentions sub-component was negatively correlated with the SIPI interpretation component ($r = -0.223$; $p = 0.68$).

Regression Analysis

According to the results obtained by the correlation analysis and in order to create the mediation models, a linear regression analysis was performed. The corresponding results allowed us to conduct the Sobel test to explore the significance of the mediation effects between the different variables.

Mediation Analysis

Mediation models were performed on the basis of the correlation and regression analyses. In order to elucidate whether the interpretation of other people's social behavior is mediated by the ability to understand intentions, beliefs, and emotions of others, four separate mediation models for the TD group and two for the ASD group were conducted.

TD Group

Firstly, the relationship between the interpretation (X) and response construction (Y) components of the SIPI, using the emotions sub-component of the CST (affective ToM) as a mediator, was explored. The Sobel test showed that this model was significant ($Z = -2.07$; $SE = 0.11$; $p = 0.03$).
Secondly, the relationship between the interpretation (X) and response construction (Y) components of the SIPI, using the CST beliefs sub-component (cognitive ToM) as a mediator, was examined. The Sobel test showed that this model was significant ($Z = -2.39$; $SE = 0.13$; $p = 0.01$).

Thirdly, we explored the relationship between the interpretation (X) and response evaluation (Y) components of the SIPI, using the CST emotions sub-component (affective ToM) as a mediator; however, this model did not show any significant indirect effect of mediation.
Finally, we explored the relationship between the interpretation (X) and response evaluation (Y) components of the SIPI, using the CST beliefs sub-component (cognitive ToM) as a mediator; however, this model did not show any significant indirect effect of mediation.
ASD Group
Firstly, we explored the relationship between the interpretation (X) and response construction (Y) components of the SIPI, using the CST intentions sub-component (cognitive ToM) as a mediator. Secondly, we explored the relationship between the interpretation (X) and response evaluation (Y) components of the SIPI, using the CST intentions subcomponent (cognitive ToM) as a mediator. The Sobel test for both models did not show any significant indirect effect of mediation.

Conclusions
The main aims of the present study were to evaluate the differences between the two groups (TD and ASD) in the used tests and to examine the role of ToM abilities in the development of social information processing capacities through mediation analysis, which is an innovative methodology to evaluate the impact of ToM components on social skills. The ANOVA analysis showed that the ASD group had difficulties in social information processes and ToM abilities compared to the TD group. Specifically, in the SIPI, children with ASD had lower scores in encoding social information efficiently (i.e., encoding component), that is, they showed difficulties in the understanding of the story. In this task, the examiner asked the children to narrate the story that they had heard. The encoding component represents the first step of social information processing because it requires the ability to understand and reproduce what other people say or narrate. According to the Crick and Dodge (1994) model [***16***], which was subsequently used by Ziv and his collaborators (2014) [***15***], another step for correct social information processing is response construction.
The children with ASD also had difficulties in response construction. In addition, they displayed impairment in response evaluation, particularly in the ability to evaluate whether other people's social behavior was right or wrong (e.g., they evaluated an aggressive behavior as being acceptable). The response evaluation component is crucial for processing social cues and to subsequently guide social behavior. An additional step is represented by the interpretation component of the SIPI. This component evaluates the tendency to attribute hostile intentions to other people in positive social situations and vice versa. The ASD group showed difficulties in this component compared to TD children.

References
[**1**] American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders: DSM-V, 5th edn. Arlington,VA: American Psychiatric Publishing, 2013.
[**2**] Mazza M, Pino MC, Mariano M, et al. Affective and cognitive empathy in adolescents with autism spectrum disorder. Front Hum Neurosci 2014;8:791-7.
[**3**] Astington JW, Edward MJ. The Development of Theory of Mind in Early Childhood Janet Wilde. Encyclopedia on early childhoon development. 2010.
[**4**] Batson CD. Prosocial motivation: Is it ever truly altruistic? In: Berkowitz L, editor. Advances in experimental social psychology, New York: Academic Press, 1987:65-122.
[**5**] Baron-Cohen S, Bowen DC, Holt RJ, et al. The "Reading the Mind in the Eyes" Test: Complete Absence of Typical Sex Difference in ~ 400 Men and Women with Autism. PLOS ONE 2015;10(8):e0136521.
[**6**] Franco F, Itakura S, Pomorska K, Abramowski A., Nikaido K, Dimitriou D. Can children with autism read emotions from the eyes? The eyes test revisited. Res Dev Disabil 2014;3(5):1015-26.

[**7**] Ziv Y, Sorongon A. Social Information Processing in Preschool Children: Relations to Sociodemographic Risk and Problem Behavior. J Exp Child Psychol 2011; 3(4):412-29.
[**8**] Lerner MD, Hutchins TL, Prelock PA. Brief report: Preliminary evaluation of the theory of mind inventory and its relationship to measures of social skills. J Autism Dev Disord 2011;41:512-7.
[**9**] Cornish K, Rinehart N, Gray K, Howlin P. Comic strip task. Melbourne: Monash University Developmental Neuroscience and Genetic Disorders Laboratory and Monash University Centre for Developmental Psychiatry and Psychology, 2010.
[**10**] Sivaratnam CS, Cornish K, Gray KM, Howlin P, Rinehart NJ. Brief report: Assessment of the social-emotional profile in children with autism spectrum disorders using a novel comic strip task. J Autism Dev Disord 2012;42:2505-12.
[**11**] Ambridge B, Bannard C, Jackson GH. Is grammar spared in autism spectrum disorder? Data from judgments of verb argument structure overgeneralization errors. J Autism Dev Disord 2015;45(10):3288-96.
[**12**] Baron RM, Kenny DA. The moderator–mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol 1986;51:1173-82.
[**13**] Lord C, Rutter M, DiLavore PC, Risi S, Gotham K, Bishop SL. Autism Diagnostic Observation Schedule (ADOS-2): Manual. 2nd ed. Los Angeles: CA: Western Psychological Services, 2012.
[**14**] Bishop, D. The Test for Reception of Grammar (TROG-2)-Version 2. London, United Kingdom: The Psychological Corporation, 2012.
[**15**] Ziv Y, Hadad BS, Khateeb Y, Terkel-Dawer R. Social information processing in preschool children diagnosed with autism spectrum disorder. J Autism Dev Disord 2014;44(4):846-59.
[**16**] Crick NR, & Dodge KA. A review and reformulation of social information: Processing mechanisms in children's social adjustment. Psychological Bulletin 1994;115(1):74-101.
[**17**] Kenny DA, Korchmaros JD, Bolger N. Lower level mediation in multilevel models. Psychol Methods 2003;8:115-28.
[**18**] Sobel ME. Some new results on indirect effects and their standard errors in covariance models. In: Tuma N, editor. Sociological Methodology. Washington D.C.: American Sociological, 1986, 159-86.

CRITICAL MERGING OF TOXICOLOGICAL AND EPIDEMIOLOGICAL DATA: THE EXAMPLE OF PFOA AND PFOS AND FETAL GROWTH

Guercio Valentina[1], Negri Eva[1], Metruccio Francesca[2], Tosti Luca[2], Bonzi Rossella[3], La Vecchia Carlo[3], Moretto Angelo[1,2]

1 *Dipartimento di Scienze Biomediche e Cliniche, Università degli Studi di Milano, Milan, Italy*
2 *ICPS-International Centre for Pesticides and Health Risk Prevention, ASST Fatebenefratelli Sacco, Milan, Italy*
3 *Dipartimento di Scienze Cliniche e di Comunità, Università degli Studi di Milano, Milan, Italy*

Introduction

An integration of evidence from toxicology and epidemiology for improving causal inference and risk assessment of chemicals has been advocated [***1, 2***]. A group of epidemiologists and toxicologists proposed the Epid-Tox framework with the aims to illustrate how epidemiological and toxicological data intersect, provide help in drawing conclusions on causal relationships, and show the influence of potential additional data [***1***]. Perfluoroalkyl acids (PFAA) have been used in a wide variety of industrial and consumer applications. The two most widely used PFAAs, perfluorooctanoic acid (PFOA) and perfluorooctane sulfonic acid (PFOS), are highly persistent in the environment and widely detected in blood samples of the general human population [***3***]. Toxicity studies in experimental animals showed that PFOA and PFOS affect fetal growth and development [***4-6***]. Several epidemiological studies on the relation between PFOA and PFOS exposures and fetal growth, including birth weight (BW), have been published in the last decades.
Integration of toxicological and epidemiological evidence on PFOA and PFOS and BW, by estimating effects separately in animals and humans, and comparing the estimated dose–response relationship.

Methods

We conducted an extensive search for available toxicological information regarding BW in experimental animals (rats, mice) administered PFOA or PFOS in Pubmed/Medline and ToxNet databases. We assessed the quality of toxicological studies using the criteria outlined by Klimisch et al. [***7***]. Only studies falling into category 1 or 2 according to Klimish's criteria were taken into consideration. Data of adequate, reliable, and relevant studies were extracted into an 'ad hoc' created Microsoft Access database. When numerical data were available, a tool to draw dose–response curves was also implemented to facilitate the evaluation of results. Furthermore, to allow a meaningful quantitative comparison between animal and human data, the relationship between PFOA and PFOS concentration in rodent serum and the administered oral dose was derived by linear extrapolation.
For epidemiological evidence we conducted a systematic literature search in Pubmed/Medline and Embase databases up to November 2015. We used eligibility criteria built on the basis of the PICOS (participants/population, intervention/exposure(s), comparator(s)/control, outcome(s), and study design) approach as follows: (i) participants; women enrolled before or during pregnancy or at delivery, (ii) exposure: PFOA and PFOS assessed in a biological sample, such as maternal or umbilical cord serum, plasma or whole blood, or maternal milk; (iii) comparators: newborns exposed to lower levels; (iv) outcomes: BW; (v) study design: cross-sectional, case–control, or cohort study. For statistical analysis, we considered the linear regression coefficient (LRC) of the linear regression of BW on PFOA or PFOS levels, and its standard error (SE). Some studies introduced untransformed PFOA/PFOS levels in the regression model, while others used log-transformed values. We performed two separate meta-analyses for actual and log-transformed values. We chose 1 ng/mL and the natural logarithm (ln) of 1 ng/ mL as measurement unit for the LRC for the untransformed and log-transformed analyses, respectively. For all meta-analyses, we used a random

effects model based on the inverse variance methods and the DerSimonian and Laird estimate of the between-study variance [***8***]. Heterogeneity was quantified by Cochran's χ2 statistic Q and by the I^2 statistics [***9***]. Publication bias was investigated by funnel plots [***10***] and by the tests proposed by Begg and Mazumdar [***11***] and by Egger [***12***]. The trim-and-fill method was also used to investigate the potential effect of publication bias on the pooled estimate.

Results

Toxicological evidence. For PFOA, 12 studies, describing 21 data sets in mice, were included. Birth and fetal weight decreased starting at doses >1 mg/kg body weight, with similar slopes up to the highest dose tested, statistically significant from 5 mg/kg body weight. For PFOS, 13 studies describing 19 datasets were included. In most datasets, birth and or fetal weight decreased following in utero exposure to PFOS, with rats being responsive at lower doses than mice. *Epidemiological evidence.* We identified 16 articles, published from 2007 to 2015 encompassing a wide range of human blood concentrations of PFOA and PFOS. The LRC for untransformed PFOA (12 studies) ranged from -213 to 10 g of BW change for a change of 1 ng/mL in PFOA maternal plasma/serum level. The pooled estimate was -12.8 g (95%CI -23.2;-2.4). Nine studies presented results for the regression of BW on log-transformed PFOA levels. The estimated LRC ranged from -142 to 5 g for an increase of 1 $\log_e$ ng/mL PFOA, i.e. for an increase of approximately 2.7 times in PFOA untransformed levels. The pooled estimate was -27.1 (95% CI -50.6;-3.6). The LRC for untransformed PFOS (8 studies) ranged from -11.3 to 5.8 g and the pooled estimate was -0.92 g (95%CI -3.4;1.6). The estimated LRC for an increase of 1 $\log_e$ ng/mL PFOS ranged from -140 to 66.1 g in 8 studies, and the pooled estimate was -46.1 (95% CI -80.3;-11.9). No consistent pattern emerged according to study location or timing of blood sampling.

The weight of evidence of epidemiology and toxicology was considered in the framework for the integration of toxicology and epidemiology for causal inference and risk assessment, as proposed previously [1]. Taking into consideration all reviewed animal data, the overall toxicological evidence for a dose dependent effect of PFOA on BW is judged plausible (***Figure 1***). Considering epidemiological data for PFOA, there was a significant inverse relationship when untransformed values were considered, although with significant moderate heterogeneity; for log-transformed values, there was a significant inverse relationship, with low heterogeneity. Overall, we evaluated the epidemiological evidence for an inverse association between PFOA maternal blood levels and BW as 'moderately likely'. In humans, the shape of the dose-response curve has not been sufficiently investigated, particularly for what concerns a possible threshold effect. Combining toxicological and epidemiological evidence in a qualitative way, the causal relationship between PFOA exposure and BW falls in the "likely" category (***Figure 1***) [***13***].

As almost all animal studies show a similar dose-response trend, biological plausibility for an inverse effect of PFOS on BW is considered ''likely''. Conversely, for epidemiological evidence, there was a non-significant inverse relationship when untransformed values were considered, with significant heterogeneity, while for log-transformed values, there was a significant inverse relationship, with low heterogeneity. Overall, we evaluated the epidemiological evidence for an inverse association between PFOS maternal blood levels and BW as 'insufficient' tending to 'moderately likely'. A casual relationship, integrating toxicological and epidemiological evidence for the effect of PFOS on BW, is less clear, falling between 'insufficient' and 'likely' (***Figure 1***) [13]. However, this is only a qualitative judgment which does not take into account information on Mode of Action (MoA), including quantitative analysis of the dose–response in animals compared with human exposure.

Conclusions

The epidemiological and toxicological evidence suggests that PFOA and PFOS interfere with intrauterine growth and elicit a decrease in BW in humans and rodents. However, the effect in animals is evident at

extrapolated serum concentrations 10^2-10^3 times higher than those observed in humans. Thus, toxicological evidence supports epidemiological one qualitatively, but not quantitatively. Pharmacokinetic differences between species hamper comparison of rodent and human data.

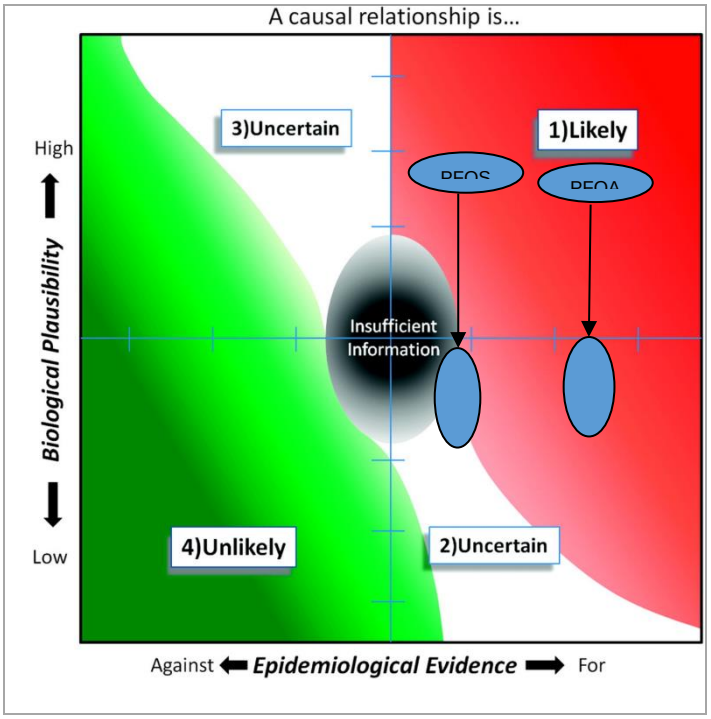

Figure 1. Graphical representation, according to Adami et al. [1,13], of the integration of toxicological and epidemiological evidence for decreased BW after exposure to PFOA and PFOS. The arrow indicates the direction of the quantitative refinement after taking into account the comparative quantitative information.

References

[1] Adami HO, Berry CL, Breckenridge CB, et al. Toxicology and epidemiology: improving the science with a framework for combining toxicological and epidemiological evidence to establish causal inference. Toxicol Sci 2011;122(2):223-34.

[2] Woodruff TJ, Sutton P. The Navigation Guide systematic review methodology: a rigorous and transparent method for translating environmental health science into better health outcomes. Environ Health Perspect 2014;122(10):1007-14.

[3] Kato K, Wong LY, Chen A, et al. Changes in serum concentrations of maternal poly- and perfluoroalkyl substances over the course of pregnancy and predictors of exposure in a multiethnic cohort of Cincinnati, Ohio pregnant women during 2003–2006. Environ Sci Technol 2014;48(16):9600-08.

[4] Lau C, Thibodeaux JR, Hanson RG, et al. Effects of perfluorooctanoic acid exposure during pregnancy in the mouse. Toxicol Sci 2006;90(2):510-18.

[5] Lau C, Thibodeaux JR, Hanson RG, et al. Exposure to perfluorooctane sulfonate during pregnancy in rat and mouse. II: postnatal evaluation. Toxicol Sci 2003;74(2):382-92.

[**6**] Thibodeaux JR, Hanson RG, Rogers JM, et al. Exposure to perfluorooctane sulfonate during pregnancy in rat and mouse. I: maternal and prenatal evaluations. Toxicol Sci 2003;74(2):369-81.
[**7**] Klimisch HJ, Andreae M, Tillmann U. A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. Regul Toxicol Pharmacol 1997;25(1):1-5.
[**8**] DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7(3):177-88.
[**9**] Higgins JP, Thompson SG. Quantifying heterogeneity in a metaanalysis. Stat Med 1992;21(11):1539-58.
[**10**] Sterne JA, Egger M.Funnel plots for detecting bias in meta-analysis:guidelines on choice of axis. J Clin Epidemiol., 2001;54(10):1046-55.
[**11**] Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics 1994;50 (4):1088-101.
[**12**] Egger M, Davey Smith G, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. BMJ, 1997;315(7109):629-34.
[**13**] Negri E, Metruccio F, Guercio V, et al. Exposure to PFOA and PFOS and fetal growth: a critical merging of toxicological and epidemiological data. Crit Rev Toxicol 2017;47(6):482-508.

TIME TRENDS IN FIRST HOSPITALIZATION FOR HEART FAILURE IN A COMMUNITY-BASED POPULATION

Lorenzoni Giulia[1], Azzolina Danila[1], Lanera Corrado[1], Gregori Dario[1], Vannuzzo Diego[2], Baldi Ileana[1]

1 *Unit of Biostatistics, Epidemiology, and Public Health, Department of Cardiac, Thoracic, and Vascular Sciences, University of Padova*
2 *Cardiovascular Prevention Center, ASS 4 and Regional Health Agency of Friuli-Venezia Giulia, Udine, Italy*

Introduction

Heart Failure (HF) is a severe and potentially life-threatening condition frequently encountered in everyday clinical practice. This is caused by a progressive aging of the population, which is in turn related with an increased incidence and severity of HF [***1-3***].
Clearly characterize the epidemiology of HF is essential in order to develop ad hoc prevention strategies and to allow a cost-effective allocation of healthcare resources in the context of heath planning. In recent years, several epidemiological studies have been delivered in this field [***4-6***], but studies on secular trends of HF incidence are scarce [***7***]. The common approach to the study of age, period of diagnosis and cohort of birth effects on HF incidence has not fully explored the separate role of these time dimensions. Few studies have reported the use of combined analysis to untangle the age–period–cohort (APC) effects. The age effect in HF incidence has been well described showing that the risk of HF increases exponentially in the elderly [***8***]. However, period and cohort effects are more difficult to understand separately. Improvements in healthcare over time, particularly the introduction of angiotensin-converting enzyme (ACE) inhibitors, spironolactone and beta-blockers may be seen as period effects, which can modify the time trends of incidence rates. Cohort effects can result from changes in wellbeing between generations.
To obtain a reliable explanation for the time trends of first hospitalization for HF, the APC dimensions should be addressed using a unique analysis that can provide a separation of the individual effects. Using a combined approach of estimating APC effects, the aim of this study is to unravel the separate effects of age, period and cohort on first hospitalizations for HF within a geographically defined community in North-East of Italy, enrolled in the "Martignacco project" [***2***] in 1977 and followed up for 37 years.

Methods

The "Martignacco project", promoted by the World Health Organization (WHO), was started in 1977 with the aim of evaluating the impact of health promotion interventions on cardiovascular health in middle-aged (40-59 years) subjects. Two comparable communities in Friuli Region, North eastern Italy, were selected: Martignacco was chosen as the intervention area and San Giorgio di Nogaro was designated as the control area.
The cohort data for the present analysis refer to the 3,066 subjects (1,324 and 1,742 drawn from the geographical area of Martignacco and San Giorgio di Nogaro, respectively) enrolled in 1977 and followed up to 31 December 2014, by means of a computerized record linkage system with administrative sources on healthcare use.
A HF event is defined as the first occurrence in the study period of a hospitalization with a primary diagnosis of heart failure (428) or hypertensive heart failure (402.01, 402.11, and 402.91), according to the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM). As reported in a literature review [***9***], this choice ensures a high positive predictive value (97%) in identifying incident HF.
Number of first hospitalizations for HF and incidence rate (per 1,000 person years), along with 95% Confidence Interval (95% CI), were calculated both overall and stratified by gender, age, period,

geographical area. The denominators of such rates are measured in years of observation time per each cohort member, starting from June 1977 until the occurrence of a HF event or the end of follow-up. Subjects who died before experiencing an HF event were censored at that time.
Age-specific incidence rates of first hospitalization for HF by birth cohort and gender were estimated. To this purpose, age at hospitalization was split into six 10-year age-groups, from 45 to 95+ years, and birth cohort was split into 5-year groups, from 1915 to 1940.
To try to disentangle the effect of age, calendar period, and birth cohort on the overall temporal trend of HF in the period 1977–2014, we fitted an APC model [***10***] to the data.
To overcome the identifiability problem of APC analysis, we set the cohort function at zero at 1930 (median of birth date) and constrained the period effects to be zero on average, with zero slope. Different parameterizations can be chosen; however, curvature of the effect is an invariant. With the chosen parameterization, age effects can be interpreted as incidence rates of first hospitalization for HF in the reference cohort of 1930, cohort effects as the rate ratios (RRs) relative to the reference cohort and the period effects as the residual rate ratios relative to the age-cohort estimation. This approach assesses whether period has the same effect on all age groups and/or whether all birth cohorts have similar patterns (decreasing, increasing or stable incidence rates).
Computations were implemented with R 3.2.3 software.

Results
A total of 427 individual patients in the study cohort (n=3,066) were discharged from the hospital for the first time with HF, over a 37-year time span. One thousand six-hundred and thirty-four subjects (52%) died during the follow-up. Of these, 123 had also a hospitalization for HF.
The incidence of first hospitalization for HF increased with age and was higher for male population. As to the calendar period component, it is evident a crude increase in HF incidence in recent years, from 2000 onwards, as the cohort ages. Martignacco intervention area exhibited HF rates comparable to that in San Giorgio di Nogaro control area.
The analysis of HF age-specific rates according to gender and birth cohort showed that HF incidence was generally higher in men compared to women. In both genders, the incidence decreased with time for the 65-75 and 75-85 age groups, whereas it remained stable below 65 years of age.
In the APC analysis, by comparing the deviance between adjacent lines (a lower p-value indicates a better fit), it was possible to identify which model provided a better fit. In both genders, the best model for HF incidence turned out to be the full APC. For the parameterization used, a cohort effect with a turning point in 1930 was observed. Following that year, a sharp decrease in the RRs occurred in both genders (statistically significant only in men). The estimated RR in the 1940 birth cohort decreased to 0.43, 95%CI (0.19- 0.92), in men and to 0.45, 95%CI (0.16- 1.26), in women. A residual effect of calendar period on RR is observed with a plateau in 1995 for women and in 2000 for men (statistically significant only in men), followed by a decline.

Conclusions
The present study described trends of first hospitalization for HF within a community based cohort in North-East of Italy. To our knowledge this one of the few studies to assess such a trend within a cohort and not through hospital episode statistics.
Findings confirm that age, period and birth cohort play an important role in the incidence of first hospitalization for HF. Similar to population trends observed in Europe [***11***], we report a plateau in the late nineties and subsequent decline in HF hospitalization rates up to 2010.
Consistently with literature [***12***], APC analysis showed that HF incidence increased markedly with age. Incidence rates were found to be slightly higher compared to those reported from previous studies [***13***].

However, comparing HF incidence with international literature is often struggling, since criteria employed for HF identification have changed over the years, and not all studies have employed standardized criteria to detect HF new cases. Moreover, the age of subjects observed may slightly change, along with data sources employed for the analysis (e.g., hospital discharge records, administrative databases, death records) [***13***].
We cannot rule out that the peak in HF incidence observed for the 1930 cohort may be related to left truncation [***14***] for the older birth cohorts (i.e. the exclusion of the presumably small portion of subjects from the 1915-1920 and 1920-1925 birth cohorts who experienced HF and died before study start may potentially lead to HF rates under-estimation).
Although an evaluation of the effectiveness of the health promotion intervention was beyond the purpose of this study, a crude comparison between the area of Martignacco (exposed) and San Giorgio di Nogaro (control) seems to suggest the absence of significant differences in HF incidence rates.
Recently it has been shown a reduction of CVD incidence rate due to advancement in prevention strategies and diagnosis [***15***], but HF disease seems to be an exception. This is probably because HF onset is influenced by several contributors; including not only lifestyle habits but also other heterogeneous risk factors (e.g., concomitant heart and cardiovascular diseases, concomitant medications, genetics).

References

[**1**] Bleumink GS, Knetsch AM, Sturkenboom MC, et al. Quantifying the heart failure epidemic: prevalence, incidence rate, lifetime risk and prognosis of heart failure. European heart journal 2004;25(18):1614-9.
[**2**] Feruglio G, Vanuzzo D, Pilotto L. The Martignacco Project. In: Puska P, Leparsky E, eds. Comprehensive Cardiovascular Community Control Programmes in Europe. DK: WHO Regional Office for Europe, Euro Reports and Studies 1988:52-8.
[**3**] Güder G, Gelbrich G, Edelmann F, et al. Reverse epidemiology in different stages of heart failure. International journal of cardiology 2015;184:216-24.
[**4**] Roger VL, Weston SA, Redfield MM, et al. Trends in heart failure incidence and survival in a community-based population. Jama 2004;292(3):344-50.
[**5**] Fang J, Mensah GA, Croft JB, et al. Heart failure-related hospitalization in the US, 1979 to 2004. Journal of the American College of Cardiology 2008;52(6):428-34.
[**6**] Chen J, Normand S-LT, Wang Y, et al. National and regional trends in heart failure hospitalization and mortality rates for Medicare beneficiaries, 1998-2008. Jama 2011;306(15):1669-78.
[**7**] Levy D, Kenchaiah S, Larson MG, et al. Long-term trends in the incidence of and survival with heart failure. N Engl J Med 2002;347(18):1397-402.
[**8**] Ziaeian B, Fonarow GC. Epidemiology and aetiology of heart failure. Nat Rev Cardiol 2016;13(6):368-78.
[**9**] Saczynski JS, Andrade SE, Harrold LR, et al. A systematic review of validated methods for identifying heart failure using administrative data. Pharmacoepidemiol Drug Saf 2012;21(1):129-40.
[**10**] Carstensen B. Age-period-cohort models for the Lexis diagram. Stat Med 2007;26(15):3018-45.
[**11**] Jhund PS, Macintyre K, Simpson CR, et al. Long-term trends in first hospitalization for heart failure and subsequent survival between 1986 and 2003: a population study of 5.1 million people. Circulation 2009;119(4):515-23.
[**12**] Vigen R, Maddox TM, Allen LA. Aging of the United States population: impact on heart failure. Current heart failure reports 2012;9(4):369-74.
[**13**] Roger VL. Epidemiology of heart failure. Circulation research 2013;113(6):646-59.
[**14**] Cain KC, Harlow SD, Little RJ, et al. Bias due to left truncation and left censoring in longitudinal studies of developmental and disease processes. Am J Epidemiol 2011;173(9):1078-84.

[**15**] Nabel EG, Braunwald E. A tale of coronary artery disease and myocardial infarction. N Engl J Med 2012;366(1):54-63.

INTERACTION OF KILLER IMMUNOGLOBULIN RECEPTORS AND HLA-C PROVIDES HIV-1 CONTROL

Monti Maria Cristina[1], Malnati Mauro[2], Pistorio Angela[2], Biassoni Roberto[2]

1 *Department of Public Health, Unit of Biostatistics and Clinical Epidemiology, University of Pavia, Pavia Italy.*
2 *Unit of Human Virology, Division of Immunolgy, transplantation and Infectious Diseases IRCCS Ospedale San Raffaele, Milan, Italy.*
3 *IRCCS Giannina Gaslini, Genoa, Italy.*

Introduction

HIV-1 infection of $CD4^+$ T lymphocytes and myeloid cells is characterized by a profound immunodeficiency leading to opportunistic infections and tumors development in most infected individuals in the absence of combination antiretroviral therapy (cART). However, a minority of people infected by HIV-1 is characterized by either long-term control of their CD4 T cell counts and good healthy conditions for several years (long-term nonprogressors) or spontaneous control of virus replication in peripheral blood (Elite Controllers) for variable periods [***1, 2***].

Among factors explaining these rare conditions (1-2% of infected individuals) of natural control of HIV-1 disease progression, there are genetic factors, mostly associated to the HLA Class I, especially the HLA-B locus. These findings have been largely interpreted and correlated with a robust $CD8^+$ cytotoxic T cell (CTL) response keeping in check the number of infected cells [***3***]. Moreover, evidence has emerged on the epistatic association between the HLA-B supratype Bw4 and the Killer Immunoglobulin–like receptors (KIRs) 3DL1/DS1 locus on human chromosome 19q13.4 in the beneficial responses to HIV-1 [***4***].

The involvement of other HLA loci, mainly HLA-C, in HIV-1 control has been long questioned, being dismissed as a simple carryover of the existing linkage disequilibrium (LD) with some protective HLA-B alleles [***1, 5***]. Only in recent years, several genetic and functional experimental evidences strongly supported an independent involvement of HLA-C molecules in HIV-1 infection outcomes and proved that the genetic variant rs9264942 C/T was associated with control of viremia, slower disease progression and to the levels of HLA-C cell surface expression in caucasoid populations [***6, 7***]. Subsequently, the causal variant responsible for these associations was identified in a SNP, rs67384697G/del that maps in the 3' UTR of HLA-C alleles affecting the binding of the microRNA Hsa-miR-148a [***2, 4***]. Noteworthy, HLA-C molecules are insensitive to the HIV-1 escape strategy mediated by the viral protein nef that down regulates only the surface expression of HLA–A and HLA-B molecules [***1, 8***]. Thus, it is likely that HLA-C molecules may participate directly in the immune response against HIV-1 mediating a CD8 T-cell selective pressure on the virus.

To analyze HIV-1 viremia control genetic determinants and their interaction among the italian caucasoid population.

Methods

Genomic DNA was extracted from circulating mononuclear cells derived from a group of 79 HIV–controllers encompassing 40 Elite Controllers (EC) featuring one or more yrs of HIV–1 plasma RNA load <50 copies/ml (median = 14; 1–3° IQR = 10–21) and 39 Long Term Non Progressors (LTNP) with a documented history of HIV–1 seropositivity >9 yrs (median = 17; 1–3° IQR = 13–22).

A control group of 189 HIV–1 Progressors (P) included 93 individuals with immunological reconstitution following cART, 61 ART–naïve viremic individuals with a documented history of HIV–seropositivity <7 yrs (median = 3; 1–3° IQR = 1–6) requiring the introduction of cART and 35 individuals with primary HIV-

infections or recent seroconverters. Genomic DNA samples were also derived from 111 HIV–seronegative healthy blood donors.
All the experiments were performed in accordance with relevant guidelines and regulations and all partecipants gave a written informed consent to study protocol prior to inclusion in the study. This study was approved by the IRCCS AOU San Martino-IST Genoa, IRCCS San Raffaele Hospital, Milan and by L. Sacco Hospital, Milan ethical Commitees.
Statistical analysis. An explorative multiple correspondence analysis (MCA) of selected chromosome 6 polymorphisms within the HLA-B and -C class I region and of KIR genes was used to verify whether combinations of the selected chromosome 6 and 19 markers could discriminate between EC and LTNP, in comparison to HIV-1 Progressors (P) or HIV-1 seronegative healthy blood donors (HD).
To assess the association between each genetic marker and HIV–controller status (EC and/or LTNP) we performed univariate allelic Pearson chi–square or Fisher exact tests when appropriate; moreover, the Cochran–Armitage trend test was applied to evaluate a possible additive dose-response effect of the considered risk allele. Adjusted Odds Ratio (OR) with 95% confidence intervals (95%CI) were derived and used as the measure of effect.
To test whether the overall haplotype variation of the chromosome 6 HLA markers influences the HIV–controller status, we performed a Likelihood Ratio (LR) based omnibus association test, considering specific haplotypes. In addition we tested whether a single marker had an independent effect (independent of the haplotypic effects formed by the remaining markers) or it was associated to EC and/or LTNP status due to Linkage Disequilibrium (LD), performing conditional LR–based tests and asymptotic p-values.
To take in account the possible role of a chromosome 6 risk haplotype, we carried out a stratified analysis on each chromosome 19 significant marker. The Mantel–Haenszel method provided pooled and separated odds ratio for each strata (presence/absence of the chromosome 6 haplotype).
The allele frequencies in controls were examined to detect any significant deviation from the Hardy–Weinberg Equilibrium (HWE) using a goodness of fit Chi–square test. LD was calculated using HAPLOVIEW software. Statistical analyses were performed using Plink 1.08 (44), XLstat (Version 2, Addinsoft) and Stata 14 (Stata Corporation, College Station, TX, USA).

Results
We found a novel signature associated with the Elite Controller but not with the long-term nonprogressors status concerning 2DS activating KIRs (chromosome 19) and HLA-C2 alleles insensitive to miRNA148a regulation (chromosome 6).
We followed the subsequent research steps.
A. Global association of the EC or LTNP status with the chromosome 6 or 19 genetic markers by multiple correspondence analysis: the presence of activating KIR genes, namely 3DS1, 2DS1 and 2DS5 on chromosome 19 clearly distinguished the 40 EC from the 111 HD, whereas the 39 LTNP segregated in an intermediate position between EC and P. Both EC and LTNP segregated independently from P and HD for chromosome 6-associated markers rs9264942C and rs67384697del as well as the presence of HLA-B alleles (both belonging to the Bw4-I80 supratype).
B. We found distinct sets of KIRs and HLA genes involved in HIV-1 disease control: the univariate associations between the status of natural HIV-1 controllers (EC and/or LTNP) and KIR genes confirmed a role for the 3DS1 gene ($p < 0.05$) and the 2DS3 gene ($p < 0.01$). Both HLA-C-associated SNPs and the Bw4-I80 supratype were significantly associated with elite controller status, while the HLA C1/C2 supratype was not associated.
C. HIV-1-controllers are characterized by a complex chromosome 6 haplotype (haplo4+): the haplotype containing the three significant alleles in the univariate analyses clearly increased the odds of being either EC (OR = 3.5, $p = 0.0013$) or LTNPs (OR = 3.5, $p = 0.0025$). Surprisingly, the addition of the C2 supratype

strongly reinforced the haplotypic association with the status of EC (OR = 6.1, p = 0.00167) but not of LTNP (OR = 3.9, p = 0.0136). These findings suggest that the presence of the HLA-C2 alleles carrying a miRNA 148a/b deleted site may play a key role in determining the EC status.
D. We assessed the distribution of each significant KIR gene in HIV-1 controllers and P, conditioning on the presence/absence of the described four marker-haplotype (Haplo4+ /Haplo4−) on chromosome 6. Among Haplo4+ carriers we described an even stronger association between activating 2DS genes and the haplotype featuring C2 alleles insensitive to the miRNA 148a/b regulation that could be considered a hallmark of EC. In contrast, LTNP did not show a relevant association with any of the activating KIRs beside a modest association with the B centromeric 2DS3 gene in the the Haplo4− individuals.

Conclusions
Our results provide evidence that a complex genotype displaying high expression of HLA-C alleles and their KIR 2DS receptors ligands is a hallmark of the EC but not of the LTNP status. These findings suggest a novel role for activating 2DS KIRs expressed on both Natural Killer and CD8+ T-cells in the control of HIV-1 viremia. Indeed, both innate and adaptive cellular immunity share HLA class I ligands and the KIRs multigene family as fine regulators of their activity.

References
[1] Deeks SG, Walker BD. Human immunodeficiency virus controllers: mechanism of durable virus control in the absence of antiretroviral therapy. Immunity 2007;27(3):406-16.
[2] Riva A, Vicenzi E, Galli M, et. al. Strenuous resistance to natural Hiv-1 disease progression: viral controllers and long-term nonprogressors. Future Virology 2011;6:521-33.
[3] Goulder PJ, Walker BD. HIV and HLA class I: an evolving relationship. Immunity 2012;37(3):426-40.
[4] Martin MP, Gao X, Lee JH, et al. Epistatic interaction between KIR3DS1 and HLA-B delays the progression to AIDS. Nat. Genet 2002;31(4):429-34.
[5] Kiepiela P, et al. Dominant influence of HLA-B in mediating the potential co-evolution of HIV and HLA. Nature 2004;432(7018):769-75.
[6] Pereyra F and the Internation HIV Controllers Study, The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. Science 2010;330(6010):1551-57.
[7] Kulkarni S, Savan R, Qi Y, et al. Differential microRNA regulation of HLA-C expression and its association with HIV control. Nature 2011;472(7344):495-8.
[8] Collins KL, Chen BK, Kalams SA, et al. HIV-1 Nef protein protects infected primary cells against killing by cytotoxic T lymphocytes. Nature 1998;391(6665):397-401.

METODI BIOSTATISTICI

CHANGES IN DIETARY FATS INTAKE IN RELATION TO MORTALITY IN US ADULTS. AN ISO-CALORIC SUBSTITUTION ANALYSIS FROM THE AMERICAN NATIONAL HEALTH AND NUTRITION EXAMINATION SURVEY LINKED TO THE US MORTALITY REGISTRY

Ricci Cristian, Smuts M. Cornelius

Centre of Excellence for Nutrition (CEN), North-West University, Potchefstroom Campus, South Africa.

Introduction

Cardiovascular disease and cancer still represents a major cause of mortality in high income countries [***1***]. In the USA it was estimated that cancer and cardiovascular diseases combined accounts for more than 50% of the overall mortality burden among adults and represents the first and second cause of mortality respectively [***2***]. It is widely acknowledged that, among others, unhealthy diet and lifestyles represents a major preventable risk factor for all-causes and cardiovascular mortality in the USA, with fat intake and unhealthy diet representing a considerable part of the risk [***3***]. Notably, in 2010 it was estimated that saturated and trans-unsaturated fats consumption was higher in the USA (11.8% and 2.8% of the overall energy intake respectively) when compared to global consumption worldwide (9.4% and 1.4% of the overall energy intake respectively). On the other hand, poly-unsaturated fats consumption was comparable between the USA and all of the other countries combined representing about 5.9% of the overall energy intake [***4***]. Population based studies showed that high intakes of saturated and trans-unsaturated fatty acids are associated with an increased risk of cardiovascular disease incidence, cardiovascular mortality and overall mortality as well [***5***]. On the contrary, a high intake of poly-unsaturated fats was consistently associated to a reduced risk of cardiovascular risk and related mortality [***6, 7***]. Moreover, scientific evidence showing the relation between dietary fat intake and cancer risk is currently accumulating. In a recent meta-analysis of observational studies a significant increased risk of gastric cancer was associated with highest dietary intakes of saturated fats while a reduced risk of gastric cancer was observed for those subjects having a higher intake of unsaturated fats [***8***].

It is admissible to suppose that adopting a healthy diet with a related increased of mono and poly unsaturated fats or the replacement of the saturated fats with poly unsaturated counterparts may lead to beneficial outcomes in terms of overall, cardiovascular and cancer mortality at the population level. Unfortunately, the question regarding how much mortality risk reduction is attributable to such a dietary improvement still needs to be fully addressed at population level. On one hand, it is acknowledged that randomized controlled trials and meta-analysis of interventional studies were conducted to evaluate such diet interventions and the effect of fat components on short term outcomes as biomarkers [***9***]. On the other hand, it is clear that these approaches cannot be technically performed when long term outcomes, such as mortality, are considered in healthy subjects. In the past, randomized control trials resulted in inconsistent outcomes since observational time, more than the sample size, is an operational limit [***10***]. There is an increasing demand of observational population based studies aimed to address the question regarding dietary improvement and mortality risk.

The aim of the present study was to estimate the association between the diet improvements due to an increase in poly and mono unsaturated fats intake or the substitution of saturated fats to poly and mono unsaturated fats and mortality reduction in US adults. To this purpose, an energy partition model and an iso-caloric substitution analysis were conducted over a prospective study resulting from the American National Health and Nutrition Examination Survey (NHANES) merged to the US national death registry.

Methods

The National Health and Nutrition Examination Survey (NHANES) is an epidemiological cross-sectional study aimed to generate a representative sample of the civilian non-institutionalized US population using a multi-stage probability sampling design [***11***]. The NHANES study started in 1999 and every year about 5,000 US citizens are involved providing data regarding their socio-economical characteristics, dietary habits, medical status and medical history. On a regular basis the NHANES participants are linked to the US death registry using a probabilistic record linkage defining a prospective mortality study. The last linkage to the US mortality registry was in 2011 [***12***].

The number of all participants from the survey conducted between 1999 and 2011 was 62,210 of which 21,290 were excluded because no available data on mortality leaving 40,870 participants. After exclusion of subjects aged between 18 and 30 years 30,509 subjects were retained. Exclusion of prevalent cases of cancer, diabetes and cardiovascular diseases reduced the sample to 21,344 participants. Finally, the exclusion of subjects with missing informations on fat intake and adjusting covariates used in the models leaded to a final sample size of 18,372 individuals resulting in a cumulative observational time of 116,398 person-years with a median observational time of 6.1 years. During the follow up 1,118 deaths for all causes, 267 deaths for cardiovascular cause and 289 deaths due to cancer were observed.

Statistical Analysis

The description of the sample was performed separately for subjects who died during the observational time and for those subjects who survived up to the end of the observational time. Medians and interquartile ranges were reported for continuous variables and percentages were reported for categorical variables.

All analysis were based on a multivariable adjusted Cox proportional hazard model having age as underlying time variable and stratified by gender and five years age classes. Raw and multivariable adjusted model considering potential confounders were evaluated in a sequential modelling approach aimed to evaluate the residual confounding bias (results not shown). After the evaluation of the results from sequential modelling a multivariate model adjusted for ethnicity (Black, white and others), body mass index (continue), alcohol intake (continue), smoking status (more than 100 cigarette/life), education (university degree), sedentariness (subjects who declared having more than 8 Hrs/day of sedentary activity or who do not declared moderate or vigorous physical activity), dietary fiber intake (continue) and blood pressure (systolic and diastolic pressure as continue) was chosen to report all of the results. The proportional hazard assumption was evaluated by Schoenfeld's residuals. Sensitivity analysis were conducted excluding death cases assessed in the first eighteen months after recruitment. Results from analyses evaluating the association between tertiles of energy density of saturated, mono-unsaturated and poly-unsaturated fats with all-cause and specific-cause mortality were reported along with trend tests performed considering one standard deviation increase. The association between changes of nutrient fat intake and all-cause and specific cause mortality was performed considering the energy partition and nutrient density paradigms as previously described by Hu and colleagues [***13***]. Briefly, in the energy partition model the total energy from the nutrients is partitioned in the different components contributing to the energy intake so that nutrients are coded in a metric of their energy contribution. In this model the relative risk estimates corresponds to a given increase of energy from that specific component when all of the other components remain constant. The analyses based on the energy partition models were performed considering all of the dietary components (carbohydrates, proteins and fats) and the fats components only. On the other hand, the nutrient density model is a substitution model in which the ratio between the energy due to the single dietary component and the overall energy from all dietary components are considered in the model so that the sum of all the components remain constant and corresponds to a given numerical value (100% if the percentage is used). This model is performed considering fats components contributing to energy intake excluding one, so that

the relative risk estimates for a given component represent the substitution of a given quantity (1% if the percentage is used) of the excluded component with this one.
All statistical analysis were performed taking into account the multi-stage probability sampling design of the NHANES study using the weights from the multi-stage probability sampling to avoid the effect of differential probabilities of selection among subgroups, and to compensate for exclusion of sampling areas in the sampling frame. The surveymean and the surveyfreq procedures of the SAS software were used to describe the sample while the Cox-models were fitted using the proc surveyphreg of the SAS software. Statistical analysis were performed using SAS package vers. 9.4 and type-I error (alpha) was set to 0.05 (5%).

Results

Participants who survived at the end of the observational time were younger, mostly women, prevalently non-smokers, had a lower systolic blood pressure and had a higher intake of fibers and energy when compared to participants who undergone death before the end of the observational time. Adiposity and percentages of energy from fat, carbohydrates and protein intakes were comparable among those participants who survived at the end of the observational time and the ones who died. The analysis of the relation between fat intake and all-causes and specific cause mortality showed that higher intakes of saturated fats resulted in an increased risk of mortality while higher intakes of poly-unsaturated fats resulted in a decreased risk of mortality. In particular, comparing the first tertile to the third tertile of energy percentage from saturated fats resulted in a significant 12% increased risk for all causes, cardiovascular and cancer mortality (HRs = 1.12 (1.07, 1.18), 1.12 (1.06, 1.17) and 1.12 (1.06, 1.17) for all causes and cancer mortality respectively). On the other hand, when comparing the first tertile to the third tertile of energy percentage from poly-unsaturated fats a significant reduction of all causes and specific cause risk of death up to 7% is observed (HRs = 0.94 (0.90, 0.98), 0.93 (0.89, 0.97) and 0.93 (0.89, 0.97) for all cause, cardiovascular and cancer mortality respectively). Notably, the association between mono unsaturated fats intake and mortality was not significant.
Partition models reported about the association between a given increase of fat intake and mortality risk when maintaining constant the other sources of energy (***Table 1***). According to this analysis it was observed that an increase of 100 kcal/day of any fats did not corresponded to a significant change of the mortality risk when energy intake from other fats, carbohydrates and proteins remained constant. On the contrary, when the analysis is performed maintaining constant the energy from fats sources only a 100 kcal/day increase of saturated fats intake resulted in a 9% significantly increased risk for all causes, cardiovascular and cancer mortality. On the other hand, 100 kcal increase of poly-unsaturated fats intake corresponded to a 2 to 3% significant reduction of the risks for all causes and cardiovascular mortality when energy intake from other fats remains constant. Again, mono unsaturated fats intake was not significantly related to the risk of death. The iso-caloric substitution analysis revealed that the substitution of 10% of energy deriving from saturated fats with an equal amount of energy from poly-unsaturated fats resulted in an 8% significant reduction of all causes and cause specific mortality risks (***Table 1***). Substitution of saturated fats with mono unsaturated fats resulted in a 3% borderline significant risk reduction for all causes mortality. A 4% decreased risk for all causes mortality was observed in the substitution model considering the substitution of 10% of energy deriving from saturated fats with an equal amount of energy from mono-unsaturated fats. Finally, an iso-caloric substitution of 10% of energy from mono unsaturated fats with an equal amount of energy from poly unsaturated fats resulted in a significant decrease of 4 to 5% for all causes and specific cause mortality. The above results are maintained when excluding subjects died in the first eighteen months after recruitment (results not shown).

Conclusions

In the present study an association between saturated fats intake and increased risk for all causes and specific cause mortality was reported. On the other hand, a reduction of all causes and specific cause mortality risks was observed for those participants having a higher intake of saturated fats compared to those having lower intakes of poly unsaturated fats. Moreover, the present study confirmed evidences from randomized control trials on the beneficial effect of poly unsaturated fats intake on the cardiovascular function [***14***]. Additionally, the present work reported about the beneficial outcomes deriving from mono and poly unsaturated fats substituting saturated fats in terms of mortality at a population level. In summary, the present findings point out for beneficial association between a diet rich in mono and poly unsaturated fats and poor in saturated fats, these findings are largely confirmed by previous studies [***6, 15, 10, 14, 16***]. Notably, the present study extended this evidence to cancer mortality.

In the present study a significant relation between mono-unsaturated fats intake and mortality risk was not found differently from previous studies [***17***]. On the other hand, the present study showed that a substitution of 10% of energy from saturated fats with an equal amount of energy from mono-unsaturated fats resulted in a borderline reduction of mortality risk pointing out the suggestive idea that substitution of saturated fats with mono unsaturated fats, more than a n increase of mono unsaturated fats intake di per se, might be associated to a beneficial outcome in terms of mortality risk [***18***]. According to our results regarding partition models an increase of unsaturated fats and a decrease of saturated fats intake has beneficial effects only when maintaining constant the energy from other fat sources excluding carbohydrates and proteins. This result points out that substitution of saturated fats with unsaturated counterparts are possibly related to an even stronger association with the mortality outcomes that were investigated. As a confirm of the previous conjecture in the present study the iso caloric substitution of saturated fats intake with unsaturated counterparts was associated to a significant reduced risk for all causes and cause specific mortality. This evidence was suggested indirectly from previous studies on cardiovascular and overall mortality [***19, 14, 16, 17***]. Again, for the first time to our knowledge the present work extend the evidence of a beneficial effect of the substitution of saturated fats intake with mono and poly unsaturated counterparts on cancer mortality.

Scientific evidence accumulated over the last decades already explained the possible mechanisms underlying the association between dietary fats intake and cardiovascular outcomes. An early identified mechanism is the mediation between cholesterol plasma level and dietary saturated fats intake. A recent meta-analysis conducted in humans and considering the association between dietary saturated fats intake with main cardiovascular outcomes and mortality reinforced this evidence by means of meta-regressions and subgroup analysis [***10***]. Among the others factors related to the relation between saturated fats and cardiovascular outcomes the adiposity certainly plays a crucial role [***20***]. Obesity and overweight are conditions normally correlated to high energy intake and sedentariness. Obese and overweight subjects normally have also impaired diet and high intakes of saturated fats. Nevertheless, obesity and overweight were associated to cytokines and other inflammatory biomarkers which are related to atherosclerosis and then to main cardiovascular outcomes [***21***]. Moreover, obesity and overweight were consistently associated to increased risk of different cancers through a mechanisms acting by impaired hormonal status and increased type-II diabetes risk [***22, 23, 24***]. It is then admissible to suppose that unhealthy diet and a related high saturated fats intake may lead to an increased cardiovascular and all causes mortality risk via adiposity related mechanisms.

Biological mechanisms regarding poly unsaturated dietary fats intake in relation to cardiovascular outcomes and adiposity has been proposed. A general beneficial effect of polyunsaturated dietary fats on cardiovascular function may be related to its activity in depressing plasma lipids, inhibiting of apoprotein synthesis and inhibiting of cyclooxygenases and lipoxygenases activity which, in turn, reduces eicosanoid synthesis decreasing thromboxane and platelets tendency to aggregate [***25***]. On the other hand it is widely acknowledged that poly unsaturated dietary fats are beneficial in terms of preventing obesity, overweight

and type-II diabetes risk [***26***]. Finally, poly unsaturated fats intake may results as beneficial in terms of all causes and cancer related mortality also because of their ability to play as antioxidants [***27, 28***].
The present work has limitations due to its observational nature that limit our results in terms of causal association. Nevertheless, underlying morbid condition at baseline could have affected the results without any possibility of control. This second limitation and a related possible reversal causation can be likely excluded since prevalent morbid conditions at baseline were excluded in the main analysis and death in the first eighteen months from follow-up were excluded in a sensitivity analysis.
In conclusion high intake of saturated fats corresponded to an increased risk of mortality in US adults, high intake of poly unsaturated fats resulted in a reduced mortality risk in US adults and high intake of mono unsaturated fats, di per se, did not reduced mortality risk in US adults. Substitution of saturated fats with mono-unsaturated and poly unsaturated counterparts leads to beneficial mortality reduction at population level.

Table 1. Partition and substitution models reporting hazard ratios for the association between dietary fats intake and all-causes, cardiovascular and cancer mortality.

Partition model$_1$	***Saturated FAs***	***Mono-unsaturated FAs***	***Poly-unsaturated FAs***
All-cause mortality	1.02 (0.99, 1.04)	1.00 (0.97, 1.03)	1.01 (0.98, 1.04)
CVD mortality	1.01 (0.98, 1.04)	1.01 (0.98, 1.04)	1.00 (0.97, 1.03)
Cancer mortality	1.01 (0.98, 1.04)	1.00 (0.97, 1.03)	1.00 (0.98, 1.03)
Partition model$_2$	***Saturated FAs***	***Mono-unsaturated FAs***	***Poly-unsaturated FAs***
All-cause mortality	1.09 (1.06, 1.12)	1.01 (0.98, 1.04)	0.98 (0.96, 1.00)
CVD mortality	1.09 (1.06, 1.12)	1.02 (0.99, 1.05)	0.97 (0.94, 0.99)
Cancer mortality	1.09 (1.06, 1.12)	1.02 (0.99, 1.05)	1.00 (0.98, 1.03)
Substitution model$_1$	***Saturated FAs***	***Mono-unsaturated FAs***	***Poly-unsaturated FAs***
All-cause mortality	-	0.97 (0.93, 1.00)	0.92 (0.90, 0.95)
CVD mortality	-	0.97 (0.93, 1.01)	0.92 (0.90, 0.95)
Cancer mortality	-	0.97 (0.94, 1.01)	0.92 (0.90, 0.95)
Substitution model$_2$	***Saturated FAs***	***Mono-unsaturated FAs***	***Poly-unsaturated FAs***
All-cause mortality	1.04 (1.00, 1.08)	-	0.96 (0.92, 1.00)
CVD mortality	1.03 (0.99, 1.07)	-	0.95 (0.91, 0.99)
Cancer mortality	1.03 (0.99, 1.07)	-	0.95 (0.91, 0.99)
Substitution model$_3$	***Saturated FAs***	***Mono-unsaturated FAs***	***Poly-unsaturated FAs***
All-cause mortality	1.08 (1.05, 1.11)	1.04 (1.00, 1.09)	-
CVD mortality	1.08 (1.06, 1.11)	1.05 (1.01, 1.10)	-
Cancer mortality	1.08 (1.06, 1.11)	1.05 (1.01, 1.09)	-

Partition model$_1$: Energy partition model considering dietary fats, carbohydrates and proteins, risks for 100 Kcal increase of dietary fats maintaining constant all of the other components. **Partition model$_2$:** Energy partition model considering only dietary fats, risks for 100 kcal increase of dietary fats maintaining constant all of the other components. **Substitution model$_1$:** Substitution of 10% of energy intake from dietary SFAs. **Substitution model$_2$:** Substitution of 10% of energy intake from dietary MFAs. **Substitution model$_3$:** Substitution of 10% of energy intake from dietary PFAs.

References

[**1**] Lozano R. Naghavi M, Foreman K, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. The Lancet 2013;380(9859):2095-128.

[**2**] Kenneth D. Kochanek SLM, Jiaquan X, Betzaida Tejada-Vera. National Vital Statisitcs Reports, Volume 65, Number 4. Available from https://www.cdc.gov/nchs/data/nvsr/nvsr65/nvsr65_04.pdf. [Accessed June 30, 2016].

[3] Danaei G, Ding EL, Mozaffarian D, et al. The preventable causes of death in the United States: comparative risk assessment of dietary, lifestyle, and metabolic risk factors. PLoS Med 2009;6(4):e1000058.
[4] Micha R, Khatibzadeh S, Shi P, et al. Global, regional, and national consumption levels of dietary fats and oils in 1990 and 2010: a systematic analysis including 266 country-specific nutrition surveys. BMJ 2014;348:g2272.
[5] de Souza RJ, Mente A, Maroleanu A, et al. Intake of saturated and trans unsaturated fatty acids and risk of all cause mortality, cardiovascular disease, and type 2 diabetes: systematic review and meta-analysis of observational studies. BMJ. 2015;11;351:h3978.
[6] Abdelhamid A, Martin N, Bridges C, Song F, HO Deane K, Hooper L. Polyunsaturated fat intake for prevention of cardiovascular disease. The Cochrane Library, 2016.
[7] Yamagishi K, Iso H, Date C, et al. Fish, ω-3 polyunsaturated fatty acids, and mortality from cardiovascular diseases in a nationwide community-based cohort of Japanese men and women: the JACC (Japan Collaborative Cohort Study for Evaluation of Cancer Risk) study. Journal of the American College of Cardiology 2008;52(12):988-96.
[8] Han J, Jiang Y, Liu X, et al. Dietary fat intake and risk of gastric cancer: a meta-analysis of observational studies. PloS one 2015;10(9):e0138580.
[9] Fattore E, Bosetti C, Brighenti F, Agostoni C, Fattore G, et al. Palm oil and blood lipid–related markers of cardiovascular disease: a systematic review and meta-analysis of dietary intervention trials. The American journal of clinical nutrition 2014;99(6):1331-50.
[10] Hooper L, Martin N, Abdelhamid A, Davey Smith G. Reduction in saturated fat intake for cardiovascular disease. The Cochrane Library, 2015. Cochrane Database Syst Rev. 2015;10(6):CD011737.
[11] Sterne JA, Egger M, Smith GD. Systematic reviews in health care:Investigating and dealing with publication and other biases in meta-analysis. BMJ 2001;323(7304):101-5.
[12] Orsini N, Bellocco R, Greenland S. Generalized least squares for trend estimation of summarized dose-response data. Stata Journal 2006:6(1):40-57.
[13] Hu FB, Stampfer MJ, Rimm E, et al. Dietary fat and coronary heart disease: a comparison of approaches for adjusting for total energy intake and modeling repeated dietary measurements. American journal of epidemiology 1999;149(6):531-40.
[14] Mozaffarian D, Micha R, Wallace S. Effects on coronary heart disease of increasing polyunsaturated fat in place of saturated fat: a systematic review and meta-analysis of randomized controlled trials. PLoS Med 2010;7(3):e1000252.
[15] Chowdhury R, Warnakula S, Kunutsor S, et al. Association of dietary, circulating, and supplement fatty acids with coronary risk: a systematic review and meta-analysis. Annals of internal medicine 2014;160(6):398-406.
[16] Siri-Tarino PW, Sun Q, Hu FB, Krauss RM. Meta-analysis of prospective cohort studies evaluating the association of saturated fat with cardiovascular disease. The American journal of clinical nutrition 2010;91(3):535-46.
[17] Wang DD, Li Y, Chiuve SE, et al. Association of specific dietary fats with total and cause-specific mortality. JAMA Internal Medicine 2016;176(8):1134-45.
[18] Kris-Etherton PM, Monounsaturated fatty acids and risk of cardiovascular disease. Circulation 1999;100(11):1253-58.
[19] de Oliveira Otto MC, Mozaffarian D, Kromhout D, et al. Dietary intake of saturated fat by food source and incident cardiovascular disease: the Multi-Ethnic Study of Atherosclerosis. The American journal of clinical nutrition 2012;96(2):397-404

[**20**] Van Gaal LF, Mertens IL, De Block CE. Mechanisms linking obesity with cardiovascular disease. Nature 2006;444(7121):875-80.
[**21**] Lau DC, Dhillon B, Yan H, Szmitko PE, Verma S. Adipokines: molecular links between obesity and atheroslcerosis. American Journal of Physiology-Heart and Circulatory Physiology 2005;288(5):H2031-41.
[**22**] Calle EE, Kaaks R. Overweight, obesity and cancer: epidemiological evidence and proposed mechanisms. Nat Rev Cancer 2004;4(8):579-91.
[**23**] Calle EE, Rodriguez C, Walker-Thurmond K, Thun MJ. Overweight, obesity, and mortality from cancer in a prospectively studied cohort of US adults. New England Journal of Medicine 2003;348(17):1625-38.
[**24**] McMillan DC, Sattar N, McArdle CS. ABC of obesity. Obesity and cancer. BMJ 2006;333(7578):1109-11.
[**25**] Kinsella JE, Lokesh B, Stone RA. Dietary n-3 polyunsaturated fatty acids and amelioration of cardiovascular disease: possible mechanisms. The American journal of clinical nutrition 1990;52(1):1-28.
[**26**] Delarue J, LeFoll C, Corporeau C, Lucas D, et al. N-3 long chain polyunsaturated fatty acids: a nutritional tool to prevent insulin resistance associated to type 2 diabetes and obesity? Reproduction Nutrition Development 2004;44(3):289-99.
[**27**] Hsieh R, Kinsella J. Oxidation of polyunsaturated fatty acids: mechanisms, products, and inhibition Intake of saturated and trans unsaturated fatty acids and risk of all cause mortality, cardiovascular disease, and type 2 diabetes: systematic review and meta-analysis of observational studies. BMJ 2015with emphasis on fish. Advances in food and nutrition research 1989;33:233-341.
[**28**] Porter NA. Mechanisms for the autoxidation of polyunsaturated lipids. Accounts of Chemical Research 1986;19(9):262-8.

A STATISTICAL METHOD TO INVESTIGATE THE CLUSTERING PATTERNS OF TYPE-SPECIFIC HPV INFECTIONS FROM PREVALENCE STUDIES

Pagan Eleonora[1], Bagnardi Vincenzo[1], Plummer Martyn[2], Franceschi Silvia[2], Vaccarella Salvatore[2]

1. *Department of Statistics and Quantitative Methods, University of Milan-Bicocca, Milan, Italy*
2. *Infections and Cancer Epidemiology Group, International Agency for Research on Cancer, Lyon, France*

Introduction

Human Papillomavirus (HPV) is the most common infectious agent sexually transmitted throughout the world. Persistent infection with one of the 13 oncogenic HPV types is considered as a necessary cause for the development of cervical cancer. Although most infections are asymptomatic and are spontaneously cleared within a period of 2 years, persistent genital HPV infections can lead to anogenital warts, cervical neoplasia, cervical cancer and its precursor lesions, and other anogenital cancers [***1***].

The prophylactic HPV vaccines, available since 2006, have the ability to prevent infections by the two genotypes that cause the majority of cancer cases (HPV16 and HPV18) and by the other two genotypes that are responsible for the majority of cases of genital warts (HPV6 and HPV11).

An important issue in public health is the evaluation of the occurrence of infections by multiple genotypes and their interactions. In fact, different HPV genotypes are frequently co-transmitted and infections with multiple HPV genotypes have been found in about 20% - 50% of HPV-infected women in many surveys [***1***].

If some HPV genotypes untargeted by the vaccine show positive, synergistic, association with the genotypes targeted by the vaccine, vaccination may lead to the reduction in the prevalence of infection caused also by untargeted HPV genotypes, enhancing the effectiveness of vaccination programs [***2***]. On the other hand, if some untargeted HPV genotypes show negative, antagonistic, association with the targeted ones, a vaccine that reduces the prevalence of one type may promote the prevalence of other carcinogenic types through a process of competitive release (the so-called "type replacement") diminishing the effectiveness of vaccination programs [***2***].

Despite only future results from long-term longitudinal studies of vaccinated women will best clarify this issue, some insights can just be gained by examining multiple infections from current prevalence studies.

The aim of this study is to present a statistical methodology to investigate the clustering patterns of type-specific HPV infections from prevalence studies, with an application to the International Agency for Research on Cancer (IARC) HPV Prevalence Surveys (IHPS) data [***3***]. Moreover, we present some considerations on the sample size required for the proposed method to detect antagonistic associations between HPV types.

Methods

Given a sample of *N* subjects and of *H* HPV types screened, let $(Y_{i1}, \dots, Y_{iH})$ be the vector of binary outcomes observed on the i^{th} individual $(i = 1, 2, \dots, N)$ indicating the presence/absence of the h^{th} HPV type $(h = 1, 2, \dots, H)$.

Let p_{ih} denote the probability for the i^{th} individual to be infected with the h^{th} HPV type: $p_{ih} = \mathrm{Prob}(Y_{ih} = 1)$.

The following mixed logistic regression model could be used to estimate subject-specific probabilities of being infected with each HPV type:

$$log[p_{ih}/(1 - p_{ih})] = \alpha + \beta Z_i + \gamma_h + u_i \qquad (1)$$

To note that this model assumes the independence between HPV types. It includes known risk factors, such as age or lifetime number of sexual partners (vector Z_i), and genotype specific effect (γ_h) as fixed effects, and the subject-specific effect u_i as a random effect, with $u_i \sim N(0, \sigma_u^2)$.
Model (1) could be estimated using standard software for generalized mixed modeling such as *PROC GLIMMIX* in SAS, *meglm* function in STATA, or *glmer* command in the lme4 R package.
The total number of joint infections E_{hk} expected under the independence model (1) could be then calculated using the model-estimated individual probabilities of being infected with HPV types *h* and *k* ($\hat{p}_{ih}, \hat{p}_{ik}$):

$$E_{hk} = \sum_{i=1}^{N} \hat{p}_{ih}\hat{p}_{ik}$$

Despite our aim is to evaluate pairwise associations between *H* HPV types and detect the presence of synergistic or antagonistic interactions, it would be not feasible to explicitly estimate all the $H \times (H-1)/2$ pairwise covariance terms in the model due to model identifiability problems.
Instead, to identify which pairs of HPV types tend to be detected together with unusually high or low frequency other than by chance only, we compare E_{hk} with the observed number of joint infections (O_{hk}) through the following observed-to-expected ratio:

$$R_{hk} = {O_{hk}}/{E_{hk}}$$

Values of R_{hk} lower than 1 indicate the presence of an antagonistic effect (i.e. type replacement) between HPV type *h* and HPV type *k*, while values higher than 1 indicate the presence of a synergistic effect.
The exact $(1-\alpha)\%$ confidence interval for R_{hk} can be calculated assuming O_{hk} to be Poisson distributed and E_{hk} to be fixed [***4***].
The methodology above described was applied to the data collected in a series of population-based HPV Prevalence Surveys (IHPS) conducted between 1993 and 2007 by IARC in 15 different areas in four continents using a common protocol, as described elsewhere [***3***].
Briefly, in each area, a random age-stratified sample of the population that included at least 100 women in each 5-year age group was obtained and study participants underwent a vaginal examination during which cervical cell samples were collected and analyzed for HPV testing. The other basic individual information on sexual behavior was collected through questionnaires.
All the analyses were performed using SAS software v. 9.4 (SAS Institute, Cary, NC, USA).
Considerations on the sample size required to identify synergistic and antagonistic interactions at varying HPV type specific population prevalences were performed through numerical examples.

Results
The total sample size of the IHPS study was 13,961 women. The analyses were restricted to the fifteen HPV genotypes with a prevalence of at least 0.5% in the study population, and included the thirteen oncogenic types. Overall, HPV prevalence was 12.3% and multiple infections were found in 32.2% of HPV-positive women. HPV16 was the carcinogenic type most often (58.5%) found alone, whereas HPV33, 35, and 45 were detected in over 70% of infected women in combination with other types [***3***].
The model (1) was fitted to the data, considering study area, age and lifetime number of sexual partners as fixed effects. Twelve potential synergistic interactions were detected for types 33/35, 33/39, 33/58, 31/35, 18/45, 39/58, 31/42, 18/31, 16/35, 16/31, 16/18, 16/58, with O/E ratios ranging from 2.0 to 8.4. No significant antagonistic interactions were detected.

To evaluate the sample size required for the proposed method to detect antagonistic associations between HPV types we simplified the problem to the case of only two genotypes A and B.
Consider the following 2x2 table:

HPVA / HPVB	1	0	Total
1	N_{11}	N_{10}	$N_{1.}$
0	N_{01}	N_{00}	$N_{0.}$
Total	$N_{.1}$	$N_{.0}$	N

where N_{11}, N_{10}, N_{01}, and N_{00} are respectively the number of subjects co-infected by A and B, infected only by A, infected only by B and no infected. The correlation between HPVA and HPVB could be expressed by the Phi-coefficient

$$\varphi = \frac{N_{11} \times N_{00} - N_{10} \times N_{01}}{\sqrt{N_{1.} \times N_{0.} \times N_{.1} \times N_{.0}}}$$

It should be noted that the lower negative limit of φ (index for antagonism) and the upper positive limit of φ (index for synergism) depend on the value of the observed prevalence p_A and p_B and these limits are not symmetrical to 0 [***5***]. In fact, when both p_A and p_B are lower than 0.5 or greater than 0.5, the range is unbalanced in favor of positive values. For example, if $p_A = 0.02$ and $p_B = 0.008$ (prevalences similar to those found in the IHPS study for HPV16, the most frequent type, and for HPV81), the lower limit is equal to -0.013 and the upper limit is equal to 0.629. In this scenario, in order to define a negative φ as statistically significant (p-value < 0.05), a sample size of at least 24,000 subjects would be necessary when zero co-infections are observed. The sample size increases to approximately 35,000, 46,000 and 55,000 in case of one, two, and three observed co-infections, respectively. Increasing prevalences (e.g., $p_A = 0.08$ and $p_B = 0.05$), and given the same patterns of co-infections described above (i.e. from zero to three), the sample sizes needed to declare as significant a negative correlation decreases to approximately 1,000, 1,400, 1,900 and 2,200 subjects, respectively.

Conclusions

Results from the present study suggest that mass vaccination may reduce the prevalence of types that are not even included in the vaccine, and that risk of type replacement (i.e. competition between two HPV types) is unlikely.
However, numerical considerations based on the lower and upper limits of the Phi-coefficient suggested that in the case of surveys conducted on the whole population, the statistical power of the proposed methodology in identifying antagonistic interactions is much lower than the power in identifying synergistic interactions. This means that, in order to reach the vast sample sizes required to identify an antagonistic interaction, it would be necessary to pooled data from numerous cohorts. Alternatively, one could consider only subgroup of subjects characterized by high prevalence of HPV infections, such as HIV-positive women [***6***].
However, these sample size considerations apply only to an oversimplified scenario where only two genotypes are evaluated and no co-factors are considered. To confirm these results and to investigate the statistical power of the methodology here proposed in identifying antagonistic interactions between multiple HPV genotypes, further simulation studies are needed.

References

[1] IARC. Monographs on the evaluation of carcinogenic risks to Humans. Volume 90: Human Papillomaviruses. Lyon: International Agency for Research on Cancer, 2007.

[2] Elbasha E.H., Galvani A.P., Vaccination against multiple HPV types. Math Biosci 2005;197:88–117.

[3] Vaccarella S., Franceschi S., Snijders P.J.F. *et al.*, Concurrent infection with multiple human papillomavirus types: pooled analysis of the IARC HPV Prevalence Surveys. Cancer Epidemiol. Prev. Biomarkers 2010;19:503–10.

[4] Ulm K., Simple method to calculate the confidence interval of a standardized mortality ratio (SMR). Am J Epidemiol 1990;131:373–5.

[5] Ekström J., The phi-coefficient, the tetrachoric correlation coefficient, and the Pearson-Yule Debate. Dep Stat UCLA, 2011.

[6] Orlando G., Tanzi E., Rizzardini G. *et al.*, Modifiable and non-modifiable factors related to HPV infection and cervical abnormalities in women at high risk: a cross-sectional analysis from the Valhidate Study. Ann Virol Res 2016;2:1–21.

MIXTURE MODELLING OF ANTIBODY COUNTS DATA FOR PRE- AND POST-VACCINATION TRENDS OF INFECTION : VARICELLA AND MEASLES IN ITALY

Del Fava Emanuele[1], Shkedy Ziv[2], Manfredi Piero[3]

1. *Carlo F. Dondena Centre for Research on Social Dynamics and Public Policy, Bocconi University, Via Guglielmo Röntgen 1, I-20136, Milan, Italy*
2. *Interuniversity Institute for Biostatistics and statistical Bioinformatics, Hasselt University, Agoralaan Building D, B-3590, Diepenbeek, Belgium*
3. *Department of Economics and Management, University of Pisa, Via Cosimo Ridolfi 10, I-56124, Pisa, Italy.*

Introduction

In the field of vaccine preventable infectious diseases, two important challenges are represented by the re-emergence of measles in industrialised countries [***1***] due to declining vaccination coverage, and the stall of varicella vaccination, due to the fear of negative effects on the incidence of herpes zoster, which is a long term consequence of varicella [***2***]. Adequate assessment of these phenomena requires high quality epidemiological data.

On this matter, serological surveys, which quantify the titres or concentrations of antibodies against a specific antigen, are among the most direct techniques available to investigate the trend of the level of immunity in a certain population [***3***]. Despite that, the resulting data remains somehow unexploited, for various reasons. One of them is that the assessment of the immunity profile, the so-called sero-prevalence of immune individuals, is still often performed by dichotomising the antibody titre measured by the serological test. This means that the information embedded in the individual antibody count is basically lost, as it is replaced by a binary variable giving the infection status, namely, whether the subject is still susceptible or immune, i.e., showing evidence of past infection or vaccination. The infection status is usually determined on the basis of a fixed cut-off associated to the employed diagnostic assay. Since these assays are primarily intended for diagnostic purposes and patient management, they are designed to have a larger specificity, i.e., a more accurate ability to identify true antibody-negative individuals. However, in a serological survey, the primary interest lies in the estimation of the proportion of seropositive individuals at each age in the population, rather than in the status of the single individual [***4***]. For this reason, the employment of a more sensitive method is warranted, that is to say, a method that is more accurate in the identification of the true antibody-positive individuals. The use of mixture models for estimating the age-specific population seroprevalence based on the analysis of serological antibody counts is indeed motivated by the fact that the method proved more appropriate from the epidemiological viewpoint and more sensitive than the fixed cut-off approach in analysing infection data, both for pre- and post-vaccination epochs [***5-8***].

The objective of this work is to provide accurate estimates of the age-specific seroprevalence, i.e., the proportion of seropositive individuals in the population in a given moment), and the force of infection (the rate at which a susceptible individual acquires infection), for varicella, and of the seroprevalence for measles. For this purpose, we employ Bayesian mixture models. These models allow to categorise all individuals' status between susceptible and immune, including situations considered as inconclusive by the fixed cut-off approach, estimating at the same time the age-specific seroprevalence, the overall population seroprevalence and the force of infection (FOI).

Methods

Serological data

Italian serodata for the varicella-zoster virus (VZV) were collected under the European sero-epidemiology network (ESEN2) project [***9***] and POLYMOD [**10**] between 2002 and 2005, when no varicella vaccination was in place in Italy. All serological tests for VZV-specific IgG were performed using a commercial enzyme linked immunosorbent assay (ELISA) assay, according to the manufacturer's guidelines. For each subject, the antibody count (evaluated quantitatively as an antibody concentration and expressed as an optical density (OD) measured in mUI/mL) and the age in completed years were collected. The age range was 0–79 years, and the sample size equal to 2517. Children under 10 years old were oversampled.

Data for measles (MeV) were collected between 2005 and 2006 in Tuscany (Italy), during routine laboratory testing, from 927 subjects, aged 1-49 years. For each subject, the antibody count and the age were collected. A commercial ELISA assay was employed for detection of measles IgG antibodies. Since the analysed sera were anonymous, it was not known whether subjects had been vaccinated or not before.

Mixture modelling of pre-vaccination data: the case of varicella in Italy

We use Bayesian mixture models [***11***] to estimate the age-specific seroprevalence and FOI for varicella directly from VZV antibody count data, thus model estimation and inference are carried out by using Bayesian Markov Chain Monte Carlo (MCMC) methods. We assume the population to be at demographic and epidemiological equilibrium. Since data were collected in absence of relevant vaccination activities, we can safely assume that each serological sample is drawn from a population consisting of just two subpopulations, one for the susceptible and one for the immune individuals. We then assume that the individual antibody count, after a logarithmic transformation, i.e., $\mathrm{Y_i} = \log_{10}(\mathrm{OD_i} + 1)$, is distributed as a mixture of two skew normal distributions, with mixture weights depending on the age a of the individuals,

$$Y_i(a) = \big(1 - \pi(a)\big)SN(Y_i|\mu_1, \sigma_1^2, \alpha_1) + \pi(a)SN(Y_i|\mu_2, \sigma_2^2, \gamma_2),$$

where $\mu_k, \sigma_k^2, \gamma_k, k = 1{,}2$ denote the mean, the variance, and the skewness parameters of the two mixture components, respectively. These parameters are assumed to be age-independent, which means that the antibody level distribution per component is assumed to be the same for all the groups. The skew normal distribution generalises the normal distribution by allowing for skewness through a specific parameter [***12, 13***] and is chosen to account for possible skewness in the two components. Indeed, a positive (negative) value of γ implies a distribution skewed to the right (left), thus a distribution with an excess of extremely high (low) antibody counts. The mixture weight of the immune component, $\pi(a)$, represents the age-specific seroprevalence, which is the expected proportion of immune individuals at exact age a in the given population [***14***]. In our analysis, we stratify the sample by one-year age groups up to age 60 (or up to the higher available age, if smaller than 60), while samples from people 60+ years are merged into one age group. Doing so, we can assess changes in seroprevalence in more detail, whilst reducing the risk of merging groups with possibly different serological patterns. Moreover, we exclude from the analysis all the individuals aged less than one year because of the presence of maternal antibodies up to six months of age and the absence of information on age in months in this group.

For the stochastic nonparametric model for the prevalence $\pi(a)$, we assume a Beta prior distribution for each age group j, namely, $\pi_j \sim Beta(\alpha_j, \beta_j)$, under the monotonically non-decreasing constraint $\pi_{j-1} \leq \pi_j \leq \pi_{j+1}$, where the hyper parameters α_j and β_j are given large non-negative prior distributions in the form of truncated normal distributions, i.e., $\alpha_j, \beta_j \sim TN_{[0,]}(0{,}10E6)$. Combining this prior function with the likelihood function in the data, we obtain the posterior distribution of the age-specific seroprevalence. Given the estimate of the seroprevalence, the FOI λ_j is successively estimated as $\pi_j'/(1 - \pi_j)$, where the first derivative of the prevalence at the numerator is approximated by $\big(4\pi_{j+1} - 3\pi_j - \pi_{j+2}\big)/2$, and then smoothed.

Mixture modelling of post-vaccination data: the case of measles in Italy.
With respect to pre-vaccination data, what usually happens with post-vaccination serological data is that the information on whether immune individuals acquired their immunity either from natural infection or from vaccination is not available. Hence, it is not possible to estimate the FOI of the natural infection, unless one makes an assumption about vaccination coverage or manages to estimate it [***15***].On the other hand, if one is not interested in estimating the FOI, there is no need to be constrained by the assumption of monotonicity for the seroprevalence. Moreover, the presence of individuals who acquired their immunity either from natural infection or from vaccination, with possibly different antibody levels, calls for the relaxation of the assumption of a mixture model with only two components, i.e., one for the susceptible and one for the immune. Rather, a mixture model with more than two components might provide a better fit to the data, as it allows multiple components to account for different measles antibody levels (even though this does not necessarily imply different immunity levels, as cell-mediated response also plays an important role in protection [***16***]).
For this scenario, we generalise the mixture model with age-dependent weights for pre-vaccination data by allowing K mixture components:

$$Y_i(a) = \sum_{k=1}^{K} \pi_k(a) SN\left(Y_i \middle| \mu_k, \sigma_k^2, \gamma_k\right),$$

where μ_k, σ_k^2, γ_k, $k = 1, \dots, K$ denote the mean, the variance, and the skewness parameters of the K mixture components, respectively. The current status of an individual is thus described by a binary latent classification random variable T_{ik} that represents the membership of the individual i to the specific component k. Since each individual can be classified only in one of the components, we have that $\sum_{k=1}^{K} T_{ik} = 1$. We also aim to classify each component k either as susceptible or immune. For this purpose, we assume that the sum of the mixture weights of the components that account for different levels of antibody response provides the overall probability that an individual is immune, therefore representing the population prevalence [***7***].
For the stochastic nonparametric model for the mixture weights $\pi_k(a)$, we do not assume any deterministic relationship between the probabilities and the age, but we rather assign a flat prior distribution for π_k at each distinct level of the age, consisting in a Uniform distribution in the range [0,1] for each component k. These K Uniform priors imply a Dirichlet prior distribution for each age group, i.e., $\left(\pi_1(a), \dots, \pi_K(a)\right) \sim Dir(\alpha_1 = 1, \dots, \alpha_K = 1)$, under the constraint that $\sum_{k=1}^{K} \pi_k(a) = 1$. Finally, model selection for the optimal number of components was based on two different measures, namely, the penalised expected deviance [***17***] and the difference in posterior deviances [***18***].
All statistical analysis is performed with R software and JAGS software.

Results

The pre-vaccination data, here from VZV in Italy, show a clear polarised distribution between the susceptible and the immune individuals, which explains why the results from the fixed cut-off are very close to those obtained from the mixture classification (***Figure 1***, top left). The estimate of the seroprevalence shows a steep linear increase in the first ten years of age, followed by a slower increase in the following years. This is reflected by the estimated FOI, which peaks between 5 and 10 years and then declines to a plateau for the following years (***Figure 1***, top right).
The post-vaccination data, here from MeV in Tuscany, show a completely different scenario. The distribution of the antibodies is not a clearly polarised as it was for the pre-vaccination data, rather a continuous distribution that increases in density as the antibody levels rise (***Figure 1***, bottom left). Model selection performed with the penalised expected deviance favours the model with three components to the one with four components, while the difference in posterior deviances does not find any statistical difference between

the two models. In ***Figure 1***, we show the fit of the 3-component mixture model to the histogram. It shows that the first component, with the lower mean, accounts for the susceptible individuals, while the following two components account for increasing levels of immunity [***6, 7***]. The seroprevalence of measles infection is also very different from the one of varicella. The vaccination against measles within 24 months of age determined the high immunity levels in the first years of age, up to 10 years. Similarly, immunity to measles is very high among the adults, from 25 years onward. Instead, among the adolescents, we notice a big drop in the prevalence, which approximately falls to 30% at around 15 years of age.

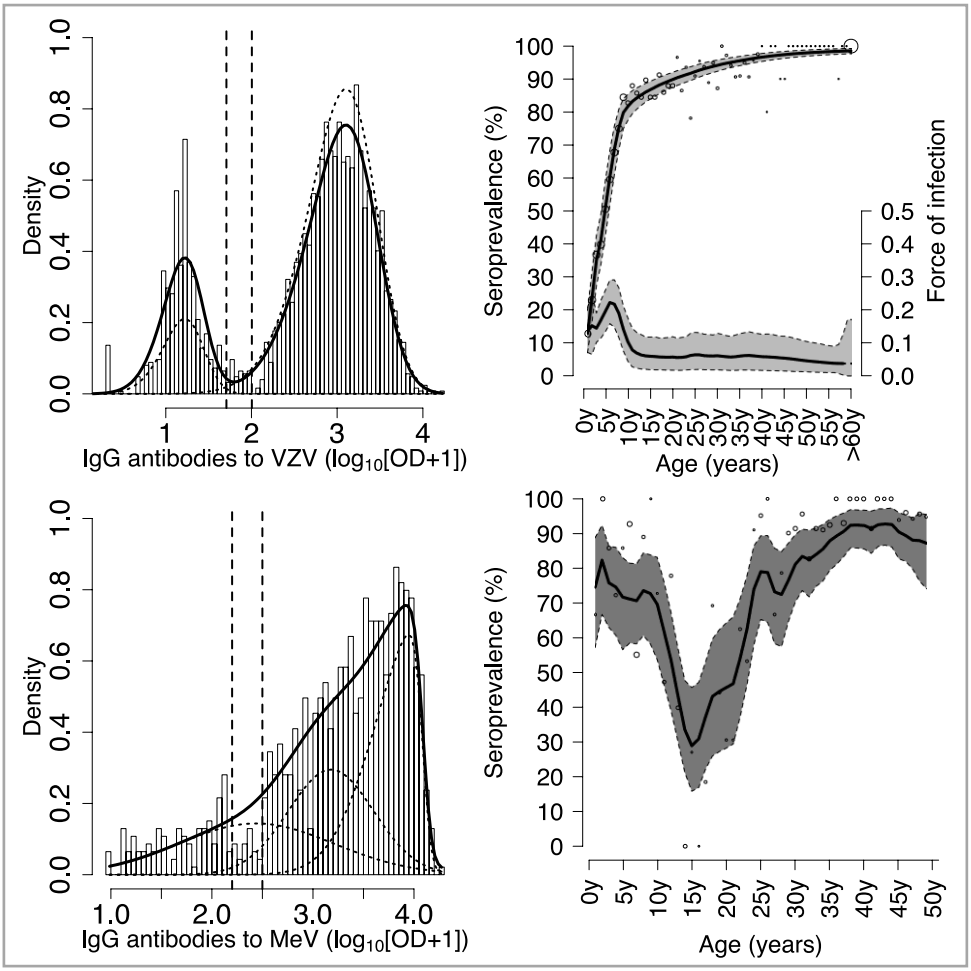

Figure 1 *Histogram of the distribution of the IgG antibodies, with overlapping fitted mixture model, and posterior mean of the epidemiological parameters of interest. Top row: Histogram of IgG antibodies to VZV in Italy (2004), with estimated mixture model (top left); posterior mean of the seroprevalence (upper curve) and of the force of infection (lower curve), with 95% credible intervals, overlapped on the estimated proportions seropositive (top right). Bottom row: Histogram of IgG antibodies to MV in Tuscany (2005-2006), with estimated mixture model (bottom left); posterior mean of the seroprevalence, with 95% credible intervals, overlapped on the estimated proportions seropositive (bottom right).*

Conclusions

In this work, we employed Bayesian mixture models to estimate key epidemiological parameters, such as the seroprevalence and the force of infection (the latter for pre-vaccination data only) directly from antibody levels. Contrarily to the fixed cut-off approach, which leads to the estimation of these parameters based on binary infection status data, the mixture model is a data-driven approach and, as such, it adapts directly to the antibody data, without losing the information contained therein. Rather, the method allows for both the classification of individuals among different serological groups (by giving each subject an age-dependent probability of belonging to a specific group), also of those cases that are classified as inconclusive under the fixed cut-off approach (those contained between the two cut-off lines in the histograms), and the estimation of the epidemiological parameters of interest. Moreover, the employment of a Bayesian MCMC approach allows to derive credible bands around all the model parameters.

The methodology is flexible enough to fit both pre-vaccination and post-vaccination serological data, which we have shown to represent quite different scenarios. Our findings show that, under the pre-vaccination scenario, the two approaches provide quite similar results in terms of detection of true seropositive individuals, as the groups of susceptible and immune are quite polarised. On the other hand, under the post-vaccination scenario, the separation between the serological groups is much more blurred due to the presence of vaccinated and naturally infected individuals. In this case, an approach based on the fixed cut-off might hide important differences in the immunity levels between individuals. Previous work showed how people immune to measles may be classified either as “high responders”, when they show high levels of antibodies, perhaps a sign of natural infection or boosting of immunity due to contact with infected individuals, or as “weak responders”, namely, people with lower immunity levels, perhaps due to lack of response to vaccination [***19***] or to a weak response to the vaccination [6,7]. Under the fixed cut-off approach, some of these latter individuals would have likely be classified as inconclusive, since lying between the two fixed cut-offs.

For all these reasons, we really think that the mixture model approach should be seriously considered as the best way of analysing data from serological surveys. As a consequence of this, we claim that more attention should be devoted to the design of the serological surveys, both for what concerns the determination of the sample size by age and the measurement of the antibody concentrations, in particular as regards the measurement of the lower antibody counts. Indeed, these latter data might contribute important information towards the discrimination between susceptible or weak responders to immunity and thus help to better inform health policies aimed to those individuals with poor protection against the infection under study.

References

[**1**] Carrillo-Santisteve P, Lopalco P. Measles still spreads in Europe: who is responsible for the failure to vaccinate? Clin Microbiol Infect 2012;18:50–6.

[**2**] Guzzetta G, Poletti P, Merler S, et al. The Epidemiology of Herpes Zoster After Varicella Immunization Under Different Biological Hypotheses: Perspectives From Mathematical Modeling. Am J Epidemiol 2016;183:765–73.

[**3**] Metcalf CJE, Farrar J, Cutts FT, et al. Use of serological surveys to generate key insights into the changing global landscape of infectious disease. Lancet 2016;388:728–30.

[**4**] Vyse AJ, Gay NJ, Hesketh LM, et al. Seroprevalence of antibody to varicella zoster virus in England and Wales in children and young adults. Epidemiol Infect 2004;132:1129–34.

[**5**] Vyse AJ, Gay NJ, Hesketh LM, et al. Interpreting serological surveys using mixture models: the seroepidemiology of measles, mumps and rubella in England and Wales at the beginning of the 21st century. Epidemiol Infect 2006;134:1303–10.

[**6**] Rota MC, Massari M, Gabutti G, et al. Measles serological survey in the Italian population: Interpretation of results using mixture model. Vaccine 2008;26:4403–9.
[**7**] Del Fava E, Shkedy Z, Bechini A, et al. Towards measles elimination in Italy: Monitoring herd immunity by Bayesian mixture modelling of serological data. Epidemics 2012;4:124–31.
[**8**] Del Fava E, Mirinaviciute G, Flem E, et al. Estimating Age-Specific Immunity and Force of Infection of Varicella Zoster Virus in Norway Using Mixture Models. PLoS ONE 2016;11:e0163636–12.
[**9**] Nardone A, de Ory F, Carton M, et al. The comparative sero-epidemiology of varicella zoster virus in 11 countries in the European region. Vaccine 2007;25:7866–72.
[**10**] Mossong JR, Hens N, Jit M, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS Med 2008;5:381–91.
[**11**] Diebolt J, Robert CP. Estimation of finite mixture distributions through Bayesian sampling. J R Statist Soc Ser B 1994;56:363–75.
[**12**] Azzalini A. A class of distributions which includes the normal ones. Scand J Stat 1985;12:171–8.
[**13**] Frühwirth-Schnatter S, Pyne S. Bayesian inference for finite mixtures of univariate and multivariate skewnormal and skew-t distributions. Biostatistics 2010;11:317–36.
[**14**] Evans RB, Erlandson K. Robust Bayesian prediction of subject disease status and population prevalence using several similar diagnostic tests. Statist Med 2004;23:2227–36.
[**15**] Goeyvaerts N, Hens N, Theeten H, et al. Estimating vaccination coverage for the trivalent measles-mumpsrubella vaccine from trivariate serological data. Statist Med 2012;31:1432–49.
[**16**] Isa MB, Pavan JV, Sicilia Don P, et al. Persistence of measles neutralizing antibody related to vaccine and natural infection acquired before HIV infection. Epidemiol Infect 2013;142:1708–12.
[**17**] Plummer M. Penalized loss functions for Bayesian model comparison. Biostatistics 2008;9:523–39.
[**18**] Aitkin M. Statistical Inference. CRC Press; 2010.
[**19**] Mossong JR, Putz L, Schneider F. Seroprevalence and force of infection of varicella-zoster virus in Luxembourg. Epidemiol Infect 1999;132:1121–7.

LINKING FOODS CONSUMED AND EATING LOCATION IN UK TEENAGERS: EVIDENCE FROM THE NATIONAL DIET AND NUTRITION SURVEY 2008-2012.

Palla Luigi[1], Chapman Andrew N[1], Pot Gerda K[2], Almiron-Roig Eva[3]

1. *Dept Medical Statistics, LSHTM, London*
2. *Diabetes & Nutritional Sciences Division Faculty of Life Sciences & Medicine, King's College London and*
3. *MRC HNR Cambridge*

Introduction

Obesity is a condition of great concern internationally and particularly in developed countries where the growing longevity of populations implies a steadily increasing burden of chronic diseases like associated to obesity. Besides increasing health-care costs, obesity reduces the quality of life and impacts on economic activity.To counter these trends, several countries have devised policies to combat the development of poor eating habits amongst young people and adults. In Britain, where the prevalence of adult obesity is higher than the OECD average, some policies were drafted as early as 2009 (" Healthy Lives, Healthy People" [***1***]). Promoting a healthy lifestyle and discouraging unhealthy habits associated to weight gain from an early age is crucial as there is evidence that, once obesity is established, it continues into adulthood. Even if obesity is a highly heritable character, the investigations conducted on overweight/obesity point at the environment as a major determinant and likely primary focus of intervention policies. The environment comprises many elements, including all influences that can be regarded as the (social, economic, psychological, physical, geographical, political) background context in which behaviour takes place and which needs to be investigated. As a contribution to characterising such context, our paper addresses the question on whether there is a relationship between different types of foods (healthy/unhealthy) and the physical location where the food is consumed. To do so we make use of the United Kingdom National Dietary and Nutrition Survey Rolling Programme (NDNS-RP) 2008-2012 database, focusing on the subgroup of respondents aged 11 to 18 years, as we assume that dietary habits acquired in teenage might settle long term and potentially be most relevant targets for intervention.

Our aims in this work were to exploit the rich dataset from NDNS-RP 2008-2012 to examine the relationship between foods and eating locations. This is thanks to NDNS collecting information on individual eating occasions for all respondents on both the food consumed and the place where it was consumed. We aim to explore the association between foods, especially when nutritionally classified as relatively healthy or unhealthy, and the location where they are eaten and then assess if there is evidence that teenage are more likely to eat such healthy or unhealthy foods in certain locations than in others.

Methods

NDNS-RP 2008-2012 data

The NDNS sample was drawn from the UK Postcode Address File, a list of all the addresses in the UK. The addresses come from small geographical areas based on postcode sectors, randomly selected from across the UK. A list of 27 addresses was then randomly selected from each postcode sector. In total, 21,573 addresses from 799 postcodes in the UK were randomly selected for the survey between April 2008 and March 2011. The randomly selected individuals were asked to complete a detailed diary of their food and drink consumption over four consecutive days and an interview was conducted to collect background information on dietary habits, socio-demographic status, lifestyle and physical activity. The response rate for completion of the diary and interview was 56%. A total of 884 teenagers aged 11 to 18 years completed four days of the food and drink diary. Strata were used in the NDNS to calibrate proportions in the sample

with the whole population [***2***]. The strata were based on age-group, sex and geographical region. The weighting system used by NDNS involved two steps designed to compensate for sampling selection probabilities and reduce bias resulting from differential non-response by age, sex and region. The 884 teenagers, aged between 11 and 18, provided a total of 62,523 diary records. The mean number of diary entries covering four days for an individual teenager is 71. The diary entries for an individual are not independent of each other. In particular we envision a two-level hierarchy originating correlation among eating occasions (measured on 4 different days), with the first level being meal-time and the second level the individual adolescent.

We conducted all analyses purely on instances, ignoring portion sizes (and consequently calories) and concentrating on whether a food was consumed at a location, not how much of it was consumed at a location. This reduced, but did not remove, the known problem of under-reporting consumption in dietary surveys [***3***].

Classification of Food Groups and Locations

The data collected in food diaries was dis-aggregated and extended by the NDNS team to provide a nutritional analysis of each diary record before being made available as the NDNS database. In particular the NDNS database classified the foods consumed and recorded into 59 main food groups [***4***], many of which contribute very little to teenagers' diets. In order to focus on the major foods for teenagers, the contributions to total teenager calorie intake of NDNS food-groups were ranked. Our analyses then concerns the food groups which contribute the top 80% of calories resulting in 25 food-groups. The 25th food-group (least contributor) contributes less than 1.5% of the total calories in the UK teenagers' diet.

We used the UK Food Standards Agency (FSA) scoring system which is detailed in the UK Government Nutrient Profiling Technical Guidance [***5, 6, 7***]. Using the nutrients breakdown provided in the NDNS diary records, we used the FSA algorithm to score the healthiness of each food diary entry by teenagers in the NDNS database. The mean FSA score for each food-group was then determined. We then sub-divided the P80 food-groups into the three categories: "healthy", "neutral" and "less-healthy" (as defined in [***8***]): healthy for FSA scores above (+4), neutral between (-2) and (+4) inclusive, and less-healthy for FSA scores below (-2)

Eating "location" were derived from the NDNS "where" codes as one of 7 categories: home, school, work, friend's/carer's/relatives' home, mobile, leisure or other. To simplify and clarify our results, we subsequently collapsed these into 3 categories: home, school or work, other.

Statistical Analysis

A contingency table was created from the food diary records giving the frequency of consumption of each food group at each location. Associations were first explored using "simple" correspondence analysis [***9, 10***] (Ringrose 2012) .Hypotheses generated via visual inspection of CA plots were then taken forward and formally tested by logistic regression via GEE [***11***]. The analyses were conducted using SAS software version 9.4 and R packages "CABOOTCRS" [***12***].

Correspondence Analysis (CA)

CA is a method for investigating the relationships between categories represented in a two-dimensional contingency table. It does this by analysing and plotting "profiles" , that is the relative frequency of consumption of a food across different locations or the relative frequency of consumption of different foods at one location. For example, if 71% of all foods consumed are eaten at Home, but 55% of all sweetened soft drinks are consumed at Home, then sweetened soft drinks will have a location "profile" different from the average food profile. CA plots show the chi squared deviation of profiles from the average profile.

To plot these multi-dimensional deviations (inertia) reduced to the two most informative dimensions, we used biplots, where row profiles are normalised (rescaled) but column profiles are not (or vice versa). The horizontal axis is the direction along which the contingency table rows and columns show their largest deviation. The vertical axis represents the direction – perpendicular to the first - having the second-largest

deviations. The percentage label for each axis is a measure of how much of the total variation (inertia) in the data has been displayed along that axis. The sum of the variation shown by the two axes is not 100%: the remaining variation would require more dimensions to display, and so is lost when we reduce to 2 dimensions. The origin in each plot represents the average profile of those points in the plot. In CA plots, the length of the vector from the origin to any profile-point represents its deviation from the average-profile. In biplots, the distance between row (food) and column (location) profile points and the direction in which they lie away from the origin is an indication of their association (greater association if closer points/similar directions).
Confidence regions (95%) for profile points have also been used that are based on bootstrap [***12***] procedures. These Confidence Regions (CRs) were useful to discard foods that were not significantly different from the average profile (origin) if the region contained the origin.
Logistic regression with GEE
We used Generalised Estimating Equations (GEE) logistic regression to estimate Odds Ratios (ORs) as we are interested in reporting population average behaviour. GEEs provide unbiased estimates of ORs despite our ignorance of true correlation (induced by the hierarchy) structure of the data [***11***]. Since the correlation structure in the model cannot be guaranteed, the GEE model provided empirical estimates of standard errors for use in inference.

Results
The majority (71%) of diary entries record eating at home, with 14% at school or work and the remainder (15%) at other locations. The teenagers were made up of 445 boys and 439 girls aged between 11 and 18 inclusive. There were 20 missing values for the socio-economic classification: the missing values were evenly distributed between the sexes and randomly spread across the ages.
Using a random process to split the diaries dataset for the P80 foods resulted in a hypothesis generating dataset of 20,567 diary records and a hypothesis testing dataset of 20,455 records. The independence of the 25 food-groups and the 7 eating locations was tested by calculating the χ^2 statistic for the P80 food-group and location contingency table. The χ^2 value exceeded 3,100 with 144 degrees of freedom, giving a p-value < 0.001 and strong evidence against independence.
Initial CA plot comprising all 25 foods showed larger deviations from the average food for Coated Chicken and Chips (french-fried potatoes) and Sweetened soft drinks which appeared in a “leisurely” direction, while Chocolates and Meat pies appeared towards Mobile and Other locations, and finally Crisps, Brown-bread and Biscuits towards School. The Home location attracts breakfast foods, and Pasta/Rice, and Vegetables.
When adding confidence regions the cluttered graph required creating separate CA plots according to the healthiness classification. This saw 3 foods classified as healthy (cooked vegetables, fruit and brown bread), 10 as neutral (pasta/rice, white bread, chips/potatoes, nondiet soft drinks, chicken/turkey dishes, semi-skimmed milk, other potato dishes, beef/veal dishes, high fibre breakfast cereals, fruit juice) and 12 as less-healthy (cheese,low-fibre breakfast cereals, sausages, coated chicken, meat pies/rolls, sugars/preserves/sweet spreads,low fat spreads, biscuits, crisps/ savoury snacks, chocolate, cakes/pastries, other unclassified food).
The CA Biplots for healthy food suggest that cooked vegetable tend to be associated with home, brown bread with school and fruit with mobile or other locations. The three “healthy” foods have CRs not including the origin. The Work location and Friend’s/Carer’s Home location both have CRs covering the origin which shows that they do not differ significantly from the average of all locations of healthy eating occasions.
The CA Biplots for neutral foods indicate that 94.1% of the inertia is summarised in the plot: with 69.9% along the horizontal axis of the plot where we find a contrast between Home and all other eating locations with Leisure and Mobile the most dis-similar to Home. The vertical axis contrasts chips and soft-drinks with white-bread and fruit-juice. These biplots for neutral foods suggest associations of School and Work with

Chicken dishes, Fruit juice and White-bread and of Leisure and Mobile locations with Chips and Sweetened Soft Drinks.
Beef and chicken dishes can be seen to have CRs which include the origin whereas chips, and non-diet soft drinks have CRs far from the origin.
The CA Biplots for less-healthy foods (***Figure 1***) indicate that 88.22% of the variation in location profiles is summarised in this plot, with 66.82% along the horizontal axis where Home is contrasted with all other eating locations, and breakfast foods are contrasted with snack foods: crisps, biscuits, cakes. The vertical axis again presents a contrast between fast foods and sandwich items.
These plots suggest that 1) cheese, less-fat spreads and biscuits are associated with School and Work; 2) that Crisps and "Cakes & sweet-pastries" are associated with non-Home locations; 3) that Coated Chicken is associated with Leisure locations; 4) and that Meat-Pies and Chocolate appear associated with Friend's & Carer's homes, Other and Mobile locations. The Chocolate CR appears entirely inside the CR for Meat-Pies.

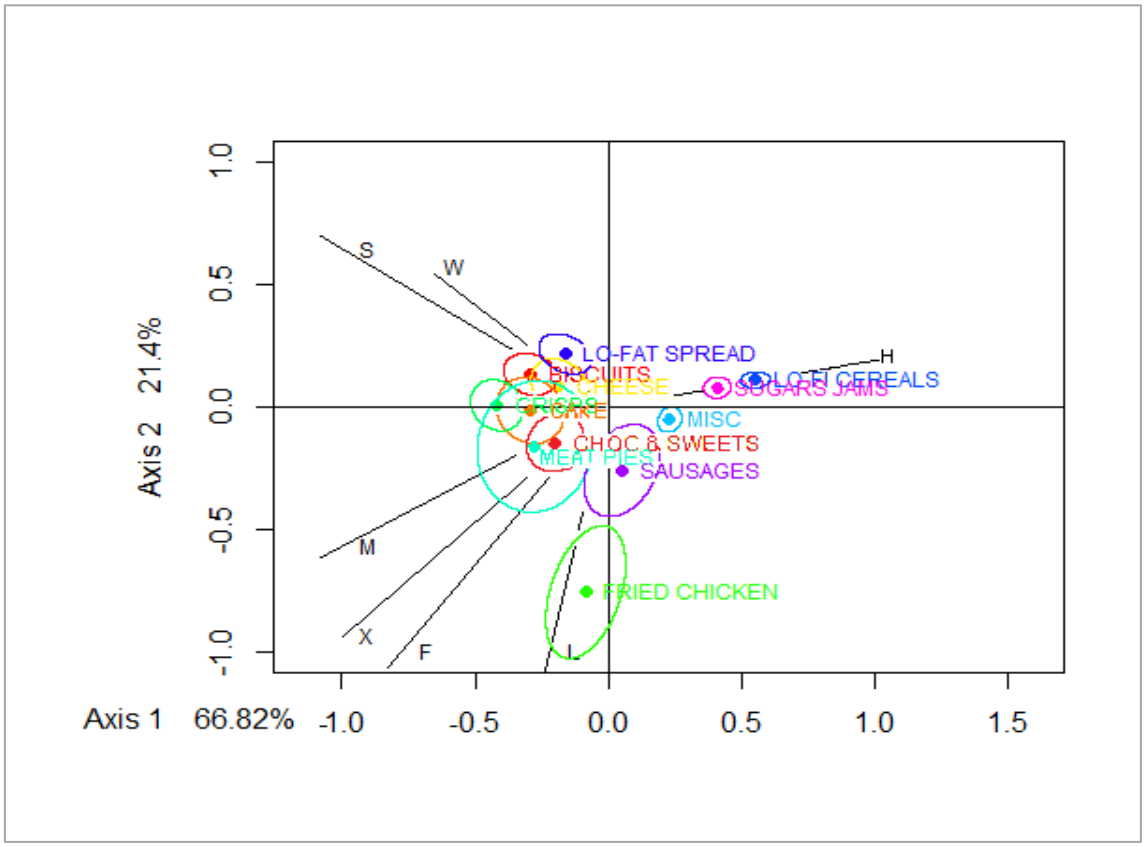

Figure 1. Biplot of locations and showing CRs for **Less-Healthy** food-groups (using CABOOTCRS)
Legend: H-Home S-School W-Work F-Friends/Carers Homes L-Leisure M-Mobile X-Other

Throughout the CA plots above, the locations appear to have aligned themselves together in three types which we have used to simplify the next stage (hypothesis testing): Home, School and Work, all Other locations. In particular Leisure, Mobile, Other, Friend's & Carer's homes often found themselves in the same quadrant of the plot. These locations are relatively less described in the literature on eating behaviour so we focused on these as opposed to Home and School/Work in order to test hypotheses on the odds ratios of

eating particular food-groups (emerged to be associated with at least some of them: meat pies, chocolate, fruit, chips and non-diet soft drinks) in those Locations compared to Home and to School/Work.
Table 1 shows that there is strong evidence that the average teenager is more likely to eat sweetened soft drink, chips, chocolate and meat pies at "Other" locations rather than at Home or at School/Work. These results are statistically significant except when comparing the odds of having meat pies at other location versus school/work.
The estimates for fruit show that the odds of eating fruit at "Other" locations are significantly lower than at Home, and less than half the odds of eating fruit at School or Work.
Besides gender and weekend days had no significant effect on the odds ratios reported while age and socio-economic status included as continuous covariates showed significant trends on consumption: it was 15% (p-value=0.002) more likely to eat meat pies and 5% less likely (p=0.018) to eat fruit at Other Location than at Home for each extra year of age and it was 6% more likely (p=0.0005) to drink sweetened soft drinks, 7% more likely (p=0.001) to eat chips and 6% less likely (p=0.004) to eat fruit at Other Location than at Home for each step lower in social class.

Table 1. Odds Ratio estimates of eating types of food in Other Locations versus at Home or versus at School/Work, adjusted by age, sex, weekend days and socio-economic status.

	Estimated, adjusted Odds Ratio for food-group at Other_locations			
Food-Group	OR vs **Home**	99% CI p-value	OR vs **School-Work**	99% CI p-value
1. **Sweetened soft drinks**	2.9	(2.3, 3.5) p<0.0001	2.3	(1.7, 3.1) p<0.0001
2. **Chips** (french-fried potatoes)	2.8	(2.2, 3.6) p<0.0001	3.4	(2.1, 5.3) p<0.0001
3. **Chocolate**	2.5	(1.8, 3.4) p<0.0001	1.8	(1.2, 2.8) p=0.0002
4. **Meat Pies**	2.8	(1.5, 5.0) p<0.0001	1.3	(0.6, 3.0) p=0.44
5. **Fruit**	0.70	(0.49, 0.98) p=0.006	0.44	(0.29, 0.67) p<0.0001

Conclusions

The use of Correspondence Analysis (CA) with confidence regions has facilitated a broad and systematic investigation of major food-groups at all locations. This allowed the generation of hypotheses concerning possible relationships, and these hypotheses have been tested using traditional statistical methods. An advantage of this approach was that of reducing the risk of obtaining significant results by chance should a multiplicity of tests be carried out. The results point at high odds ratios of eating unhealthy foods in other locations compared to home. Interpretation should be careful and take into account that these are relative results and don't provide any information on the absolute number of eating occasions in various locations (for example 55% - that is most- of the soft drink eating occasions occurs at home anyway). Besides the data based on eating occasions rather than quantity consumed. This however has the advantage that the negative effects of underreporting are likely to be reduced as it is more likely that smaller amounts of consumption are reported than reporting is omitted altogether.
As concerns the specific findings, we noticed that the chocolate CR is plotted entirely within the CR for meat-pies. The Meat-pies food group includes sausage-rolls, Cornish-pasties, and meat-pastries which are all convenient foods to carry and eat at any location. The same applies to chocolate confectionery. Besides both meat-pies and chocolate are available at low prices. Individual sausage rolls and small meat-pastries are available on every high street at a price-point below ≤1: the same applies to chocolate. A group of four teenagers at lunchtime or after school can put together ≤5 to share a promotion pack of chocolate, and share a promotion bundle of warm sausage rolls from a well-known bakery-chain retailer, and also share a large 2 litre bottle of sweetened soft-drink – all within their budget; Patterson [***13***] noted that teenagers aged

11-14 years in a deprived London Borough, spent a median of ≤2 and an upper quartile of ≤3 when buying fast food, and that taste, quick access and peer influence were major factors in the choice of food.
However, one difference between meat-pies and chocolate is that there is a rising trend in meat-pie consumption as age increases.
Sausage rolls and meat-pastries are also available from school lunch counters, but chocolate at school will not be at a promotional price if available at all. This may explain why the odds for eating meat-pies at School/Work locations is comparable to Other locations, whereas Chocolate is not (at the 1% level).
We also found that teenagers from families with a lower socio-economic class had higher odds of consuming chips and sweetened soft-drinks at Other locations. This is consistent with previous findings: Patterson [***13***] has noted that more fast food outlets are found in areas of lower socio-economic status, and Pechey [***8***] found that lower socio-economic classes purchased a greater proportion of their energy from less healthy foods and beverages.
In conclusion, we provided quantitative results about the relationship between foods and eating location that may help the policymakers to inform appropriate interventions. They have tools to influence the cost of foods, the advertising of foods and the promotion of healthy eating, and could use this evidence to consider incentives to make more healthy choices available and attractive to teenagers when they are "grazing" for food at "Other" locations away from adult supervision.

References

[**1**] HM Government (2009). Our Health and Wellbeing Today. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/215911/dh_122238.pdf. [Accessed 2017].
[**2**] NDNS_Appendix_B (2011). Weighting the Core Sample. Available from: www.gov.uk/government/uploads/system/uploads/attachment_data/file/216487/dh_130786.pdf. [Accessed 2017].
[**3**] NDNS_Appendix_X (2011). Mis-reporting in the NDNS: DLW Study. Available from: www.food.gov.uk/sites/default/files/ndns-appendix-x.pdf. [Accessed 2017].
[**4**] NDNS_Appendix_R (2011). Main and Subsidiary Food Groups. Available from: www.food.gov.uk/sites/default/files/ndns-appendix-r.pdf. [Accessed 2017].
[**5**] Nutrient Profiling Technical Guidance (2011). F.F.S. Agency. Available from: http://www.food.gov.uk/sites/default/files/multimedia/pdfs/techguidenutprofiling.pdf. [Accessed 2017].
[**6**] Scarborough P, Rayner M, Stockley L. Developing nutrient profile models: a systematic approach. Public Health Nutrition 2007;10(4):330-6.
[**7**] Rayner M, Scarborough P, Kaur A. Nutrient profiling and the regulation of marketing to children. Possibilities and pitfalls. Appetite 2013;62:232-5.
[**8**] Pechey R, Jebb SA, Kelly MP, et al. Socioeconomic differences in purchases of more vs. less healthy foods and beverages: Analysis of over 25,000 British households in 2010. Social Science & Medicine 2013;92:22-6.
[**9**] Greenacre MJ. Correspondence Analysis in Practice. 1993 London, Academic Press.
[**10**] Beh EJ, Lombardo R. Correspondence Analysis: Theory, Practice and New Strategies, 2014 Wiley.
[**11**] Liang KY, Zeger SL. Longitudinal data-analysis using generalized linear-models." Biometrika 1986;73(1):13-22.
[**12**] Ringrose TJ. Bootstrap confidence regions for correspondence analysis. Journal of Statistical Computation and Simulation 2012;82(10):1397-1413.
[**13**] Patterson R, Risby A, Chan MY. Consumption of takeaway and fast food in a deprived inner London Borough: are they associated with childhood obesity? BMJ Open 2012;2(3):e000402.

ASSESSMENT OF ADHERENCE TO MEDITERRANEAN DIET WITH DIFFERENT *A PRIORI* SCORES IN ITALIAN BREASTFEEDING WOMEN

Bravi Francesca[1], Turati Federica[1], Decarli Adriano[1], Wiens Frank[2], Ferraroni Monica[1]

1. *Laboratory of Medical Statistics, Biometry and Epidemiology "G.A. Maccacaro", Department of Clinical Sciences and Community Health, Università degli Studi di Milano, Milan, Italy*
2. *Nutricia Research, Utrecht, Netherlands*

Introduction

The traditional Mediterranean diet typical of the populations from areas bordering the Mediterranean Sea is characterized by a frequent consumption of fruit and vegetables, cereals, legumes, and fish, a limited consumption of meat and dairy products, a moderate consumption of alcoholic beverages, and the use of olive oil as preferred fat. A favourable role of this dietary pattern on different health outcomes, including cardio and cerebrovascular diseases, and various types of cancers has been observed in the last decades in different populations [***1***]. More recently, the Mediterranean diet has been proposed as a model of sustainable diet, although studies on this issue are still scant [***2***].

Various *a priori* scores have been proposed to quantify adherence to Mediterranean diet, differing in terms of foods/nutrients considered, type and number of cut-offs discriminating adherence for each component, and mathematical structure of the score. To date, there has been no systematic attempt to evaluate the degree of agreement of different scores in measuring the adherence to Mediterranean diet, and, in the absence of an objective gold standard, no consensus has been reached on the best score to be adopted. Nonetheless, these scores are largely used in epidemiological studies and most of them were favourably related to various health outcomes.

The aim of the present work is to assess how much different Mediterranean diet scores agree in ranking and classifying subjects according to their adherence to Mediterranean diet.

Methods

This work was conducted within the MediDiet study, an Italian multicentric observational study carried out to compare human milk composition of breastfeeding mothers who were adherent to the Mediterranean diet with that of non-adherent breastfeeding mothers. Between October 2012 and June 2014, 300 women were recruited during the perinatal visits in five hospital centres (Torino N=110, Florence N=23, Rome N=46, San Giovanni Rotondo N=101, and Palermo N=20). About 6 weeks post partum, the participants provided information on a number of socio-demographic and lifestyle factors, including their dietary habits during the post partum period, and donated a sample of their breast milk for biochemical analysis. In particular, we used a food frequency questionnaire (FFQ) with proven validity and reproducibility [***3, 4***] to assess usual maternal dietary intake. The FFQ investigated the average weekly consumption of 78 foods and beverages. Further questions concerned alcohol consumption, and types of fat and oil used. Occasional consumptions (i.e. at least once per month and less than once per week) were coded as a frequency of 0.5 per week. Macro- and micronutrient intakes, as well as total energy intake, were estimated using an Italian food composition database and other sources when needed [***5, 6***]. Answers to questions on fat intake pattern, as well as portion size, were used to modulate the composition of recipes.

From on a preliminary review of the literature, we identified 16 scores measuring adherence to the Mediterranean diet, differing with respect to: 1) the choice of specific components of the score (nutrients, foods, foods groups); 2) cut-offs discriminating the level of adherence in each component (number of levels of adherence; *a priori* defined versus study-specific thresholds); 3) scoring system for the level of adherence

(0/1, or finer scoring systems, depending on the number of the levels) in each component; 4) algorithm used to calculate the overall score (e.g. sum, ratios of absolute or standardized components).
We selected and applied on our data 5 of these 16 scores, with different characteristics in terms of the aforementioned points: the Dietary Score (DS) proposed by Panagiotakos [***7***], the Mediterranean Diet Score (MDS) proposed by Trichopoulou and colleagues [**8**], the Italian Mediterranean Index (IMI) proposed by Agnoli [***9***], the Mediterranean Adequacy Index (MAI) proposed by Alberti-Fidanza [***10***], and the Mediterranean Dietary Pattern (MDP) proposed by Sanchez-Villegas [***11***].
The DS is the sum of 11 components, each scored 0 to 5, based on *a priori* defined cut-offs of frequencies of consumption: 7 Mediterranean components (i.e. whole cereals, potatoes, fruit, vegetables, legumes, fish, and olive oil), and 3 non Mediterranean components (i.e. red and processed meat, poultry, dairy products), plus moderate alcohol consumption. Thus, the DS ranges from 0 to 55.
The MDS is the sum of 9 binary components, based on study-specific cut-offs: 1 point is assigned for a consumption above the study-specific median for components typical of the Mediterranean Diet (i.e. cereals, fruit and nuts, vegetables, legumes, fish, monounsaturated/saturated fatty acids ratio), and 0 otherwise; the reverse applies to non-Mediterranean components (i.e. meat, and dairy products); for alcohol, *apriori* defined cut-offs were used to credit moderate consumption (between 5 and 25 g of ethanol per day for women). Thus, the MDS ranges from 0 to 9, with higher scores indicating higher adherence to MD.
The IMI is the sum of 11 binary indicators defined according to study-specific cut-offs: it assigns 1 point for consumption above the highest tertile of 6 Mediterranean components (pasta, typical Mediterranean vegetables, fruit, legumes, olive oil, fish) and below the lowest tertile of 4 non-Mediterranean components (soft drinks, butter, red meat, potatoes), and 1 point for moderate alcohol intake (>0 to 12 g/day of ethanol). Thus, the IMI ranges from 0 to 11.
The MAI is the ratio between the sum of the percentage of energy derived by typical Mediterranean foods (bread, cereals, legumes, potatoes, vegetables, fruit, fish, red wine, and vegetables oils) and the sum of percentage of energy derived by non-Mediterranean foods (milk, cheese, meat, eggs, animal fats and margarine, soft drinks, cakes, cookies, and sugar). Thus, the MAI ranges from 0 to $+\infty$.
The MDP is the sum of standardized energy-adjusted intakes of Mediterranean foods/nutrients (legumes, cereals, fruit, vegetables, monounsaturated to saturated fatty acids ratio) minus the standardized intakes of non-Mediterranean foods/nutrients (meat, dairy products, and trans-fatty acids), plus standardized moderate alcohol intake. For alcohol, a transformation is applied before standardization, to obtain the highest value for men consuming 30 gr/day and 20 gr/day for women, and lower values for progressively increasing/decreasing intakes. The MDP is then expressed as a relative percentage of adherence using the range of the values in the sample. Thus, the score ranges between 0% and 100%.
All the scores define *a priori* the components to be included; only one score categorize the components according to *a priori* thresholds, whereas the scoring systems of the remaining scores relies on the intake distribution of the specific population.
We calculated for each subject the 5 selected scores. In order to assess the degree of agreement of these score, we computed Spearman's correlation coefficients between pairs of scores. Moreover, for each score we computed (approximate) tertiles to classify the subjects as low/moderate/high adherent to the Mediterranean diet. Weighted Kappa and the proportion of observed perfect agreement (i.e. women in the diagonal cells of the 3x3 table) and partial agreement (i.e. women in the cells adjacent to the diagonal) were calculated between pairs of categorized scores. In a further analysis, simple Kappa was calculated for a binary classification, i.e. adherent versus non-adherent, corresponding to the third tertile and first plus second tertiles, respectively.

Results

Participants had a median age of 33 years (Q1-Q3: 30-36), and a median pre-pregnancy BMI of 21.7 kg/m^2 (Q1-Q3: 20.2-24.0). Forty-seven percent were primiparae, 40% gave birth to the second-born, and 13% had 3 or more children. Over 60% of women never smoked, 24% were ex-smokers, and 15% stopped smoking when became pregnant.

The ***Figure 1*** shows the distribution of the selected Mediterranean diet scores, scatter plots and Sperman's correlation coefficients between pairs of scores, the weighted Kappa, and the proportion of observed perfect and partial agreement between pairs of scores categorized in tertiles.

All the 5 scores were not normally distributed. Medians (ranges) were 25 (13-37) for DS, 4 (0-8) for MDS, 4 (0-10) for IMI, 1.9 (0.3-20.4) for MAI, and 53.5 (0-100) for MDP. Most scores did not reach the theoretical maximum, since almost all the participants did not report alcohol drinking, which is discouraged during breastfeeding, while a moderate alcoholic intake is credited by the scores. For MDS, DS, and IMI, subjects were not distributed equally across tertiles because of clustering of scores around certain values. Spearman's correlation coefficients were all positive and ranged from 0.24 for DS-MAI to 0.64 for MDP-MDS, indicating weak to moderate concordance between pairs of scores in ranking individuals according to their adherence to Mediterranean diet. When considering categorized scores, modest agreement was found for most pairs of scores, being the weighted Kappa 0.14 for MAI-DS, 0.19 for MAI-IMI, 0.24 for IMI-DS, 0.27 for MAI-MDS, 0.30 for IMI-MDP, 0.33 for MAI-MDP, and 0.34 for MDS-DS and IMI-MDS. Higher Kappa values were found for MDP-DS (0.40) and MPD-MDS (0.46). The percentage of subjects ranked into the identical tertile (perfect observed agreement) ranged between 38.7% when considering DS-MAI to 56.3% when considering MDP-MDS (mean perfect agreement: 48.6%). Focusing on the different tertiles, for all the pairs of scores we observed a higher perfect agreement within the first and third tertiles (varying between 15% and 24% for the first and between 15% and 22% for the third), and a lower one within the second tertiles (varying between 7% and 12%).

When we evaluated the 5 categorized scores overall, about 7% and 5.7% of women were consistently classified respectively in the top and bottom tertiles of adherence by all the Mediterranean diet scores. When the MAI score was not considered, corresponding percentages were 10.7% and 12.0%.

Furthermore, when women were classified as adherent (third tertile) and non adherent (first and second tertiles combined) to Mediterranean diet, and pairs of different scores were compared, still the degree of agreement measured by Kappa appeared modest for most pairs (ranging from 0.11 for MAI-DS to 0.51 for MDP-MDS).

Discussion

Mediterranean diet scores have been developed to measure, at least in theory, the same concept, i.e. adherence to the traditional dietary profile of the Mediterranean populations; thus, one would expect a high correlation between scores and a good agreement in classifying the level of adherence to Mediterranean diet. However, we observed weak to moderate concordance in most cases. This firstly reflects the lack of a precise and quantified definition of Mediterranean diet, as pointed by the different foods/nutrients considered in the scores' formulation. Reaching a consensus on this definition is however difficult, since Mediterranean diet is not a homogeneous model within the Mediterranean area and across time period. Part of the observed disagreement is further explained by the different scoring systems and algorithms used to compute the overall scores. In any case, in our current application, scores with better concordance, i.e., MDS, MDP and DS, appear to share most of the components, with minor variations, despite major differences in computational aspects. In particular, MDS relies on the sum of binary components defined by study-specific medians, MDP uses the sum of 6-levels components defined according to *a priori* cut-offs, while MDP has a more complex algorithm based on sums/subtractions of standardized (and hence depending on study-specific distributions) energy-adjusted intakes.

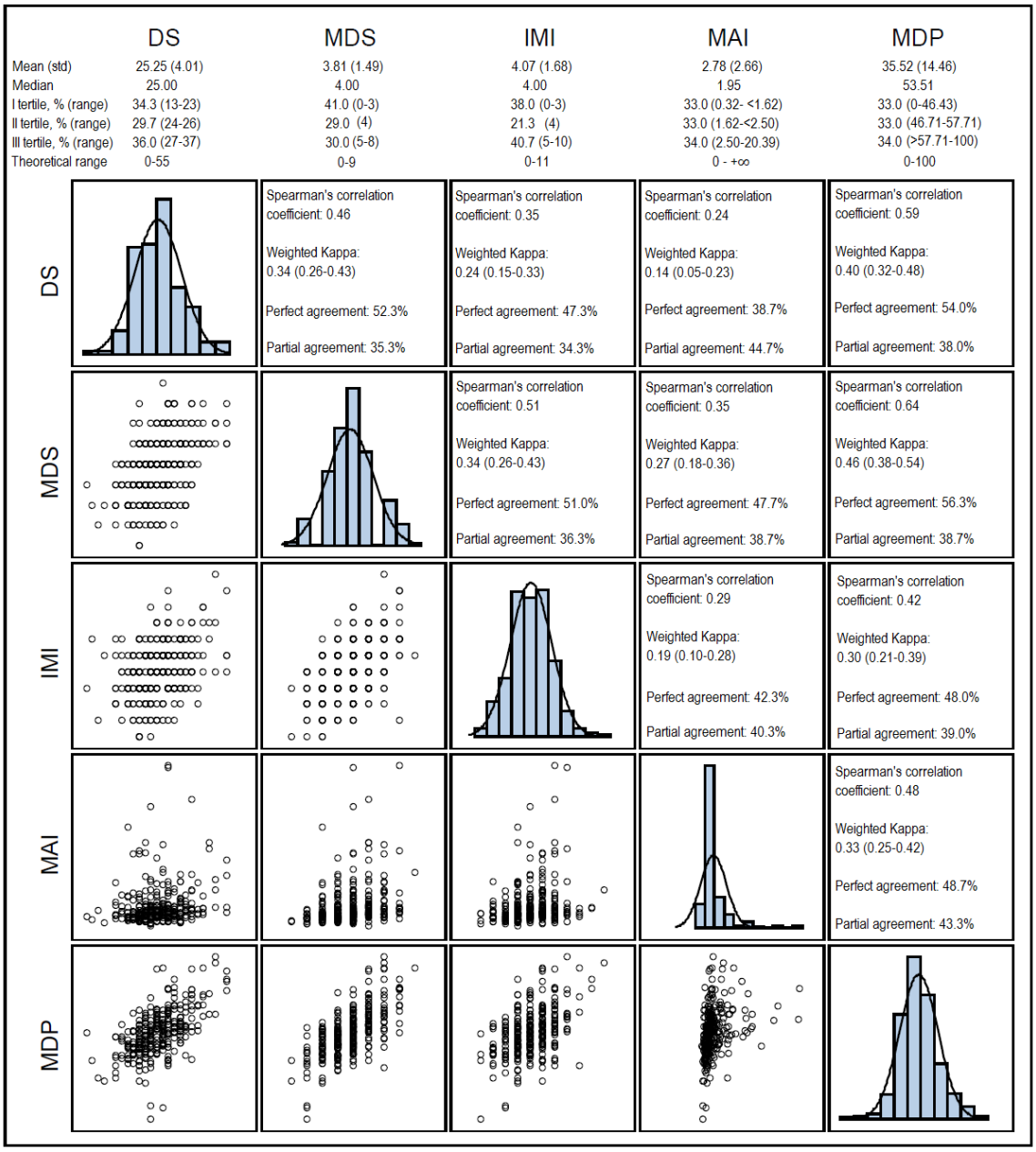

Figure 1. The distribution of the selected Mediterranean diet scores, scatter plots and Sperman's correlation coefficients between pairs of scores, the weighted Kappa, and the proportion of observed perfect and partial agreement between pairs of scores categorized in tertiles.

On the other hand, MAI and IMI were poorly concordant with each other and with the aforementioned scores. In particular, MAI introduced additional foods/nutrients among the non Mediterranean items (e.g. eggs, soft drinks, cakes) and is based on the ratio between the percentage of energy from Mediterranean foods and that from non-Mediterranean foods. Also IMI relies on a somewhat different choice of the components, restricting some food groups (e.g. pasta instead of cereals, only typical Mediterranean vegetables),

excluding dairy products, and considering potatoes among the non-Mediterranean items. Of note, each Mediterranean component is positively scored for the highest tertile, and the reverse applies to non-Mediterranean ones, giving therefore more emphasis to a stricter adherence to the Mediterranean diet.
The issue of measuring Mediterranean diet is challenging, because selected scores have shown only low to moderate concordance, and there is no recognized gold standard to estimate their accuracy. All the scores have different advantages and limitations, which should be considered in the choice of the score to apply, taking into account the characteristics of the population under investigation. Scores based on study-specific cut-offs are relative measures, and high values of these scores indicate a high adherence to the Mediterranean diet "relative" to the population's level of adherence to this pattern. Thus, their application in non-Mediterranean populations may be tricky since the intake of several components is likely lower (or higher) than in a typical Mediterranean population, and the score may be less able to discriminate between beneficial or harmful levels of intakes. On the other hand, relying on *a priori* defined cut-offs is evenly problematic in epidemiological studies using FFQs for dietary assessment, since these instruments are useful to rank individual intakes, while their validity for measuring absolute nutrient intakes has not been demonstrated. Moreover, *a priori* thresholds on Mediterranean components may be difficult to be reached in non-Mediterranean populations. The distribution of these scores may have therefore limited variability, resulting in a low discriminatory power. Conversely, scores based on study-specific cut-offs have the appealing property of classifying for each component pre-specified fractions of subjects as adherent/non-adherent (e.g. halves in the case of median).
As concerns the choice of the score components, it is noteworthy that there is bound to be heterogeneity across populations in types of foods eaten within most Mediterranean components, in particular the broader ones (e.g. vegetables, fruit, and cereals), as well as alcohol. On the other hand, scores with a finer characterization of the Mediterranean components targeted on a specific population (such as IMI for the Italian one), may have limited generalizability to other settings.

Conclusions

Our application indicated that the choice of the Mediterranean score influences the ranking and classification of the subjects according to their adherence to the dietary pattern. A precise definition of the Mediterranean diet is needed to achieve a more accurate quantification of the adherence to the Mediterranean dietary profile. Further developments beyond the present application will be to evaluate the degree of agreement of different Mediterranean diet scores across various populations and to assess the impact of the choice of the score in the evaluation of the relationship between Mediterranean diet and health outcomes. In previous studies of ours, 3 selected Mediterranean scores yielded similar inverse associations with gastric and pancreatic cancers, despite low to moderate concordance.

References

[1] Sofi F, Macchi C, Abbate R. et al. Mediterranean diet and health status: an updated meta-analysis and a proposal for a literature-based adherence score. Public Health Nutr 2014;17(12):2769-82.
[2] Dernini S, Berry EM, Serra-Majem L, et al. Med Diet 4.0: the Mediterranean diet with four sustainable benefits. Public Health Nutr 2017;20(7):1322-30.
[3] Decarli A, Franceschi S, Ferraroni M, et al. Validation of a food-frequency questionnaire to assess dietary intakes in cancer studies in Italy. Results for specific nutrients. Ann Epidemiol 1996;6(2):110-8.
[4] Franceschi S, Negri E, Salvini S, et al. Reproducibility of an Italian food frequency questionnaire for cancer studies: results for specific food items. Eur J Cancer 1993;29A(16):2298-305.
[5] Gnagnarella P, Parpinel M, Salvini S, et al. The update of the Italian food composition database. J Food Comp Analysis 2004;17:509-22.

[**6**] Salvini S, Parpinel M, Gnagnarella P, et al. Banca di composizione degli alimenti per studi epidemiologici in Italia. Milano, Italia: Istituto Europeo di Oncologia; 1998.

[**7**] Panagiotakos DB, Pitsavos C, Stefanadis C. Dietary patterns: a Mediterranean diet score and its relation to clinical and biological markers of cardiovascular disease risk. Nutr Metab Cardiovasc Dis 2006;16(8): 59-68.

[**8**] Trichopoulou A, Costacou T, Bamia C, et al. Adherence to a Mediterranean diet and survival in a Greek population. N Engl J Med 2003;348(26):2599-608.

[**9**] Agnoli C, Grioni S, Sieri S, et al. Italian Mediterranean Index and risk of colorectal cancer in the Italian section of the EPIC cohort. Int J Cancer 2013;132(6):1404-11.

[**10**] Alberti-Fidanza A, Fidanza F. Mediterranean Adequacy Index of Italian diets. Public Health Nutr 2004;7(7):937-41.

[**11**] Sanchez-Villegas A, Martinez JA, De Irala J, et al. Determinants of the adherence to an "a priori" defined Mediterranean dietary pattern. Eur J Nutr 2002;41(6):249-57.

REPRODUCIBILITY OF DIETARY PATTERNS ACROSS POPULATIONS: A NOVEL APPROACH

De Vito Roberta[1], La Vecchia Carlo[2], Parmigiani Giovanni[3], Edefonti Valeria[2] and the INHANCE consortium investigators*

1. *Department of Computer Science, Princeton University, Princeton, NJ, USA.*
2. *Department of Clinical Sciences and Community Health, Università degli Studi di Milano, Milan, Italy.*
3. *Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute and Department of Biostatistics, Harvard T. H. Chan School of Public Health, Boston, MA, USA.*

Introduction

Dietary patterns (DPs) have long been recognized as a useful tool for assessing overall diet in free-living individuals, as they synthetize multiple related dietary components in a single variable. Despite these advantages, little research in nutritional epidemiology has focused on statistical methods specific to DPs. Most of the recent studies use standard multivariate statistical methods [principal component analysis, factor analysis (FA), cluster analysis] to empirically identify *a posteriori* DPs, or consider *a priori*-defined DPs or indexes, where scores are assigned to an individual's adherence to pre-specified dietary recommendations/behavior. The lack of consistent methodology to derive DPs has severely limited the ability to draw firm conclusions about the health risks or benefits associated with DPs [***1***].

Reproducibility and validity of DPs have not been extensively assessed [***2***]. In 2012, the Dietary Patterns Methods Project was initiated to strengthen research evidence on *a priori* and *a posteriori* DPs and health. The collaborating research groups conducted standardized and parallel analyses on selected *a priori* DPs and mortality outcomes in 3 large cohorts in the United States [***1***]. Other studies assessed the same issues on a smaller scale. A major drawback of DPs is their limited applicability to different populations, which is particularly critical when different cultures are involved in the comparison. This is especially true for the *a posteriori* DPs, which are meant to reflect existing dietary behavior in a population and may, therefore, be difficult to replicate in other settings [***3***]. To our knowledge, five studies only assessed the so called "external reproducibility" of DPs across different populations [***3-7***].

De Vito et al. [***8***] recently proposed multi-study factor analysis (MSFA) as a generalization of FA able to handle multiple studies simultaneously [***8***]. Multi-study factor analysis learns the so-called shared DPs, which are common to all studies, as well as additional study-specific DPs for some of the studies. It considers all data at once in an integrated approach based on the maximum likelihood approach [***9***]. The choice of the number of shared and total (shared and specific) DPs to include in the model, as well as the final model selection, are handled within MSFA in a principled statistical way. Multi-study factor analysis tackles the issue of DP reproducibility from a different perspective: the reproducible DPs are those that each study population shares with the others, within a model that integrates information from all studies.

The International Head and Neck Cancer Epidemiology (INHANCE) consortium [***10, 11***] was established in 2004 to elucidate the etiology of head and neck cancer (HNC) through pooled analyses of individual-level data from several studies on a large scale. Dietary habits have been previously investigated within the consortium, using both single-component and the DP approaches [***12-14***]. A *posteriori* DPs were identified with a standard principal component FA based on a common list of nutrients and food components from 5 studies (~7500 subjects) [***14***]. In more recent versions of the INHANCE dataset, two other studies (~3200 extra subjects) provide comparable information on nutrients and food components [***13***]. In addition,

standardized definitions of exposures, confounding factors, and outcomes of interest reduce possible variability from other sources in the assessment of reproducibility and validity of *a posteriori* DPs.[1]
The main objective of this paper is to assess the reproducibility of *a posteriori* DPs derived with the MSFA approach within the INHANCE Consortium. We will identify shared and study-specific DPs. Each study population is free to share one or more reproducible DPs with the others, and to also have one or more study-specific DPs. This will address the following research questions: 1) are there consistent and empirically estimable eating patterns across these populations?; 2) are there additional study- or country-specific eating behaviors?

Materials and Methods
In version 1.5 of the INHANCE consortium pooled data set, ten case-control studies collected information on several nutrient intakes at the individual level. Seven of these studies [***15-22***] provided a common list of twenty-three nutrients that we used for MSFA in this paper.
Cases were included if their cancer had been originally classified as an invasive cancer of oral cavity, oropharynx, hypopharynx, oral cavity or pharynx not otherwise specified, larynx, or HNC unspecified. Cases with cancers of the salivary glands or of the nasal cavity/ear/paranasal sinuses were excluded. We also removed from our analysis: 1. cases without information on the site of origin of cancer; 2. subjects with implausible (<500 or >5500 kcal) or missing values of daily non-alcohol energy intake. Thus, our analysis included a total of 10,668 subjects, with 3844 HNC cases and 6824 controls.
Intakes of total energy, several nutrients and food components were derived by combining information from study-specific FFQs - assessing subjects' usual diet during a reference period preceding cancer diagnosis for cases or interview for controls - with information from country-specific food composition databases [***23, 24***]. From the study-specific lists, we selected twenty-three major nutrients and food components to provide a comprehensive representation of the dietary habits of a Western population and to assess their potential joint role in HNC risk. All the nutrient intakes were expressed on a daily base.
We log-transformed the study-specific data to improve adherence to the assumption made by MSFA that the shared and study-specific factors and the study-specific errors are normally distributed.
As a preliminary step, we explored whether the seven study-specific correlation matrices of the log-transformed data were factorable, using both visual inspection and statistical procedures, including Bartlett's test of sphericity, the Kaiser-Meyer-Olkin (KMO) measure, and individual measures of sampling adequacy [***25***]. Similarly, we evaluated the factorability of the overall correlation matrix generated using log-transformed data from all the available studies combined. Given the reassuring results obtained (see Results section), we applied MSFA to identify *a posteriori* shared and study-specific DPs for the overall set of HNC cases and controls.
We carried out MSFA [***8***] to describe the variance-covariance structure among nutrients in terms of a few underlying unobservable and randomly varying *shared* and *study-specific* factors, or DPs. Specifically, we considered S=7 studies, each represented by the same set of P=23 nutrients. Study s has n_s subjects, each represented by a P-dimensional log-transformed and standardized data vector, $x_{is,}$ with i=1,…,n_s, s=1,…,S. The x_{is} are expressed by MSFA in terms of K shared factors and J_s additional study-specific factors, giving a total T_s=K+J_s factors. Let f_{is} be the (K × 1) shared latent factor vector for subject i in study s, and Φ be the (P × K) shared factor loading matrix. Moreover, let l_{is} be the (J_s × 1) study-specific latent factor vector and Λ_s be the (P × J_s) specific factor-loading matrix. Multi-study factor analysis assumes that the P-dimensional vector x_{is} can be written as:

$$x_{is} = \Phi f_{is} + \Lambda_s l_{is} + e_{is} \quad i=1,\ldots, n_s \quad s=1,\ldots, S.$$

[1] See the full list of authors at the end of the manuscript.

where the error term e_{is} has a diagonal covariance matrix $\Psi_s = diag(\psi_{s1},\ldots, \psi_{sp})$.
The corresponding likelihood is a product over studies of the usual product over subjects found in standard maximum likelihood FA. In addition, each product term includes the corresponding study-specific correlation matrix that reflects the simultaneous presence of the shared and study-specific factors: $\Sigma_s = \Phi\,\Phi^T + \Lambda_s\,\Lambda_s^T + \Psi_s$.
Multi-study factor analysis is fitted using the Expectation Conditional Maximization Algorithm (ECM) [***26***], which is a generalization of the Expectation-Maximization (EM) Algorithm [9] used to derive the maximum likelihood estimates of the common and study-specific factor loadings and of the covariance of the error term. In MSFA, the EM maximization step normally done for FA is replaced by a set of sequential conditional maximization steps for each parameter or groups of parameters.
We chose the number of factors using the following criteria. First, we estimated the total number of factors T_s for each of the S studies, using a combination of standard techniques for FA, including Horn's parallel analysis, Cattell's scree plot and the Steiger's root mean square error of approximation index [***27***]. Next, we used the Akaike Information Criterion (AIC) [***28***] on the overall MSFA model to select the number of shared factors, K. The number of study-specific factors J_s for each study s was then found by difference as T_s - K, s=1,…,S. A global AIC measure was also used as a final criterion to identify the optimal pair (K, J_s).
We then applied a varimax rotation to the shared factor-loading matrix to achieve a better-defined loading structure.
Factor scores indicate the degree to which each subject's diet conforms to one of the identified patterns. We calculated factor scores in MSFA by adapting the standard Bartlett and Thurstone methods for FA [***29, 30***]. In detail, we calculated a factor score for each subject and factor within each study by using the study-specific correlation matrix of the log-transformed data and the overall factor-loading matrix [$\Phi \mid \Lambda_s$]. The correlations between the two types of scores were 0.99 for all factors, so we continued with the Bartlett method, since its scores have zero sample mean vector and zero sample covariances [***29***].
We also evaluated internal consistency and internal reproducibility of the identified DPs using Cronbach's coefficient alpha and split-half techniques. Finally, we applied our MSFA to a subset of our data including the five studies that were originally used to derive *a posteriori* DPs with principal component FA within the INHANCE consortium [***14***], to assess if MSFA and standard FA approaches identify similar shared DPs with a similar performance.

Results
Studies from Europe contributed approximately 50% of cases of oral and pharyngeal cancer combined, 60% of cases of laryngeal cancer, and over 60% of controls. Over 90% of the subjects were white. Cases were more often exposed to tobacco smoking and alcohol drinking than controls.
Correlations among individual nutrients were strong enough to suggest that the study-specific correlation matrices were factorable.
The AIC-based results from the model selection procedure for fixed pre-selected values of T_s, pointed to a MSFA model that presented three DPs shared among all studies and one study-specific DP for each of the four studies from the United States. No additional specific factors were found for the European studies.
The shared factors explained 74.7% of the total variance in each study-specific dataset. The rotation had the effect of making loadings positive for the three shared factors, in such a way that only the magnitude of each loading (and not its sign) was used to name the factors. The first factor, *Animal products and cereals*, had the greatest loadings on phosphorus, riboflavin, zinc, total protein, calcium, niacin, thiamin, vitamin B6, sodium, potassium, iron, cholesterol, and total carbohydrates. The second factor, *Anti-oxidant vitamins and fiber*, had the greatest loadings on vitamin C, total fiber, total carotene, total folate, vitamin E, and potassium. The third factor, *Fats*, had the greatest loadings on monounsaturated, polyunsaturated, and saturated fatty acids, and vitamin E.

The study-specific factors explained 3.85%, 2.76%, 4.86%, and 3.25%, respectively, of the total variance in each of the four original datasets from the United States. The four factors, named *Dairy products and breakfast cereals*, showed a very similar pattern. The Los Angeles-specific DP had the largest positive loading on calcium and largest negative loadings on niacin and vitamin B6. The Boston-specific DP had the largest positive loading on calcium and a negative one on niacin. The MSKCC-specific DP had the largest positive loadings on calcium and phosphorus and the largest negative loadings on niacin, thiamin, vitamin E, and zinc. Finally, the North Carolina (2002-2006)-specific DP had the largest positive loading on calcium, whereas niacin had the largest negative loading at -0.23.
Checks of internal consistency and internal reproducibility of DPs were reassuring. Finally, when we carried out the MSFA on the subset of five studies analyzed in Edefonti et al. [***14***] (Boston, Italy Multicenter, Los Angeles, MSKCC, and Switzerland), the three shared DPs of Edefonti et al. (*Animal products and cereals*, *Anti-oxidant vitamins and fiber*, and *Fats*) were satisfactorily reproduced [***14***]. In addition, the MSFA estimated one extra DP for each of the studies from the United States: the American study-specific DPs were similar to the corresponding ones from the more recent analysis on the entire set of seven studies. Percentages of explained variances were similar for the corresponding DPs in both the analyses (***Figure 1***).

Discussion

We introduce MSFA in nutritional epidemiology to give insight into cross-study reproducibility and validity of *a posteriori* DPs. In our analysis of INHANCE data, we found that study populations from Italy, Switzerland, and the United States share three eating patterns characterized by consumption of animal products and cereals, vitamin-rich foods, and fats, respectively, (~75% of the total variance in each study-specific dataset). In addition, the American studies are characterized by a somewhat similar additional DP, which opposes calcium and niacin as core nutrients in all the studies.
Our analysis had several strengths. The INHANCE consortium offers a natural set-up to apply MSFA. Besides the harmonization of exposure, potential confounders, and outcomes that is typical of well-conducted consortia, we have information from different studies within the same countries, some of which (the European ones and two of the American ones) share the same FFQ. We previously carried out a standard principal component FA on the five INHANCE studies that had provided detailed information on nutrient intakes in 2007 [***14***]. We attempted to assess the reproducibility of our MSFA DPs on a subset of our current dataset including the five studies analyzed in Edefonti et al. [***14***]. The MSFA results are consistent with the standard FA. The factor-loading matrix of the shared factors is similar to the one already published in Edefonti et al. [***14***]. Our MSFA also adds to the previous (now called 'shared') DPs the three extra study-specific DPs for the American studies. In addition, results from the MSFA on five studies were in agreement with those from the MSFA applied on seven studies: the added study from the United States expressed an extra study-specific DP, whereas the added study from Europe was represented at his best by the shared DPs only. We acknowledge that the INHANCE consortium includes case-control studies and this may limit the strength of our conclusions. However, our MSFA is, in principle, applicable to any consortium or network of consortia, including cohort data.
The proposed approach is very attractive because it offers a statistical solution to the issue of external reproducibility of DPs that is currently unsolved in nutritional epidemiology except for three very recent attempts [***3-5***]. The key issue of how many shared and study-specific DPs to retain is integrated within the procedure of DP identification and relies on objective criteria, including the AIC, and not only on the visual inspection of scree plots and the (less stringent) eigenvalue>1 criterion, as it usually happens with standard FA. Moreover, we integrated MSFA with standard checks of internal reproducibility and internal consistency of the identified DPs, with reassuring results [***31***].

Future efforts will deal with generalizations of MSFA to deal with the identification of DPs that are common to a subset of studies only or to deal with situations where some studies provide single nutrients (i.e. monounsaturated fats) and others aggregated information on a family of nutrients (i.e. total fats).

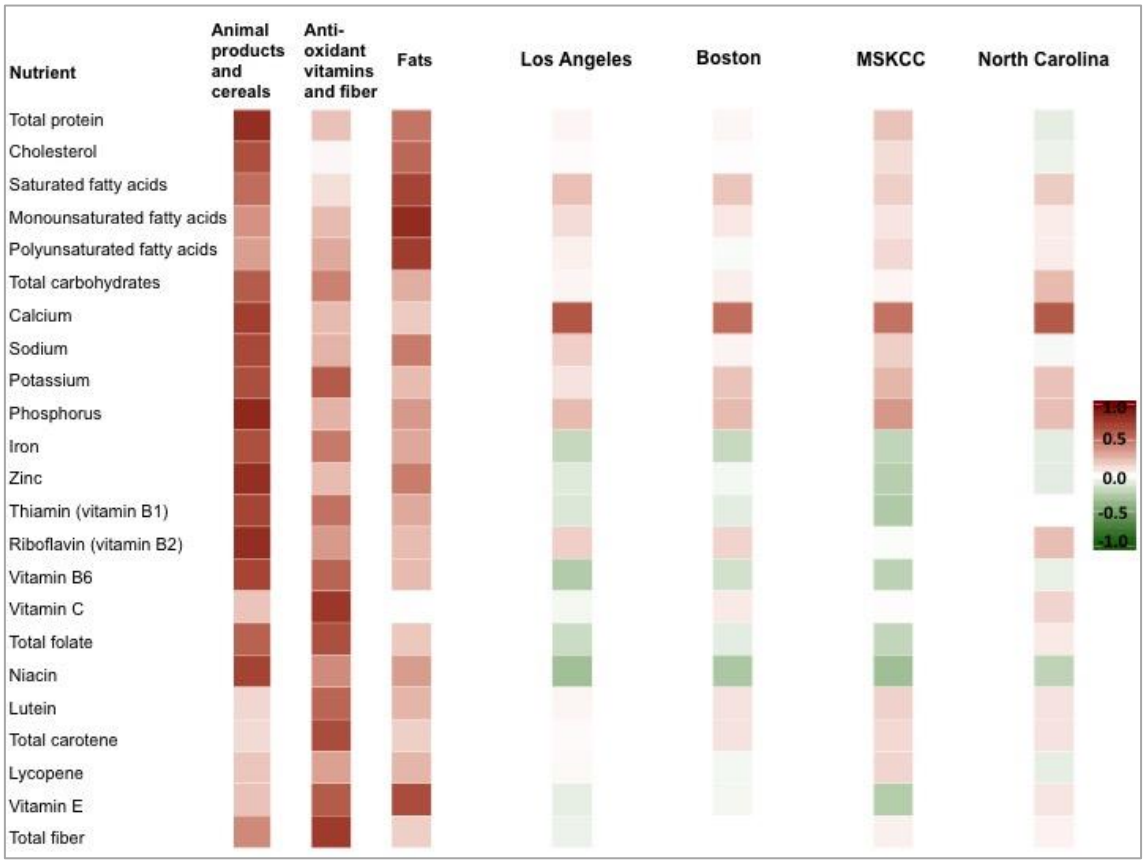

Figure 1. - Heatmap of the estimated factor loadings and cancer-specific odds ratios for the shared and study-specific dietary patterns identified with the multi-study factor analysis. International Head and Neck Cancer Epidemiology (INHANCE) consortium.
ABBREVIATIONS: MSKCC: Memorial Sloan Kettering Cancer Center.

Conclusions

In conclusion, the use of MSFA in nutritional epidemiology allowed us to show that there exist general eating patterns across INHANCE consortium populations, as well as US-specific eating behaviors, which are potentially associated with risk of oral and pharyngeal cancer combined and laryngeal cancer.

References

[1] Liese AD, Krebs-Smith SM, Subar AF et al. The Dietary Patterns Methods Project: synthesis of findings across cohorts and relevance to dietary guidance. J Nutr 2015;145: 393-402.

[2] Edefonti V, Randi G, La Vecchia C et al. Dietary patterns and breast cancer: a review with focus on methodological issues. Nutr Rev 2009;67:297-314.

[3] Castello A, Buijsse B, Martin M et al. Evaluating the Applicability of Data-Driven Dietary Patterns to Independent Samples with a Focus on Measurement Tools for Pattern Similarity. J Acad Nutr Diet 2016;116:1914-1924 e1916.
[4] Castello A, Lope V, Vioque J et al. Reproducibility of data-driven dietary patterns in two groups of adult Spanish women from different studies. Br J Nutr 2016;116:734-42.
[5] Moskal A, Pisa PT, Ferrari P et al. Nutrient patterns and their food sources in an International Study Setting: report from the EPIC study. PLoS One 2014;9:e98647.
[6] Varraso R, Garcia-Aymerich J, Monier F et al. Assessment of dietary patterns in nutritional epidemiology: principal component analysis compared with confirmatory factor analysis. Am J Clin Nutr 2012;96:1079-92.
[7] Huijbregts P, Feskens E, Rasanen L et al. Dietary pattern and 20 year mortality in elderly men in Finland, Italy, and The Netherlands: longitudinal cohort study. BMJ 1997;315:13-7.
[8] de Vito R, Bellio R, Trippa L, Parmigiani G. Multi-study Factor Analysis. arXiv:1611.06350 [stat.AP]. in press.
[9] Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. Journal of the royal statistical society. 1977;Series B (methodological):1-38.
[10] Conway DI, Hashibe M, Boffetta P, et al. Enhancing epidemiologic research on head and neck cancer: INHANCE - The international head and neck cancer epidemiology consortium. Oral Oncol 2009;45:743-6.
[11] Hashibe M, Brennan P, Benhamou S, et al. Alcohol drinking in never users of tobacco, cigarette smoking in never drinkers, and the risk of head and neck cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. J Natl Cancer Inst 2007;99:777-89.
[12] Chuang SC, Jenab M, Heck JE, et al. Diet and the risk of head and neck cancer: a pooled analysis in the INHANCE consortium. Cancer Causes Control 2012;23:69-88.
[13] Edefonti V, Hashibe M, Parpinel M, et al. Natural vitamin C intake and the risk of head and neck cancer: A pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. Int J Cancer 2015;137:448-62.
[14] Edefonti V, Hashibe M, Ambrogi F, et al. Nutrient-based dietary patterns and the risk of head and neck cancer: a pooled analysis in the International Head and Neck Cancer Epidemiology consortium. Ann Oncol 2012;23:1869-80.
[15] Bravi F, Bosetti C, Filomeno M, et al. Foods, nutrients and the risk of oral and pharyngeal cancer. Br J Cancer 2013;109:2904-10.
[16] Schantz SP, Zhang ZF, Spitz MS, et al. Genetic susceptibility to head and neck cancer: interaction between nutrition and mutagen sensitivity. Laryngoscope 1997;107:765-81.
[17] Levi F, Pasche C, La Vecchia C, et al. Food groups and risk of oral and pharyngeal cancer. Int J Cancer 1998;77:705-9.
[18] Bosetti C, Gallus S, Trichopoulou A, et al. Influence of the Mediterranean diet on the risk of cancers of the upper aerodigestive tract. Cancer Epidemiol Biomarkers Prev 2003;12:1091-4.
[19] Peters ES, McClean MD, Liu M, et al. The ADH1C polymorphism modifies the risk of squamous cell carcinoma of the head and neck associated with alcohol and tobacco use. Cancer Epidemiol Biomarkers Prev 2005;14:476-82.
[20] Cui Y, Morgenstern H, Greenland S, et al. Polymorphism of Xeroderma Pigmentosum group G and the risk of lung cancer and squamous cell carcinomas of the oropharynx, larynx and esophagus. Int J Cancer 2006;118:714-20.
[21] Hashibe M, Morgenstern H, Cui Y, et al. Marijuana use and the risk of lung and upper aerodigestive tract cancers: results of a population-based case-control study. Cancer Epidemiol Biomarkers Prev 2006;15:1829-34.

[**22**] Divaris K, Olshan AF, Smith J, et al. Oral health and risk for head and neck squamous cell carcinoma: the Carolina Head and Neck Cancer Study. Cancer Causes Control 2010; 1:567-75.
[**23**] US Department of Agriculture (USDA), Agricultural Research Service. USDA National Nutrient Database for Standard Reference, Release 26 and previous versions. Nutrient Data Laboratory Available from: http://www.ars.usda.gov/Services/docs.htm?docid=8964. [Accessed 2017].
[**24**] Gnagnarella P, Salvini S, Parpinel M. Food Composition Database for Epidemiological Studies in Italy. Version 1.2015. Available from: http://www.bda-ieo.it/. [Accessed 2017].
[**25**] Pett MA, Lackey NR, Sullivan JJ. Making sense of factor analysis: the use of factor analysis for instrument development in health care research. CA: Sage, 2003.
[**26**] Meng X-L, Rubin DB. Maximum likelihood estimation via the ECM algorithm: A general framework. Biometrika 1993;80:267-78.
[**27**] Mulaik SA. Foundations of factor analysis. CRC press, 2009.
[**28**] Akaike H. A new look at the statistical model identification. IEEE transactions on automatic control 1974;19:716-23.
[**29**] Johnson RA, Wichern DW. Applied multivariate statistical analysis. Upper Saddler River, NJ: Prentice Hall, 2002.
[**30**] DiStefano C, Zhu M, Mindrila D. Understanding and using factor scores: Considerations for the applied researcher. Practical Assessment, Research & Evaluation 2009;14:1-11.
[**31**] Edefonti V, Bravi F, Garavello W, et al. Nutrient-based dietary patterns and laryngeal cancer: evidence from an exploratory factor analysis. Cancer Epidemiol Biomarkers Prev 2010;19:18-27.

Full list of authors from the INHANCE consortium:

R. De Vito[1], Y. C. A. Lee[2], M. Parpinel[3], D. Serraino[4], A. F. Olshan[5], J. P. Zevallos[6], F. Levi[7], Z. F. Zhang[8], H. Morgenstern[9], K. Kelsey[10], M. McClean[11], S. Schantz[12], G. P. Yu[13], P. Boffetta[14], S. C. Chuang[15], M. Hashibe[16], C. La Vecchia[17], G. Parmigiani[18,19], and V. Edefonti[17]

[1] Department of Computer Science, Princeton University, Princeton, NJ, USA;[2] Division of Public Health, Department of Family & Preventive Medicine, University of Utah School of Medicine, Salt Lake City, UT, USA;[3] Department of Medicine, University of Udine, Udine, Italy;[4] Epidemiology and Biostatistics Unit, CRO Aviano National Cancer Institute, IRCCS, Aviano, Italy;[5] University of North Carolina School of Public Health, Chapel Hill, NC, USA;[6] Department of Otolaryngology/Head and Neck Surgery, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA;[7] Institute of Social and Preventive Medicine (IUMSP), Lausanne University Hospital (CHUV), Lausanne, Switzerland;[8] Department of Epidemiology, UCLA School of Public Health, Los Angeles, CA, USA;[9] Departments of Epidemiology and Environmental Health Sciences, School of Public Health and Comprehensive Cancer Center, University of Michigan, Ann Arbor, MI, USA;[10] Department of Epidemiology and Pathology and Laboratory Medicine, Brown University, Providence, RI, USA;[11] Department of Environmental Health, Boston University School of Public Health, Boston, MA, USA;[12] Department of Otolaryngology, New York Eye and Ear Infirmary, New York, NY, USA;
[13] Medical Informatics Center, Peking University, Peking, China; [14] The Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA;[15] Institute of Population Health Sciences, National Health Research Institutes, Miaoli, Taiwan;[16] Division of Public Health, Department of Family & Preventive Medicine and Huntsman Cancer Institute, University of Utah School of Medicine, Salt Lake City, UT, USA;[17] Branch of Medical Statistics, Biometry and Epidemiology "G. A. Maccacaro", Department of Clinical Sciences and Community Health, Università degli Studi di Milano, Milano, Italy;[18] Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Boston, MA, USA;[19] Department of Biostatistics, Harvard T. H. Chan School of Public Health, Boston, MA, USA.

MODELLING THE EFFECT OF GESTATIONAL AGE AND SEX ON ENCEPHALIC GROWTH IN THE 5TH MONTH OF PREGNANCY

Zolin Anna[1], Spada Elena[1], Gervasi Federico[1], Conte Giorgio[2], Milani Silvano[1].

1. *Laboratory of Medical Statistics, Biometry and Epidemiology "G.A. Maccacaro", Department of Clinical Sciences and Community Health – Università degli Studi di Milano, Milan, Italy*
2. *Postgraduation school in radiagnostics, Università degli Studi di Milano, Milan, Italy.*

Introduction

Prenatal Magnetic Resonance (MR) imaging has an important role in the evaluation of fetal brain development, being usually performed as second-look investigation when suspected brain abnormalities are detected by prenatal ultrasound (US). Prenatal MR imaging has been demonstrated to improve the diagnostic accuracy, leading to changes in clinical management in a large number of cases [***1***]. The evaluation of prenatal MR images usually consists in the assessment of cranial/intracranial biometry and morphology, and normative data of the main fetal brain dimensions are necessary to detect possible disorders in brain development. Several studies have reported normative MR biometric data of fetuses in the late 2nd and 3rd trimesters of gestation but few studies reported them in fetuses in the early 2nd trimester and they had many methodological limitations, such as small number of fetuses, inadequate descriptive analysis based on maximum and minimum values rather than on centiles of each traits; volumetric measures difficult to be applied in a clinical contest.
The aim of this study is to model encephalic growth and provide growth reference charts of 14 MR traits on the basis of the MR biometric data of a cohort of 169 fetuses from 20 to 24 weeks of gestation.

Data

Data were derived from the databases of two centers of Fetal Medicine, accounting for the examinations performed between 2005 and 2016: "Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico" (Milano, Italy) and "Ospedale Pediatrico Vittore Buzzi" (Milano, Italy). Inclusion criteria: fetuses from 20 to 24 weeks of gestational age (GA) followed to term and found to be normal at birth. Exclusion criteria: poor image quality, clear or suspected central nervous system (CNS) abnormalities at prenatal MR, extra-CNS malformations frequently associated with CNS malformations, pregnancies complicated with infections, chromosomopathy. GA was assessed on the basis of the last menstrual period and US criteria, and expressed as completed weeks.
The following MR traits were considered: posterior cranial fossa (latero-lateral diameter), cerebrum (antero-posterior and latero-lateral diameters), cerebellum (latero-lateral diameter, antero-posterior and superior-inferior vermian diameters), clivus-supraocciput angle, corpus callosum length, cranium (antero-posterior and latero-lateral diameters), mesencephalic axis (anterior-posterior diameter), pons (antero-posterior and supero-inferior diameter), lateral ventricular atrium (width). All MR measures were expressed in mm, with the only exception of the clivo-supraocciput angle (degree). Each measure was taken twice or thrice on the same or different acquisitions and the average of the measures made by each assessor was regarded as his best (i.e. most reliable) measure. The average of these best measures was regarded as the best estimate of the value of the MR trait, and used to trace the growth reference charts.

Statistical methods

The estimates of the size attained at 22 wks of GA and of the mean weekly increase in the interval from 20 to 24 weeks were derived from the following linear model:

$$E(\text{MR trait}) = \mu + \alpha \times s + \beta \times t + \gamma \times s \times t$$

where E(MR trait) is the expected value of the MR trait, s is 0 for females and 1 for males, t is =GA-22, μ is the mean size attained by females at 22 wks, α is the "male *vs* female" difference in size attained at 22 wks, β is the average weekly increase shown by females from 20 to 24 wks, γ is the "male *vs* female" difference in average weekly increase from 20 to 24 wks. The above general linear model was fitted using PROC GLM of SAS/STAT software v.9.2, 2008 (SAS Institute, Cary, NC, USA).
To trace the reference centiles, the CG-LMS method [***2***] was adopted, using the software LMS program version 1.29 (Medical Research Council, UK). This is a model that expresses the centiles in terms of GA-specific curves called L(t), M(t) and S(t). The M(t) and S(t) curves correspond to the GA-dependent median and coefficient of variation of the MR trait, whereas the L(t) curve allows for the GA-dependent skewness of the distribution of the same trait. The value (y) of the MR trait at a given age can be transformed into a standard deviation score (SDS)

$$SDS = \frac{(y/M(t))^{L(t)} - 1}{L(t) \times S(t)}$$

The value y(p,t) of the pth centile at GA t can be calculated as

$$y(p,t) = M(t) \times (1 + z_p \times L(t) \times S(t))^{(1/L(t))}$$

where z_p is the standard normal deviate corresponding to probability p.

Results
Encephalic growth. At 22 wks of GA, the MR traits under study attained, on the average, a size ranging from 4.5 mm (mesencephalic axis: anterior-posterior diameter) and 7.7 mm (pons: supero-inferior diameter) to 58.5 mm (cerebrum: antero-posterior) and 65.2 mm (cranium: antero-posterior). The clivus-supraocciput angle was 72.5 degree. The average weekly increase in the period from 20 to 24 weeks was highly significant for all traits except for clivus-supraocciput angle and mesencephalic axis, and ranged from 0.31 mm/wk (pons: antero-posterior and supero-inferior diameters) to 3.67 mm/wk (cranium: antero-posterior diameter) and 3.81 mm/wk (cranium: antero-posterior diameter). Growth rate (i.e. the percent ratio "weekly increase/size" at 22 wks) ranged from 4.0% and 4.6% (pons: supero-inferior and antero-posterior diameters) to 6.3% (cerebrum: antero-posterior) and 9.2% (corpus callosum: length). The only MR trait showing a significant decrease with increasing GA was lateral ventricular atrium width (-0.20 mm/wk, -3.0%).
Difference between sexes in encephalic growth. At 22 wks of GA males fetuses are already larger for all MR traits, with the only exception of cerebellum (antero-posterior diameter): 6 out 14 MR traits were significantly larger in males. The maximum differences were observed in the posterior cranial fossa (latero-lateral diameter), in the cerebellum (latero-lateral diameter) and in clivus-supraocciput angle, where males are larger than females by 5.2%, 3.2% and 3.1%, respectively.
As for the weekly increase of MR traits, males appeared to grow faster than females in the period from 20 to 24 wks of GA for 6 out of the 11 MR traits found to increase significantly, with maxima of 17.9%, 13.8% and 11.8% in the posterior cranial fossa (latero-lateral diameter), in the pons (antero-posterior diameter) and in the cerebellum (latero-lateral diameter), respectively. Anyhow, possibly because of the limited width of the GA interval considered, none of the differences between sexes in growth velocity was statistically significant.
Reference centiles. The sex-differences reported above, although biologically interesting, were not large enough to suggest the need of reference charts different for females and males. For this reason, MR traits were fitted with a CG-LMS model independently of sex. In the narrow GA interval under study, M(t) and S(t)

were assumed to change linearly with GA, whereas L(t) was assumed to be constant. The use of more flexible CG-LMS models did not improve consistently the goodness of fitting. All the diameters of cerebrum, cerebellum (latero-lateral diameter), and cranium were found to have a positively skewed distribution, as well as the clivus-supraocciput angle. This may suggest that in the 5th month there are, for these MR traits, a few fetuses growing faster than the majority of the others. The remaining MR traits were negatively skewed. This may suggest that in the 5th month there are, for these MR traits, a few fetuses growing slower than the majority of the others. The coefficient of variation appeared to increase slightly with increasing GA in all MR traits with the only exceptions of the cerebrum (antero-posterior diameter) and of the mesencephalic axis (anterior-posterior diameter). At 20 wks of GA, the inter-individual variability, in terms of width of the reference interval expressed as percent of the median value, ranges from 14.0% (cerebrum: latero-lateral diameter) and 15.1% (cranium: antero-posterior diameter) to 36.4% (mesencephalic axis: anterior-posterior diameter) and 38.6% (lateral ventricular atrium: width). At 24 wks of GA, the inter-individual variability of 12 MR traits out of the 14 under study is wider than at 20 wks, with a 24.4% average increase, the only exceptions being of cerebrum (latero-lateral diameter), -7.3%, and mesencephalic axis (anterior-posterior diameter), -16.3%. The increments range from 17.6% (cerebrum: anterior-posterior diameter) and 19.2% (cranium: antero-posterior and latero-lateral diameters) to 32.4% (mesencephalic axis: anterior-posterior diameter) and 63.5% (lateral ventricular atrium: width). These findings clearly indicate that individual characteristics of fetuses tend to diverge with increasing GA. A graphical representation of how the CG-LMS model fits a GA-dependent distribution is given in ***figure 1***.

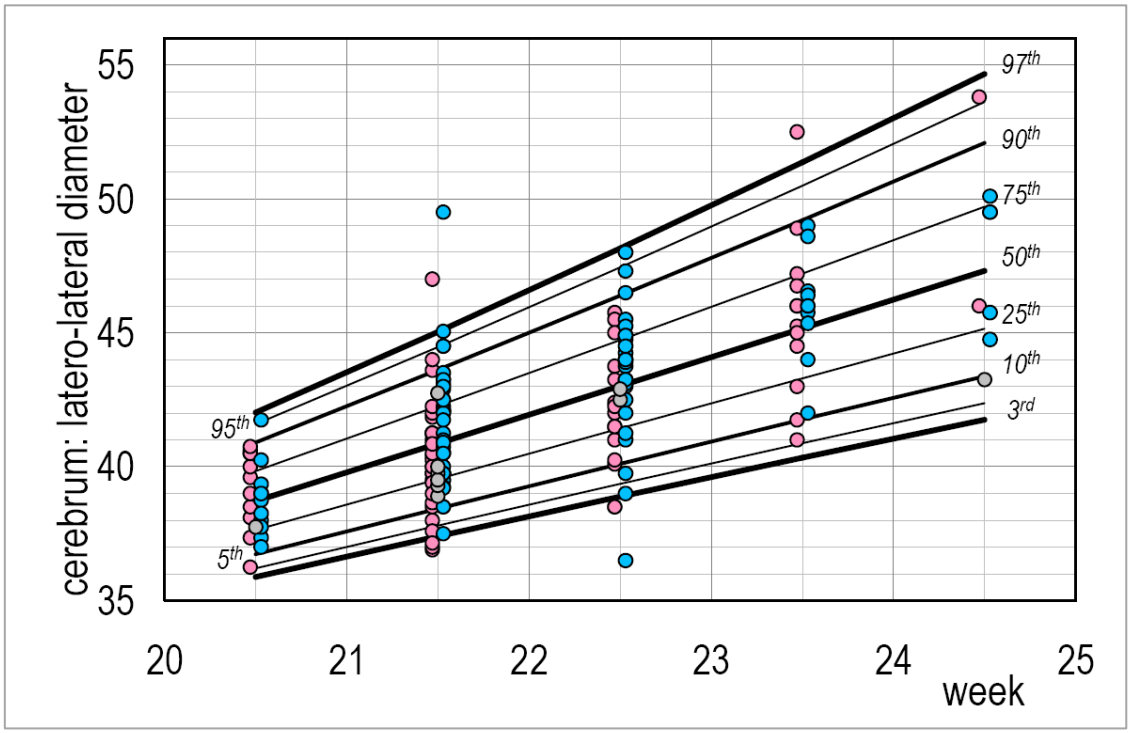

Figure 1. Reference charts for cerebrum growth (latero-lateral diameter) traced with L(t)=-0.171, dM(t)/dt=1.24 mm/wk, dS(t)/dt=0.52/wk. Pink dots refer to females, cyan dots to males, gray dots to sex unknown.
(Note: the mean GA of fetuses of x completed weeks is x+0.5 weeks).

Conclusions

The linear model here used to analyze encephalic growth appeared to be particularly suitable for describing three main characteristics of the encephalic growth in the 5th month of pregnancy: (1) at 2 wks of GA males are already larger than females, (2) in this GA period growth velocity tends to be higher in males, (3) growth rate varies considerably between MR traits.

The CG-LMS model, even in the parsimonious form here adopted, since M(t) and S(t) were assumed to change linearly with GA, and L(t) was assumed to be constant, was proved to fit satisfactorily the GA-dependent distribution of the MR traits.

References

[*1*] Cole TJ, Green PJ. Smoothing reference centile curves: the LMS method and penalized likelihood. Statistics in Medicine 1992;11:1305-19.

[*2*] Griffiths PD, Bradburn M, Campbell MJ, et al. Use of MRI in the diagnosis of fetal brain abnormalities in utero (MERIDIAN): a multicentre, prospective cohort study. Lancet 2017;389:538-46.

COMPARISON OF BAYESIAN MACHINE LEARNING TECHNIQUES ON BIOMEDICAL DATASETS

Bottigliengo Daniele[1], Lanera Corrado[2], Baldi Ileana[2], Cavallo Franco[1], Gregori Dario[2], Berchialla Paola[3]

1. *Department of Public Health and Pediatrics Sciences, University of Torino*
2. *Department of Cardiac, Thoracic and Vascular Sciences, University of Padova*
3. *Department of Clinical and Biological Sciences, University of Torino*

Introduction

Making prediction and classification when dealing with clinical data has always been a challenging task. Toughness of statistical analysis is mostly due to the features of clinical data: limited size of the available samples, number of variables larger than number of observations and big amount of missing data.
Machine Learning techniques are popular tools for dealing with classification and prediction, because of their ability to handle high-dimensional data, to allow for non-linear interaction effects without specifying any semi-parametric model assumption [***1***].
Bayesian methods have become very popular in many research fields, because of their robustness to overfitting in small sample sizes analysis and their ability to handle with missing data [***2***]. Since they rely on the definition of a probabilistic model, they are able to make inference on the parameters of the model integrating external information through the use of prior distribution probabilities [***3***].
Among Machine Learning techniques there exists a subgroup of techniques that take advantage of a Bayesian approach. Bayesian Machine Learning techniques incorporate the advantages of Bayesian methods, since learning process can be performed by simply applying Bayes' rule. Results of Bayesian learning are expressed in terms of a probability distribution over all unknown quantities. This probabilities are used to express all forms of uncertainty [***4***]. Traditional Machine Learning techniques instead usually provide only point estimates, ignoring the uncertainty surrounding these estimates [***2***].
Naïve Bayes [**5**] and Bayesian Network (BN) [***6***] are the most used Bayesian Machine Learning techniques in clinical setting. Recently, Bayesian Additive Regression Trees (BART) has been employed in many clinical applications [***7***].
At our knowledge, there is a lack of works comparing different Bayesian Machine Learning (BML) methods in terms of accuracy performances for classification tasks in clinical setting. In order to fill this gap, we carried out this study to provide a comparison of some of the most popular Bayesian Machine Learning algorithms previously mentioned. Their evaluation was carried out with reference to two clinical datasets with different features: one representing a typical dataset with a limited sample size and one representing a typical high-dimensional dataset.

Methods

First dataset was obtained from the larger series of Chron's Disease (CD) and Ulcerative Colitis patients enrolled in the observational study of Inflammatory Bowel Disease (IBD) genetics conducted by Giachino et al. (2007) [***8***] in collaboration with three gastroenterology Units in Torino, Italy.
Second dataset was obtained from Parkinson's Progression Markers Initiative (PPMI) database. PPMI is a public-private partnership funded by Michael J. Fox Foundation for Parkinson's Research and funding partners. PPMI consists in an observational multi-center study whose mission is to identify one or more biomarkers of Parkinson's disease (PD) progression, a critical next step in the development of new and better treatments for PD. PPMI database consists in a collection of databases on Parkinson's disease. Each database includes data on status and socio-economics information, tests, assessment questionairres, trials,

examinations on patients included in the study, collected during the study in order to identify one or more biomarkers of Parkinson's disease progression.
Naïve-Bayes was described by Duda and Hart (1973) [***9***]. It is a simple network which possess the following peculiarities: outcome variable as the parent node of all other nodes and no connections between the variables. This kind network is easy to construct, since the structure is given a priori and no structure learning procedure is needed. Due to its simplicity, Naïve Bayes is easy to use and it allows to get a good result especially in case of small databases [***10***]. Moreover, naïve classifiers can be extremely fast in comparison to more sophisticated methods of data mining. The performance of Naïve Bayes is somewhat surprising, since the above assumption is clearly unrealistic [***10***].
In order to avoid the unwarranted assumptions about independence underneath Naïve -Bayes, Bayesian Networks were considered. Bayesian Networks are Directed Acyclic Graphs (DAGs) [***11***] that allow efficient and effective representation of the joint probability distribution over a set of random variables. Each vertex in the graph represents a variable, and edges represent direct correlations between the variables. More precisely, the network takes into account the following conditional independence statements: each variable is independent of its non-descendants in the graph given the state of its parents. These independencies are then exploited to reduce the number of parameters needed to characterize a probability distribution, and to efficiently compute posterior probabilities given evidence. Probabilistic parameters are encoded in a set of tables, one for each variable, in the form of local conditional distributions of a variable given its parents.
BART is a Bayesian approach to nonparametric function estimation using regression trees. It was proposed by Chipman et al. (2010) [***12***]. In order to approximate an unknown function f, regression trees depend on recursive binary partitioning of predictor space into a set of hyper-rectangles. Predictor space has dimension equal to the number of variables, which we denote p. Tree-based regression models have an ability to consider interactions and nonlinearities. BART can be distinguished from other ensemble-of-trees models due to its under- lying probability: because of its Bayesian nature, BART consists of a set of priors for the structure and the leaf parameters and a likelihood for data in the terminal nodes. The aim of the priors is to provide regularization to prevent that any single regression tree dominates the total fit.

Results

Three different Bayesian Machine Learning models were estimated for each dataset: Naïve-Bayes, Bayesian Network and BART. Every estimated model was cross-validated in order to get a more accurate estimated of his prediction performance, computing distance between estimated output from the actual output, and to assess how statistical analysis results could be generalized to an independent data set. Model prediction performances were tested in term of sensitivity, specificity, positive predicted value, negative predicted value and accuracy. Indicator values are shown in ***Table 1***.

Table 1. Sensitivity, specificity, positive predictive value (PPV) or precision, negative predictive value (NPV) and accuracy of the models used during the analysis.

	Sensitivity	Specificity	PPV	NPV	Accuracy
IBD Data					
BART	93%	46%	70%	82%	73%
Bayesian Network Chow-Liu	96%	3%	57%	33%	56%
Naïve-Bayes	67%	20%	54%	31%	47%
PPMI Data					
BART	93%	89%	96%	82%	93%
Bayesian Network Tabu	96%	68%	89%	87%	89%
Bayesian Network Hill-Climbing	96%	68%	89%	87%	89%
Naïve-Bayes	83%	77%	91%	63%	82%

Conclusions

BART was the model that generally showed the best prediction performances during the analysis, especially on PPMI dataset. BART brilliant performances were probably due to the fact that it incorporates advantages of ensemble-of-trees models, which create a strong learner from a set of weak learners, and Bayesian models, which consist of set of priors for the structure and the leaf parameters and a likelihood for internal nodes, allowing the priors to provide regularization and preventing that a single tree dominates to total fit.
Bayesian Network showed quite good performances, especially on PPMI dataset. This was probably due to the fact that Bayesian network structure learning and inference algorithms often assume all random variables are discrete and all the variables contained into PPMI dataset were discrete.
Naïve-Bayes showed very poor performances on IBD dataset, but it presented quite good prediction ability on PPMI dataset.
One of the limitations of the study lies in Naïve-Bayes structure: its unrealistic assumption (conditional independency of the attributes given the outcome variable value) doesn't keep in consideration potential relationships and correlations between covariates. Another limitation lies in the choice of performing IBD analysis without considering missing data, because of their probable "non-random" nature, leading to information loss.

References

[1] Wu J, Roy J, Stewart WF. Prediction Modeling Using EHR Data: Challenges, Strategies, and a Comparison of Machine Learning Approaches. Medical Care 2010;48:106-13.
[**2**] Hernandez B, Raftery AE, Pennington SR, et al. Bayesian Additive Regression Trees using Bayesian Model Averaging. arXiv:1507.00181v2 [stat.CO] 2015.
[**3**] Kruschke KJ. Doing Bayesian Data Analysis: A Tutorial with R, JAGS and Stan. Accademic Press; 2015.
[**4**] Chakraborty S. Bayesian Machine Learning. [Gainesville, FL, USA]: University of Florida; 2005.
[**5**] Wei W, Visweswaran S, Cooper GF. The application of naive Bayes model averaging to predict Alzheimer's disease from genome-wide data. J Am Med Inform Assoc 2011;18(4):370-75.
[**6**] Scutari M. Learning Bayesian Networks with the bnlearn R Package. J Stat Softw 2010;35(3):1–22
[**7**] Sparapani RA, Logan BR, McCulloch RE, et al. Nonparametric survival analysis using Bayesian Additive Regression Trees. BART. Stat Med 2016;35:2741-53.
[**8**] Giachino DF, Regazzoni S, Bordessono M, et al. Modeling the role of genetic factors in characterizing extra-intestinal manifestations in Crohn's disease patients: does this improve outcome predictions? Curr Med Res Opin 2007;23(7):1657-65.
[**9**] Duda R, Hart P. Pattern classification and scene analysis. Wiley; 1973
[**10**] Rish I. An empirical study of the Naïve Bayes classifier. IJCAI 2001 workshop on empirical methods in artificial intelligence, 2001;41-6.
[**11**] Jensen FV. Introduction to Bayesian Networks. New York: Springer-Verlag; 1996.
[**12**] Chipman HA, George EI, McCulloch RE. BART: Bayesian Additive Regression Trees. Ann Appl Stat 2010;4(1):266-98.

MISSING IMPUTATION STRATEGIES IN A BAYESIAN NETWORK META-ANALYSIS FRAMEWORK

Minto Clara[1], Azzolina Danila[1], Baldi Ileana[1], Berchialla Paola[2], Gregori Dario[1]

1. *Unit of Biostatistics, Epidemiology and Public Health, Department of Cardiac, Thoracic and Vascular Sciences, University of Padua, Italy.*
2. *Department of Clinical And Biological Science University of Turin*

Introduction

A Bayesian network meta-analysis (NMA) model is often used to make direct and indirect comparisons across multiple treatment options. In several cases, a NMA is associated to a loss of information due to incomplete study data retrieved through a systematic review, which are therefore excluded from the analysis. Previous literature on meta-analysis suggests two main approaches to manage missing data: researchers can choose to select only articles with complete information or, otherwise, to adopt a specific imputation method. Although the first approach could seem more rigorous and accurate, it is not free of potential bias. The deletion of relevant papers causes a reduction in sample size, resulting in a remarkable reduction in test power. Moreover, when a researcher chooses this type of approach, implicitly chooses to consider all data as Missing Completely at Random (MCAR): tenability of this assumption is uncommon, and when MCAR assumption is violated the analysis will produce biased estimates [***1***]. Several statistical methods are provided in literature to handle missing or incomplete data in a NMA. In some cases, only baseline and follow-up measurements are available and it could be necessary to consider pre-post study correlation in order to obtain data about mean change. In this case, some authors suggest imputation strategies of pre-post correlation, following Bayesian approach: the correlation extracted from another study is used to calculate the variability measure of mean change [***2***]. Moreover, a variability measure associated to mean change score might be unavailable. In this case, different imputation methods are suggested, as those based on maximum standard deviation (MSD) imputation: the highest variability measure extracted from another study is used to complete missing data.
The purpose of this study is to verify the robustness of Bayesian NMA with respect to two different imputation strategies: the pre-post correlation and the MSD imputation methods. Hypotheses are verified through simulation analyses.

Methods

Fifty trials are simulated in full databases by including baseline, follow-up and Delta variation information. Baseline data are obtained by sampling from bounded 0-100 normal distributions ($X \sim N(41.8, 21.5)$) (Cannon, 2000), to mimick the support of Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC). Delta variation data are simulated from normal distributions with parameters provided by a review on 6 Non-Steroidal Anti-Inflammatory Drugs (NSAIDs). Follow up variability data are provided from generated Delta and baseline variability measures setting hypotheses on pre-post correlation and considering, in each scenario, a sequence from 0.3 to 0.95 by 0.05. Sample size are obtained by sampling from an uniform 50-100 distribution. Between trial heterogeneity has been included as a variability measure by following, for each simulation setting, a sequence from 0.1 to 5 by 0.1. Each scenario provides different combinations of heterogeneity between trials and pre-post correlation creating 700 scenarios. For each scenario 2 imputed databases are generated. In the first case, information about Delta variation are randomly removed, from full database, leaving only baseline end follow up data, then variability of mean change is imputed using the correlation method. In the other case, also information at baseline and follow up are removed, then Delta variability is imputed with maximum standard deviation method. On simulated

dataset, NMA with random effect and Uniform (0,5) prior on heterogeneity parameter, has been performed (MCMC method, 200000 iterations, 4 chains). To investigate robustness of NMA, under several scenarios and different imputation methods, the bias of rank probabilities estimates has been computed in order to check models performance in ranking treatments. For each scenario, the mean, bias and the standard deviation of the first rank probability, for full and imputed databases, have been computed.

Results

The results show that the bias is very small for every scenario, then ranking provided by models is robust with respect to different imputation methods. The method is more robust to impute in a low heterogeneity framework, especially if considered trials are conducted on similar population. Small bias is observed also for heterogeneity values similar to expectation of NMA heterogeneity prior, indicating more robustness if a priori knowledge is well specified.

Conclusions

Management of missing data in NMA is a debated issue, especially for scientific topic with few and low-quality evidence. In studies on pharmacological treatments, this aspect has high importance, and should be taken into consideration in order to give exhaustive and correct clinical information. A meta-analysis on treatments comparison that includes just complete data, automatically excludes all papers in which incompleteness of data is probably due to non-significant results. An effective solution to this problem, is offered by imputation of missing data with methods of correlation and maximum standard deviation.

References

[**1**] Fu R, Vandermeer BW, Shamliyan TA, et al. Handling continuous outcomes in quantitative synthesis. AHRQ Methods for Effective Health Care 2013.

[**2**] Abrams KR, Gillies CL, Lambert PC. Meta-analysis of heterogeneously reported trials assessing change from baseline. Stat Med 2005;24(24):3823-44.

HANDLING POOR ACCRUAL IN ADAPTIVE TRIAL SETTING: BAYESIAN INTERIM ANALYSIS OF RESCUE TRIAL

Azzolina Danila[1], Baldi Ileana[1], Berchialla Paola[2], Bressan Silvia[3], di Leo Valentina[3], Da Dalt Liviana[3], Gregori Dario[1]

1. *Unit of Biostatistics, Epidemiology and Public Health, Department of Cardiac, Thoracic and Vascular Sciences, University of Padua, Italy.*
2. *Department of Clinical and Biological Sciences, University of Turin, Italy.*
3. *Department of Pediatrics, University Hospital, Padua, Italy.*

Introduction

In several clinical trial settings, it is very difficult to recruit the overall sample provided at the design stage, and different problems may occur in patient's enrolment.
The amount of information conveyed by a trial terminated prematurely for poor accrual may be minimal. A Bayesian analysis of such a trial may salvage this information, by providing a framework in which to combine prior with current evidence.
In this work, we propose the Bayesian analysis [***1***] of a trial terminated and designed in a frequentist framework for poor accrual.

Methods

RESCUE trial is a randomized controlled trial evaluating the effect of adjunctive oral steroids to prevent renal scarring in young children with febrile urinary tract infections.
Primary outcome is the difference in scarring proportion between standard antibiotic therapy versus standard therapy + corticosteroids. By study protocol, a frequentist approach to sample size calculation require 92 randomized patients per arm, considering 20% lost follow-up. After 2 years, only 8 patients completed the follow up to determine the study outcome (3 in corticosteroids therapy group and 5 in control group).
The sample size was recalculated with the Bayesian Worse Outcome Criterion [***2***] for differences in proportions (length=0.3 and coverage=0.9) applying a 0.5% down-weight to prior parameters. An informative prior on scar proportions was derived from literature considering a scar probability equal to 0.33 and 0.66 respectively in treatment and control group [***3***].
An interim Bayesian analysis on recruited patients has been performed [***4***]; having a few data to estimate the likelihood, inference was expected to be seriously conditioned by the prior choice. To assess robustness of conclusion [***5***], a sensitivity analysis on prior definition has been performed considering 1) informative Beta prior as in sample size estimation 2) informative Beta with 0.5% down-weight 2) uninformative Beta (1,1) prior.
Results are compared in term of posterior probability that $\vartheta_{treat} - \vartheta_{control} < \Delta$, *where* $\Delta = -0.2$ *as indicated by protocol. Also, has been considered the simple reduction* $\vartheta_{treat} - \vartheta_{control} < \Delta$ *with* $\Delta = 0$

Results

The estimated Bayesian sample size is 41 infants per arm, leading to a reduction of 51 patients compared with frequentist one.
The Bayesian interim analysis shows that the probability to observe a reduction in scar proportion less than Δ, with $\Delta = -0.2$, is 0.81, 0.57 and 0.69 respectively for analysis conducted in informative prior setting, for

informative Beta with down-weight and using uninformative distribution. Also the probability to observe reduction, setting $\Delta = 0$, are 0.99, 0.83, and 0.98 for different prior as shown above.
It is possible to assess a general robustness of inferential conclusion on scar probability absolute reduction ($\Delta = 0$) respect to different prior definition. However, evaluating the effect in term of entity of reduction ($\Delta = -0.2$), results are more influenced by informative prior and weakly by likelihood.

Conclusions

The Bayesian inference is a flexible tool, compared to frequentist one, taking into account of a-priori knowledge about treatment effect.
The informative inference, on small sample, may be weakly influenced by data. However, sensitivity analysis lead to consider the robustness of inferential conclusion.
Nevertheless, we advocate to choose beforehand a Bayesian design and not to switch to a Bayesian analysis method that produces a more favourable outcome after observing the data.

References

[1] Box GE, Tiao GC. *Bayesian inference in statistical analysis* vol 40: John Wiley & Sons)
[2] Joseph L, du Berger R and Bélisle P 1997 Bayesian and mixed Bayesian/likelihood criteria for sample size determination. Statistics in medicine 2011;16:769-81.
[3] Huang YY, Chen MJ, Chiu N-, Chou HH, Lin KY, Chiou YY. Adjunctive oral methylprednisolone in pediatric acute pyelonephritis alleviates renal scarring. Pediatrics 2011;128(3):e496-504.
[4] Berry DA. Bayesian clinical trials. Nature reviews. Drug discovery 2006;5(1):27-36.
[5] Berger JO. Robust Bayesian analysis: sensitivity to the prior. Journal of statistical planning and inference 1990;25:303-28.

MODELLING THE PROGRESSION OF CERVICAL DILATION IN SPONTANEOUS LABOUR

Plebani Maddalena[1,2], Vener Claudia[2], Ferrazzi Enrico[3], Paganelli Andrea[3],Milani Silvano[2]

1. *Unit of Medical Statistics, Biometry and Bioinformatics, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy*
2. *Laboratory of Medical Statistics, Biometry and Epidemiology "G.A. Maccacaro", Department of Clinical Sciences and Community Health – Università degli Studi di Milano, Milan, Italy*
3. *Department of Woman, Mother and Neonate, Buzzi Hospital, Biomedical and Clinical Sciences School of Medicine, Università degli Studi di Milano, Milan, Italy*

Introduction

The first attempt to model the progression of cervical dilation during spontaneous labour of pregnant women dates back to 1955 [***1***]. From the graphical analysis of the time profiles of dilation observed in 500 women aged 13 to 42 years, the Author derived a mean labour sigmoid curve, consisting in a latent phase followed by an active phase ending in a deceleration phase. This cervimetric graph, or cervicogram, is still reported in current manuals of obstetrics. Zhang *et al* [***2, 3***] fitted cervical dilation profiles with mixed models based on 10th or 8th degree polynomials, and concluded that the duration of latent and active phases largely differs from woman to woman, and that there is no deceleration at the end of the active phase. For this reason, these Authors supply charts reporting the empirical 95th centile of the distribution of labour duration for different values of cervical dilation at admission to labour room, thus abandoning the idea of modelling the progression of cervical dilation. The aim of this presentation is to show the use of parsimonious nonlinear mixed models to trace cervimetric charts, reporting current cervical dilation *vs* time to full dilation.

Data and methods

Data here used derive from an observational study including 328 low-risk women (146 primiparae and 182 multiparae), who delivered at Buzzi Children's Hospital between April and June 2013 [***4***]. All women delivered vaginally at term, after uncomplicated single pregnancy and spontaneous labour managed by midwives, without any kind of medical intervention.

We had to take into account a lot of difficulties and hindrances to construct a model for the progression of cervical dilation. The time of the beginning of labour is unknown (1); women are admitted to labour room at different degrees of dilation (2); cervical dilation measures are taken at irregular intervals as required by clinical practice (3); midwives usually assess cervical dilation with fingers, though the measure is reported in cm, from 0 (no dilation) to 10 (full dilation) (4); full cervical dilation is not indicated with a measure, but with a value of 10 cm, arbitrarily assigned to all women (5); individual dilation profiles are largely incomplete, a maximum of 5 measures per woman was recorded before full dilation, and only 60 women (18%) were assessed twice or more (6); the progression of cervical dilation is extremely erratic (7).

Since dilation cannot be related to the unknown time from the onset of labour, we considered, as already suggested by Zhang *et al* [2], the time remaining to the attainment of full cervical dilation (t, time to full dilation), and expressed current dilation ($y_i(t)$, from 1 to 9 cm) observed in the ith woman as a function of a parsimonious nonlinear model (3 parameters only), instead of the 9 to 11 parameters of the polynomials used by Zhang *et al* [***2, 3***]:

$$logit(y_i(t)/10) = \alpha_i + \beta_i \times log(t + \delta_i) + \varepsilon_{i(t)} \qquad \{1\}$$

In {1}, $logit(y_i(t)/10)$ is a linear function of the log-transformation of time: parameter β_i is the dilation velocity constant, the ratio α_i/β_i determines the time at maximum dilation velocity, and δ_i modulates the shape of the

log-transformation of time; intra-individual random terms $\varepsilon_{i(t)}$ were assumed to approximate a normal distribution, with variance proportional to $\{[E(y_i(t))/10][1-E(y_i(t))\}^{-1}$. When back-transformed to the original scale, expression {1} defines a family of strongly asymmetrical never-decreasing (since cervical dilation is an irreversible process) sigmoid curves, with a slight slowdown of dilation velocity a little before, or at full dilation, in this latter case the curve presents an exponential shape.
Because cervical dilation profiles were largely incomplete, we could not resort to the usual two-stage models [***5***] to trace cervimetric charts, but we were forced to adopt a nonlinear mixed model [***6***], which can obtain estimates of the parameters of the individual cervical dilation curves also for the women with profiles made up by a number of observations lower the number of paramtes (in our case 1 or 2 assesments only):

$$logit(y_i(t)/10) = \alpha_0 + \alpha_i + \alpha_P x + (\beta_0 + \beta_i + \beta_P x)\, log(t + \delta_0 + \delta_P x) + \varepsilon_{i(t)} \qquad \{2\}$$

Parity (x=0 for primiparae, x=1 for multiparae) was included as a covariate into model {2}, since multiparae are known to progress somewhat faster in active-phase labour [***7***]. Parameters α_0 and β_0 refer to primiparae, α_P and β_P refer to the difference between multiparae and primiparae, whereas δ_0 and δ_P modulate the log-transformation of time by parity. Random terms α_i and β_i, which model inter-individual differences, are assumed to have bivariate normal distribution with $E(\alpha_i)=E(\beta_i)=0$ and unstructered covariance matrix $Cov(\alpha_i,\beta_i)=[\sigma^2_\alpha, \sigma_{\alpha\beta}, \sigma^2_\beta]$.

Models {1} and {2} were fitted using PROC NLIN and NLMIXED of SAS/STAT® software (SAS Institute, Cary, NC; v.9.4, 2013).

Results

Although based on 3 parameters only, model {1} proved to be flexible enough to describe cervical dilation profiles of rather different shape. As shown in ***figure 1*** (left), concerning primiparous women, profiles may present an inflection point (i.e. a maximum dilation velocity) already 4 hours before the attainment of full dilation (orange curve) or at about 1 hour (dark red and green curves) or in the last half hour of the dilation process (the remaining profiles). Ten hours before the attainment of full dilation there are women with no more than 1 cm dilation and women with so much as 4 cm dilation. During the labour, dilation velocity may vary considerably from woman to woman: 10 hours before the attainment of full dilation, dilation velocity is always less than 0.5 cm/hr, but maximum velocity may be more than 3.5 cm/hr when initial dilation is 1 cm (red and blue curves) or be about 1 cm/hr when initial dilation is 4 cm (green curve). It is worth noting that women with different initial dilation (red curve: 1.7 cm, olive green curve: 2.7 cm) may present the same maximum velocity (1.7 cm/hr), since their dilation profiles differ in convexity.
Figure 1 (right) shows, plotted on the cervicometric charts traced with mixed model {2}, the cervical dilation profiles (green lines) of the 95 primiparous women with 2 or more assessments and the dilation values (green dots) of the 51 women with 1 assessment only, predicted on the basis of model {2}.
Green dots As expected, though individual profiles differ largely, the large majority of them lies completely within the interval 3rd – 97th centile of the cervicometric charts. At 10 hours, the distribution of dilation values conditional on time to full dilation is highly right skewed, then positive skewness decreases and the distribution becomes symmetrical when median dilation is 5 cm, and left skewed subsequently. The same results were observed in the 182 multiparous women included in the study.
We observe that a 4 cm dilation is the 97th centile 10 hours before the attainment of full dilation, and is the 3rd centile at 0.5 hours. This means that, at 10 hours, 3% of women present more than 4 cm dilation, but that another 3% of women still present a 4 cm dilation half an hour before the end of dilation process. So a 4 cm dilation observed from 10 to 0.5 hours before the attainment of full dilation cannot be regarded as unusual. Analogously, at 10 hours, 10% and 25% of women present more than 3 and 2 cm dilation,

respectively, but 10% and 25% of women still present a 3 or 2 cm dilation, 1 and 3 hours before the end of dilation process.
Our data confirm that labour progresses faster in multiparae than in primiparae: e.g. a 5 cm dilation is achieved 45 min *vs* 1 hour and half before the end of dilation process in multiparae. Primiparae present a wider variability (+30% in terms of interquartile range) in the distribution of dilation values conditional on time to full dilation.

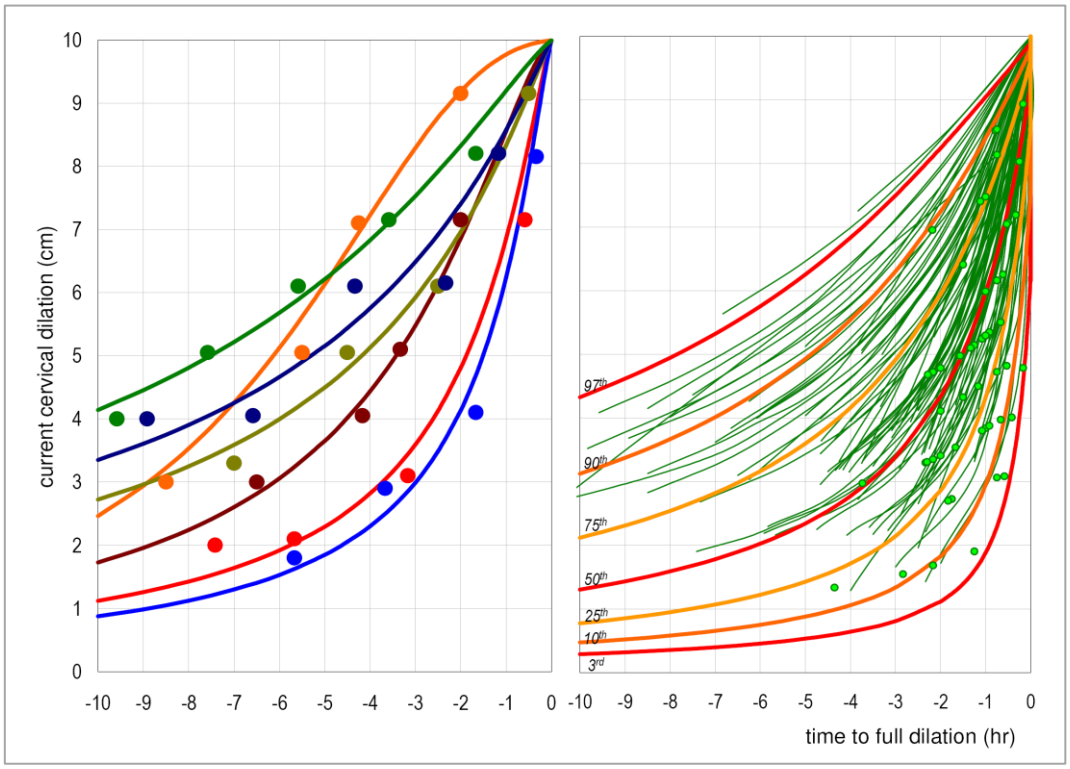

Figure 1. Left: examples of individual cervical dilation profiles, observed (solid circles) in 7 primiparous women, and predicted by model {1} (continuous lines). Right: cervical dilation profiles (green lines) of the 95 primiparous women with 2 or more assessments and dilation values (green dots) of the 51 women with 1 assessment only, predicted on the basis of mixed model {2}, and plotted on the cervicometric charts traced with model {2}. Time to full dilation denotes the time remaining to the attainment of full cervical dilation.

Conclusions

In 1955 Friedman [***1***] wrote: "The dynamic nature of parturitional change has, in the past, rendered exceedingly difficult the detailed and critical analysis of its vagaries". Actually, try to model the kinetics of the cervical dilation process turned out still to be a very hard task. Nonetheless some important points were established: (1) a parsimonious nonlinear parametric model is suitable to describe very different shapes of the dilation process, (2) mixed nonlinear models allow to trace plausible cervimetric charts even in the case of dilation profiles largely incomplete, (3) the above model provide quantitative estimates of the inter-individual variability and of the difference in the progression of cervical dilation between primiparae and multiparae.
Unfortunately, the classical cervicograms give an unrealistic and useless picture of labour progression when referred to a single patient, the course of cervical dilation being largely erratic even in the case of

spontaneous labour in uncomplicated pregnancies. In the first place, during labour the time remaining to the attainment of full cervical dilation is unknown, so the charts cannot be used in obstetric practice; in the second place a subsequent value of dilation cannot be predicted on the basis of the previous assessments. For these reasons, Ferrazzi *et al* [***4***] proposed cervimetric charts reporting the distribution of time needed to gain 1 cm in cervical dilation as a function of current dilation. Although theoretically usable, the practical value of these charts remains very low. In primiparous women, the time needed to gain 1 cm in dilation ranges from 10 (10th centile) to 110 min (90th centile) when current dilation is 1 cm, and from 5 to 70 min when current dilation is 9 cm. In multiparous women, the reference interval is somewhat narrower, the time needed ranging from 3 to 70 min when current dilation is 1 cm, and from 2 to 45 min when current dilation is 9 cm. On the basis of these results they concluded that the progression of cervical dilatation in normal human labour is unpredictable.

References

[**1**] Friedman EA. Primigravid labor. A graphicostatistical analysis. Obstetrics and Gynecology 1955;6:567-89.

[**2**] Zhang J, Troendle JF, Yancey MK. Reassessing the labor curve in nulliparous women. American Journal of Obstetrics and Gynecology 2002;187:824-28.

[**3**] Zhang J, Landy HJ, Branch DW, *et al*. Contemporary patterns of spontaneous labor with normal neonatal outcomes. Obstetrics and Gynecology 2010;116:1281-7.

[**4**] Ferrazzi E, Milani S, Cirillo F, *at al*. Progression of cervical dilatation in normal human labor is unpredictable. Acta Obstet Gynecol Scand 2015;94:1136-44.

[**5**] Laird NM, Ware JH. Random-effects models for longitudinal data. Biometrics 1982;38:963-74.

[**6**] Goldstein H. Nonlinear multilevel models, with an application to discrete response data. Biometrika 1991;78:45-51.

[**7**] Friedman EA. An objective approach to the diagnosis and management of abnormal labor. Bull N Y Acad Med 1972;48:842-58.

MODELLING BIRTHWEIGHT AS A FUNCTION OF GESTATIONAL AGE AT DELIVERY IN HYPERTENSIVE DISORDER OF PREGNANCY

Vener Claudia[1], Ferrazzi Enrico[2], Zullino Sara[2],Milani Silvano[1]

1. *Laboratory of Medical Statistics, Biometry and Epidemiology "G.A. Maccacaro", Department of Clinical Sciences and Community Health – Università degli Studi di Milano, Milan, Italy*
2. *Department of Woman, Mother and Neonate, Buzzi Hospital, Biomedical and Clinical Sciences School of Medicine, University of Milan, Milan, Italy*

Introduction

Pre-eclampsia (PE) is a multifactorial syndrome with different clinical phenotypes. To address this heterogeneity, the syndrome has been subdivided, according to gestational age (GA), into early and late PE, but this classification is not completely satisfying [***1***]. The most-studied clinical phenotype is that caused by shallow trophoblastic invasion of the spiral arteries, with oxidative stress and release of placental factors, which lead to endothelial dysfunction and organ damage [***2***]. Doppler velocimetry of the uterine arteries (UtA-PI) and measurement of vascular growth factors can be used to screen for early vascular damage, associated with reduced placental growth and intrauterine growth restriction (IUGR) later in pregnancy [***3***]. This clinical phenotype is predominant in cases prior to 34 wks of GA, thus early PE can be regarded as a proxy for cases affected by this sequence of events. Anyhow, the disease does not end at 34 wks of GA, and less severe cases represent a large proportion of syndromic manifestations of PE in late gestation [***1***]. Cases affected by hypertensive disorders of pregnancy (HDP) including PE are frequently found to have normal neonates and placentas [***4***]. In these cases, endothelial dysfunction can be ascribed to the placental pro-inflammatory response, dyslipidemia or other pro-inflammatory conditions resulting in low-grade inflammation due to pre-existing maternal conditions, such as maternal metabolic syndrome and cardiovascular risk factors. These risk factors are not suppposed to operate only from 34 wks of GA onwards.

The aim of this presentation is to verify whether a classification based only on fetal abdomen circumference (AC) and UtA-PI is suitable for defining two internally homogeneous subgroups of women affected by HDP, characterized by different effects on fetal growth, as can be inferred from the outcome, in terms of birthweight conditional on GA, and whether GA at the onset of HDP has a role in determining fetal growth.

Data and methods

This multicentre study includes all eligible cases of HDP consecutively admitted over a 12-month period (from December 2013) to seven Italian tertiary referral centres for maternal-fetal medicine. To be eligible for inclusion, the clinical records of each patient had to report two "key items": fetal assessment of abdomen circumference (AC) and mean UtA-PI. Women with multiple pregnancy and with chromosomal or structural fetal abnormalities were excluded. A total of 902 patients met these inclusion/exclusion criteria. Data concerning maternal characteristics, pregnancy complications and outcome were retrieved from the archived clinical records [***5***].

Pregnant women were grouped into two phenotypes, independently of GA at the onset of HDP. The first phenotype, which was named HDP-IUGR (n=124) [***5***], is characterized by maternal HDP associated with fetal AC at admission below the 5th centile of Todros *et al* reference [***6***], and an abnormal UtA-PI, defined as mean of right and left UtA-PI>95th centile of Gomez *et al* reference [***7***]. The second phenotype, which was named HDP-AGA, i.e. fetuses whose AC is appropriate-for-gestational age (n=205), is characterized by maternal HDP associated with fetal AC≥5th centile and UtA-PI≤95th centile. The group was then subdivided

in early and late-onset HDP, depending whether GA at the onset of disease was <34+0 or ≥34+0 wks+days. The remaining patients were excluded because of missing data on either fetal AC or UtA Doppler velocimetry (n=461), or because discordant values of fetal AC and UtA Doppler velocimetry (n=112), i.e. AC at admission ≥5th centile and UtA-PI>95th centile, or AC at admission <5th centile and UtA-PI≤95th centile.
To model birthweight as a function of GA at delivery we resorted to the generalization of the logistic function introduced by von Bertalanffy [***8***], which can be written as [***9***]:

$$E(y(t)) = \frac{\kappa}{\left(1+\alpha e^{-\beta(t-\tau)}\right)^{1/\alpha}}$$

where E(y(t)) is the value of the birthweight y expected at GA t on the basis of the model, κ is the upper asymptote (ideally birthweight at term), β is a velocity constant and τ is GA at the inflection point, i.e. the age at which maximum growth velocity occurs, α determines the degree of asymmetry of the function. On this function was based the model used to trace the INeS (Italian Neonatal Study) reference charts [***10***].
The effect of sex (x_1), phenotype (x_2) and gestational age at admission (x_3) were modelled introducing the proper multiplicative terms into the nonlinear model:

$$E(y(t)) = \frac{\kappa(1+\vartheta x_1)(1+\xi x_2)(1+\varphi x_3)}{\left(1+\alpha e^{-\beta(t-\tau)}\right)^{1/\alpha}}$$

where x_1, x_2, x_3, are 1 when sex is male, clinical phenotype is HDP-IUGR, GA at the onset of HDP is <34+0 wks+days, respectively, and is 0 otherwise; parameters ϑ, ξ and φ are the proportional increase in birthweight due to the effect of these conditions. Least-squares estimates of the parameters of the above model were obtained with Marquardt algorithm, resorting to SAS PROC NLIN of SAS/STAT® software (SAS Institute, Cary, NC; v.9.4, 2013).

Results

Figure 1 shows birthweight as a function of GA at delivery in neonates classified by sex, maternal phenotype and GA at the onset of HDP. When the model includes only the term accounting for the effect of sex, birthweight of males is estimated to be higher than that of females by 4% (95%CL: 1%, 7%; p=0.0158). The introduction of the term expressing the effect of phenotype into the model consistently improves the goodness of fit: the coefficient of determination increased from 0.86 to 0.91 and the residual standard deviation decreased from 366 to 294 g. The children of women with phenotype HDP-IUGR show a birthweight lower by 23% (95%CL: 20%, 26%; p<0.0001) than those of women with phenotype HDP-AGA. No further improvement in model fit was observed when GA at admission was added to the model. This result was confirmed when birthweight was expressed as standard deviation score (SDS), on the basis of the INeS reference charts [10], so as to remove the effect of gestational age at delivery, birth order and sex. The mean birthweight SDS (±standard error) of babies delivered by HDP-IUGR women was -1.53±0.05 (corresponding to the 6th centile) *vs* -0.05±0.06 (corresponding to the 48th centile) for babies delivered by HDP-AGA women (p<0.0001). As emerges from figure 1, GA at the onset of HDP did not appear to exert any effect on birthweight. At each GA, the children born to women with late onset of HDP (≥34 wks) show birth weights close to those of children born to women with early onset of HDP (<34 wks), the difference being about 1% (95%CL: -2%, 4%; p=0.5116).

Conclusions

The use of the parametric nonlinear model presented in this study yielded a synthetic and faithful description of the relationship between birthweight and GA in over 3 hundred babies born to women suffering with hypertensive disorders of pregnancy, and delivered between the 25th and the 42nd week of GA. Furthermore

the model provided estimates of the separate effects that sex, HDP phenotype, and GA at the onset of HDP exert on birthweight.
Among babies born to HDP mothers, birthweight of males was found to be higher than that of by 4%: an identical difference was observed in the reference set used to trace INeS charts [10], this suggest that HDP impairs fetal growth independently of sex. The babies born to mothers with phenotype HDP-IUGR were found to have birthweight lower by about 25% than those born to mothers with phenotype HDP-AGA: this is not unexpected, fetal weight being partly determined by abdomen volume. Nonetheless, since delivery occurred, on the average, 3 weeks after the assessment of abdomen circumference, such a difference suggests the persistence of conditions unfavourable to fetal growth up to end of pregnancy. The third hypothesis that the model could test was whether GA at the onset of HDP affects the severity of fetal growth impairment. In this case we would expect that birthweight of babies born to mothers with early onset of HDP is considerably lower. Results showed that fetal growth follows the same pattern, independently of GA at the onset of HDP.

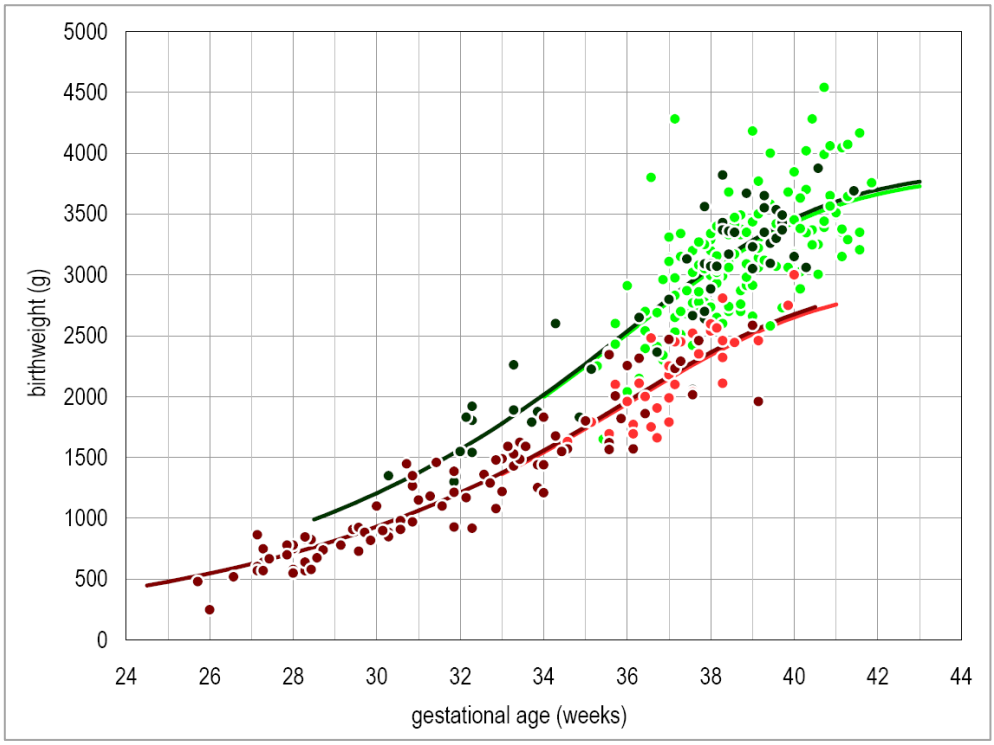

Figure 1. Birthweight as a function of GA at delivery in cases of hypertensive disorder of pregnancy (HDP) with intrauterine growth restriction (IUGR) or appropriate-for-gestational age (AGA) fetuses, with early (<34 wks) or late (≥34 wks) GA at the onset of HDP. Individual data points on fitted curve are plotted for each subgroup: dark green, HDP-AGA<34 wks (n=52); light green, HDP-AGA≥34 wks (n=153); dark red, HDP-IUGR<34 wks (n=87); light red, HDP-IUGR≥34 wks (n=37). Continuous curves were traced using a nonlinear model based upon the generalized logistic function [8], and are adjusted for the effect of sex.

References

[1] Verlohren S, Melchiorre K, Khalil A, et al. Uterine artery Doppler, birth weight and timing of onset of pre-eclampsia: providing insights into the dual etiology of late-onset pre-eclampsia. *Ultrasound* Obstet Gynecol 2014;44:293-98.

[2] Cindrova-Davies T. Gabor Than Award Lecture 2008: pre-eclampsia – from placental oxidative stress to maternal endothelial dysfunction. Placenta 2009;30:S55–S65.

[3] Levine RJ, Maynard SE, Qian C, et al. Circulating angiogenic factors and the risk of pre-eclampsia. N Engl J Med 2004;350:672-83.

[4] Conde-Agudelo A, Belizan JM. Risk factors for pre-eclampsia in a large cohort of Latin American and Caribbean women. BJOG 2000;107:75-83.

[5] Ferrazzi E, Zullino S, Stampalija T, et al. Bedside diagnosis of two major clinical phenotypes of hypertensive disorders of pregnancy. Ultrasound Obstet Gynecol 2016;48:224-31.

[6] Todros T, Ferrazzi E, Nicolini U, et al. Fitting growth curves to head and abdomen measurements of the fetus: A multicentric study. J Clin Ultrasound 1987;15:95-105.

[7] Gomez O, Figueras F, Fernandez S, et al. Reference ranges for uterine artery mean pulsatility index at 11–41 weeks of gestation. Ultrasound Obstet Gynecol 2008;32:128-32.

[8] von Bertalanffy L. Quantitative laws in metabolism and growth. Q Rev Biol 1957;32:217-31.

[9] Marubini E. Mathematical handling of long-term longitudinal data. In Human Growth, vol.1: Principles and Prenatal Growth; Falkner F, Tanner JM(eds). Plenum Press: New York 1978; 209-25.

[10] Bertino E, Spada E, Occhi L, et al. Neonatal anthropometric charts: the Italian neonatal study compared with other European studies. J Pediatr Gastroenterol Nutr 2010;51:353-61.

UN MODELLO DINAMICO PER LA VALUTAZIONE DELL'ANDAMENTO DELLA GLICEMIA IN ADOLESCENTI CON DIABETE DI TIPO 1 IN TERAPIA CON UN SISTEMA INTEGRATO

Skrami Edlira[1], Gesuita Rosaria[1], Cherubini Valentino[2], Ferrito Lucia[2], Carle Flavia[1], Ferrante Luigi[1]

1. *Centro di Epidemiologia, Biostatistica ed Informatica medica, Università Politecnica delle Marche, Ancona, Italia*
2. *Diabetologia Pediatrica, Dipartimento Materno-Infantile, Azienda Ospedaliero-Universitaria Ospedali Riuniti, Ancona, Italia*

Introduzione

L'introduzione di nuove tecnologie per il controllo della glicemia e la somministrazione dell'insulina nei pazienti diabetici ha permesso di ottenere misurazioni continue della glicemia. Il controllo dei valori di glicemia è di fondamentale importanza per ridurre il rischio di complicanze del Diabete di Tipo 1 (DT1) negli adolescenti soprattutto durante l'attività fisica. Essa richiede un adattamento personalizzato della terapia insulinica al fine di ridurre il rischio delle ipoglicemie e delle complicanze cardiovascolari [***1***]; i campi di educazione terapeutica (CET) rappresentano il contesto ideale per fare acquisire ai giovani pazienti competenze nella gestione della patologia durante l'attività fisica, anche in relazione all'educazione all'uso delle nuove tecnologie.

MiniMed 640G™ è uno dei nuovi sistemi integrati per la somministrazione dell'insulina che monitora la glicemia prevedendone l'andamento dei valori e conseguentemente permette per esempio la sospensione automatica dell'infusione dell'insulina, evitando così gli episodi di ipoglicemia [***2***]. Il sistema genera così una serie temporale di valori di glicemia e permette di fissare diversi livelli soglia; la possibilità di analizzare la dinamica glicemica con modelli statistici adeguati consente di ottenere utili informazioni per l'ottimizzazione della gestione della malattia.

Le elaborazioni e le interpretazioni delle informazioni relative all'andamento del controllo glicemico in funzione della soglia fissata, nelle condizioni di vita quotidiana, sono ancora limitate; l'analisi delle serie temporali rappresenta il metodo che meglio consente di seguire il profilo glicemico e ricavare informazioni sulla variabilità circadiana della glicemia, soprattutto in corrispondenza dei pasti e dell'attività fisica.

Valutare l'andamento della glicemia in adolescenti con diabete di tipo 1 in terapia con il sistema integrato MINIMED640G in funzione della soglia glicemica stabilita per la sospensione dell'insulina.

Metodi

Trentuno adolescenti con età compresa tra 15 e 20 anni, con diagnosi di DT1 da almeno 12 mesi, che usavano il microinfusore e il glucometro da almeno 3 mesi, in buon compenso metabolico (emoglobina glicata - HbA1c <10%), hanno partecipato ad un CET della durata di 4 giorni, nel mese di giugno 2016. Sono stati esclusi gli adolescenti che hanno avuto ipoglicemie severe o episodi di DKA (Diabetic KetoAcidosis) nei due mesi precedenti il campo, o che presentavano patologie concomitanti che potevano influenzare il controllo metabolico. I soggetti provenivano da 12 centri di diabetologia pediatrica di 10 Regioni Italiane (Piemonte, Lombardia, Liguria, Lazio, Toscana, Marche, Abruzzo, Puglia, Sicilia e Sardegna). Durante il campo, tutti i partecipanti hanno utilizzato la tecnologia MiniMed 640G (sensore glicemico, trasmettitore, microinfusore, glucometro dedicato, software di scarico dati) che rileva i valori di glicemia ogni 5 minuti. Per tutti i soggetti la soglia glicemica di sospensione è stata fissata a 70 mg/dl durante il giorno (ore 8:00-22:00) e 90 mg/dl durante la notte. I ragazzi sono stati assegnati in modo casuale a due gruppi in funzione della soglia stabilità durante l'attività fisica: gruppo A con soglia durante l'attività fisica innalzata a 90mg/dl, gruppo B con soglia mantenuta a 70mg/dl.

Il primo giorno del campo sono state rilevate in tutti i soggetti età, genere, Body Mass Index - BMI, HBA1c, glicemia, durata del diabete; il monitoraggio ogni 5 minuti della glicemia è stato effettuato per 72 ore. *Analisi statistica*. Per confrontare la dinamica del glucosio plasmatico {G_t, t=1,…,N=865} nei due gruppi abbiamo utilizzato l'analisi delle serie temporali. Il modello proposto per ciascun gruppo assume la seguente scomposizione per G_t:

$$G_t = T_t + C_t + u_t$$

dove T_t indica la componente del trend; C_t la componente ciclica; u_t la parte stocastica del modello.
Poiché le rilevazioni della glicemia sono risultate incomplete (percentuale media di dati mancanti per soggetto = 7%), per il confronto della dinamica della glicemia nei due gruppi sono state utilizzate le serie temporali dei valori medi delle misurazioni ad ogni tempo in ciascun gruppo.
La stima del trend in entrambi i gruppi è stata ottenuta utilizzando un polinomio di grado n, stimato sulla base della bontà di adattamento e significatività statistica dei coefficienti. Il confronto del trend fra i due gruppi è stato effettuato utilizzando la media integrale, M:

$$M = \frac{1}{L}\int_0^L T_t dt$$

dove L=72 ore (h) è l'intervallo di tempo di osservazione.
La serie ottenuta depurando la serie temporale dalla componente T_t, è stata utilizzata per stimare la componente ciclica del modello, C_t. L'analisi di Fourier è stata impiegata per valutare il periodo del ciclo e la componente ciclica è stata ottenuta come riportato da Brockwell P.J and Davis R.A. [***3***].
Infine, la serie è stata depurata anche dalla componente C_t e la serie stocastica (u_t) è stata analizzata con un modello ARMA (p, q):

$$u_t = \varphi_1 u_{t-1} + \cdots + \varphi_p u_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

dove ε_t è un *white noise,* ossia un processo con media 0, varianza costante e autocorrelazione nulla.
Per tutte le analisi, la significatività statistica è stata stabilita ad un livello di probabilità del 5%.

Risultati

I 31 soggetti valutati (18 maschi) avevano un'età mediana pari a 15 anni, una durata mediana del diabete di 7 anni e HbA1c mediana del 7%. Il gruppo A era composto da 18 soggetti e il gruppo B da 13. I due gruppi non sono risultati significativamente diversi in termini di età, genere, BMI, durata del diabete e HbA1c.
I risultati dell'analisi della serie temporale sono riportati nella ***tabella 1***. La regressione polinomiale ha individuato per il trend un polinomio di terzo grado per entrambi i gruppi.
La differenza tra le glicemie medie stimate dal trend in ciascun gruppo non è risultata statisticamente significativa.
L'analisi di Fourier sui dati depurati dal trend ha evidenziato una componente ciclica di periodo 24 h in entrambi i gruppi. La stima della componente ciclica ha permesso di valutare le ampiezze delle oscillazioni glicemiche, con una differenza tra i due gruppi di circa 20 mg/dl. Inoltre, i tempi in corrispondenza del massimo e del minimo dell'oscillazione glicemica nei due gruppi sono risultati sovrapponibili.
Per il gruppo A è stato individuato un modello ARMA con un numero di ritardi pari a 4 (20 minuti) mentre nel gruppo B il numero di ritardi è risultato pari a 2 (10 minuti); per entrambi i modelli la funzione di autocorrelazione (ACF) e autocorrelazione parziale (PACF) dei residui hanno indicato un processo *white-noise*, evidenziando la correttezza dei modelli individuati.

Tabella 1. Confronto della serie storica glicemica tra i due gruppi

	Gruppo A n=18	**Gruppo B** n=13	p
Valutazione del trend			
Media Integrale (h·mg/dl)	162.15	155.11	0.922
Dev.St (h·mg/dl)	43.58	56.85	
Valutazione del ciclo			
Periodo (h)	24	24	
Ampiezza (mg/dl)	93.28	113.20	
t_{max}	09:15	09:25	
t_{min}	12:10	12:40	
Modello ARMA			
(p, q)	(4, 3)	(2, 3)	
φ_1	1.69	1.8	
φ_2	-0.27	-0.81	
φ_3	-0.85	-	
φ_4	0.41	-	
θ_1	-0.52	-0.84	
θ_2	-0.27	0.34	
θ_3	0.29	-0.18	
σ^2	6.35	8.62	
AIC	4072.76	4333.64	

t_{max}: il tempo in corrispondenza al massimo valore glicemico;
t_{max}: il tempo in corrispondenza al minimo valore glicemico;
AIC: Akaike Information Criterion

Conclusioni

Il modello individuato sembra adeguato a descrivere la dinamica delle glicemia media nei due diversi processi di controllo della somministrazione insulinica durante l'attività fisica. La struttura del modello ha permesso di valutare le diverse componenti della serie temporale: trend, ciclicità e componente stocastica. L'innalzamento della soglia di glicemia da 70 mg/dl a 90 mg/dl durante l'attività fisica, non sembra apportare variazioni di rilievo nella dinamica glicemica, sia per quanto riguarda l'andamento dei valori medi (esposizione glicemica) che per la variabilità in termini di ampiezza e momenti in cui si verificano le oscillazioni. Inoltre la glicemia a un determinato tempo è condizionata da valori appartenenti ad un breve lasso temporale precedente, paragonabile per entrambe le soglie definite.

Il modello dinamico proposto permette di sfruttare al meglio le informazioni ottenute con il monitoraggio glicemico continuo, e quindi può rappresentare un utile strumento per identificare i determinanti dell'andamento del controllo metabolico e per personalizzare l'intervento educativo e terapeutico negli adolescenti con diabete di tipo 1.

Bibliografia

[**1**] Herbst A, Kordonouri O, Schwab KO, et al. Impact of physical activity on cardiovascular risk factors in children with type 1 diabetes: a multicenter study of 23,251 patients. Diabetes Care 2007;30:2098-100

[**2**] Choudhary P, Olsen BS, Conget I, et al. Hypoglycemia Prevention and User Acceptance of an Insulin Pump System with Predictive Low Glucose Management. Diabetes Technol Ther 2016;18(5):288-91

[**3**] Brockwell PJ, Davis RA. Introduction to time series and forecasting. Springer-Verlag, 2002

PENALIZED ESTIMATION IN LATENT MARKOV MODELS, WITH APPLICATION OF MONITORING SERUM CALCIUM LEVELS IN END-STAGE KIDNEY INSUFFICIENCY

Farcomeni Alessio, Nofroni Italo, Spagnoli Alessandra, Vestri Annarita

Dipartimento di Sanità Pubblica e Malattie Infettive Sapienza – Università di Roma

Introduction

Latent Markov (LM) models are now a well established tool for the analysis of longitudinal regression models. They can be seen as generalized linear models with flexible non-constant random effects.
Outcomes are possibly multidimensional and extensions are available for several complex situations, including the case of informative drop-out. See Bartolucci *et al.* [***1***] for a general account and *R* package *LMest,* freely available on *CRAN for download*.
With longitudinal categorical outcomes, as in our motivating example, the response can be seen as a direct or indirect measure, with error, of the latent categorical predictor. This flexible but parsimonious formulation allows the user to take into account time-varying unobserved heterogeneity. Additionally, patients can be clustered according to their latent trajectories, making random effects interpretable and directly useful for identifying subgroups. Similarly, when there are several repeated measures, one can cluster time-occasions. There are several application fields in medical statistics, often based on a single or few independent, relatively long, time series.
A general limitation of LM models, especially in biomedical applications, is that often a large number of observations or a large number of repeated measures are needed to obtain reliable estimates. With small data sets the likelihood might be unbounded due to separation problems, and very unstable estimates might be obtained. Parameter estimates or their standard errors might be very far from zero, making them completely unreliable.
A typical example where the LM model could be very informative is our motivating application: serum calcium (Ca) of patients with severe kidney insufficiency was measured before surgery, and then 1 month, 1 year and 3 years after parathyroidectomy (surgical removal of parathyroids). In this example we have n=52 patients and 4 measurement occasions. A LM model would be very appropriate for the data at hand. Given the long time horizon and the very few subject-specific covariates collected it is natural to expect that there might be unobserved factors influencing the outcome, and that these are changing over time. Additionally, clustering patients might be useful to identify those who benefit the most or the least from surgery. Unfortunately, the MLE of most LM model formulations is unbounded, or at least empirically breaks down, even for small numbers of latent components.
In this work we adapt the well-known penalized likelihood approach of Firth [***2***] for LM models. This is based on an ad-hoc inline penalty to reduce bias of maximum likelihood estimates. The resulting estimates have a lower bias than classical maximum-likelihood ones, and do not suffer from lack of convergence even with small data matrices.

Methods

We do not maximize the likelihood to obtain parameter estimates. Instead, we formulate an objective function which is the sum of the log-likelihood and half the log determinant of the information matrix. A difficulty with LM models is that the information matrix is seldom available in closed form. We demonstrate

that computation of the penalized likelihood can be performed with a simple modification of the usual forward recursion targeting the observed likelihood, therefore having very little additional computational cost. The method therefore proceeds as follows: for an initial starting solution perform *n* forward recursions to compute the individual contributions to the objective function and the information matrix. Use then a numerical maximization routine to improve the current solution, until convergence. For our application we have used the Nelder-Mead algorithm, which proceeds by approximating a Newton-Raphson iteration.
A related problem is that LM models are usually expressed as a function of a fixed number of latent states *k*. This parameter is usually chosen by repeatedly fitting the model for several values of *k* (say, *k*=1,2,3,4,5) and then minimizing the Akaike Information Criterion or the Bayesian Information Criterion. AIC for instance is the sum of minus twice the log-likelihood and twice the number of free parameters. These methods are based though on the assumption that the full likelihood has been used to estimate parameters. In our case, we have maximized a misspecified model, and AIC should be substituted with its generalization, the Takeuchi Information Criterion (TIC). In TIC, the penalty of twice the number of parameters is substituted with twice the trace of the cross product of the squared score and the inverse of the observed information at convergence. We derive an explicit formula for TIC in our context and demonstrate with extensive simulation studies that it works better than AIC for cardinality selection of the number of clusters.
Additional methodological details and *R* code are available with the accompanying paper Farcomeni [***3***].

Results

NKF-K/DOQI guidelines define safe serum calcium levels in the very limited range 8.5-9.5 mg/dl. While in absence of diseases the actual values can be expected to be almost always well within this interval, in end-stage kidney insufficiency calcium excretion is reduced, often leading to persistent hypercalcemia. Hypercalcemia can lead to stones, bone pain, nausea and vomiting, and even psychiatric overtones and cardiac arrhytmias. Acute hypercalcaemic crises can lead to coma. One possibility for treatment in this scenario is removal of parathyroids, whose hormones have the effect of increasing Ca levels. In our example we follow 52 patients who have undergone PTX for three years. We have very few predictors, and end up using only three occasion indicators and basal Phosphate serum levels. It is natural to expect a strong unobserved heterogeneity, as many important factors were not measured and the time range of the experiment is quite long. Some of these factors (e.g., co-morbidities, vitamin D intake) might have changed during the observation period, hence unobserved heterogeneity might easily be time-varying.
Since the primary endpoint of the analysis is connected with NKF-K/DOQI ranges, we work with the response coded as an ordinal variable with three levels: below, within, and above the recommended ranges. This corresponds to hypocalcemia, good control of serum Ca, and hypercalcemia.
In the following ***Table 1*** we compare parameter estimates for the usual MLE for a LM model with *k* latent groups, and our penalized estimates (pMLE). The first three parameters are associated with the time dummies (1 month, 1 and 3 years) and the last with basal P.

Table 1. Parameter estimates for the usual MLE for a LM model with *k* latent groups, and our penalized estimates (pMLE).

	MLE					pMLE				
	***k*=1**	***k*=2**	***k*=3**	***k*=4**	***k*=5**	***k*=1**	***k*=2**	***k*=3**	***k*=4**	***k*=5**
β1	-3.6	-6.3	-59.2	-38.0	-359.8	-3.4	-5.9	-6.0	-9.0	-7.9
β2	-2.5	-4.2	-35.8	-30.6	-213.3	-2.4	-4.2	-3.6	-5.8	-5.7
β3	-2.5	-4.3	-43.9	-32.6	-276.9	-2.3	-4.3	-3.4	-5.6	-5.1
β4	0.1	0.3	0.4	1.3	13.0	0.1	0.4	0.2	0.3	0.2

It is apparent that as soon as k>2 the MLE breaks down and is not reliable any more. On the other hand, the pMLE can be reliably estimated for any value of *k*. Our TIC criterion leads to choose *k*=3, where it can be seen that the risk of hypercalcemia decreases abruptly after surgery, but tends to increase again over time. Higher basal P values are also associated with higher risk of hypercalcemia.

References

[**1**] Bartolucci F, Farcomeni A, Pennoni F. Latent Markov Models for Longitudinal Data, Chapman & Hall/CRC Press 2013.

[**2**] Firth D. Bias reduction of maximum likelihood estimates. Biometrika 1993;80:27-38.

[**3**] Farcomeni A. Penalized estimation in latent Markov models, with application to monitoring serum Calcium levels in end-stage kidney insufficiency. Biometrical Journal 2017;*in press.*

THE MUSIC THERAPY – SESSION ASSESSMENT SCALE (MT-SAS): CONCEIVING, DEVELOPING, REFINING AND VALIDATING A NEW TOOL FOR PSYCHOMETRIC ASSESSMENT IN MUSIC THERAPY

Gnesi Marco[1], Monti Maria Cristina[1], Oasi Osmano[2], Gianotti Marta[3], Imbriani Chiara[4], Montomoli Cristina[1], Raglio Alfredo[4]

1. *Department of Public Health, Experimental and Forensic Medicine, Section of Biostatistics and Clinical Epidemiology, University of Pavia, Pavia, Italy*
2. *Department of Psychology, Catholic University of Milan, Milan, Italy*
3. *Fondazione Istituto Ospedaliero di Sospiro, Cremona, Italy*
4. *Istituti Clinici Scientifici Maugeri, Music Therapy Laboratory, Pavia, Italy*

Introduction

Active Music therapy (MT) interventions aim to develop a close relationship between patient and therapist; this mainly happens in a non-verbal context through sonorous-music production, and makes it possible for the therapist to "explore" the patient's personality, emotional world, resources, and even personal history [***1***].

As any other form of treatment, music therapy needs assessment protocols in order to evaluate its relevant outcomes, which is crucial in planning and following an intervention (American Music Therapy Association's Standards of Clinical Practice, American Music Therapy Association, 2015). Nevertheless, there is a huge lack of tools specifically developed and/or validated for MT: this could lead to information bias [***2***]. Moreover, to our knowledge none of those tools is designed to assess non-verbal communication in MT.

The aim of this study was to develop and validate a new observational tool for MT assessment, namely the *Music Therapy – Session Assessment Scale* (MT-SAS), for the evaluation of non-verbal communication between patient and therapist in videotaped MT sessions.

Methods

Pilot Study: Grid Design and Refinement

MT-SAS has been conceived in the frame of a strong expertise in MT, and particularly in investigating the relationship involving therapist and patient [***3***].

A first set of items was developed during a series of meetings held by the five therapists involved in the project. The definition of those items took into account both theoretical concepts and practical issues as they rose from those brainstormings, attempting to balance the needs for clinical and research applications. The assessment process was designed to be carried out on videotaped session, in order to avoid recall bias and to observe from a phenomenological point of view music therapy interactions. The first version of the grid was used by the five therapists to blindly assess the videotapes of the same 10 sessions. Then, critical issues were discussed in a new meeting, aiming to improve the items. A preliminary statistical analysis of those pilot data also contributed in identifying and addressing the needs for refining (Krippendorff's Alpha index for ordinal variables and Multiple Correspondence Analysis, see below). As a result of this pilot study, the final version of MT-SAS was proposed, together with a detailed guideline for assessment.

Final Version of MT-SAS

The final MT-SAS consists of seven items regarding countenance, non-verbal communication and sound-music communication, namely: 1) Eye contact; 2) Body reciprocity; 3) Countenance and body signals indicating emotional engagement; 4) Refusal/Disturbed behaviors; 5) Sonorous-music productions; 6) Attuned sonorous-music productions; 7) Dynamism/Variations.

The scoring system is binary, so each item can be assessed either as "predominantly absent" or "predominantly present". The evaluation takes into consideration the whole therapeutic session and is made *indirectly* (i.e. on videotapes or by a person behind a mirror, rather than by the therapist itself during the session).
It should be noticed that the Italian version of MT-SAS and related guidelines were used during the entire validation process. Although a translated English version has been produced, and even though the items are non-linguistical, the wording of their descriptions could influence a rater and we do not guarantee for the use of the translated version.
Validation Study
The validation process of the final MT-SAS started from the evaluation of inter-rater reliability. For this purpose, the videotapes of 10 sessions from different patients were assessed by each of the five therapists (raters) involved in the study, blindly one from another. The concordance in their scorings was evaluated by applying Krippendorff's Alpha index for nominal variables [***4***]. Inter-rater reliability was considered optimal if the index was 0.80 or higher, suboptimal if at least 0.65, non-optimal otherwise.
The sample size for the construct validation sample was determined by enrolling at least 5 patients for each score of each item (minimum 70, actually enrolled 100). Patients had to be in the middle phase of a MT treatment, and they (or their legal tutor) had to sign a written informed consent.
The videotape of one session per each patient was assessed by one of the five therapists (20 sessions per therapist). After suitable descriptive analyses, a tetrachoric correlation matrix of scorings was produced, in accordance with the dichotomous nature of the items. An Explorative Factor Analysis (EFA) was applied to the tetrachoric correlation matrix in order to unveil the latent structure of data, thus looking at construct validity; Principal Component Factors was chosen as extraction method. The number of factors to be retained was decided according to either Kaiser's criterion (eigenvalues ≥ 1), Cattel's criterion (screeplot), or Parallel Analysis. Pattern matrix of the factors was interpreted after applying a suitable rotation, preferring orthogonal to oblique methods when correlations among non-orthogonal factors were weak. In the end, the orthogonal method VARIMAX was adopted. Items were assumed to be entirely loaded by the factor for which they had the highest saturation. Each factor individually underwent a second-level EFA to assure monodimensionality; saturations and uniquenesses from the second-level EFA were used for interpretation of each single factor. Finally, given that it is a widespread habit in Psychometrics, Cronbach's Alpha for standardized items was computed to assess internal consistency of each factor, even though this approach has been criticized from both the mathematical (underestimation) and conceptual standpoint (what is "internal consistency"? how does it relate to reliability?), raising concerns on is meaning and trustworthiness [***5***]. Therefore, we recommend a cautious interpretation of the reported index.
Finally, in order to look at MT-SAS also in a perspective of concurrent validity, Multiple Correspondence Analysis (MCA) with Burt's matrix approach was applied in order to explore if the scorings tended to be gathered in archetypical profiles, thus discerning groups of subjects on relevant and interpretable characteristics. MCA focuses on single categories rather than entire items (like EFA would do), and produces dimensions from the associations between items' categories. The number of dimensions to be retained was decided in order to reach an explained inertia (i.e. variability) of 80% or more; results were then interpreted graphically on the coordinate plot, in which categories of items are represented (the closer, the more strongly associated).
All the statistical analyses were performed using Stata 13 (StataCorp. 2013. Stata Statistical Software: Release 13. College Station, TX: StataCorp LP).

Results

Pilot study

The first version of the grid consisted of 11 items with 4-levels Likert scores. Its use in a first round of assessment highlighted hurdles in certain items (reflected by a low inter-rater reliability). Moreover, MCA showed that the two lowest and the two highest scores tended to gather, suggesting that the multilevel scoring system was even too subtle and the loss of information would not be relevant with a binary one. An *ex-post* dichotomization of the scores showed that inter-rater reliability could be improved by eliminating the difference between "extreme" and "moderate", perhaps because it reduced the impact of therapists' different attitude to attribute extreme scores.

Taking into account all these considerations, the grid was refined and the final version (see Methods) was proposed. To further improve consistency in the assessment protocol, the final grid was completed with a detailed guideline, and a shared re-training was carried out with all the therapists.

Validation Study

The validation sample consisted of 100 patients with the following clinical conditions: Alzheimer's Diseases (46 patients, Females: n=31), Frontotemporal Dementia (3 patients, Females: n=2), Autism Spectrum Disorders (13 children, Females: n=4), Multiple Sclerosis (1 adult, Female: n=1), Amyotrophic Lateral Sclerosis (2 patients, Female: n=1), Stroke (14 patients, Females: n=8), Parkinson's Disease (12 patients, Females: n=6), Intellectual Disabilities (9 adults patients, Females: n=6).

As regards observed scores, Items "Eye contact", "Body reciprocity", "Sonorous-musical productions", and "Attuned sonorous-musical productions" had a marked prevalence of 'Predominantly present', oppositely to "Countenance and body signals indicating emotional engagement", "Non-acceptance/Disturbance behaviors", and "Dynamism/Variation". The occurrence of missing data was negligible, so subsequent analyses of MT-SAS were performed after a listwise deletion of subjects with missing MT-SAS data.

The inter-rater reliability, evaluated by means of Krippendorff's Alpha, was optimal or suboptimal for most of the items, excepting Items 2 and 3 (alpha of 0.56 and 0.60 respectively); in the former case, in particular, it might be explained by a low variability. Therapists participated in a cooperative retraining to further refine their assessment skills.

The matrix of tetrachoric correlations showed moderate or strong correlations among all items, with the exception of "Eye contact", and suggested a reverse coding for "Non-acceptance/Disturbance behaviors" (as expected by design). Coherently, EFA unveiled two latent factors according to Kaiser's criterion (given that neither Cattel's criterion nor Parallel Analysis were informative), with an explained variance of 76%. In a PROMAX-rotated solution factors had very low correlations, so the VARIMAX method was adopted to provide the rotated pattern matrix for interpretation. Almost all the items had their highest saturations on the first factor, with "Eye contact" essentially alone on the second one.

Monodimensionality of the first factor was verified by a subsequent individual-factor EFA. Although saturations of "Countenance and body signals…" and "Non-acceptance/Disturbance behaviors" were slightly lower, and uniquenesses slightly higher, the factor was pretty homogeneous. The counter-coding of "Non-acceptance/Disturbed behaviors" was confirmed. The overall internal consistency is at least sub-optimal, as suggested by a Cronbach's Alpha of 0.70 (with all the concerns on the interpretation of this index).

Finally, in order to explore the discriminant (profiling) capability of MT-SAS, we applied Multiple Correspondence Analysis on all the items. The first dimension alone was able to explain more than 83% of inertia (i.e. "observed variability"), while the second dimension only accounted for 2.6%; thus, only the first dimension was assumed relevant in splitting the potential profiles (see ***Figure 1***). A profile clearly appears on the left: it includes all the positive scores, thus suggesting an archetypical subject characterized by an *attuned behavior* on our non-verbal items. The negative scores are all placed in the right part of the plot, representing a general profile of *disattuned behaviors* in which three further "subclusters" can be identified:

one gathers Items 1, 3 and 7, another Items 4, 5 and 6, while Item 2 tends to be alone. It is useful to notice that Item 1 ("Eye contact") is the less-separated on the horizontal dimension, while being the most-separated on the vertical axis; also Items 3 and 4 ("Countenance and body signals…" and "Non-acceptance/Disturbed behaviors"), albeit well integrated in the profiles on the horizontal dimension, are quite spread vertically.

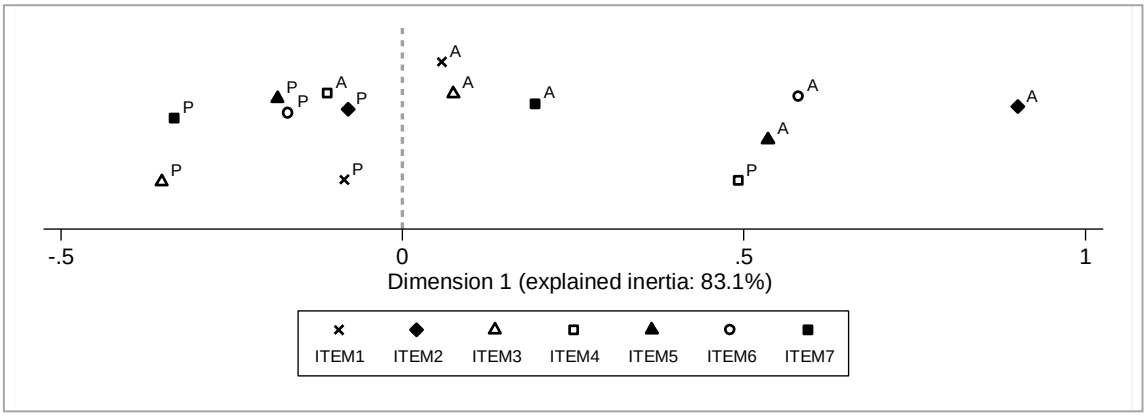

Figure 1. Coordinate plot after MCA. Normalization: principal coordinates.

Conclusions

Following previous expertise in the field and according to Stern's theoretical framework, we developed a new tool called MT-SAS. Carrying out the assessment on videotapes of entire MT active sessions make it a versatile tool with potential applications both in research and clinical practice. The validation analyses proves that MT-SAS is appropriate for assessment (construct validity), and the concurrent validity – i.e. the coherence of the structure (EFA) with the theoretical framework and the profiling capability (MCA) – confirms that it is appropriate as regards sonorous-music and non-verbal behaviors in active MT sessions. MT-SAS could also allow a pseudoquantitative assessment. Given the homogeneity of the main factor of MT-SAS (as suggested by the individual-factor EFA), we propose the creation of an ordinal total score by summing Items 2 to 7 (adding 1 for each item marked 'Predominantly present', or 'Predominantly absent' for Item 4). The item of "Eye contact" can be reported as a +/- added to the previous score.

We have to point out that the evaluation of convergent validity was not possible, due to the lack of a valid gold standard. Furthermore, test-retest reliability was not applicable: the trait is supposed to be not stable over time, so any change between two sessions can reflect actual variability in the trait rather than biases or reliability issues. Another potential limitation arises from the choice of a composite sample (as regards gender, age strata, clinical diagnosis), notwithstanding that the patients in common MT practice are a composite population. In any case, we also looked for differences in items' distributions by gender, age or clinical diagnosis, and coherently with the theoretical background only clinical groups had relevant associations to items' score distribution. This may reflect actual differences in the behavioral outcomes of those different conditions, and is another proof of concept for concurrent validity.

To sum up, MT-SAS can be considered an easy and valid tool that music therapists can use in active MT individual treatment to consistently assess a relational trend and to correlate MT process outcomes with clinical results.

References

[1] Wigram T, Gold C. Music therapy in the assessment and treatment of autistic spectrum disorder: clinical application and research evidence. Child: care, health and development 2005;32(5):535-42.

[2] Wigram T. Improvisation. Methods and techniques for music therapy clinicians, educators and students. Jessica Kingsley, London 2004.

[3] Raglio A, Traficante D, Oasi O. The evaluation of music therapy process in the intersubjective perspective: the Music Therapy Rating Scale. A pilot study. Pragmatic and Observational Research, 2011; 2:19-23.

[4] Hayes AF, Krippendorff K. Answering the call for a standard reliability measure for coding data. Communication methods and measures 2007;1(1):77-89.

[5] Sijtsma K. On the use, the misuse, and the very limited usefulness of Cronbach's alpha. Psychometrika 2009;74(1):107-20.

AN INVERSE PROBABILITY WEIGHTING APPROACH TO DEAL WITH INFORMATIVE CENSORING WITH APPLICATION TO CHILDHOOD LEUKEMIA

Bernasconi Davide Paolo, Blanco Lopez Jessica, Rossi Emanuela, Antolini Laura, Valsecchi Maria Grazia

School of Medicine and Surgery, University of Milano-Bicocca, Via Cadore 48, 20900 Monza, Italy

Introduction

In middle- and low-income countries average survival rates of children affected by Acute Lymphoblastic Leukemia (ALL) are much lower than those observed in the western part of the world. This gap can be explained by several factors including delayed diagnosis and limitations in health care resources. Another cause of poor outcome in developing countries is abandonment of treatment, which occurs when patients discontinue treatment. AHOPCA (Asociacion de Hemato-Oncologia Pediatrica de Centro America) is a network of Hospital Units specialized in childhood cancer treatment in Central America (Guatemala, Honduras, El Salvador, Nicaragua, Costa Rica), Panama and Dominican Republic [***1***]. In all the studies conducted on children treated in AHOPCA centers, the clinicians had to deal with a high rate of abandonment of treatment varying across different countries. In early studies, abandonment was occurring in up to 50% of children and was the first cause of treatment failure. Recent studies showed that abandonment is associated with the disease course and treatment toxicity and with the socioeconomic situation of the family. In this context, the need of comparing treatment protocols adopted in different countries and periods becomes challenging because abandonment differently affects each specific cohort. We analyzed the data of Hondurans children enrolled in two subsequent protocols for ALL treatment, GHS 2000-2007 and BFM 2008-2015, with a different abandonment rate: 21.2% and 13.4% respectively. Our goal was to assess and compare the potential outcome of the two protocols in the counterfactual situation where abandonment is not present and given all patients exposed to both treatment protocols. The outcome endpoint was the event free survival (EFS) defined as the time from diagnosis to the first event among relapse, resistance to treatment, second malignant neoplasm and death.

Methods

The naïve approach to estimate the potential EFS is the protocol-specific Kaplan-Meier (KM) estimator considering abandons as censored observations. However, this method will typically provide biased results, as it assumes that the censoring mechanism (which includes abandons) is independent form the potential event time. This assumption is obviously unrealistic. The comparison between the KM curves of the two protocols cannot be interpreted as an average causal effect in the absence of randomization [***2***]. Marginal structural models based on the inverse probability of treatment and censoring (IPTC) weighting, under specific assumptions, allow the estimation of potential outcomes as if no informative censoring occurred and as if the whole cohort was exposed, or not, to a factor. We adopted an IPTC weight-adjusted Kaplan-Meier method [***3***]. The estimation of the time-dependent censoring (abandon) weights was carried out using the Aalen additive model [***4***], while logistic regression was adopted for the treatment (protocol) weights. In both models, we included information on patients and disease characteristics as well as on their socio-economic conditions (living conditions, nutritional state, family type, parental education, family income, possession of a phone, time to reach the hospital from home). Normalization and truncation of the weights was also considered. Pointwise 95% confidence intervals for the potential EFS were computed using bootstrap resampling. In addition, we investigated the consistency of the results obtained by performing an alternative counterfactual analysis based on the application of G-computation to the Aalen model [***5***].

Results

The factors showing a stronger association with abandonment were: GHS protocol, younger age, undernourishment, absence of a phone at home and longer time to reach the hospital. Children enrolled in the GHS protocol had less central nervous system involvement and malnutrition, belonged to better educated families and lived more frequently in a urban context than those of the BFM protocol. According to the naïve analysis (***Figure 1A***), the estimated potential EFS at 5 years from diagnosis (with 95% confidence interval) was 0.567 (0.521-0.618) in the GHS protocol and 0.522 (0.473-0.575) in the BFM protocol with overlapping confidence intervals. Using the IPTC weighted KM we obtained 0.542 (0.511-0.627) and 0.554 (0.474-0.581), respectively (***Figure 1B***), again with overlapping confidence intervals but with and earlier tendency to better outcome of the more recent BFM protocol.

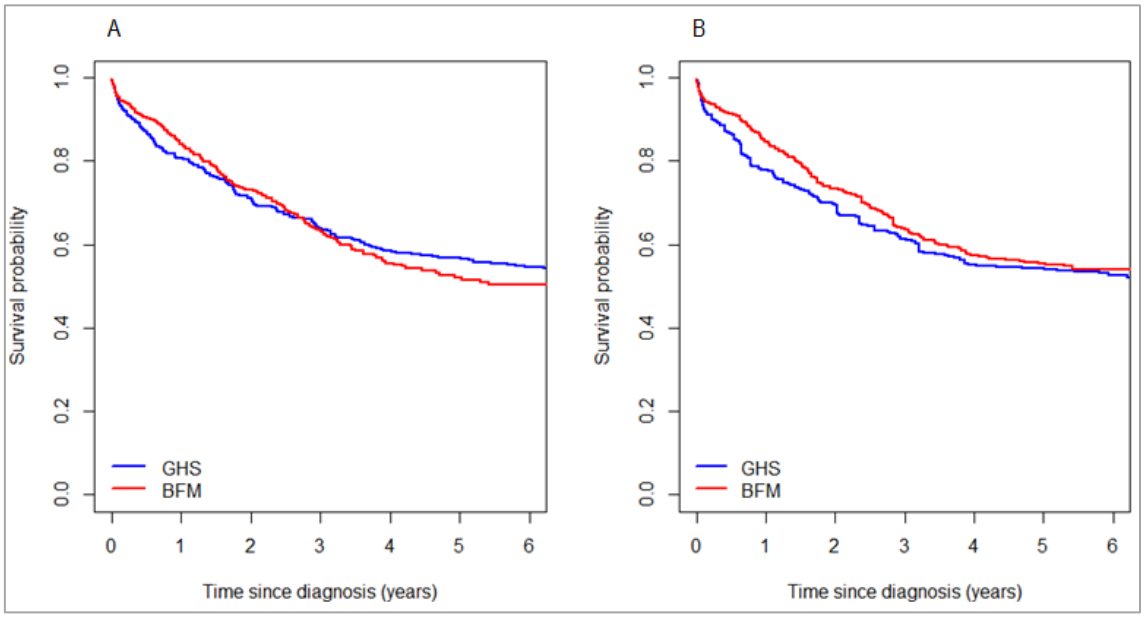

Figure 1. Comparison of the potential EFS in the two treatment protocols estimated using: the protocol-specific KM estimator censoring observations at abandon (A), the IPTC weighted KM estimator (B).

Conclusions

Using an IPCT weighted method may be useful to compare the efficacy of treatment protocols avoiding the unbalance due to the possibly different amount of observed abandons and also eradicating the possible selection bias due to the different case-mix of patients on which each protocol was administered.
This is possible only under some untestable assumptions (i.e identifiability, conditional exchangeability, coarsening at random and correct model specification for the weights). However, it allows to analyze survival data in the presence of informative censoring.

References

[1] Barr RD, Antillon F, Baez F, et al. Asociacion de Hemato-Oncologa Pediatrica de Centro America (AHOPCA): Model for Sustainable Development in Pediatric Oncology. Pediatr Blood Cancer 2014;61:345-54.

[2] Robins J, Hernan M, Brumback B. Marginal Structural Models and Causal Inference in Epidemiology. Epidemiology 2000;11(5):550-60.

[**3**] Edwards JK, Cole SR, Lesko CR, et al. An illustration of inverse probability weighting to estimate policy-relevant causal effects. Am. J. Epidemiol 2016;184(4):336-44.
[**4**] Martinussen T, Scheike T. Dynamic regression models for survival data, New York: Springer, 2006.
[**5**] Snowden JM, Rose S, Mortimer KM. Implementation of G-computation on a simulated data set: demonstration of a causal inference technique. Am. J. Epidemiol 2011;173(7):731-38.

INTERVAL BASED PHASE I DESIGNS: AN APPLICATION IN A RARE SUBGROUP OF PATIENTS

Galimberti Stefania, Magri Maria Chiara, Valsecchi Maria Grazia

Center of Biostatistics for Clinical Epidemiology, School of Medicine and Surgery, University of Milano-Bicocca

Introduction

Innovative therapies to face rare or very rare diseases are the new challenge of clinical research. The concern about very rare diseases is very specific and implies ethical, methodological and practical issues. In this perspective, the traditional trial designs fail to achieve reliable evidence, mainly because they generally refer to large populations. In phase I clinical trials that address safety definition and determine the maximum tolerated dose (MTD) of new drugs, the rarity of the disease further reduces the number of participants involved.

The majority of phase I clinical trials still use the rule based 3+3 design, owing to its simplicity, despite it has been proved to lack of precision and accuracy in MTD selection [***1***]. Thus, when a 3+3 trial is conducted, current patients are more likely to be treated outside the therapeutic window during the dose escalation process of the phase I trial, but, more importantly, this will happen for future patients in trials involved in the following phases of clinical research. The Continual Reassessment Method (CRM) is a model based design that adopts a Bayesian approach to lead to a phase I dose escalation. The CRM design has largely shown good operative characteristics, as accuracy in MTD selection, suitable allocation of patients among doses and correct toxicity management [***2***]. Nevertheless, physicians hardly refer to the CRM, as it is not easy to set and carry out.

More recently two interval based designs were introduced within the Bayesian framework: the modified Toxicity Probability Interval (mTPI) design [***3***] and the Bayesian Optimal Interval (BOIN) design [***4-5***]. They represent valid alternatives to CRM, as they can be easily implemented, and show attractive operative characteristics, at least in simulations involving more than 25 patients [***6***]. Specifically, mTPI and BOIN designs let physicians inspect the rules leading to the dose escalation process in advance, while CRM is completely blinded about cohort allocation before the study starts.

The motivating clinical context is represented by a phase I trial on an innovative therapy in a rare subgroup of patients with Acute Lymphoblastic Leukemia (ALL). Patients who relapsed or were refractory (e.g. never achieving a Complete Remission) after Hematopoietic Stem Cell Transplantation (HSCT) have a very dismal prognosis, and they might potentially benefit of the killing activity against leukemic cell targets of modified Cytokine Induced Killer (CIK) cells. The experimental treatment consists in donor-derived CIK cells modified with anti-CD19 chimeric antigen receptor (CAR) by transposon.

Two are the main aims of this work:

1) define a strategy that uses interval based designs in the development of phase I trials in the context of rare diseases or rare subgroups of patients;
2) design a phase I trial to determine the safety of Allogeneic (donor derived) CIK cells transduced with a transposon CD19-CAR gene (CARCIK-CD19) in adult and pediatric transplanted patients with relapsed or refractory B-cell precursor ALL.

Methods

The first step in implementing both designs is the definition of an interval of acceptable toxicity around the target Dose Limiting Toxicity (DLT) probability p_T $[p_T-\varepsilon_1, p_T+\varepsilon_2)$ and two additional sub-intervals of low $[0, p_T-\varepsilon_1)$ and high toxicity rates $[p_T+\varepsilon_2, 1]$.

In the mTPI design DLTs follow a binomial distribution. Whenever a new cohort of patients is assigned to a dose level i (d_i, with i=1…K), the posterior distribution of the toxicity probability $p_i \sim Beta(1+x_i, 1+n_i-x_i)$ is derived by the product of the binomial likelihood and the beta prior Beta(1,1), assuming x_i and n_i being the observed DLTs and the patients treated at dose d_i, respectively. The unit probability mass (UPM) is then calculated for each sub-interval, and is defined as the posterior probability that p_i is in the subinterval, divided by the width of the subinterval. The interval with the highest UPM dictates the dose decision for the next cohort of patients, where escalate, stay at the current dose or de-escalate are associated to the intervals $[0, p_T-\varepsilon_1)$, $[p_T-\varepsilon_1, p_T+\varepsilon_2)$ and $[p_T+\varepsilon_2, 1]$, respectively .
The BOIN design selects optimal interval boundaries (λ_1 and λ_2) to refine the pre-fixed interval of acceptable toxicity around p_T, minimizing the decision errors of dose assignment under a bayesian framework. Dose transition is based on the observed toxicity rate at the current dose d_i ($\hat{p}_i = x_i/n_i$): if $\hat{p}_i < \lambda_1$ escalate to d_{i+1}, if $\lambda_1 < \hat{p}_i < \lambda_2$ stay at dose d_i, and if $\hat{p}_i > \lambda_2$ de-escalate to d_{i-1}. In addition, for both designs, stopping rules are recommended to take into account extra safety:

1) early study termination: $Pr(p_1 > p_T \mid data) > p$
2) dose exclusion: $Pr(p_i > p_T \mid data) > p$

where p is predefined and usually fixed at 95%.
The trial is terminated when the maximum sample size is reached. At the end of the trial, all observed data are used for MTD selection. MTD is the dose, between the admissible doses satisfying the safety criterion (i.e. $Pr(p_i > p_T \mid data)$ p), with the isotonic estimates of toxicity rate ($\hat{p}_i^*$) closest to the target p_T: $\min_i |\hat{p}_i^* - p_T|$, where isotonic transformation are obtained applying the Pooled Adjacent Violators Algorithm [**7**] and assure that: $\hat{p}_1^* \leq \cdots \leq \hat{p}_i^* \leq \cdots \leq \hat{p}_K^*$.
In order to select the interval based design more suitable in our clinical context, the performances of mTPI and BOIN were assessed through simulations. Eighteen scenarios of toxicities were considered for 4 doses, starting from a situation of no DLT occurrence to the one with all doses highly toxic. We considered cohorts of 3 patients, different sample sizes (n=12,15,18) and 2000 simulations for each configuration. As for the design parameters, the target toxicity rate was set at 30% with a tolerance toxicity interval between 20 and 40% and two different risk levels for the control of excessive toxicity were assumed (e.g. 90% and 95%).

Results

The operative characteristics of the two interval based designs did not reveal any relevant difference in the configuration we explored neither in terms of percentage of correct MTD selection nor in percentage of patients treated at MTD. However, the Trial Monitoring Table which should guide the decisions on dose transitions in the BOIN design showed more flexibility compared to the one obtained for mTPI.
The dose escalation phase I trial on CAR-CIK was thus planned based on the BOIN design with parameters equal to the ones used in the simulations (p_T=0.30, $\varepsilon_1 = \varepsilon_2$=0.10). Due to rarity of the condition 18 patients will be included in the dose escalation phase of the study and 3 patients will be further enrolled at the MTD in a dose expansion phase. Based on simulation results we decided for a more stringent control of toxicity (p=90%) and we modified the algorithm for MTD selection. MTD will be defined as the dose with a toxicity rate closest to the target of 30%, but not superior to 40%, that is the upper bound of the tolerance interval of toxicity we fixed (p_T+□$_2$).

Conclusions

The application of the 3+3 design is the established practice in phase I trials mainly due to the difficulties in the implementation of CRM that is considered the gold standard in dose escalation studies. The interval based designs, which are emerging Bayesian phase I designs, are good alternatives to CRM and represent a compromise between 3+3 and CRM. We have appreciated their flexibility in exploring the spectrum of

possible doses still assuring patients safety and the fact that the algorithm for MTD selection and the stopping rules can be easily adapted to specific situations. Moreover, in this work we have shown that the application of interval based designs is feasible even in the context of rare diseases. Overall, our experience in the use of these bayesian phase I designs is positive and we think that they should deserve more consideration in the earliest phase of clinical research.

References

[**1**] Le Tourneau C, Lee JJ, Siu LL. Dose escalation methods in phase I cancer clinical trials. J Natl Cancer Inst 2009;101(10):708-20.

[**2**] Iasonos A, Wilton AS, Riedel ER. et al. A comprehensive comparison of the continual reassessment method to the standard 3 + 3 dose escalation scheme in Phase I dose-finding studies. Clin Trials 2008;5(5):465-77.

[**3**] Ji Y, Liu P, Bekele BN. A modified toxicity probability interval method for dose finding trials. Clinical Trials 2010;7:653-63.

[**4**] Liu S, Yuan Y**.** Bayesian Optimal Interval Designs for Phase I Clinical Trials. Journal of the Royal Statistical Society: Series C 2015;64:507-23.

[**5**] Yuan Y, Hess KR. Hilsenbeck S.G. et al., Bayesian Optimal Interval Design: a simple and well performing design for phase I Oncology Trials. Clin Cancer Res 2016;22(17):4291-301.

[**6**] Horton BJ, Wages NA, Conaway MR. Performance of toxicity probability interval based designs in contrast to the continual reassessment method. Stat in Med 2017;36(2):291-300.

[**7**] Barlow RE, Brunk HD. The isotonic regression problem and its dual. Journal of the American Statistical Association 1972;67:140-47.

ADHERENCE TO EVIDENCE-BASED DRUG THERAPIES AFTER MYOCARDIAL INFARCTION: IS GEOGRAPHIC VARIATION RELATED TO HOSPITAL OF DISCHARGE OR PRIMARY CARE PROVIDERS? METHODOLOGICAL CHALLENGES AND POLICY PERSPECTIVES.

Di Martino Mirko[1], Alagna Michela[2], Lallo Adele[1], Davoli Marina[1], Fusco Danilo[1].

1. *Department of Epidemiology, Lazio Regional Health Service, Roma, Italy.*
2. *Faculty of Education - Free University of Bolzano, Italy.*

Introduction

Patients who have had an acute myocardial infarction (MI) are at increased risk of repeated MI and death. Evidence-based prevention strategies include changes in life style and drug therapy. International guidelines agree on the use of combinations of drugs belonging to specific anatomical therapeutic chemical (ATC) groups: platelet aggregation inhibitors (antiplatelets), beta-blocking agents (β-blockers), agents acting on the renin-angiotensin system (ACEI/ARBs) and HMG-CoA reductase inhibitors (statins) [***1***]. The benefits of chronic poly-therapy in reducing cardiovascular disease have been clearly shown [***2***].

The gap in clinical practice

However, observational studies reported poor adherence to chronic poly-therapy. Therefore, therapies with proven benefit for MI are underused despite strong evidence that their use will result in better patient outcomes. Moreover, substantial geographic variation exists in the treatment of patients with acute myocardial infarction, and these gaps between knowledge and practice have important consequences in terms of equity in access to optimal care [***3***]. Unfortunately, from the current scientific evidence it is not possible to quantify how much of the "distance from clinical guidelines" is attributable to the patient behavior, to the therapeutic approach recommended at hospital discharge or to the primary care providers, such as local health districts. The local health district is a body delegated by the National Health System to provide health care to a specific area. Each local health district is composed of a well-defined group of general practitioners sharing the same clinical guidelines and participating in the same learning interventions, coordinated by a district director. The analysis of these "components of variation" may be a useful tool to define areas for more targeted interventions aimed at improving adherence to guidelines and equity in health care.

The objectives of this study are as follows: to measure the adherence to chronic poly-therapy after MI in clinical practice; to quantify and compare the proportions of variation attributable to the hospitals of discharge and to the primary care providers; to identify determinants of adherence to poly-therapy.

Methods

Setting and study cohort

The study was based on the population living in the Lazio region of Italy, which comprises approximately 5 million persons. Using data from the regional Hospital Information System (HIS), the study included a cohort consisting of all patients discharged from hospitals between January 1, 2007 and October 31, 2010 with a diagnosis of MI (index admission). An MI was defined as either a primary diagnosis of ICD-9-CM codes 410.xx or as a primary diagnosis of an MI-related condition along with a secondary diagnosis of 410.xx. Patients aged 35-100 years at discharge were considered for inclusion in the analysis. Patients with hospitalizations for MI or related causes in the 9 years before index admission were not considered eligible for the study. Patients who were not registered in the regional health assistance file throughout the whole study period were excluded, because they could not be retrieved from the regional health information system. Patients with a duration of the index admission > 21 days (95th percentile) were excluded from the

analyses as they were considered “statistical outliers”, probably representing extremely complex or instable patients. Finally, patients who received an outpatient regimen for less than 30 days were excluded, in order to allow a long enough time for consistently estimating the adherence to poly-therapy.

Follow-up

Individual follow-up for measuring drug exposure was considered to start on the first day after discharge from the index admission. The end of the observation period was defined as either the end of two-year follow-up, the time of death or the date of any hospitalization following discharge from the index admission, whichever occurred first. The last “censoring” criterion allows to measure the net impact of the hospital that has discharged the patient without the potential interference of subsequent hospitalizations.

Drug exposure: the adherence to medication

Drug exposure information was collected from the regional registry of all drugs dispensed by public and private pharmacies. Information about prescriptions of antiplatelets, β-blockers, ACEI/ARBs, and statins were retrieved for all patients. Adherence to medication was measured through the medication possession ratio (MPR), calculated as the number of days of medication supplied during the follow-up on the basis of defined daily doses (DDDs) divided by the number of calendar days in the follow-up. Adherence to individual medications was defined as a MPR ≥ 0.75. Adherence to chronic poly-therapy was defined as a MPR ≥ 0.75 for at least three of the four evidence-based drugs [***2***].

Statistical analysis

A map of the Lazio region was produced in order to show and compare the proportions of adherent patients by local health district. The classes used in the maps have been calculated applying the Jenks natural breaks optimization algorithm, which reduces the variance within classes and maximizes the variance between classes [***4***]. It is important to note that the data structure is not purely hierarchical. In fact, patients are nested within local health districts and within hospitals of discharge. However, the nesting structure may be less clear when we consider health districts and hospitals of discharge. In other words, we can say that patients are nested within the “cross-classification of health districts and hospitals”. Therefore, cross-classified logistic multilevel models were performed in order to analyze geographic variation, by measuring and comparing the proportions of variability attributable to hospitals of discharge and primary care providers [***5***]. The variance components were expressed in terms of Median Odds Ratios (MORs). The MOR quantifies the variation between clusters by comparing two persons from two randomly chosen, different clusters. Consider two persons with the same covariates, chosen randomly from two different clusters. The MOR is the median odds ratio between the person of higher propensity and the person of lower propensity. This measure is always greater than or equal to 1.00. If the MOR is 1.00, there is no variation between clusters. If there is considerable between-cluster variation, the MOR will be large [***6***]. The MORs were estimated controlling for patients' characteristics. In fact, explanatory variables that are divided very selectively across the groups can often explain a fair amount of group level variance. The interpretation would generally be that this does not reflect a real contextual effect, but rather the unequal composition of the groups. A cross-classified model was also applied to identify determinants of adherence to poly-therapy, properly taking into account the correlation within the specified clusters. Determinants of adherence were identified in two steps. First, the following factors were selected based on a priori knowledge [***7***]: gender and age of patient, educational level, discharge ward, length of stay of the index admission, percutaneous coronary intervention (PCI) during the index admission, use of antiplatelets, β-blockers, ACEI/ARBs or statins during the 12 months prior to the index admission (defined as at least 2 prescriptions), 15 co-morbidities retrieved from the hospital records for both the index admission and the 9 previous years; gender, age and organizational arrangement (none, association, network, group practice) of general practitioner. Second, the potential determinants were further selected using a bootstrap stepwise procedure to determine which factors were actually associated with the outcome of interest [***8***]. The Akaike Information Criterion (AIC) was used to

determine the model that provided the best account of the data. Odds ratios (ORs), 95% confidence intervals (95% CIs) and p-values were reported.

Results

The study cohort

From the initial number of 13,571 patients discharged from hospital with an incident diagnosis of MI, 9,606 patients were enrolled in the cohort. Approximately 58% of patients underwent PCI during the index hospitalization. The prevalence of PCI decreased as age increased: 67% in the age group "35 - 54", 65% in the group 55 - 69, 54% in the group 70 - 84 and 28% in the age group "≥ 85" years. About 10% of patients were not discharged from cardiology wards. A total of 6,532 MI patients (68%) were men. The mean age was 64 ± 12 years for men and 72 ± 12 years for women. The impact of comorbid health conditions increased with age. Hypertension (21%), arrhythmia (16%), vascular diseases (13%), and heart failure (10%) were the most common comorbidities. Overall, more than 50% of MI patients had at least one concomitant disease. The median follow-up time was 730 days.

Adherence to medication

Antiplatelets were characterized by the highest adherence (77%), followed by statins (73%), ACEI/ARBs (67%), and β-blockers (53%). It is worth noting that, for both males and females, the adherence to each of the recommended medications markedly decreased moving from the age group "70 - 84" to the group "≥ 85" years. Overall, in the Lazio region, about 63% of MI patients were adherent to chronic poly-therapy, defined as a MPR ≥ 0.75 for at least three of the four evidence-based drugs. If we consider the full combination therapy (i.e. MPR ≥ 0.75 for all of the four evidence-based drugs), the percentage drops to 28%. The cross-classified logistic multilevel model showed that the probability of adherence to chronic poly-therapy (at least three out of four drugs) was strongly influenced by the patient and general practitioner characteristics. With regard to patient characteristics, female gender was associated with a lower probability of adherence. The effect of age was not linear: with respect to the reference category (age less than 55 years) the probability of adherence significantly increased in the age class [55-69] years (OR=1.15, p=0.031) and substantially decreased in the older age group (age greater than or equal to 85 years; OR=0.42, p<0.001). The effect of the educational level was not significant. Moreover, adherence was significantly higher for patients discharged from cardiology wards, for patients with a length of stay longer than 7 days (the median value), for patients who underwent PCI during the index hospitalization and for patients who were prescribed with β-blockers, ACEI/ARBs or statins in the 12 months before admission. On the contrary, all comorbidities were associated with lower adherence to chronic poly-therapy. As regards the general practitioner characteristics, adherence was higher for patients with younger physicians and for patients with general practitioners working in group practice, i.e. sharing facilities, electronic patient records, administrative and clinical staff.

The geographic variation

The "hierarchical" healthcare system was composed as follows: 2,156 general practitioners, 55 local health districts, and 93 cross-classified hospitals of discharge. A high geographic variation was observed between the local health districts of the region. The percentages of adherence to poly-therapy ranged from 49% to 74%. In ***table 1*** the proportions of variation attributable to the hospitals of discharge and to the primary care providers were measured and compared. When analyzing the variation among primary care providers, after controlling for patients' characteristics, a relevant variation between local health districts was detected (MOR=1.24, p<0.001). However, when introducing the hospital of discharge as a cross-classified level, the variation between local health districts decreased (MOR=1.13, p=0.020). When introducing the hospital level, the variation between local health districts can be seen as the variability between districts as if all patients were discharged from the same hospital. Therefore, a portion of the variability in primary care is attributable to the hospital that has discharged the patient. Moreover, the variability in patient adherence

attributable to the hospital of discharge was statistically significant (p<0.001) and substantially higher. In fact, the MOR associated to the hospital of discharge was 1.37 whereas the MOR attributable to the local health district was 1.13.

Table 1. The trade-off between hospitals of discharge and primary care providers.

	Median Odds Ratio * (*p-value*)	
Levels of health care system	Logistic multilevel model including the "***local health district***" level	Cross-classified model including both the "***local health district***" and the "***hospital of discharge***" levels
Local health district	1.24 p < 0.001	1.13 p = 0.020
Hospital of discharge	-	1.37 p < 0.001
AIC	*11,500.38*	*11,431.07*

* The analyses were performed controlling for patients' characteristics.
AIC: Akaike Information Criterion.

Conclusions

In a study of 9,606 patients, we found that after a hospital discharge for MI, only 63% of patients were adherent to poly-therapy in the following two years. Treatments with proven benefit for MI are underused despite strong evidence that their use will result in better patient outcomes. This result is even more alarming if we consider that our definition of adherence was not restrictive. In fact, adherence to poly-therapy was defined as a MPR ≥ 0.75 for at least three of the four evidence-based drugs. Our findings are consistent with the results of other investigations, which reported unsatisfactory prescribing rates of secondary prevention drugs after MI in different time-periods and in different countries [***7***]. Among patient determinants of adherence to poly-therapy, we found that older age (greater than or equal to 85 years) and comorbidities played an important, negative role. A hypothesis for this finding may be related to the cumbersomeness of therapy, which increases with age and number of comorbidities. The longer and more complex is the list of drugs prescribed the lower is the adherence of the patient. The impact of the type of discharge ward was very impressive: patients discharged from cardiology wards were much more likely to be adherent to evidence-based medications. As regards general practitioner determinants, adherence to poly-therapy was higher with younger general practitioners and with physicians working in teams, sharing facilities, electronic patient-records, and clinical staff.

The geographic variation

A relevant geographic variation in adherence to guidelines was observed between the local health districts of the Lazio region. The "spatial" heterogeneity raises equity concerns in access to optimal care. This kind of unwarranted and avoidable variation in healthcare delivery is not unique to the Italian context [***3***]. This study focus on the "trade-off" between hospital and primary care in determining variation. The median odds ratios estimated by the cross-classified multilevel models are very interesting. They allow to measure and compare the amount of variation attributable to the "discharge phase" and to the following "primary care phase". The reduction of the variability among local health districts after entering the hospital level in the model proved that the differences we observe in primary care partially "reproduce" the clinical and organizational approach of the hospital of discharge, whose aims are both the correct setting of drug therapy, and the planning of the subsequent visits for patient monitoring. Really, adherence to chronic poly-therapy in people with previous MI was influenced more by the hospital that discharged the patient (MOR = 1.37) than by the primary care providers (MOR = 1.13).

The potential "plans of action"

According to the study results, it is possible to formulate hypotheses about the potential "plans of action" for health policies aimed at improving adherence to poly-therapy, such as, 1) to organize prescribing upskilling sessions for general practitioners, focusing on the most recent clinical guidelines; 2) to promote education on doctor-patient relationships, underlining the effectiveness of systematic motivational support; 3) to stimulate association for primary care physicians, in order to improve the continuity of care; 4) to improve the organizational processes within the hospital, in order to discharge MI patients from specialist wards and plan the subsequent visits for patient monitoring.

Study strengths and limitations

The population-based design, large numbers and robustness of analytical procedures are the main strengths of this study. Moreover, cross-classified multilevel models proved to be a useful tool for identifying the priority lines of action to improve adherence and define areas for more targeted health-care interventions. However, there are some study limitations to be considered. First, the results come from a single region in Italy and may be not generalizable to other geographic areas. However, our findings are in line with results of other studies carried out in other regions of Italy [***9***]. Second, our pharmaceutical database does not contain information on the prescribed daily doses and adherence to drug treatment was estimated on the basis of the DDDs. Although this is a useful instrument for comparing the results from different studies, misclassification of drug utilization may have occurred. Third, although the relative efficiency of using claims databases for studies of adherence in large populations in a 'real-word' setting is highly advantageous, we are unable to determine if patients actually consumed the dispensed medication. Thus, medication adherence based on claims data may be overestimated and the study results should be interpreted with appropriate caution.

References

[**1**] National Institute for Clinical Excellence. Secondary prevention in primary and secondary care for patients following a myocardial infarction. NICE guidelines [CG172]. Published date: November 2013.

[**2**] Kirchmayer U, Di Martino M, Agabiti N, et al. Effect of evidence-based drug therapy on long-term outcomes in patients discharged after myocardial infarction: a nested case–control study in Italy. Pharmacoepidemiol Drug Saf 2013; 22(6):649-57.

[**3**] Brooks JM, Cook EA, Chapman CG, et al. Geographic variation in statin use for complex acute myocardial infarction patients: evidence of effective care? Med Care 2014;52 Suppl 3:S37-44.

[**4**] Jenks GF, The Data Model Concept in Statistical Mapping. International Yearbook of Cartography 1967;7:186-90.

[**5**] Hox J, Multilevel Analysis: techniques and Applications. Mahwah, NJ: Lawrence Erlbaum Associates, 2002. Page 123.

[**6**] Larsen K, Merlo J. Appropriate Assessment of Neighborhood Effects on Individual Health: Integrating Random and Fixed Effects in Multilevel Logistic Regression. American Journal of Epidemiology 2005;161:81-8.

[**7**] Ho PM, Bryson CL. Rumsfeld J.S., Medication adherence: its importance in cardiovascular outcomes. Circulation 2009;119(23):3028-35.

[**8**] Austin PC, Tu JV. Automated variable selection methods for logistic regression produced unstable models for predicting acute myocardial infarction mortality. J Clin Epidemiol 2004;57(11):1138-.

[**9**] Filippi A, D'Ambrosio G, Giustini SE, et al. Pharmacological treatment after acute myocardial infarction from 2001 to 2006: a survey in Italian primary care. J Cardiovasc Med 2009;10:714-8.

A DISCRETE CHOICE EXPERIMENT TO UNDERSTAND PREFERENCES OF COLORECTAL CANCER PATIENTS TOWARDS ORGANIZATIONAL CHARACTERISTICS OF THE CHEMOTHERAPY OUTPATIENT SERVICE. PRELIMINARY RESULTS

Rosato Rosalba[1,3], Di Cuonzo Daniela[3], Fanchini Laura[2], Zanini Marcello[2], Ritorto Giuliana[2], Racca Patrizia[2], Pagano Eva[2]

1. *Università di Torino, Dipartimento di Psicologia*
2. *AOU Città della Salute e della Scienza di Torino, Colorectal Cancer Unit – Oncologia*
3. *AOU Città della Salute e della Scienza di Torino, Epidemiologia Clinica e Valutativa – CPO Piemonte*

Introduction

Although the outcomes of treatment are very important to patients, recent years have seen a growing interest in patient satisfaction within oncology. Healthcare providers and regulators have declared patient satisfaction a measure of healthcare quality [***1, 2***] and as such an important patient-focused outcome to measure and improve. Dissatisfaction with care can compromise compliance with treatment beneficial recommendations, and thereby affect therapeutic effectiveness. Various methodologies are in use to evaluate patient satisfaction in oncology settings, including interviews, focus groups, panel studies and surveys. However, direct questions to evaluate the service received for a serious illness, using closed format questions, are likely to be biased as the responders may unconsciously be induced to gratify their doctors. A different approach for studying the factors that make the greatest contribution to patient satisfaction is based on estimating individual preferences through consolidated economic data analysis techniques (standard gamble, time trade-off or discrete choice methods) [***3***]. In 2000 Piedmont Region established the Regional Network for Cancer Care (PCCN) with the aim of guarantee homogeneous and uniform care all over the region and to support patients in every phase of the disease, with a patient-centered approach. For this purpose, two organizational figures were implemented: the “Service and Reception Centre”, for the first reception of the patient, and the "Interdisciplinary Care Group", where different specialists discuss the case and plan together the therapeutic approach. How much this RCCN organization meets the patients preferences has not yet been assessed.

Using a discrete choice experiment (DCE) [***4***], the aims of the present study were to: 1) estimate the patients’ preferences of cancer care service attributes in a sample of colorectal cancer patients receiving chemotherapy; 2) explore preference heterogeneity between patients subgroups; and 3) assess the impact on preferences of the satisfaction for the service previously received.

Methods

The study was conducted at the “Città della Salute e della Scienza” hospital of Turin, interviewing colorectal cancer patients treated with chemotherapy in charge at the “Service and Reception Centre” of the Oncology Unit, between April 2015 and October 2016.

The protocol was approved by the hospital Ethical Committee and written informed consent was obtained from participants. Before performing the DCE experiment, respondents filled out a questionnaire on the satisfaction for the service received during their previous treatment (EORTC-INPATSAT32) and a questionnaire on quality of life (EORTC QLQ-C30). Sociodemographic and clinical information were also collected. Cancer service attributes were identified using results of a previous focus groups involving two separate groups, one with patients and caregivers and one with professionals (physicians and nurses). Five key attributes were identified: continuity of care, interpersonal skills, information, treatment choice, and time for therapy. These attributes were scaled on two/three levels as reported in the ***Table 1***. All three-level and

two level attributes gave the full factorial design with 108 possible combinations of levels. To provide a manageable task for responders the D-optimality criteria was used to maximize the efficiency of the design. Eighteen choice sets with two alternatives (i.e. hypothetical cancer service) were constructed. To make the questionnaire more manageable, two sets of 9 choices were randomly blocked. For each scenario respondents were asked which hypothetical oncology service was, in their opinion, preferable. The scenarios were randomized to prevent ordering effects bias. One additional control scenario was used to assess whether respondents choice the alternative with the best (favorite) level of all attributes. The additional scenario was excluded from the main analysis. For each block we need to survey 50 patients. Choice data were analyzed by using a conditional logit regression, allowing for clustering due to multiple observations from individuals. Service attribute levels were put in the analysis as dummy variables. Utility weight will be derived for the final model. Further analyses will be directed to assess statistical interaction between attributes and demographic or clinical characteristics. All statistical analysis were performed using SAS 9.4 and Limdep [***5***].

Results

Fifty-eight respondents were enrolled at May 2016 (mean age=61.14 yrs, sd=10.5, 50.9% male), 70.2% had colon and 29.8% a rectum cancer. Patients in early-stage (I or II) were n=12 (21.8%). One patient was excluded from analysis because failed the additional control scenario. All attributed had the expected sign. “Time for therapy” was the only attribute with coefficients not significantly different from 0 for all the levels. “Continuity of care” and “Interpersonal skills” had significant coefficients as well as the “Information” and “Treatment choice” more extreme attributes level (see ***Table 1***), indicating an influence in decision making. “Interpersonal skills” was the most important attribute, followed by “Information” and “Continuity of care”. Random effect models will be applied on the final sample to provide more accurate estimates of coefficients and to assess the impact on preferences of the satisfaction for the service previously received.

Table 1. Discrete choice experiment attributes: descriptions, levels and coefficients.

Attributes	Description	Levels	β	SE	p-value
Continuity of care	The patient has a trusted doctor	Yes	0.46	0.12	0.0002
		No	REF		
Interpersonal skills	Individual ability of the personnel for caring the patient	High ability	1.21	0.13	<0.0001
		Low ability	REF		
Information	Provision of info on the disease, prognosis and treatments by the medical doc tors	Detailed and complete	1.14	0.16	<0.0001
		Generic	0.19	0.17	0.2658
		Not complete	REF		
Treatment choice	Who makes the final choice of the patient’s treatment	The physician and the patient together	0.84	0.16	<0.0001
		The patient after consulting the physician	0.11	0.18	0.5281
		The physician alone	REF		
Time for therapy	Waiting time for starting therap y	Less than 4 hours	0.29	0.17	0.0908
		Around 4-5 hours	0.13	0.17	0.427
		More than 5 hours	REF		

Conclusions

Most attributes examined in this experiment had an influence on patient preference. Results suggest that "Interpersonal medical skills" was the most important attribute, followed by "Information" and "Continuity of care". Understanding which attributes of cancer care influence patient preference could potentially improve outcomes and treatment adherence in patients with cancer. Moreover, study results are of great interest in re-organizing cancer care at regional level in order to better catch patients preferences.

References

[1] Donabedian A. The quality of care: how can it be assessed?. Jama 1988;260:1743-48.
[2] Drain M. Quality improvement in primary care and the importance of patient perceptions." The Journal of ambulatory care management 2001;24:30-46.
[3] Ryan M, Karen G, Mabel AA, eds. Using discrete choice experiments to value health and health care. Vol. 11. Springer Science & Business Media, 2007.
[4] Hensher DA, John MR, William HG. Applied choice analysis: a primer. Cambridge University Press, 2005.
[5] Greene WH. LIMDEP version 6.0: user's manual and reference guide. No. 005.14 GRE 1992. CIMMYT. 1992.

REGRESSION MODELS IN THE PRESENCE OF SEMI-COMPETING RISKS

Orenti Annalisa[1], Biganzoli Elia Mario[1, 2], Ambrogi Federico[1], Boracchi Patrizia[1]

1. *Laboratory of Medical Statistics, Epidemiology and Biometry G. A. Maccacaro, Department of Clinical Sciences and Community Health, University of Milan, Italy*
2. *Unit of Medical Statistics, Biometry and Bioinformatics, IRCCS Istituto Nazionale dei Tumori, Milan, Italy*

Introduction

Medical research frequently yields multiple event times, thus each patient is exposed to the risk of different types of events during the follow-up. The observation and recording of all failure times from the start till the end of the study or death ("event history") is very important to characterize the progression of the disease and the association among different events.

The occurrence of different events have a specific impact on the progression of the disease and a specific meaning on the efficacy of the therapeutic strategy. Depending on the research concern and the study aim, different endpoints can be of interest, for example the first event of any type (interpreted as a treatment failure) or a subset of events, specifically related to therapeutic strategy. The latter are often intermediate non terminal events (death is not included), which cannot always be observed in real world, since death can occur before them. This is a typical situation of "semi-competing risks", as the occurrence of terminal event (i.e. death) can prevent the subsequent observation of non terminal event, but not vice versa [***1***].

An example of this situation in oncologic field, is relapse free survival, where the interest is focused on oncologic events (local and distant relapses, metastases, other primary tumours) and death without any previously recorded oncologic event is the terminal "competing" event. It is worth of note that estimating the probability of surviving free of relapse (relapse free survival) is possible only in a situation where relapses always occur for each patient before death. The occurrence of death before relapse makes the estimate of relapse free survival difficult to obtain in the case of non severe disease with long follow-up and in this case relapse free survival can be interpreted as the survival probability in an hypothetical situation were relapses can occur before death (net survival). Methods based on censoring times to terminal event can be applied only if times to occurrence of terminal and non terminal events are independent. Otherwise specific statistical methods assuming specific bivariate distributions taking into account the structure of association between times must be considered. With semi-competing risks data, the association parameter is often of biological interest to understand the disease progression, because it measures the extent to which the occurrence of an intermediate non terminal event hastens the occurrence of the more severe terminal event. Problems in estimating the survival probability of the non terminal event arises because bivariate distribution of time to events is unknown and can be only partially estimated from available data. Several bivariate distribution for survival times have been proposed, but the most used in semi-competing risks context are Copulas, because of their flexibility, in particular Clayton copula is the one that has been adopted even in regression models [***2***]. Few proposal of regression model on semi-competing risks can be found in medical statistical literature, most recent are the one of Peng and Fine [***3***] and the one of Hsieh and Huang [***4***]. These models based on hazard of net survival are useful, but software ad hoc is needed which has not been implemented in standard statistical software for survival analysis and this limits the application of such methodology for the analysis of clinical data.

This work aims to present an innovative regression model for net survival in semi competing risks which can be implemented by standard software based on pseudo-values. The advantage is that this is based on double steps: first to estimate the net survival under a defined copula and as second, using this net survival estimates to obtain pseudo-values as response for regression models.

Being a "new method" his performance is compared to that of other method already presented in the literature using a Clayton Copula.

Methods

The distribution of survival function of a non terminal event (X) is non-parametrically identifiable only in the context where it always precedes terminal event (Y), otherwise it is necessary to identify the joint distribution of times to multiple events. In semi competing risks setting the Clayton Copula is usually expressed [***2***]:

$$S(x,y) = P(X > x; Y > y) = \left[S_X(x)^{1-\theta} + S_Y(y)^{1-\theta} - 1\right]^{\frac{1}{1-\theta}} \quad 0 \le x \le y$$

where $S_X(x)$ and $S_Y(y)$ are the marginal survival functions of the non terminal and terminal event respectively and the parameter θ is a measure of the association between X and Y in the observable region (0≤x≤y). $S_y(y)$ is always estimable because the time to terminal event is recorded even after the occurrence of a non terminal event. When a specific time point t is considered, x=t and y=t, thus S(x,y)=S(t,t)=$S_F(t)$, which is the survival function for events of any type. Based on the relationship among $S_F(t)$, $S_Y(t)$ and $S_X(t)$, the marginal survival function for non terminal event can be obtained as follows:

$$S_X(t) = \left[S_F(t)^{1-\theta} - S_Y(t)^{1-\theta} + 1\right]^{\frac{1}{1-\theta}},$$

Where $S_F(t)$ and $S_Y(t)$ can be estimated by Kaplan Meier method.
Specific approach have to be applied in the presence of semi-competing risks to estimate θ [***1, 5***].
Furthermore when the interest is to evaluate the effect of different therapeutic strategies or covariates on the occurrence of a non terminal event in a semi-competing risks setting, specific regression model have to be used [***3, 4***]. We propose here to adopt the methodology based on pseudo-observations.
Pseudo-observations [***6***] provide a common approach to various kinds of survival models by replacing the incompletely observed outcome by the following quantity:

$$\hat{p}_i(t) = n \cdot \hat{p}(t) - (n-1) \cdot \hat{p}(t)^{-i}$$

where $\hat{p}(t)$ is the survival estimator for the non-terminal event Sx(t) computed on the whole dataset and $\hat{p}(t)^{-i}$ is the survival estimator for the non-terminal event applied to the sample of size n−1 obtained by eliminating the i-th individual from the data set. In the absence of censoring, at each time t pseudo observations can assume only two values: 1 if the subject is still alive at the time t and 0 if the subject has died before t. In the presence of censoring, pseudo-observations can be lower than 0 or greater than 1.
After computing the pseudo-observations for every individual at predefined time points, a generalized linear model with a chosen link function is fitted by applying a generalized estimating equations approach. It enables estimating the effect of covariates Z on the occurrence of the non-terminal event. We adopted a complementary log-log link function, to compare our results to those available in statistical literature:

$$\log\left(-\log\left(p(t|Z)\right)\right) = \log\left(-\log\left(S_X(t|\mathbf{Z})\right)\right) = \boldsymbol{\beta}'\mathbf{Z}$$

A binomial error distribution cannot be adopted as censoring is present, thus a Gaussian error distribution is used, as suggested by [***6***].

Results

Firstly the same simulation proposed by Peng and Fine and by Hsieh and Huang was used to compare model results of our approach. The following Monte Carlo simulation scheme was adopted.
Consider the models: $\log(X_i/3) = -\beta_X Z_i + e_{Xi}$ and $\log(Y_i/3) = -\beta_Y Z_i + e_{Yi}$, for i = 1, ... , n, where Z_i is a normal random variable with mean 1 and variance 0.5 constrained in [0,2], $Pr(e_{Xi} > x)$ and $Pr(e_{Yi} > x)$ both follow e^{-e^x}. The dependence structure of (e_{Xi}, e_{Yi}) follows the Clayton model.

The parameter settings in this study include θ=1.5 (corresponding to a Kendall's τ=0.2) and β_x=1, β_y=0 β_y=0.2 ; and sample sizes n of 200 and 500. For β_y=0, the independently censoring time C_i is generated from U(1,10), in which the censoring percentages for X and Y are 27% and 23%, respectively. For β_y=0.2, the independent censoring time C_i is generated from U(0,1) if γ = 1 and from U(1,1.2) if γ = 0, where γ is from Bernulli (0.2). In this case, the censoring percentages of X and Y are 52% and 67%, respectively. All simulations are based on 1000 replicates.
For the models based on pseudo-observations, we fitted a generalized linear model with Gaussian error and complementary log-log link function. A proportional hazard regression models on pseudo-observations is considered, which includes the treatment covariate and time point used to generate pseudo values by dummy variables.
Table 1 presents the mean, bias, empirical standard deviation (EmpSD) of the covariate effect on the non terminal event estimated by Hsieh's method and pseudo-observations method respectively. Hsieh regression model gives almost unbiased estimates of the covariate coefficient, obtained at different time points previously determined and small empirical standard deviation. Model based on pseudo-observation has similar values, even though little higher biases and empirical standard deviations.
We analyze data regarding a trial involving women with small, non-metastatic primary breast cancer and comparing quadrantectomy, axillary dissection and radiotherapy (QUART, 299 women) and quadrantectomy, axillary dissection without radiotherapy (QUAD, 280 women) [**7**].
In our analysis we will focus attention on relapse free survival (local relapse, regional or distant metastases, contralateral breast carcinoma and other primaries). This is a typical setting of semi-competing risks where a terminal event (death without previous relapse) can censor a non terminal event (relapse) but not vice-versa and this censoring cannot be considered independent, as the association between relapse and death cannot be excluded.
The following analysis will concentrate on evaluating the impact of treatment (QUART vs QUAD) on relapse free survival.
To estimate the association between times to relapse and times to death Fine's method is adopted. The estimated concordance parameter of the Clayton's copula is 7.35, corresponding to a Kendall's tau of 0.76. It can be interpreted as a predictive hazard ratio: the hazard of death for women who had experienced a neoplastic event is 7.35 times bigger that the hazard of death for women who had not experienced a neoplastic event.
In order to evaluate the effect of treatment (QUART vs QUAD), on relapse free survival, a complete regression analysis based on pseudo-values can be carried out. We considered 2-years intervals for time and we computed pseudo-values for each interval. The association parameter used to compute pseudo-values is that of the whole population θ=7.35. A generalized linear model with complementary log log link function and Gaussian error distribution is adopted. In order to compare the estimates of treatment effect obtained by pseudo-value method and the estimates obtained by Hsieh & Huang method on the same time intervals, a time varying coefficient for treatment is considered. The results obtained by pseudo-values method and Hsieh and Huang method are similar, identifying a significant smaller hazard of relapse for women treated with QUART. The pseudo values method, considering QUAD as reference category, gives a minimum net hazard ratio of 0.432 (95% CI: 0.146 - 1.279) at 2 year and a maximum value of 0.683 (95% CI: 0.496 - 0.940) at 12 years. Hsieh and Wang method gives a minimum net hazard ratio of 0.446 (95% CI: 0.344 - 0.579) at 2 year and a maximum value of 0.610 (95% CI: 0.473 - 0.788) at 10 years.

Table 1. Simulation results for the covariate effect on the occurrence of non terminal event estimated with pseudo-observations and Hsieh and Wang regression models.

n=200, β_x=1		β_y=0						β_y=0.2				
		mean	bias	EmpSD	ModSD	CP		mean	bias	EmpSD	ModSD	CP
Pseudo values method	*0.22*	1.029	0.029	0.388	0.364	0.956	*0.2*	1.034	0.034	0.422	0.393	0.951
	0.47	1.022	0.022	0.273	0.268	0.952	*0.4*	1.026	0.026	0.385	0.303	0.958
	0.72	1.027	0.027	0.241	0.246	0.969	*0.6*	1.028	0.028	0.282	0.275	0.958
	0.97	1.036	0.036	0.249	0.248	0.964	*0.8*	1.046	0.046	0.283	0.276	0.954
	1.22	1.045	0.045	0.270	0.263	0.966	*1.0*	1.059	0.059	0.308	0.290	0.951
Hsieh and Huang regression method	*0.22*	0.994	-0.006	0.317			*0.2*	1.006	0.006	0.358		
	0.47	0.979	-0.021	0.255			*0.4*	0.979	-0.021	0.284		
	0.72	0.983	-0.017	0.230			*0.6*	0.976	-0.024	0.265		
	0.97	0.984	-0.016	0.224			*0.8*	0.982	-0.018	0.261		
	1.22	0.955	-0.045	0.257			*1.0*	0.966	-0.034	0.268		
n=500, β_x=1		β_y=0						β_y=0.2				
		mean	bias	EmpSD	ModSD	CP		mean	bias	EmpSD	ModSD	CP
Pseudo values method	*0.22*	1.021	0.021	0.238	0.227	0.947	*0.2*	1.012	0.012	0.247	0.241	0.953
	0.47	1.018	0.018	0.168	0.168	0.960	*0.4*	1.010	0.010	0.187	0.186	0.958
	0.72	1.022	0.022	0.150	0.154	0.962	*0.6*	1.018	0.018	0.167	0.171	0.962
	0.97	1.022	0.022	0.146	0.154	0.969	*0.8*	1.019	0.019	0.164	0.170	0.959
	1.22	1.028	0.028	0.157	0.162	0.964	*1.0*	1.033	0.033	0.171	0.177	0.966
Hsieh and Huang regression method	*0.22*	0.998	-0.002	0.208			*0.2*	0.997	-0.003	0.227		
	0.47	0.985	-0.015	0.150			*0.4*	0.972	-0.028	0.175		
	0.72	0.968	-0.032	0.167			*0.6*	0.969	-0.031	0.177		
	0.97	0.898	-0.102	0.163			*0.8*	0.940	-0.060	0.159		
	1.22	0.945	-0.055	0.153			*1.0*	0.952	-0.048	0.153		

Conclusions

In biomedical studies, it is often of interest to evaluate the efficacy of treatment or the effect of specific covariates on the occurrence of intermediate non terminal events in a “semi-competing risks” setting, where the occurrence of the terminal event can prevent the observation of the non terminal event, but not vice versa. This type of censoring cannot be considered independent, since non terminal and terminal event are usually associated. Thus it is informative to assess the degree of association between the two events and then apply specific survival method to evaluate survival function of the non terminal event and corresponding covariates effects.

The difficulties in taking into account the presence of semi-competing risks and properly estimating the net event free survival function, rely in the lack of appropriate statistical software function. This is why competing risks methods or Kaplan Meier methods are often used to estimate net survival from intermediate events without correctly take into account the association between intermediate and fatal events. On the contrary we implemented ad hoc R functions in order to correctly compute net survival functions and corresponding survival regression methods in the presence of semi-competing risks.

As regards the simulation results it is worth noting that the method proposed here based on pseudo-observation has biases and empirical standard deviations slight higher than the ones obtained by the method based on conditional likelihood proposed by Hsieh and Huang. This suggest a future work to improve the estimating procedure. As regards the results obtained on the trial dataset, similar net hazard ratios are obtained by the two regression methods, with pseudo-observation regression presenting wider confidence intervals. Pseudo-values regression model has the advantages of enabling adjusting covariates effect by all

other covariate of interest and enabling flexibly modelling the effect of time on survival from severe events and the time dependent effect of covariates through interaction between time and covariates.
Further work is needed in order to accurately evaluate pseudo-values properties which guarantee a correct use of them in GEE model.

References

[**1**] Fine J P, Jiang H, Chappell R. On semi-competing risks data. Biometrika 2001;88(4):907-19.
[**2**] Clayton DG. A model for association in bivariate life tables and its application to epidemiological studies of familial tendency in chronic disease epidemiology. Biometrika 1978;65(1):141-51.
[**3**] Peng L, Fine JP. Regression Modeling of Semicompeting Risks Data. Biometrics, 2007 Mar; **63**(1): 96-108.
[**4**] Hsieh JJ, Huang YT. Regression analysis based on conditional likelihood approach under semi-competing risks data. Lifetime Data Analysis 2012;18(3):302-20.
[**5**] Lakhal L, Rivest LP, Abdous B. Estimating Survival and Association in a Semicompeting Risks Model. Biometrics 2008;64(1):180–8.
[**6**] Andersen PK. Pohar Perme M. Pseudo-observations in survival analysis. Statistical Methods in Medical Research 2010;19(1):71–99.
[**7**] Veronesi U, Marubini E, Mariani L, et al. Radiotherapy after breast-conserving surgery in small breast carcinoma: Long-term results of a randomized trial. Annals of Oncology 2001;12(7):997-1003.

DEALING WITH SYSTEMATICALLY MISSING CONFOUNDERS IN INDIVIDUAL PARTICIPANT DATA META-ANALYSIS – AN APPLICATION TO THE RELATIONSHIP BETWEEN SOCIOECONOMIC STATUS AND GASTRIC CANCER RISK IN THE STOMACH CANCER POOLING (STOP) PROJECT

Rota Matteo[1], Rumi Federica[1], Pelucchi Claudio[2], Lunet Nuno[3,4], Boccia Stefania[5,6,7], Negri Eva[1], La Vecchia Carlo[2] on behalf of the StoP Project study group.

1. *Department of Biomedical and Clinical Sciences L. Sacco, University of Milan, Milan, Italy.*
2. *Department of Clinical Sciences and Community Health, University of Milan, Milan, Italy.*
3. *ISPUP-EPIUnit, Universidade do Porto, Porto, Portugal.*
4. *Departamento de Epidemiologia Clínica, Medicina Preditiva e Saúde Pública, Faculdade de Medicina, Universidade do Porto, Porto, Portugal.*
5. *Section of Hygiene - Institute of Public Health; Università Cattolica del Sacro Cuore, Fondazione Policlinico Universitario "Agostino Gemelli", L.go F. Vito, 1 – 00168 Rome, Italy.*
6. *Icahn School of Medicine at Mount Sinai, 17 East 102 St, New York, NY, USA.*
7. *IRCCS San Raffaele Pisana, Via della Pisana, 235, 00163 Rome, Italy.*

Introduction

Despite incidence and mortality have been falling in most areas of the world over the last several decades, there are still about one million new diagnoses of gastric cancer (GC) per year worldwide, and GC remains the third leading cause of cancer mortality. The earlier reduction of the burden of GC in high income countries - largely explained by downwards trends in the prevalence of *Helicobacter pylori (H. pylori)* infection and by improvements in diet and food conservation – led to neglect this neoplasm in terms of research and development efforts. However, the study of aetiological factors remains a global priority for the prevention and control for GC.

The Stomach Cancer Pooling (StoP) Project is an international consortium of case-control (CC), including nested CC, studies on GC started in 2012 [***1***]. This consortium joins together scientists from several areas of the world aimed to study the role of lifestyle, dietary habits and genetic determinants in GC aetiology, through an individual participant data meta-analysis (IPD-MA).

IPD-MA is considered to be a gold standard approach in evidence synthesis as allows access to raw data from each participant, data checking, verification and centralized recoding of data according to common definitions, thus minimizing publication and reporting bias typical of aggregated data meta-analysis [***2***].

At the time of writing, 31 case-control studies agreed to participate to the StoP project, for a total of more than 14,000 GC cases and 33,700 healthy controls (StoP dataset release 2). This data represented a unique opportunity to better quantify the association between alcohol drinking and tobacco smoking on the risk of GC [***3, 4***]. Moreover, it is an opportunity for medical statisticians to explore and apply IPD-MA related methods, and their drawbacks [***5***]. Among them, a hot discussed topic in the literature is the choice of the analysis method, i.e. one-stage or two-stage, and the management of systematically missing confounding factors, i.e. variables not collected in one or more studies of the IPD-MA.

To give a brief overview of the one-stage and two-stage analyses methods, we applied some recently developed advanced techniques to address the problem of systematically missing variables in an IPD-MA. Among them, we dealt with i) a multivariate random effect model in a two-stage framework (hereafter called "Fully and Partially Adjusted Meta-Analysis" - FPAMA) [***6***] and ii) a one-step analysis based on Multiple Imputation (MI) by Chained Equations (MICE), hereafter called multilevel multiple imputation (MLMI) [***7***]. These methods have been applied within the StoP project to study the relationship between socioeconomic

status and GC risk while allowing for adjustment of *H. Pylori infection*, a variable that has not been collected in some case-control studies of the StoP consortium.

Methods

Statistical methods for IPD-MA should take into account the hierarchical structure of the data in which individuals are nested within studies. Two approaches have been described: a one-stage and a two-stage approach [***8***]. In the one-stage approach data are analysed in a single step through a random-effects model, while in the two-stage approach each study is analysed separately, and results combined through standard meta-analytic techniques. Several studies investigating differences between one-stage and two-stage IPD-MA either theoretically, empirically or via simulation concluded that "largely the two approaches produce similar results" [***5, 8***].

In the presence of one, or more, systematically missing confounders, researchers could choose to carry out a complete data analysis by including only studies without systematically missing variables, or a complete confounder analysis including all the studies but adjusting only for those confounders available in all the studies. Both these approaches are biased, the former because a data discharge leads to a loss of information and precision, and it is further biased if the systematically missing variables are not missing completely at random (MCAR), the latter because results remains potentially confounded. The widely used two-stage approach, although flexible as allows to use the entire IPD-MA dataset, also potentially suffers from residual confounding in studies with systematically missing confounders.

Let us consider an IPD-MA of i=1,…, N studies with j=1,….,N_i subjects in the i-th study. We denote the observed dichotomous outcome (Y) for subject j in study i as y_{ij} and, without loss of generality, we consider a scenario with two binary covariates X_1 and X_2, where X_1 is the exposure variable and X_2 a confounder that is systematically missing in M studies, with 0<M<N. We anticipate some degree of heterogeneity between the studies, and therefore consider hereafter random-effects models.

The FPAMA method [***6***] is a two-stage approach based on the borrowing of strength concept among studies that can provide both fully and partially adjusted estimates, and those who can give only partially adjusted estimates. This is not a MI based methods, and thus missing values of X_2 are not replaced through imputed values, but the relationship between the fully and partially adjusted estimates for the N-M complete studies is assumed to apply to the M studies where only partially adjusted estimates can be obtained.

In the first step, for N-M studies, i.e. those without the systematically missing confounder X_2, the following fully and partially adjusted models can be fitted:

$$y_j \sim \text{Bernoulli}(\pi_j)$$

$$\text{fully adjusted:} \quad \eta_j = \text{logit}(\pi_i) = \beta_0 + \beta_1^f x_{1j} + \beta_2 x_{2j} \quad \text{for} = 1,\ldots,N$$

$$\text{partially adjusted:} \quad \eta_j = \text{logit}(\pi_i) = \beta_0 + \beta_1^p x_{1j} \quad \text{for} = 1,\ldots,N.$$

For M studies where X_2 is systematically missing, only the partially adjusted model can be fitted to the data. In the second step, it can be assumed for each study the following bivariate model:

$$\begin{pmatrix} \hat{\beta}_1^f \\ \hat{\beta}_1^p \end{pmatrix} \sim N\left(\begin{pmatrix} \beta_1^f \\ \beta_1^p \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \right)$$

where σ_1^2, σ_2^2 and ρ are assumed to be known. In practice, variances of fully and partially adjusted regression coefficients β_1^f and β_1^p can be simply obtained using any statistical packages, while the correlation coefficient $\rho = \text{Corr}(\beta_1^f, \beta_1^p)$ for the N-M studies can be obtained by bootstrapping [6]. Under the missing at random (MAR) assumption and for computational convenience, the M studies where X_2 is systematically missing can be

included in the multivariate model by giving arbitrary values to regression coefficients, i.e. $\hat{\beta}_1^f=0$, with very large within-study variances, i.e. $\hat{\sigma}_1^2=10^6$, and null within-study correlations ρ [***9***].

Being heterogeneity between studies a matter of fact in IPD meta-analysis, the couple $\left(\beta_1^f,\beta_1^p\right)$ can be modeled as:

$$\begin{pmatrix}\beta_1^f\\ \beta_1^p\end{pmatrix} \sim N\left(\begin{pmatrix}\beta_1^f\\ \beta_1^p\end{pmatrix}, \begin{pmatrix}\tau_1^2 & \rho\sigma_1\sigma_2 k\tau_1\tau_2\\ k\tau_1\tau_2 & \sigma_2^2+\tau_2^2\end{pmatrix}\right),$$

leading to the marginal random-effects bivariate model:

$$\begin{pmatrix}\hat{\beta}_1^f\\ \hat{\beta}_1^p\end{pmatrix} \sim N\left(\begin{pmatrix}\beta_1^f\\ \beta_1^p\end{pmatrix}, \begin{pmatrix}\sigma_1^2+\tau_1^2 & \rho\sigma_1\sigma_2+k\tau_1\tau_2\\ \rho\sigma_1\sigma_2+k\tau_1\tau_2 & \tau_2^2\end{pmatrix}\right).$$

This model can be fitted either by ML or REML methods. Inference is done on the fully adjusted regression coefficient $\hat{\beta}_1^f$ but the model, unlike methods based on MI, does not provide the pooled estimates of the systematically missing confounder(s) X_2. We used R package "mvmeta" with maximum likelihood (ML) estimation method.

The MLMI approach by Jolani et al. [***7***] is based on a generalized linear mixed model (GLMM) for specifying the conditional distributions (fully conditional specification, FCS) by using the Wishart distribution to draw estimates of uncertainty around the variance component of the model. The rationale beyond MI is to create D copies of the IPD-MA dataset where missing values are replaced by random draws made from the posterior distribution of missing values giving the observed data, under the MAR assumption.

For N-M studies where X_2 is observed, the following GLMM model is fitted:

$$x_{2ij} \sim \text{Bernoulli}(\pi_{ij})$$
$$\xi_{ij}=\text{logit}(\pi_{ij})=\alpha_{0i}+\alpha_{1i}x_{1ij}+\alpha_{2i}y_{ij}$$

where

$$\alpha_{ki}=\alpha_k+a_{ki} \quad a_{ki} \sim N(0,\psi_k^2) \quad \text{for } k=0,1,2.$$

After obtaining estimates of $\boldsymbol{\alpha}$ (fixed component) and $\boldsymbol{\Psi}$ and $\mathbf{a_i}$ (random component), we get a random draw $\gamma_k^* \sim MVN(\hat{\boldsymbol{\alpha}}_k, var(\hat{\boldsymbol{\alpha}}_k|\hat{\boldsymbol{\Psi}}_k))$ and compute $\boldsymbol{\Lambda_k}=\sum_{i=1}^{N-M}\mathbf{a_{ki}}\mathbf{a_{ki}^T}$ for k=0,1,2.

For the M studies where X_2 is systematically missing, for D times draw $\boldsymbol{\Psi}_k^{*-1} \sim Wishart(\boldsymbol{\Lambda}_k^{-1}, N-M)$ - being the Wishart the posterior distribution of $\boldsymbol{\Lambda}$ - get $\mathbf{a}_{ki}^* \sim MVN(\mathbf{0},\boldsymbol{\Psi}_k^*)$ for k=0,1,2, and finally obtain the imputed values for X_2 as $x_{2ij}^*=\text{logit}^{-1}\left[(\gamma_0^*+a_0^*)+(\gamma_1^*+a_1^*)x_{1ij}+(\gamma_2^*+a_2^*)y_{ij}\right]$.

After imputation, the D copies of the IPD-MA dataset can be analyzed through the one-stage approach. This means that for each imputed dataset we get D fixed and random regression coefficients estimates. The final combined estimates of the fixed and random components can be obtained through the Rubin's rule [***10***].

We used the extension developed by Jolani et al. [***7***] to the "mice" R package to fit the model, adopting ML estimation through Laplacian approximation.

We used data from the first release of the StoP project dataset, including 23 case-control studies for a total of 10,290 GC cases and 26,153 controls [***1***]. Socioeconomic status was defined using educational level and categorized in each study in low, intermediate and high according to the International Standard Classification of Education (ISCED). Models included terms for age (<40, 40-44, 45-49, 50–54, 55–59, 60–64, 65–69, 70–74, ≥75 years), sex, tobacco smoking (never, former, current ≤10 cigarettes/day, >10 to ≤20 cigarettes/day, and >20 cigarettes/day), fruit and vegetable consumption (study-specific tertiles), study centre (for multicentric studies) and *H. pylori* infection. This latter variable was systematically missing in some studies

of the consortium. To overcome this limitation, we applied the FPAMA and MLMI approaches described above to estimate the association between socioeconomic status and GC risk while allowing for *H. Pylori* infection.

Results

We found a strong association between socioeconomic status and risk of GC. The two investigated methods, FPAMA and MLMI, gives similar results (***Figure 1***). Intermediate (MLMI method: odds ratio (OR) 0.76, 95% confidence interval(CI) 0.64-0.90, FPAMA method: OR 0.75, 95% CI 0.60-0.92) and higher socioeconomic status (MLMI method: OR 0.55, 95% CI 0.46-0.65, FPAMA method: OR 0.52, 95% CI 0.42-0.64) were associated with a reduced risk of GC. Standard errors for fixed effects were somewhat smaller for the MLMI methods as compared to FPAMA (0.089 vs 0.109 for intermediate socioeconomic status and 0.165 vs 0.2 for high socioeconomic status).

Conclusions

From a methodological point of view, both FPAMA and MLMI are feasible to account for systematically missing confounders, leading to very similar results. However, the MLMI method is more flexible than the FPAMA method as it can be used to deal with sporadically missing values and can also handle more complex missing patterns. The value of the methods here presented and discussed is greater when the dataset is smaller as compared to the StoP dataset.

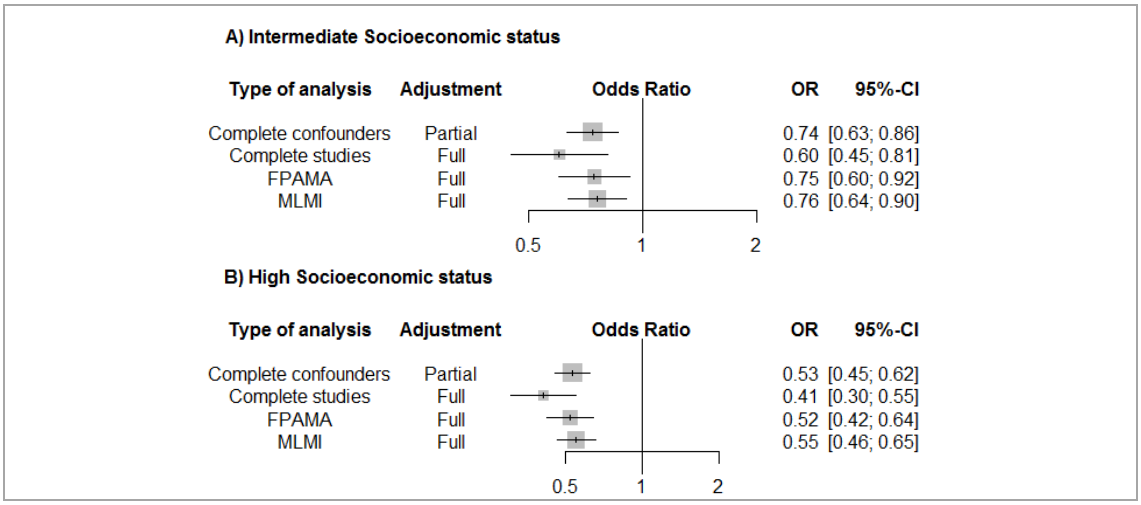

Figure 1. Odds ratio estimates for the association between socioeconomic status and GC risk in the StoP project consortium deriving from different approaches (FPAMA and MLMI) to deal with systematically missing confounders.

References

[**1**] Pelucchi C, Lunet N, Boccia S, et al. The Stomach Cancer Pooling (StoP) project: study design and presentation. Eur J Cancer Prev 2015;24(1):16-23.

[**2**] Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. BMJ 2010;340:c221.

[**3**] Praud D, Rota M, Pelucchi C, et al. Cigarette smoking and gastric cancer in the Stomach Cancer Pooling (StoP) Project. Eur J Cancer Prev 2016 Aug 24 [Epub ahead of print]. DOI: 10.1097/CEJ.0000000000000290

[**4**] Rota M, Pelucchi C, Bertuccop P, *et al*., Alcohol consumption and gastric cancer risk – A pooled analysis within the StoP project consortium. Int J Cancer 2017 [in press].

[**5**] Burke DI, Ensor J, Riley RD. Meta-analysis using individual participant data: one-stage and two-stage approaches, and why they may differ. Stat Med 2017;36(5):855-75.

[**6**] Fibrinogen Studies Collaboration, Jackson D, White I, et al. Systematically missing confounders in individual participant data meta-analysis of observational cohort studies. Stat Med 2009;28(8):1218-37.

[**7**] Jolani S, Debray TP, Koffijberg H, et al. Imputation of systematically missing predictors in an individual participant data meta-analysis: a generalized approach using MICE. Stat Med 2015;34(11):1841-63.

[**8**] Debray TP, Moons KG, Abo-Zaid GM, et al. Individual participant data meta-analysis for a binary outcome: one-stage or two-stage? PLoS One 2013;8(4):e60650.

[**9**] Riley RD, Price MJ, Jackson D, et al. Multivariate meta-analysis using individual participant data. Res Synth Methods 2015 ;6(2):157-74.

[**10**] Rubin DB. Multiple imputation for Nonresponse in Surveys. Wiley Online Library: New York, USA, 1987.

CHALLENGES IN COMBINING CLINICAL AND OMICS DATA

Ambrogi Federico, Boracchi Patrizia

Dipartimento di Scienze Cliniche e di Comunità, Università degli Studi di Milano

Introduction

In last years biomedical research has devoted much efforts in the generation of high-throughput data to enhance the knowledge and the prognosis evaluation of patients with major diseases. The field of omics (genomics, proteomics, lipidomics, etc.) is continuously growing getting more and more complex, while other examples of massive data production are coming for example from imaging. The next stage in the omics is that of integrating the measurements coming from genes to proteins in a system biology approach. Prognostic models may take advantage from the huge amount of information that is potentially available to characterize patients bioprofiles.

However many observational studies highlight how the added contribution of genomic data to clinical information may be difficult to be appreciated [***1***]. One of the key elements of the success of the combination is linked to the magnitude of the effect to be detected and, obviously to the sample size (in terms of number of events). In fact, frequently, the hazard ratios associated with genomic data are not comparable to those of clinical data. The new biomarkers (or scores combining multiple biomarkers) should in particular add information with respect to traditional/clinical prognostic factors. This topic has been investigated in an observational context, for example, by De Bin and colleagues [***2***]. In their study, based on the analysis of two datasets (one in breast cancer and the other on neuroblastoma) they found that it is very hard to add information to a clinical model, especially when the clinical information is really rich, such as in breast cancer. The consideration about the added value of the omic information with respect to the clinical information implies the adoption of a measure of prognostic value. De Bin and colleagues used the prediction error [***4***] and the C-Index [***5***] to perform such an evaluation. However the quantification of the added usefulness of a biomarker is a very broad area of research and many other possibilities do exist. In fact it is generally accepted that prediction error or the C-index are too rough measures to correctly highlight the added usefulness of a biomarker. One possible alternative measure is the integrated discriminant improvement [***6***]. However this index was criticized and its application should be carefully considered [***7***].

In this paper we want to investigate using a simple simulation approach the possible added prognostic impact of omic information in terms of prediction error and C-index. In particular the focus will be on the magnitude of the effects to be detected in order to appreciate a change in the considered measures of prognostic value.

Methods

The plan is to compare two advanced regression methods to combine clinical and omics information. When the number of covariates exceeds the number of subjects, standard regression tools cannot be used. The first problem is therefore to adopt some regression method able to deal with high dimensions. In this work we focus the attention on offset boosting [***8***] and on Lasso regression [***10***], both applied to the Cox regression model. Both methods are regularized regression techniques. In boosting a so-called weak learner is applied iteratively until a satisfying model is obtained (in terms of compromise between bias and variance). The optimal compromise is calculated using cross validation. The same principles apply with Lasso regression, which uses a L^1 penalty to shrink unimportant coefficients to 0 thus performing model selection. The amount of penalty (and therefore the number of non-0 coefficients) is determined using cross-validation. Both techniques allow for the exclusion of standard covariates from the selection process. This is one way to

perform the combination, which keeps the individual information from each single covariates and adopt a principled framework of analysis. Other methods can be considered to combine clinical and omic covariates. Some simple methods are described in [***1***] with a more advanced method proposed for classification problems, while other strategies are outlined in [***2***].

Prediciton Error

To evaluate the models under study, we estimated the prediction error according to the method of Gerds and Schumacher [***11***]. The method can be used to evaluate any regression model estimating a survival probability and is based on a weighted version of the Brier score. The IPCW weights are used for assuring the uniform consistency of the proposed estimator. The computation of the prediction error on the same data used for estimation provides however an optimistic estimate. This is especially true for complex modeling techniques requiring the tuning of ancillary parameters, for example the amount of penalty in penalized regression. To avoid such an optimistic estimate, cross validation or bootstrap cross-validation can be used. In bootstrap cross-validation, for each bootstrap sample, the left-out subjects are used to evaluate the prediction error. The estimators based on cross-validation can be too pessimistic as the model is trained on only a portion of the data actually available. To obviate to this problem the .632+ estimator was proposed.

The computation of all these estimators is available in the R package pec, [***12***], which handles many regression models (from the standard Cox regression model to random forest). The package was adapted to work with the function glmnet [***13***] used for estimating the Cox model with lasso penalization and with the function CoxBoost [***3***] for cox model boosting.

Simulation

The work for properly setting up the simulation framework is still ongoing. We report here some preliminary results for Lasso regression and Cox boosting. The simulation of survival times will follow the structure proposed by Bender et al. [***14***] to simulate from the Cox proportional hazard model. Two independent standard normal covariates were generated with log hazard ratio equal to 0.5 to mimic important clinical covariates. We then simulated 1000 covariates to mimic high-throughput omic experiments with only 8 of them with an impact on survival times. The 1000 ”omic” covariates were generated as independent and then with a specific correlation structure following the proposal of Bender et al. [***15***]. Specifically block-correlated standard normal covariates were generated: covariates 1-50 have 0.5 correlation; covariates 51-100 have correlation equal to 0.35 correlation; covariates 101-200 have 0.05 correlation and covariates 201-300 have 0.32 correlation. The remaining covariates are i.i.d.. Four pairs of covariates with an impact on survival time, were randomly chosen from each block of correlated variables (with log hazard ratios: b ; -b). The value of the parameter b was chosen to simulate different levels of associations: low 0.1; medium 0.2; high 0.3. In this setting it is assumed that the omic variables, even in the scenario with with ”high” impact, have a lower hazard ratio compared to that of the clinical covariates.

For each scenario a Cox regression model with all the relevant covariates (the 2 ”clinical” and the 8 ”omic” covariates) was fitted to be used as a benchmark in the simulation. A Cox model with Lasso penalty on all the available 1002 covariates was used as a practical method for variable selection. The penalty of the for Cox model was set up to only act on ”omic” covariates, while the two ”clinical” covariates were always selected. For each simulated data the models were fitted and the prediction error was calculated using bootstrap .632+ cross-validation. The prediction error was also computed for the marginal Kaplan-Meier estimator to give an idea of the added information of the included covariates. Results are expressed as the percent difference in prediction error with respect to the benchmark model. In this way it is possible to rank the different models.

Results

The obtained results for the three investigated scenarios in terms of prediction error are reported in ***figure 1***. Sample size was fixed at 500 times with 30% censoring. Considering an effect size of 0,3 for the 8 relevant "omic" covariates, as expected, the distance from the benchmark prediction error of the model with only the clinical covariates is, by far, greater than the difference obtained using Lasso or boosting variable selection (excluding the "clinical" covariates from the selection). However, when the impact of the "omic" covariates is equal to 0.2, the distance among the 3 models is reduced although the ranking is preserved. When the coefficient of the "omic" covariates is set to a small "0.1", the performance of the regression models is reversed: the smallest difference from the benchmark model is now for the model with only the clinical covariates. In this situation the regression strategies with variable selection adopted here, are not effective in selecting the covariates with impact on time to event and seems to actually select noise with a negative impact on generalization performance. It is to be remarked that in this last scenario, boosting is in any case more effective than Lasso. The false positive rate is about 3.5% for both but the true positive rate is 28% for Lasso and 32& for boosting.

The simulation with 200 times shows that the performances of the regression strategies with selection (Lasso and boosting) are outperformed by the model with only the clinical covariates when the effect size is 0.2. When the sample size is 200, the Lasso selection is outperformed by boosting which shows a similar performance to the model with only the clinical covariates when the impact of the "omic" covariates is 0.2. When the coefficients of the omic covariates is 0.1 the model with only the clinical covariates has almost the same prediction error as the true model: in this case the models which tries to select the omic covariates cannot be actually effective.

Conclusions

The first studies of prognostic models with genomic information tried to use the new biomarkers to replace clinical data. Starting from that first attempts, a more realistic goal appears to be that of adding information to clinical data. However combining clinical data and omic information is challenging from many point of view. One obvious difficulty is the large number of covariates from high throughput experiments used to describe patients bioprofiles. Appropriate regression tools or specific strategies for model selection are then to be used. Another open issue is how to combine the information coming from different sources. In this work we used penalised regression techniques able to incorporate without penalisation some of the covariates included in the regression model.

One very important aspect about the combination, regards the magnitude of the effects under study. Many authors highlight how the effects from biomarkers from high-throughput experiments are frequently of lower magnitude than the effect for clinical prognostic factors. Moreover such studies are characterised by small sample sizes, at list with respect to the number of information to deal with. In this perspective the problem of small effects is more daunting. In fact, as outlined in [***9***], Type S and type M errors could be quite important: the Authors call type M error when it happens that a statistically significant finding results from a huge overestimate of the true effect, while type S error refers to an estimate in the opposite direction of the true effect.

In this preliminary applications, the prediction error was used to quantify the usefulness of the regression models. Another possibility would be to use the integrated discrimination improvement (IDI). However the IDI was criticised, so the debate about how to measure the added usefulness of biomarkers is still open as more sensitive measures are actually needed to judge the added contribution of biomarkers with respect to well known clinical covariates. Further simulations are needed to evaluate predictive genes, that is including genes by treatment interactions.

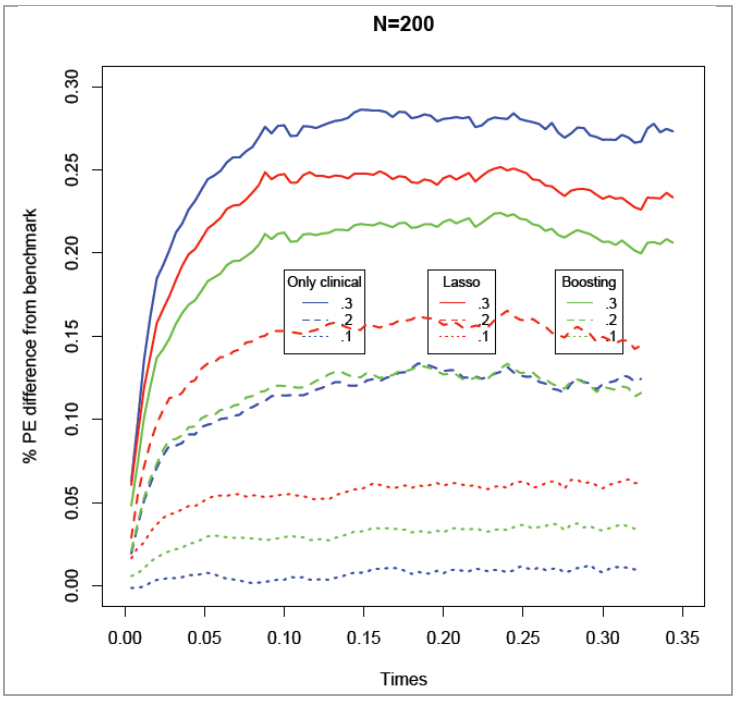

Figure 1. Percent prediction error difference from the benchmark Cox regression are reported for each regression model considered: (1) The Cox model with only the two ”clinical” covariates; (2) the Lasso Cox regression with L^1 penalty; (3) The Cox boosted regression. The benchmark Cox regression is the one including the 2 ”clinical” covariates and the 8 ”omic” covariates. Prediction Error curves were calculated using bootstrap cross validation 0.632 +. The scenario

References

[1] Boulesteix AL, Porzelius C, Daumer M.. Microarray-based classification and clinical predictors: On combined classi-fiers and additional predictive value. Bioinformatics 2008;24:1698-706.

[2] De Bin R, Sauerbrei W, Boulesteix AL. Investigating the prediction ability of survival models based on both clinical and omics data: two case studies. Stat Med 2014;30:5310-29.

[3] Harald Binder. CoxBoost: Cox models by likelihood based boosting for a single survival endpoint or competing risks. R package version 1.4 2013. https://CRAN.R-project.org/package=CoxBoost

[4] Schumacher M, Binder H, Gerds T. Assessment of survival prediction models based on microarray data. Bioinfor-matics 2007; 14:1768-74.

[5] Gerds TA, Kattan MW, Schumacher M, Yu C. Estimating a time-dependent concordance index for survival prediction models with covariate dependent censoring. Stat Med 2013;13:2173-84.

[6] Pencina MJ, D'Agostino RB, D'Agostino RB, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med 2008;2:157-72; discussion 207-12.

[7] Hilden J, Gerds TA. A note on the evaluation of novel biomarkers: do not rely on integrated discrimination improvement and net reclassification index. Statist. Med 2014;33:3405-14.

[8] Binder H, Schumacher M. Allowing for mandatory covariates in boosting estimation of sparse high-dimensional survival models. BMC Bioinformatics 2008;9:14.

[**9**] Gelman A, Lee D, Park D, Shor B, Su Y, Weakliem D. Of beauty, sex, and power: Statistical challenges in estimating small effects 2007

[**10**] Tibshirani R. The lasso method for variable selection in the Cox model. Stat Med 1997;4:385-95.

[**11**] Gerds TA, Schumacher M. Efron-type measures of prediction error for survival analysis. Biometrics 2007;4:1283-7.

[**12**] Mogensen UB, Ishwaran H, Gerds TA. Evaluating Random Forests for Survival Analysis Using Prediction Error Curves. Journal of Statistical Software 2012;50(11):1-23.

[**13**] Simon N, Friedman J, Hastie T, Tibshirani R. Regularization Paths for Cox's Proportional Hazards Model via Coordinate Descent. Journal of Statistical Software 2011;39(5):1-13. URL http://www.jstatsoft.org/v39/i05/.

[**14**] Bender, R, Augustin, T, Blettner, M. Generating survival times to simulate Cox proportional hazards models. Stat Med 2005;11:1713-23.

[**15**] Binder H, Schumacher M. Adapting prediction error estimates for biased complexity selection in high-dimensional bootstrap samples. Stat Appl Genet Mol Biol 2008;7(1):Article12

[**16**] Freidlin B, Simon R. Adaptive signature design: an adaptive clinical trial design for generating and prospectively testing a gene expression signature for sensitive patients. Clin. Cancer Res 2005;21:7872-8.

[**17**] Janes H, Brown MD, Pepe MS. Designing a study to evaluate the benefit of a biomarker for selecting patient treatment. Stat Med 2015;27:3503-15.

NUMERICAL METHODS FOR PROCESSING LONGITUDINAL DATA: A CLINICAL APPLICATION

Migliaretti Giuseppe[1], Perracchione Emma[2], Stura Ilaria[1], Ditaranto Serena[1], Cavallo Franco[1]

1 *Dipartimento di Scienze della Sanità Pubblica e Pediatriche, Università di Torino*
2. *Dipartimento di Matematica T. Levi-Civita, Padova, Italy*

Introduction

Processing of longitudinal data is of interest in several fields: for instance in medicine, where repeated measurements are used to monitor the patients' health status and to adjust therapies accordingly. These data are commonly analysed using parametric models such as generalized estimation equations or generalized linear mixed effect models. However, they are often irregularly-spaced, i.e. scattered, since sampling in clinical practice depends on the specific clinical needs. Moreover, their accuracy critically depends on the available clinical equipment. Thus, a more flexible and robust approach to longitudinal data analysis is needed.
The aim of this study is to assess alternative statistical methods to be used as robust and accurate tools for the prediction of individual trajectories from sparse longitudinal data.

Methods

Functional Data Analysis (FDA) [***1***] and mixed models that use Radial Basis Function in combination with optimization methods, such as Particle Swarm Optimization [***2***] and Cuckoo algorithms [***3***], are here presented. As motivating example, the height measurements of 141 male paediatric patients affected by Growth Hormone Deficiency (GHD) are considered and discussed.

Results

FDA showed particularly suitable to describe longitudinal data, while the mixed methods allowed good predictions on the final height reached by the patients starting from few initial height measurements. Moreover, valuable aside information can be extrapolated by the outputs of the methods (see ***Figure 1***). Actually FDA provides both mean and point-by-point growth velocities, which are useful for evaluating the treatment effectiveness. Conversely, the mixed methods provide the expected growth curve to be compared with real follow-up data.

Conclusions

FDA estimates the time evolution of some parameters, such as growth velocity and acceleration, which can be correlated with clinical treatment effectiveness.
The proposed mixed method produced useful personalized output. For instance, starting from the first 1-2 height measurements of a single child, the models provide a (continuous) growth curve (see Figure 1) from the beginning of the therapy to the adulthood (18 years old), allowing the ' ad hoc' estimation of the final height for each single patient.

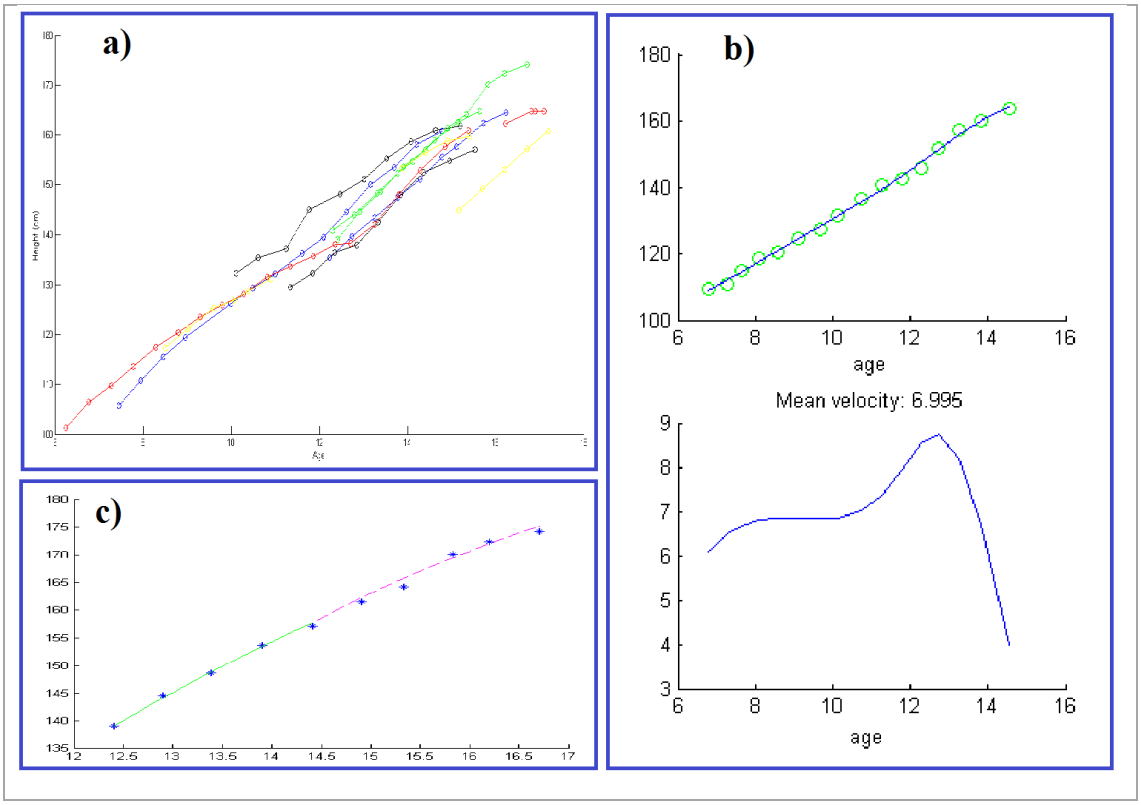

Figure 1. a) Some examples of longitudinal data; **b**) the output of FDA method; **c**) the output of the mixed model

References

[1] Ramsay JO, Silverman BW. Applied Functional Data Analysis. New York 2002: Springer-Verlag.

[2] Perracchione E, Stura I. RBF kernel method and its applications to clinical data, Dolomites Res. Notes Approx 2016;9:13-8.

[3] Yang X, Deb S. Cuckoo search via lévy flights. In 2009 World Congress on Nature and Biologically Inspired Computing, NABIC 2009 - Proceedings.

THE USE OF TIME DEPENDENT COX MODEL AND JOINT MODEL TO EVALUATE THE RELEVANCE OF PTX3 AS AN EARLY MARKER OF ACUTE GRAFT-VERSUS-HOST DISEASE.

Rebora Paola, Arisido Maeregu Woldeyes, Andreano Anita, De Lorenzo Paola, Valsecchi Maria Grazia

Center of Biostatistics for Clinical Epidemiology, School of Medicine and Surgery, University of Milano-Bicocca, Monza, Italy

Introduction

Acute graft-versus-host disease (GVHD) is one of the major causes of morbidity and mortality associated with allogeneic stem cell transplants (HSCT), resulting from immune-mediated attack of recipient tissues by donor T cells contained in the transplant [***1***]. Currently, the diagnosis of GVHD is based on clinical signs and symptoms and requires invasive biopsies of disease target organs in uncertain cases, which are sometimes unfeasible. To improve diagnosis and prognosis of GVHD recent research focus on specific biomarkers measured in the plasma or serum of HSCT patients as a new tool for detecting GVHD prior to clinical manifestation and for GVHD management. The long pentraxin 3 (PTX3) is a component of the humoral arm of innate immune system with a recognized non-redundant role in the modulation of inflammation [***3***]. PTX3 is an acute-phase protein, rapidly produced by vascular endothelial cells, mesenchymal cells and fibroblasts, as well as by innate immune response cells upon stimulation with pro-inflammatory cytokines, damaged tissue-derived signals and microbial antigens. Differently from other acute phase proteins, such as the C-Reactive Protein, PTX3 is considered a rapid marker for primary local activation of innate immunity and inflammation due to its peculiar pattern of production.
The aim of this study is to evaluate whether repeated PTX3 plasma dosage after HSCT could be used as a marker for early detection of GVHD onset.

Methods

One-hundred sixteen patients with haemato-oncological diseases, who underwent allogeneic HSCT at the Pediatric Clinics of “San Gerardo Hospital” (Monza) and “Regina Margherita Hospital” (Torino, Italy) were enrolled in the study [***4***]. Peripheral blood samples were collected before the beginning of conditioning regimen, on the day of HSCT (day 0), weekly after HSCT until day +100, and at the development of symptoms consistent with GVHD. Plasma was obtained after centrifuging whole blood and frozen until PTX3 was evaluated by ELISA assay.
A Cox model for time-dependent variables was applied to evaluate the association between PTX3 levels (ng/mL) and aGVHD onset. The lagged value of PTX3, i.e. the last weakly measurement before GVHD ascertainment, was included in the model to evaluate role in predicting disease onset. A joint model for longitudinal and time to event data was also applied to estimate the association of the longitudinal profile of PTX3 and GVHD [***5***]. The joint model allows to investigate the association between the longitudinal PTX3 and GVHD occurrence by combining two submodels: the linear mixed model for the longitudinal PTX3 and the Cox model for the GVHD occurrence. We specified a random intercept and random slope model to estimate the underlying unobserved true longitudinal PTX3, which was then used to fit the Cox model. To satisfy the normality assumption of the linear mixed model, the square root of PTX3 was used, and the transformation was consistently applied in all models.

Results

The time dependent Cox model including the updated value of the square root of PTX3 plasma levels revealed that it had a highly significant diagnostic relevance, with a hazard ratio (HR) of 1.07 (95%

confidence interval, CI, 1.03; 1.11, p-value < 0.0001). When the lagged value of PTX3 (i.e. the one measured in the previous week, as per planned sampling) was considered in the time dependent Cox model, the relationship between the pattern of PTX3 plasma level and the risk of GVHD lost statistical significance (p-value = 0.13, HR=1.04, 95% CI: 1.00; 1.08).
To better account for the measurement error in the profile of PTX3 value in time, a joint model was applied finding a significant association between the PTX3 value and GVHD risk (HR 1.16, 95%CI: 1.04; 1.30), p-value=0.007). The lagged value of PTX3 was also significantly associated with the event, showing a similar HR (1.16, 95%CI: 1.03; 1.31, p-value=0.01).

Conclusions

The relevance of PTX3 pattern on early GVHD diagnosis was evaluated with different models. The time dependent Cox model suggested that PTX3 is not useful to predict the risk of GVHD onset during follow-up, but that it has a significant diagnostic relevance, suggesting that the protein level rose concurrently with the clinical manifestations of GVHD. However, the shared parameter joint model revealed a significant association between the marker and GVHD both considering the most updated value and the lagged one, suggesting the possibility to predict GVHD onset with regular PTX3 dosage after HSCT. The joint model offers the benefit of effectively accounting for the measurement error in the longitudinal PTX3 profile, whereas the time dependent Cox model might be influenced by PTX3 measurement error. Future work will include the evaluation of the advantages of the joint model as compared with the time dependent Cox model, which in this case would have led to a different clinical strategy.

References

[**1**] Ferrara JL, Levine JE, Reddy P, Holler E. Graft-versushost disease. Lancet 2009;373:1550-61.
[**2**] Harris AC, Ferrara JLM, Levine JE. Advances in predicting acute GVHD. Br J Haematol 2013;160(3):288-302.
[**3**] Mantovani A, Valentino S, Gentile S, Inforzato A, Bottazzi B, Garlanda C. The long pentraxin PTX3: a paradigm for humoral pattern recognition molecules. Ann N Y Acad Sci 2013;1285:1-14.
[**4**] Dander E, De Lorenzo P, Bottazzi B, et al. Pentraxin 3 plasma levels at graft-versus-host disease onset predict disease severity and response to therapy in children given haematopoietic stem cell transplantation. Oncotarget 2016,7(50),82123-38.
[**5**] Rizopoulos D. Joint Models for Longitudinal and Time-to-Event Data: With Applications in R. Chapman and Hall/CRC 2012.

GAMLSS REGRESSION ANALYSIS IN ASSESSMENT OF REFERENCE INTERVALS OF OXIDATIVE STRESS BIOMARKERS IN HEALTHY ITALIAN POPULATION.

Chamitava Liliya[1], Cazzoletti Lucia[1], Degan Paolo[2], Pasini Andrea[3], Fratta Pasini Anna[3], Nicolis Morena[1], Olivieri Mario[4], Corsico Angelo[5], Pirina Pietro[6], Bono Roberto[7], Zanolin Maria Elisabetta[1]

1 *Unit of Epidemiology and Medical Statistics (SESM), Department of Public Health and Community Health, University of Verona, Italy*
2 *Epidemiology, Prevention and Special Functions, National Institute of Cancer Research AOU S. Martino IST, Genova, Italy;*
3 *Department of Internal Medicine, University of Verona, Verona, Italy;*
4 *Unit of Occupational Medicine, University of Verona, Verona, Italy;*
5 *Devision of Respiratory Diseases, ERCS, S. Matteo, Hospital University of Pavia, Pavia, Italy;*
6 *Institute of Respiratory Diseases, University of Sassari, Sassari, Italy;*
7 *Department of Public Health and Pediatrics, University of Turin, Turin, Italy;*

Introduction

Recent interest to investigation of oxidative stress (OS) linked disorders is due to its high involvement in many chronic and acute pathologies: cardiovascular, lung, hematological, diabetes, cancer and many others [***1***]. Oxidative stress is an imbalance between production of reactive oxygen species (ROS) and antioxidant defensive capacity of organism [***2***]. Normally ROS are generated in human body and are considered to be natural by products of metabolism of oxygen [***3***]. But in the framework of oxidative stress, ROS suppress antioxidant capacity in vivo, damaging DNA, lipids, and proteins [***4***]. ROS are highly reactive molecules with a short half-life and therefore it is very difficult to detect them in human biological fluids. Instead, the products of their oxidizing reactions, denoted as biomarkers of oxidative stress, are used widely to detect the level of OS in a human body [***5, 6***].

In this work we aimed at identifying main demographic and laboratory determinants of urinary biomarkers of oxidative stress (DNA-derived 8-oxodG and lipid membrane-derived 8-isoprostane), and at defining their reference intervals, adjusted for main determinants, in a sample of healthy adults from the general Italian population, using generalized additive models for location, scale and shape (GAMLSS) regression models.

Methods

In current study, the data on 281 subjects from the general Italian population, gathered during the Gene Environment Interactions in Respiratory Diseases (GEIRD) project who were never-and ex- (not smoking over the last year) smokers and who did not report respiratory symptoms at the screening questionnaire and at the clinical survey, either other comorbidities (heart disease, ictus, high blood pressure, diabetes and cancer) at the clinical interview and with normal lung function and allergologic test, have been used. Adjusted for determinants reference intervals of 8-oxodG and of 8-isoprostane were predicted using GAMLSS (generalized additive models for location, scale and shape) regression analysis.

Results

The main determinants of both biomarkers of oxidative stress were 'distance from collection' (DFC, the period from the moment of urine collection and its laboratory processing), and season, the period of year when urine was collected and split into warm (*April - September*) and cold (*October - March*). The reference intervals of 8-oxodG and of 8-isoprostane stratified by season and adjusted for DFC showed slight but statistically significant degradation of both biomarkers with increase of DFC in both seasons, except 8-oxodG biomarker during the warm season, which provided unchanged values with increased DFC. Thus, a

statistically significant mean negative association of DFC with log(8-oxodG) presented in the cold season (coef.α=-0.003, p=3.52*10^{-6}); no statistically significant association between DFC and log(8-oxodG) in warm season (coef.α=1.502*10^{-5}, p=0.971); a statistically significant negative association was seen in this model between DFC and log(8-isoprstane) (coef.β=-0.002, p=0.038); a statistically significant negative association was found in the warm season between 8-isoprostane and DFC (coef.β=-0.002; p=0.025).
The 95% reference intervals were estimated using the best developed GAMLSS model for each biomarker per each season, per each two weeks of DFC. During the cold season, 8-oxodG values decreased from 1.68 in first two weeks (36 days) to 0.08 ng/mg$_{creat}$ in last two weeks (756 days) for lower 2.5% limit of 95% RI and from 28.85 in first two weeks (36 days) to 14.19 ng/mg$_{creat}$ in last two weeks (756 days) for upper 97.5% limit of 95% RI. During the warm season 8-oxodG values were quite constant: 95% RI = 0.70 - 21.52 ng/mg$_{creat}$ in first two weeks (36 days) and 95% RI = 0.71 – 21.75 ng/mg$_{creat}$ in last two weeks (756 days). There was also a decrease of 8-isoprostane values during the cold season: 95% RI = 0.05 – 5.17 ng/mg$_{creat}$ in first two weeks (230 days) and 95% RI = 0.02 – 1.55 ng/mg$_{creat}$ in last two weeks (755 days). During the warm season 8-isoprostane values decreased as well: 95% RI = 0.06 – 6.54 ng/mg$_{creat}$ in first two weeks (230 days) and 95% RI = 0.03 – 2.96 ng/mg$_{creat}$ in last two weeks (755 days).

Conclusions

To our knowledge, it is for the first time when it was shown that both OS biomarkers, 8-oxodG and 8-isoprostane, should be evaluated in association with DFC and season when urine has been collected. It is especially important in large epidemiological studies when long-term conservation of urine is stipulated. *(Semi)*parametric GAMLSS regression analysis is a new useful technique that can be used for estimating adjusted for appropriate determinants reference intervals of urinary biomarkers (8-oxodG and 8-isoprostane) from general adult population.

References

[**1**] Wu LL, Chiou CC, Chang PY, et al. Urinary 8-OHdG: a marker of oxidative stress to DNA and a risk factor for cancer, atherosclerosis and diabetics. Clin Chim Acta 2004;339(1-2):1-9.
[**2**] Tkaczyk J, Vizek M. Oxidative stress in the lung tissue--sources of reactive oxygen species and antioxidant defence. Prague Med Rep 2007;108(2):105-14.
[**3**] Devasagayam TP, Tilak JC, Boloor KK. et al. Free radicals and antioxidants in human health: current status and future prospects. J Assoc Physicians India 2004;52:794-804.
[**4**] Sakano N, Wang DH, Takahashi N, et al. Oxidative stress biomarkers and lifestyles in Japanese healthy people. J Clin Biochem Nutr 2009;44(2):185-95.
[**5**] Ciencewicki J, Trivedi S, Kleeberger SR. Oxidants and the pathogenesis of lung diseases. J Allergy Clin Immunol. 2008;122(3):456-68; quiz 69-70.
[**6**] Sakano N, Takahashi N, Wang DH, et al. Plasma 3-nitrotyrosine, urinary 8-isoprostane and 8-OHdG among healthy Japanese people. Free Radic Res 2009;43(2):183-92.

HIGHLY UNBALANCED TEXT CLASSIFICATION IN HEALTHCARE DATA, THE HUTCH STUDY

Lanera Corrado[1], Sharma Abhinav[2], Minto Clara[1], Gregori Dario[1], Berchialla Paola[3], Baldi Ileana[1]

1 *Unit of Biostatistics, Epidemiology and Public Health, Department of Cardiac, Thoracic and Vascular Sciences, University of Padua, Italy.*
2 *Department of Biological Sciences & Bioengineering, Indian Institute of Technology Kanpur, India.*
3 *Department of Clinical and Biological Science, University of Turin, Italy.*

Introduction

Systematic reviews of clinical trials are an important means of synthesizing medical evidence. With the growing number of published studies and increasing amount of textual information in biomedical data, text mining has been considered as a potential solution to make use of the knowledge encoded in the text documents. It reduces the reviewer time by automating some of the screening process. But one of the major challenge in biomedical text classification is related to the high disproportion in amount of relevant and non-relevant papers which are available for the user query [***1***]. Data unbalance occurs when one of the classes is highly under-represented (minority class), that is only small number of examples are represented by that class whereas the majority class has very high number of examples. In our training sample the ratio of the relevant papers (minority class) to non-relevant ones (majority class) is 23:400. Machine learning algorithms used as classifiers generally are biased towards the majority class [***2***] and show very poor results towards the minority class and hence their performance often decreases [***1***].
Various data handling strategies and algorithmic approaches have been used to deal with the problem of unbalanced data in text classification. Resampling strategies like oversampling and undersampling have been widely used in the past to handle unbalanced data [***3***]. Ensemble methods which are based on popular techniques like Boosting and Bagging have drawn more attention because of their flexible characteristics. One of the reason being is their similar learning scheme: resampling, base learning algorithm, voting but different strategies in each phase [***4***]. Another popular approach is Cost sensitive learning, this approach uses different cost matrices for the misclassification of examples and tries to minimize the cost errors [***5***].
The aim of this study is to determine the best overall Machine Learning strategy to deal with problem of highly unbalanced text classification, particularly focused on healthcare data. To achieve this objective we simulate a literature search conducted on a indexed platform, such as PubMed, on a different one without using its own interface, i.e. WHO ICTRP. We investigated the application of different Machine Learning algorithms trained after the application of selected methods to manage the problem of unbalanced data.

Methods

The classifications are performed in two steps: a first step which is a method to deal with the unbalanced data and the second one which is an application of a selected classifier trained on the balance data output of the first step. Balancing of the data was performed by methods selected from sampling strategies like Random Under Sampling (RUS) with reduced bias to the majority class [***6***], SMOTE [***7***], with ensemble methods like RUS Boost, SMOTE Boost, SMOTE Bagging and cost sensitive learning with different misclassification costs.
For each of these methods, used in the first step, we have considered for the second step the following classifiers: SVM, kNN, Naïve Bayes, RF, Logit Boost, Maximum Entropy and GLMnet. The training of each classifiers include a 10-fold cross-validation step to assess the fine tuning of the relative parameters, with the aim to achieve the best non over-fitted performance according to the “one standard error from the best”

criteria [***8***]. Finally, different performance measures like precision, recall, F-measure and AUC are stated and compared for each combination of the methods and classifiers.
For the motivating example, we selected one of the published systematic review [***9***] included in Baudard's study [***10***]: here, author reanalyse systematic review to identify missed relevant papers included in registries but not included in original search. We converted this investigation into an automatic investigation conducted using Machine Learning algorithms. Hence, the PubMed database was searched for articles to provide a training set with "positives" and "negatives" samples reproducing the results of Liu's study [***9***]. For the negatives, we randomly selected a set of clinical trials from the results of the original search string. The following text pre-processing procedures were applied to the title and abstract fields of the retrieved records in the following order: conversion to lowercase, removing numbers, removing punctuation, removing stop-words, stemming words, stripping white space, and building a sequence of two adjacent words from the text (bigrams). The whole collection was finally tokenized in a document-term matrix (DTM) which was used as input for each methods of the first step of our procedure for the balancing of classes. Then, the outcome dataset was used to train and validate with each of the considered classifiers. Finally, the trained classifiers was tested on the WHO ICTRP data comparing the results with the one in [***10***].

Results

At the moment one ML tool (10-fold cross validated Logit Boost) finished the training on the DTM (423 documents × 12151 terms, i.e. title + abstract) from Pubmed and has been used to predict which trials in WHO ICTRP should be included in the review.
Training sample contains 23 positives Pubmed trials retained in the review and 400 negatives, randomly sampled from Pubmed trials not included in the string search output.
Results of the internal validation are given in the ***Figure 1*** which represents the level of accuracy we reached.
In this PubMed training set, the sensitivity of the classifier is 81.2% and the specificity is 99.3%.
Results for the prediction on WHO ICTRP lead to the identification of 15 additional studies (1 true and 14 false).

Conclusions

Routine examination of registry databases deserves further consideration since it may allow a more accurate characterization of publication and outcome reporting biases and improve the validity of systematic reviews. The approaches described here provides an automated solution that can be used to address a variety of clinical trial-related questions by building a comprehensive search on both literature and registry databases. Results from the techniques which are still on training will provide the necessary information to asses a more accurate selection of strategies to better identify the best procedure able to manage highly unbalanced data in text classification.

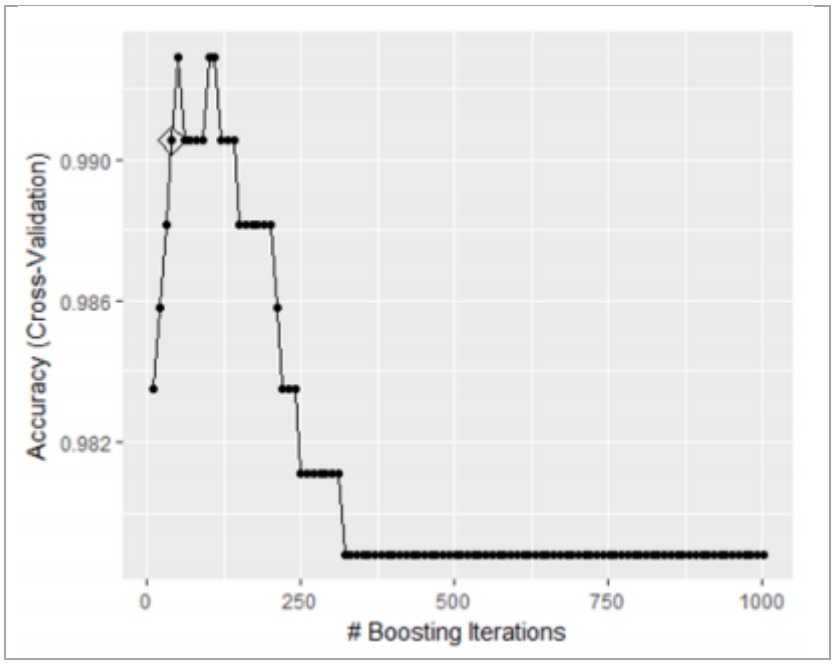

Figure 1. Accuracy of Logit Boost at each step of the internal 10-fold cross-validation procedure

References

[1] Borrajo L, Romero R, Iglesias EL, et al. Improving imbalanced scientific text classification using sampling strategies and dictionaries. Journal of Integrative Bioinformatics 2011;8(3):90-104.

[2] Longadge R, Dongre SS, Malik L. Class imbalance problem in data mining: Review. International Journal of Computer Science and Network 2013;2(1):www.ijcsn.org.

[3] Liu AY. The effect of oversampling and undersampling on classifying imbalanced text datasets. The University of Texas at Austin, 2004.

[4] Wang S, Yao X. Diversity analysis on imbalanced data sets by using ensemble models. in Computational Intelligence and Data Mining 2009. CIDM'09. IEEE Symposium on. 2009. IEEE.

[5] Laza R, Pavón R, Reboiro-Jato M, et al. Evaluating the effect of unbalanced data in biomedical document classification. Journal of Integrative Bioinformatics 2011;8(3):105-17.

[6] Kocbek S., Cavedon L, Martinez D, et al. Text mining electronic hospital records to automatically classify admissions against disease: Measuring the impact of linking data sources. Journal of Biomedical Informatics 2016;64:158-67.

[7] Chawla NV, Bowyer KW, Hall LO, et al. SMOTE: synthetic minority over-sampling technique. Journal of artificial intelligence research 2002;16:321-57.

[8] Breiman L, Friedman J, Stone CJ, et al. Classification and Regression Trees 1984: Taylor & Francis.

[9] Liu X, Xiao Q, Zhang L, et al. The long-term efficacy and safety of DPP-IV inhibitors monotherapy and in combination with metformin in 18 980 patients with type-2 diabetes mellitus—a meta-analysis. Pharmacoepidemiology and drug safety 2014;23(7):687-98.

[10] Baudard M, Yavchitz A, Ravaud P, Perrodeau E, Boutron I. Impact of searching clinical trial registries in systematic reviews of pharmaceutical treatments: methodological systematic review and reanalysis of meta-analyses. BMJ 2017;17:356-j448

GRAPHICAL REPRESENTATIONS AND SUMMARY INDICATORS TO ASSESS THE PERFORMANCE OF RISK PREDICTORS

Antolini Laura, Tassistro Elena, Bernasconi Davide Paolo, Valsecchi Maria Grazia

Center of Biostatistics for Clinical Epidemiology, School of Medicine and Surgery, University of Milano-Bicocca, Monza, Italy

Introduction

The availability of novel biomarkers in several branch of medicine opens room for refining prognosis by adding factors on top of those having an established role. It is accepted that the impact of novel factors should not rely solely on regression coefficients and their significance, but on predictive power measures, such as Brier Score and ROC based quantities. However, novel factors who are promising at the explorative stage often results in disappointingly low impact in the predictive power and without reaching significance. This motivated the proposal of the net reclassification index and the integrated discrimination improvement, as direct measures of gain due to additional factors based on the concept of reclassification tables. This measures became extremely popular in cardiovascular disease and cancer applications. However, recent contributions in the biostatistical literature enlightened strong limitations, including an apparent easy interpretation and the tendency to indicate advantageous models obtained adding unrelated factors. A further measure proposed a decade ago, namely the net benefit, is starting being used. This measure appears to be promising in assessing the consequences in terms of costs and benefits when using a risk predictor in practice for classification.
This work starts reviewing the conceptual formulations of the available graphical methods and summary measures for evaluating risk predictor models. The aim is to provide guidance in the evaluation process that from the model development brings the risk predictor to be used in practice.

Methods

A large simulated data set (N =10000) with prevalence of disease ρ = 10.2% will be used as an entire population. The binary disease status is generated depending on continuous covariate X and Y. Risk predictors are obtained from the correctly specified model including only X or also Y on top of X. The risk prediction is intended for defining a binary classification rule according to a risk threshold r. If the risk prediction is greater or equal than r we assume the patient will be treated, if the risk prediction is lower than one we assume the patient will not be treated.
On the simulated data set, the following graphical representations are obtained: Predictivness curve, true positive fraction (TPF) and false positive fraction (FPF) against the risk threshold r graphs, Receiver operating characteristic curve (ROC), Decision/utility curve (DUC). The summary indicators: Brier score measure, Mean risk difference (MRD), Integrated discrimination improvement (IDI), Above average risk difference (AARD), Youden Index, area under the Receiver Operating Characteristic curve (AUC), net reclassification index (NRI) are discussed and linked to the aforementioned graphical representations (***Table 1***).

Results

This work started from the observation that calibration is a basic requirement a risk predictor should satisfy to be used in practice. Assessing this aspect carefully becomes crucial especially when an externally fitted model is applied on novel data coming from a different population from that of development, to assess generalizability. If a satisfactory calibration is not achieved, the model should be considered not generalized.

Nevertheless, some investigators proceed to evaluate the discrimination and classification performance. This is however not appropriate since in this context the real numerical value of the risk prediction will be used to drive medical decision, e.g. observing if is greater than some threshold r and submitting subsequently the patient to a treatment. A better strategy perhaps is to derive a revised model through recalibration procedures. Thus, the discrimination assessment should start from a calibrated model.
The MRD indicator is a very simple measure of discrimination defined as the difference between the average risk prediction in diseased and non-diseased. This indicator summarizes the graphical representation of TPF and FPF versus the possible r risk threshold values. Of note, the use of these graphs is recommended as an alternative to the ROC curve since in the plot of TPF across FPF the risk threshold r for classification is not shown explicitly. In addition, the ROC curve is invariant with respect to monotonic transformations of the risk predictor, which is in this context is a disadvantage. Moreover, the AUC does not reflect the clinical use of the risk predictor and it can be dominated by differences in risk distribution who are not clinically relevant. The classification performance when an operative risk threshold r is set, can be assessed through simple distance measures between TPF and FPF in the plots against the risk threshold r, namely AARD and Youden index. The NRI of nested models, although very simple to calculate being based on reclassification tables, can be dominated by differences in risk distribution who are not clinically relevant. Moreover, the tendency of NRI to indicate advantageous models obtained adding unrelated factors and mis-calibrated models was proved in simulations. This research was corroborated by subsequent contributions on the behaviour of the NRI which appeared positively inflated even comparing nested logistic regression models under correct specification on independent validation data sets. These unexpected behaviour was not found for traditional performance measures, such BS, ROC quantities. Of note, it was proved, however, that classification rules with the same marginals could result in improvement in NRI. This serious flaw is probably due to the definition of the measure given directly on the contrast between the nested risk predictors and not as difference between single risk predictor performance measures.
The net benefit measure definition requires the specification of the cost C of over-treatment of a nondiseased subject assuming equal to 1 the gain of treating a diseased subject. This enables to derive an optimal risk threshold r for classification in order to maximize the net benefit count at the single patient level. Then the average benefit, namely NB, is calculated by multiplying the fraction of patients who are both - diseased and with a risk prediction greater than the optimal r – discounted by the product of C by the fraction of patients who are both – non diseased and with a risk prediction greater than the optimal r. The concept of net benefit enables also to derive indirectly the cost of over treatment C when a risk threshold r is specified relying in clinical (or even patient specific perceptions) and assumed as optimal. This again enables to calculate the average NB, which can be displayed against C and against the risk threshold r. leading to the DUC curve. The NB multiplied by 100 is interpretable ad the number of patients correctly treated, any 100 patients submitted to the risk prediction, accounting for the cost of over treatment.

Table 1. Main characteristics of graphical methods and summary measures

	Numerical value	Disease prevalence	Risk threshold	Clinical Intepretatio n	Costs and Benefits
Predictivness curve	Y	Y	Y	Y	N
BS	Y	N	N	N	N
TPF and FPF graphs	Y	Y	Y	Y	N
MRD	Y	N	N	Y	N
AARD, Youden	Y	N	Y	Y	N
ROC curve	N	N	Y	Y	N
AUC	N	N	N	N	N
NB, DUC	Y	Y	Y	Y	Y

Conclusions

The main characteristics of the graphical methods and summary measures described in this work are represented in the table below. They are listed in the ordering that, in our opinion, one should follow in assessing the model performance from the model development to the risk use in practice. For each graphical representation/measure it is indicated with Y/N if the key aspects listed in the columns are accounted. The key aspects are: numerical value of risk prediction (meaning that there is not invariance with respect to rank preserving transformations), Prevalence of cases ρ (if accounted in the measure definition), Risk threshold r (if displayed in the graph or present in the interpretation of the measure), Clinical interpretation (if satisfactory), Benefits and costs (if accounted). Besides the popularity of NRI measure, it was not included since the serious flaw presented in the literature are really discouraging. Of note, NB accounts for all the key aspects, and among all the available measures is the one more related to the consequences of the risk predictor used in practice.

References

[1] Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med 2008;27(2):157-72.

[2] Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006;26(6):565-74.

[3] Kerr KF, Wang Z, Janes H, McClelland RL, Psaty BM, Pepe MS. Net reclassification indices for evaluating risk prediction instruments: a critical review. Epidemiology 2014;25(1):114-21.

[4] Vickers AJ, Pepe M. Does the net reclassification improvement help us evaluate models and markers? Ann Intern Med 2014;160(2):136-37.

[5] Pepe MS, Fan J, Feng Z, Gerds T, Hilden J. The Net Reclassification Index (NRI): a misleading measure of prediction improvement even with independent test data sets. Stat. Biosci 2014;7(2):282-95.

[6] Baker SG, Schuit E, Steyerberg EW, et al. How to interpret a small increase in AUC with an additional risk prediction marker: decision analysis comes through. Stat. Med 2014;33(22):3946-59.

[7] Kerr KF, Brown MD, Zhu K, Janes H. Assessing the Clinical Impact of Risk Prediction Models With Decision Curves: Guidance for Correct Interpretation and Appropriate Use. J Clin Oncol 2016;34(21):2534-40.

[8] Lee MLT, Gail M, Pfeiffer R, Satten G, Cai T, Gandy A. (2013). Risk assessment and evaluation of predictions. Springer New York.

USO DI BANCHE DATI

L'IMPATTO DEL RITARDO AL TRATTAMENTO SULLA SOPRAVVIVENZA DEI TUMORI DI TESTA E COLLO

Zucchetto Antonella[1], Furlan Carlo[2], Birri Silvia[1], Barzan Luigi[3], Taborelli Martina[1], Panato Chiara[1], Serraino Diego[1], Franchin Giovanni[2], Polesel Jerry[1]

1 *Società di Epidemiologia Oncologica, Centro di Riferimento Oncologico (CRO), IRCCS Aviano (PN)*
2 *Società di Oncologica Radioterapica, Centro di Riferimento Oncologico (CRO), IRCCS Aviano (PN)*
3 *Società di Oncologia Chirurgica, Centro di Riferimento Oncologico (CRO), IRCCS Aviano (PN)*

Introduzione

La stadiazione veloce ed il trattamento tempestivo sono cruciali per una buona prognosi dei pazienti affetti da carcinomi squamosi di testa e collo (CSTC) [***1***], considerando che il tempo di raddoppio del volume del tumore è di circa 90 giorni. Tuttavia, negli ultimi decenni l'intervallo tra diagnosi e trattamento si sta allungando, come conseguenza dell'uso sempre più frequente di procedure diagnostiche ad alta definizione e della personalizzazione del trattamento [***2, 3***].
Diventa quindi prioritario definire con precisione il ritardo accettabile al trattamento, che medi tra le necessità competitive di sfruttare le migliori tecnologie per la stadiazione e personalizzazione delle cure da una parte e, dall'altra, di iniziare il trattamento il prima possibile. A fianco della rilevanza clinica per il paziente, l'identificazione di un benchmark per il ritardo al trattamento è essenziale per indirizzare le scelte di politica sanitaria.
Un recente studio condotto negli Stati Uniti [***1***] ha mostrato come un ritardo al trattamento superiore a 60 giorni sia associato ad una riduzione significativa della sopravvivenza nei pazienti con CSTC. Tuttavia, tale effetto può variare in relazione a molteplici fattori ambientali, come le caratteristiche del sistema sanitario di riferimento e la pratica clinica per tali tumori.
L'obiettivo di questo lavoro è quello di valutare l'impatto del ritardo al trattamento sulla sopravvivenza dei pazienti con CSTC, tenendo in considerazione le caratteristiche cliniche del paziente e del tumore, nonché delle caratteristiche del sistema sanitario.

Metodi

I dati di questo studio sono stati ottenuti incrociando diverse banche dati sanitarie. Fin dal 1995, nella regione Friuli Venezia Giulia è attivo un registro tumori su base regionale che raccoglie informazioni su tutte le nuove diagnosi di tumore nella popolazione residente. Questo registro copre l'intero territorio regionale (circa 1,2 milioni di abitanti), riportando elevati standard di qualità e completezza [***4***].
Il registro tumori del Friuli Venezia Giulia è integrato nel Data Warehouse sanitario regionale, un sistema informatico che, fin dai primi anni '80, raccoglie in modo organizzato tutti i principali dati sanitari delle persone residenti nel territorio regionale. In particolare, il Data Warehouse raccoglie tutte le informazioni sullo stato in vita, movimenti migratori, ricoveri (schede di dimissione ospedaliera, SDO), specialistica ambulatoriale, prescrizioni farmaceutiche. Le varie banche dati sanitarie sono legate da un codice alfanumerico univoco per paziente che rende anonimo l'intero Data Warehouse, nel rispetto delle normative nazionali e regionali in tema di tutela della Privacy, e che ha permesso di effettuare un incrocio tra le banche dati.
Ai fini di questo studio, sono stati inclusi tutti pazienti con diagnosi di carcinoma squamo cellulare (codici morfologici ICD-O-3: 8000-8082, 8430, 8480, 8560, 8801) nel periodo gennaio 2003-dicembre 2009 nelle seguenti sedi anatomiche: cavo orale (topografia ICD-O-3: C00.3-C00.9, C02-C06, C14), orofaringe (C01, C09-C10), ipofaringe (C13), e laringe (C32). Lo stato in vita è stato aggiornato al 31 dicembre 2014.

In totale sono stati identificati 2178 pazienti. Tuttavia, in relazione agli obiettivi dello studio, sono stati esclusi 53 pazienti con diagnosi autoptica, 398 pazienti con età >75 anni e 111 pazienti persi al follow-up. Lo studio ha quindi incluso 1616 pazienti che hanno ricevuto un trattamento con intento curativo. La quasi totalità dei pazienti (99.1%) ha avuto una diagnosi istologicamente confermata.
Gli anni-persona a rischio sono stati calcolati dalla data di diagnosi del tumore fino alla data di morte, di migrazione o al 31 dicembre 2014. Il tempo all'inizio del trattamento (TIT) è stato definito come i giorni intercorsi tra la data di diagnosi ed il primo trattamento (chirurgia, radioterapia o chemioterapia). Il TIT è stato quindi categorizzato in quattro gruppi: 0-29 giorni, 30-44 giorni, 45-89 giorni, ≥90 giorni [***1***]. I trattamenti effettuati da ogni pazienti e le strutture dove sono stati effettuati sono stati recuperati dal Data Warehouse regionale.
La sopravvivenza (overall survival, OS) è stata stimata attraverso il metodo di Kaplan-Meier e le differenze tra curve sono state valutate attraverso il log-rank test [***5***]. Il rischio di morte (Hazard ratio, HR) ed i corrispettivi intervalli di confidenza (IC) al 95% sono stati stimati attraverso il modello a rischi proporzionali di Cox [***5***], aggiustando per sesso ed età. Gli assunti di proporzionalità sono stati verificati attraverso l'analisi dei residui di Schoenfeld. Nel modello multivariato sono state incluse le seguenti covariate: provincia di residenza, sede topografica del tumore, storia di tumori precedenti, stadio TNM, approccio terapeutico, tipo di struttura del primo trattamento e mobilità di cura. L'approccio terapeutico è stato definito come sola chirurgia (CH), sola radioterapia (RT), radio-chemioterapia concomitante radicale (RCT), chemioterapia neo-adiuvante seguita da radioterapia (neoCT-RT). La struttura del primo trattamento è stata definita come istituto di ricerca o universitario, ospedale di rete, cliniche privati ad altri (inclusi hospice e strutture fuori regione); la mobilità di cura è stata definita come nessuna (cura nella stessa struttura di diagnosi o all'interno di istituti di ricerca/universitari), mobilità all'interno di ospedali di rete, mobilità da ospedali di rete verso istituti di ricerca/universitari.

Risultati

La maggior parte dei pazienti (52,9%) è stata trattata entro 30 giorni dalla diagnosi, mentre il 6,4% ha atteso il trattamento per ≥90 giorni (TIT mediano: 28 giorni; range interquartile, Q1-Q3: 13-45 giorni). La mediana di TIT è significativamente più bassa per i tumori della laringe (22 giorni) rispetto ai tumori dell'orofaringe (33 giorni), ipofaringe (32 giorni) e cavo orale (28 giorni). Si è inoltre osservata un'associazione diretta tra il TIT e lo stadio alla diagnosi, con una mediana di 21 giorni per gli stadi I fino a 29 giorni per gli stadi IV. I pazienti diagnosticati in un ospedale di rete che si sono spostati per il trattamento in un istituto di ricerca/universitario hanno una mediana di TIT maggiore (41 giorni) rispetto a coloro che sono stati trattati all'interno degli ospedali di rete (16 giorni).
L'OS a 5 anni è inversamente associata al ritardo al trattamento, essendo pari a 62,1% per i pazienti trattati entro 30 giorni dalla diagnosi, 51,1% per TIT tra 30 e 44 giorni, 50,4% per TIT tra 45 e 89 giorni e 39,4% per TIT≥90 giorni (log-rank test: p<0,01). Il rischio di morte cresce con l'aumento del ritardo al trattamento anche se, dopo l'aggiustamento per le covariate, solo i pazienti con trattamento dopo 90 giorni mostrano un rischio significativamente più elevato di quelli trattati entro 30 giorni (HR=1,47; $IC_{95\%}$: 1,05-2,05 – ***Tabella 1***). Tuttavia, emerge un rischio apprezzabile, anche se non statisticamente significativo, per i pazienti trattati dopo 45 giorni dalla diagnosi.
Tali rischi sono al netto delle altre covariate che risultano modificare il rischio di morte, tra cui il sesso (HR per donne vs uomini=0,72; $IC_{95\%}$ 0,59-0,87), l'età (HR per 70-75 anni vs 18-54=1,63; $IC_{95\%}$ 1,32-2,00), la sede topografica del tumore (HR per laringe vs cavo orale=0,74; $IC_{95\%}$ 0,61-0,89), un tumore precedente (HR per Si vs No=1,53; IC 95% 1,32-1,77) e lo stadio (HR per stadio IV vs stadio I=2,66; $IC_{95\%}$ 2,07-3,40).
La mobilità per trattamento dagli ospedali di rete verso gli istituti di ricerca/universitari è associata ad una migliore prognosi (HR=0,73; $IC_{95\%}$ 0,60-0,88), seppure il tempo di attesa per il trattamento sia più lungo

rispetto ai pazienti che rimangono negli ospedali di rete per le cure. Infatti, i pazienti che si muovono per il trattamento sono generalmente quelli con i fattori prognostici peggiori (tumore del cavo orale o tumore

Tabella 1. Rischio di morte (Hazard ratio, HR) ed i corrispettivi intervalli di confidenza (IC) al 95% per tempo all'inizio del trattamento, caratteristiche sociodemografiche, caratteristiche del tumore, trattamento.

	Pazienti	Decessi		HR univariato	HR multivariato
	n	n	(%)	(IC 95%)[a]	(IC 95%)[b]
Tempo all'inizio del trattamento (giorni)					
0-29	855	428	(50,1)	Riferimento	Riferimento
30-44	356	204	(57,3)	1,32 (1,12-1,56)	1,08 (0,90-1,30)
45-89	301	185	(61,5)	1,43 (1,20-1,70)	1,13 (0,92-1,39)
≥90	104	70	(67,3)	1,55 (1,20-1,99)	1,47 (1,05-2,05)
Sesso					
Uomini	1318	759	(57,6)	Riferimento	Riferimento
Donne	298	128	(43,0)	0,67 (0,56-0,81)	0,72 (0,59-0,87)
Età (anni)					
18-54	354	169	(47,7)	Riferimento	Riferimento
55-59	314	150	(47,8)	0,97 (0,78-1,21)	0,95 (0,76-1,18)
60-64	311	166	(53,4)	1,16 (0,94-1,44)	1,15 (0,93-1,44)
65-69	338	194	(57,4)	1,27 (1,03-1,56)	1,32 (1,07-1,63)
70-75	299	208	(69,6)	1,75 (1,43-2,15)	1,63 (1,32-2,00)
Sede topografica					
Cavo orale	462	236	(51,1)	Riferimento	Riferimento
Orofaringe	346	218	(63,0)	1,34 (1,12-1,62)	0,85 (0,69-1,05)
Ipofaringe[c]	212	148	(69,8)	1,57 (1,27-1,93)	0,87 (0,69-1,10)
Laringe	596	285	(47,8)	0,75 (0,63-0,89)	0,74 (0,61-0,89)
Tumore precedente					
No	1216	608	(50,0)	Riferimento	Riferimento
Sì	400	279	(69,8)	1,46 (1,27-1,69)	1,53 (1,32-1,77)
Stadio TNM					
I	297	86	(29,0)	Riferimento	Riferimento
II	105	50	(47,6)	1,80 (1,27-2,55)	1,67 (1,17-2,37)
III	104	59	(56,7)	2,45 (1,76-3,41)	2,03 (1,44-2,86)
IV	424	311	(73,4)	4,26 (3,35-5,42)	3,31 (2,54-4,31)
Sconosciuto	686	381	(55,5)	3,12 (2,46-3,96)	2,66 (2,07-3,40)
Trattamento					
Solo chirurgia	564	237	(42,0)	Riferimento	Riferimento
Solo RT	165	93	(56,4)	1,51 (1,19-1,92)	1,51 (1,10-2,07)
RCT	492	348	(70,7)	2,56 (2,17-3,03)	1,96 (1,58-2,42)
Neo- CT-RT	395	209	(52,9)	1,45 (1,21-1,75)	1,20 (0,98-1,46)
Struttura del trattamento					
Ist. Ricerca/Università	1082	600	(55,5)	Riferimento	Riferimento
Ospedale di rete	432	243	(56,3)	0,98 (0,84-1,14)	0,96 (0,64-1,45)
Privata/altro	97	43	(44,3)	0,68 (0,50-0,93)	0,90 (0,60-1,33)
Mobilità verso istituti di ricerca/universitari[d]					
No	464	252	(54,3)	Riferimento	Riferimento
Sì	283	154	(54,4)	0,95 (0,78-1,17)	0,73 (0,60-0,88)

[a]Aggiustato per sesso ed età. [b]Aggiustato per tutte le variabili in tabella. [c]Incluso faringe NOS. [d]Solo per pazienti diagnosticati in ospedali di rete.

avanzato) che maggiormente beneficiano dei trattamenti tecnologicamente avanzati disponibili presso gli istituti di ricerca/universitari.

Conclusioni

I risultati di questo studio confermano, anche nella realtà italiana, un'associazione inversa tra tempo al trattamento e sopravvivenza nei pazienti affetti da carcinoma squamoso di testa e collo. In particolare, suggeriscono che per tali tumori sarebbe ottimale iniziare il trattamento entro 45 giorni dalla diagnosi. Tuttavia, tale tempistica può essere in contrasto con la transizione verso gli istituti di ricerca/accademici che possono offrire maggior possibilità terapeutiche.
Tutti i dettagli dello studio sono disponibili in Polesel J et al. [***6***].

Bibliografia

[**1**] Murphy CT, Galloway TJ, Handfor EA, et al. Survival impact of increasing time to treatment initiation for patients with head and neck cancer in the United States. J Clin Oncol 2016;34(2):169-78.

[**2**] Primdahl H, Nielsen AL, Larsen S, et al. Changes from 1992 to 2002 in the pretreatment delay for patients with squamous cell carcinoma of larynx or pharynx: a Danish nationwide survey from DAHANCA. Acta Oncol 2006;45(2):156–61.

[**3**] Murphy CT, Galloway TJ, Handorf EA, et al. Increasing time to treatment initiation for head and neck cancer: An analysis of the National Cancer Database. Cancer 2015;121(8):1204-13.

[**4**] Forman D, Bray F, Brewster DH, et al (Eds). Cancer Incidence in Five Continents Vol. X IARC Sci Pub No. 164; IARC: Lyon, 2014.

[**5**] Kalbfleisch JD, Prentice RL. The statistical analysis of failure time data. Wiley, New York, 2002.

[**6**] Polesel J, Furlan C, Birri S, et al. The impact of time to treatment initiation on survival from head and neck cancer in north-eastern Italy. Oral Oncol 2017;67:175-82

UTILIZZO DEI DATABASE AMMINISTRATIVI PER LA STIMA DELLA PREVALENZA DI EMBOLIA POLMONARE IN PAZIENTI GIUNTI IN PRONTO SOCCORSO PER SINCOPE.

Casazza Giovanni[1], Russo Antonio[2], Solbiati Monica[3], Quinn James[4], Costantino Giorgio[3]

1. *Dipartimento di Scienze Biomediche e Cliniche "L. Sacco" Università degli Studi di Milano;*
2. *Unità di Epidemiologia ATS della Città Metropolitana di Milano;*
3. *Dipartimento di Medicina Interna e Specializzazioni mediche, Fondazione IRCCS Ca' Granda, Ospedale Maggiore Policlinico, Università degli Studi di Milano;*
4. *Department of Emergency Medicine, Stanford University, Stanford, US;*

Introduzione

La sincope è una condizione abbastanza frequente con una occorrenza di eventi avversi relativamente bassa. L'embolia polmonare (EP), la più comune manifestazione di tromboembolismo venoso (TEV), è uno fra questi eventi che, se non riconosciuto e curato per tempo, può avere esiti anche letali. Da dati di letteratura risulta che la prevalenza di EP in pazienti che si presentano in Pronto Soccorso (PS) in seguito ad episodio sincopale è molto bassa [***1***]. Tuttavia, uno studio recente [***2***] ha riportato un'elevata prevalenza (17%) di EP in pazienti ospedalizzati in seguito ad episodio sincopale, corrispondente al 4% circa di prevalenza di EP fra i pazienti giunti in PS per sincope. Se questo dato fosse confermato, si renderebbe necessaria l'introduzione nella pratica clinica di un'accurata valutazione dei pazienti con sincope per ricercare la presenza di EP. Un approccio del genere, se applicato indiscriminatamente, porterebbe ad un incremento di costi per il sistema, di rischi per il paziente e ad una sovra-diagnosi di embolia. Prima di introdurre una procedura diagnostica costosa e non esente da rischi per il paziente è quindi opportuna un'accurata valutazione del rischio di EP nei pazienti con sincope, ed i database amministrativi possono essere utili a questo scopo.
Valutare, mediante l'utilizzo di database amministrativi, la prevalenza di EP in tutti i pazienti giunti in PS per sincope.

Metodi

Sono stati utilizzati i dati provenienti da due fonti.
ATS Città Metropolitana di Milano. Dati di tutti i residenti nei territorio dell'ATS (3.2 milioni), estratti da registro degli accessi in PS, database dei ricoveri ospedalieri e registro di mortalità. Il periodo considerato va dal 01/01/2014 al 30/09/2016.
Optum database. Clinformatics™ DataMart Database (OptumInsight, Eden Prairie, MN, USA) è un database anonimizzato, creato da una compagnia di assicurazione che opera su tutto il territorio degli USA. L'intero database comprende informazioni relative a 58 milioni di membri provenienti dai 50 stati. Sono riportati dati relativi ai rimborsi (*claims*) richiesti da circa 12-14 milioni di membri per ogni anno, di cui 2.5-3.5 milioni sono coperti dal programma *Medicare*. La rappresentatività dei pazienti in questo database è considerata buona per molte caratteristiche individuali. I database analizzati comprendono visite in PS e ricoveri ospedalieri, nel periodo dal 01/01/2004 al 31/12/2015.
Per entrambi i database, sono stati inclusi tutti i pazienti giunti in PS, con età≥18 anni e con un codice diagnosi di sincope (ICD-9-CM: 780.2, ICD-10: R55). Questi codici sono stati validati in precedenza [***3***]. Per pazienti con più episodi sincopali nel periodo esaminato è stato considerato solo il primo episodio (evento indice).

Abbiamo poi considerato quale *outcome* un episodio di EP, definito come presenza di uno fra i seguenti codici:

- 415, 453.4, 453.5, 453.8, 453.9 (ICD-9-CM);
- I26, I801-I803, I808, I809, I821-I823, I828, I829 (ICD-10)

in una qualunque delle diagnosi riportate nella visita in PS o nelle ospedalizzazioni avvenute entro 90 giorni dall'evento indice. Questo termine è stato scelto in quanto il follow-up a 90 giorni è considerato *reference standard* per la diagnosi di EP, e quindi possiamo considerare gli eventi avvenuti entro tale temine come indicazione di episodio di EP già presente al momento dell'evento indice [***4***]. L'utilizzo dei codici ICD elencati sopra comporta una sovrastima della prevalenza di EP che si traduce, ai fini dell'obiettivo dello studio, in risultati più conservativi. La prevalenza di EP è stata quindi calcolata come proporzione di pazienti con un evento embolico registrato nei 90 giorni successivi all'evento indice, rispetto al totale dei pazienti giunti in PS per sincope. I risultati sono riportati come proporzioni con intervalli di confidenza (IC) al 95%. Infine, per ottenere risultati maggiormente comparabili con il dato di letteratura citato precedentemente [***2***], la prevalenza è stata stimata anche per il solo sottogruppo dei pazienti con ospedalizzazione immediatamente successiva (entro un giorno) alla dimissione da PS (sottogruppo *"pazienti ospedalizzati"*). I due database, ATS ed *Optum*, sono stati analizzati separatamente.

Risultati

ATS Milano. Sono stati inclusi 29543 pazienti di età ≥18 anni giunti in PS per sincope dal 01/01/2014 al 30/09/2016. La prevalenza di EP è stata complessivamente di 100/29543 (0.34%; $IC_{95\%}$ 0.28%-0.41%), e di 69/5598 (1.23%; $IC_{95\%}$ 0.96%-1.56%) nel sottogruppo "*pazienti ospedalizzati*". Nessuno dei pazienti è deceduto per EP.

Optum database. Sono stati selezionati 296045 pazienti di età ≥ 18 anni giunti in PS per sincope dal 01/01/2004 al 31/12/2015. La prevalenza di EP in questi pazienti è stata di 3576/296045 (1.21%; $IC_{95\%}$ 1.17 %-1.25%), mentre nel sottogruppo "*pazienti ospedalizzati*" è stata pari a 2894/108294 (2.67%; $IC_{95\%}$ 2.58%-2.77%). I dati di mortalità non erano, al momento delle analisi, disponibili.

Conclusioni

Pur tenendo in considerazione tutte le limitazioni note, i database amministrativi in ambito sanitario si confermano uno strumento utile nella realizzazione di studi che permettono di produrre evidenza con una buona validità esterna. Tenendo anche conto della sovrastima del fenomeno dovuta ai codici di diagnosi utilizzati, i risultati ottenuti ci permettono di concludere che la prevalenza di EP in pazienti con sincope (sia visti in PS che ospedalizzati) sembra essere estremamente bassa. Eventuali modifiche del percorso diagnostico di questi pazienti non sembrano quindi giustificate.

Bibliografia

[**1**] Blanc JJ, L'her C, Touiza A, et al. Prospective evaluation and outcome of patients admitted for syncope over a 1 year period Eur Heart J 2002;23(10): 815-20.

[**2**] Prandoni P, Lensing AWA, Prins MH, et al. Prevalence of Pulmonary Embolism among Patients Hospitalized for Syncope. N Engl J Med 2016;375(16):1524-31.

[**3**] Ruwald MH, Hansen M, Lamberts M, et al. Accuracy of the ICD-10 discharge diagnosis for syncope. Europace 2013;15(4):595-600.

[**4**] Konstantinides SV, Torbicki A, Agnelli G, et al. 2014 ESC Guidelines on the diagnosis and management of acute pulmonary embolism. Eur Heart J 2014;35(43):3033-80.

USE OF ANTIDEPRESSANT MEDICATION IN PREGNANCY AND ADVERSE NEONATAL OUTCOMES: A POPULATION-BASED INVESTIGATION

Cantarutti Anna[1], Merlino Luca[2], Giaquinto Carlo[3], Corrao Giovanni[1]

1 *Department of Statistics and Quantitative Methods, Division of Biostatistics, Epidemiology and Public Health, Laboratory of Healthcare Research and Pharmacoepidemiology, University of Milano-Bicocca, Milan, Italy*
2 *Operative Unit of Territorial Health Services, Region of Lombardy, Milan, Italy*
3 *Department of Women's and Children's Health, University of Padova, Padova, Italy*

Introduction

Untreated depression and antidepressant use during pregnancy may have negative consequences for births. On the basis of previous studies, the incidence of women experiencing depression during pregnancy ranges from 15% to 20%, and those treated with an antidepressant medication ranges from 5% to 13% [***1-3***]. Despite the widespread use of antidepressants during pregnancy, their effectiveness in pregnancy remains unknown due to inconsistent and sometimes conflicting findings, making the clinical management of depression a significant challenge. To date, it still remains unclear whether there is a causal relationship between exposure to antidepressant medication in utero and poor neonatal adaptation, or whether maternal depression and its associated behavior is itself responsible for this negative neonatal outcome. Several studies have investigated the relationship between the use of antidepressants during pregnancy and the risk of several adverse neonatal outcome. However, they were unable to control for confounding factors that might alter the relationship between exposure to antidepressants during pregnancy and the development of neonatal outcomes such as mother's socioeconomic status, mother's educational attainment, mental and psychiatric comorbidities, sub- stance dependence, and the use of other medication.
To assess the potential association between the use of antidepressants during pregnancy and several neonatal outcomes. Several confounders such as depression, other maternal psychiatric disorders, and mother's sociodemographic profile were take into account. The potential relation between the specific classes of antidepressant together with their timing of exposure were also investigated.

Methods

A population-based cohort study including 9825 deliveries exposed to an antidepressant between 9 months before last menstrual date through to delivery, from January 2005 to December 2010 in the Lombardy region of Italy, was conducted. We evaluated the prevalence ratio (PR) of low Apgar score, small for gestational age, intrauterine hypoxia and birth asphyxia, neonatal convulsion, and other respiratory conditions, among infants born to mothers exposed to an antidepressant during pregnancy using log-binomial regression analysis.
Results are presented according to 2 levels: an unadjusted analysis and an analysis performed with the use of propensity score stratification to further control for all predefined covariates (i.e., sociodemographic characteristics, maternal covariates, and concomitant medication) categorized as stated in ***Table 1*** (PSS[1]). Propensity score stratification was performed in 4 stages. First, propensity scores (PS) were derived from the predicted probability of treatment estimated in a logistic-regression model that contained all the covariates described above. Second, observations in non-overlapping areas of the PS were omitted. Third, after ranking only the exposed patients based on the PS, we created 10 equally sized PS strata, assigning unexposed patients to these strata based on their PS. Fourth, weighted regression models were used to derive an adjusted exposure effect after stratification, in which each exposed patient received a weight of 1 and unexposed patients were weighted in proportion to the distribution of the exposed in the stratum into

which they fell.
Furthermore, we performed specified exposure subgroup and sensitivity analysis to evaluate the robustness of the main analysis. Sensitivity analysis restricted to women (a) filling 2 or more prescriptions of antidepressants during pregnancy, or (b) having days of supply that overlap with the exposure window of interest, was conducted to evaluate the effect of potential misclassification of exposure. To study timing of exposure in relation to specific outcomes, we redefined exposure as (a) having at least one prescription of an antidepressant during the first trimester of pregnancy, with or without exposure before pregnancy but not during the second and third trimester, vs women exposed only before pregnancy; (b) having at least one prescription of an antidepressant during the second trimester of pregnancy, with or without exposure before the second trimester but not during the third trimester, vs women without this exposure; and (c) having at least one prescription of an antidepressant during third trimester of pregnancy, with or without exposure before the third trimester, vs women without such exposure. Moreover, since caesarean delivery, as well as premature birth, may be involved in the stage from exposure to the selected outcomes, we conducted subgroup analyses restricted to vaginal deliveries and full-term births (weeks' gestation ≥ 37). Finally, we performed specified subgroup analysis for type of antidepressant (i.e., SSRIs, tricyclic antidepressant, serotonin–norepinephrine reuptake inhibitor, and other), and in the analysis of SSRIs, we stratified the analysis according to the use of just that drug class (monotherapy with SSRIs).
The robustness of estimates regarding potential bias introduced by unmeasured confounders was investigated by using the rule-out approach described by Schneeweiss [***4***]. Briefly, this approach involves detecting the extension of the overall confounding required to fully account for the exposure-outcome association, thus moving the observed point estimate to null. We set the possible generic unmeasured confounder (e.g., smoking during pregnancy): (a) to have a 10% prevalence of exposure among pregnant women, (b) to increase the neonatal outcome onset up to 10-fold more in mothers exposed during pregnancy than in those exposed only before pregnancy to the confounder, and (c) to be up to 10-fold more common among women exposed during pregnancy than in those exposed only before pregnancy.

Results
We identified 354 735 eligible singleton births. Of these, mothers who had never been exposed to an antidepressant were not included in the analysis, considering just the 9825 women exposed during the observational period (from 9 months before LMP through to delivery). During pregnancy, 3283 women (33.4%) used at least one antidepressant: 2664 (27%) were exposed to an SSRI, 273 (2.8%) to a tricyclic antidepressant, 403 (4.1%) to a serotonin–norepinephrine reuptake inhibitor, and 110 (1.1%) to another antidepressant. Psychiatric comorbidities (particularly for depression, bipolar disorder, personality disorders, sleep disorder and/or anxiety, and other psychiatric disorders), diabetes, and benzodiazepines concomitant medication were more likely to be diagnosed in women exposed during pregnancy than in women exposed only before pregnancy. Sociodemographic characteristics were virtually the same between the 2 groups of exposures. Women exposed to an antidepressant during pregnancy were slightly older (mean age 33 ± 5 y), Italian, poorly educated, unmarried, and unemployed.
Infants born to mothers with exposure to antidepressants during pregnancy showed an increase in the prevalence ratio of low Apgar score in the unadjusted analysis (PR = 1.63, 95% CI, 1.01-2.61), of intrauterine asphyxia and birth asphyxia both in unadjusted (PR = 1.42, 95% CI, 1.12-1.81) and in the adjusted analyses (PSS[1] PR = 1.39, 95% CI, 1.08-1.73), of neonatal convulsion (PR = 2.85, 95% CI, 1.08-7.47; PSS[1] PR = 2.81, 95% CI, 1.07-7.36), and of other respiratory conditions (PR = 1.28, 95% CI, 1.04-1.59; PSS[1] PR = 1.24, 95% CI, 1.00-1.52). There was no increase in prevalence ratio of small for gestational age. Restricting the cohort to women exposed to an SSRI during pregnancy markedly increased the unadjusted and adjusted prevalence ratio of low Apgar score (SSRIs PSS[1] PR = 1.69, 95% CI, 1.02-2.79), of intrauterine asphyxia

and birth asphyxia (SSRIs PSS[1] PR = 1.39, 95% CI, 1.07-1.81), and of other respiratory conditions (SSRIs PSS[1] PR = 1.37, 95% CI, 1.08-1.74) (**Figure 1**). Results were similar in the sensitivity and subgroup analyses. Interesting, it was reported a clear association in an increased prevalence ratio for neonatal convulsion in infants born to mothers with exposure to antidepressants during the first trimester of pregnancy. Whereas, a low Apgar score, intrauterine asphyxia and birth asphyxia, and the other respiratory conditions seem to be associated with exposure during the third trimester of pregnancy.
The effect of a generic unmeasured confounder that might overinflate the observed harmful effect of antidepressants use during pregnancy was investigated by using the rule-out approach. We set smoking during pregnancy as the possible unmeasured confounder with a 10% prevalence of exposure among pregnant women. For example, we consider that antidepressant users during pregnancy had 3-fold higher smokers' prevalence than antidepressants users only before pregnancy (exposure-confounder odds ratio = 3). In these conditions, pregnancy smoking should increase the risk of low Apgar score by 7-fold (confounder-outcome relative risk = 7), of intrauterine hypoxia and birth asphyxia by 4-fold, and of other respiratory condition by 3-fold to nullify the observed harmful effect of antidepressants use during pregnancy. On the other hand, admitting that smoking during pregnancy increases the risk of neonatal convulsions by 5-fold, prevalence of smokers among antidepressants users during pregnancy should be 9-fold higher than antidepressant users only before pregnancy to nullify the observed effect.

Conclusions

Our data on drug utilization patterns in the real-world setting offer evidence that exposure to an antidepressant during pregnancy increases the prevalence ratio of low Apgar score, intrauterine hypoxia and birth asphyxia, neonatal convulsion, and other respiratory conditions. Importantly, the increased prevalence ratio of neonatal convulsion was due to exposure to an antidepressant during the first trimester of pregnancy, compared to antidepressant use only before pregnancy, while the increased prevalence ratio of low Apgar score, of intrauterine hypoxia and birth asphyxia, and of other respiratory condition was as a result of exposure during the third trimester, compared to women without this exposure. These effects seem to be attributed to the treatment rather than to the disease itself. It remains to be determined whether maternal antidepressant medication use is more beneficial or has adverse effects beyond the underlying depression. In the meantime, the clinician and the woman herself need to balance the degree of severity of the depressive disorder and the risk of relapse, with the emerging safety profile of antidepressant drugs.

References

[**1**] Ververs T, Kaasenbrood H, Visser G, Schobben F, de Jong-van den Berg L, Egberts T. Prevalence and patterns of antidepressant drug use during pregnancy. Eur J Clin Pharmacol. 2006;62(10):863-70.

[**2**] Cooper WO, Willy ME, Pont SJ, Ray WA. Increasing use of antide- pressants in pregnancy. Am J Obstet Gynecol. 2007;196(6):544.e1- 544.e5.

[**3**] Oberlander TF, Warburton W, Misri S, Aghajanian J, Hertzman C. Neonatal outcomes after prenatal exposure to selective serotonin reuptake inhibitor antidepressants and maternal depression using population-based linked health data. Arch Gen Psychiatry. 2006;63 (8):898-906.

[**4**] Schneeweiss S. Sensitivity analysis and external adjustment for unmea- sured confounders in epidemiologic database studies of therapeutics. Pharmacoepidemiol Drug Saf. 2006;15(5):291-303.

Table 1. Selected Cohort Characteristics of Women among the Exposure Groups. Region of Lombardy, 2005-2010

	Unadjusted				Adjusted†				
	Exposed during pregnancy		Exposed before pregnancy only		Exposed during pregnancy		Exposed before pregnancy only		*Standardized difference in %*
Characteristics	**(N=3283)**		**(N=6542)**		**(N=3282)**		**(N=6536)**		
Socio-demographic characteristics‡									
Maternal Age - mean (sd), yr	33	(5)	32	(4.8)	33	(5.1)	33	(4.9)	0
Nationality									
Italy - no. (%)	2905	(88.49)	5695	(87.06)	2904	(88.5)	5776	(88.3)	0.5
Other - no. (%)	378	(11.51)	847	(12.94)	378	(11.5)	763	(11.7)	-0.5
Marital status									
Married - no. (%)	2345	(71.44)	4741	(72.47)	2345	(71.5)	4665	(71.3)	0.3
Unmarried - no. (%)	938	(28.56)	1801	(27.53)	937	(28.5)	1874	(28.7)	-0.3
Employment									
Employed - no. (%)	2376	(72.36)	4861	(74.3)	2376	(72.4)	4745	(72.6)	-0.4
Unemployed - no. (%)	907	(27.64)	1681	(25.7)	91	(27.6)	1794	(27.4)	0.4
Educational attainment				()					
Low - no. (%)	1149	(35)	2229	(34.07)	1148	(53.5)	3496	(53.5)	0
Intermediate - no. (%)	1527	(46.52)	3092	(47.27)	1527	(46.5)	3043	(46.5)	0
High - no. (%)	607	(18.48)	1220	(18.65)	607	(18.5)	1225	(18.7)	0.6
Parity									
Nulliparous- no. (%)	1731	(52.74)	3429	(52.42)	1730	(52.7)	3460	(52.9)	-0.4
Multiparous - no. (%)	1552	(47.26)	3113	(47.58)	1552	(47.3)	3079	(47.1)	0.4
Maternal covariates∫									
Previous miscarriages - no. (%)‡	551	(16.79)	1155	(17.65)	551	(16.8)	1091	(16.7)	0.3
Depression - no. (%)	245	(7.46)	263	(4.02)	244	(7.4)	463	(7.1)	1.4
Epilepsy - no. (%)	30	(0.91)	42	(0.64)	29	(0.9)	52	(0.8)	1
Hypertension - no. (%)	158	(4.81)	279	(4.26)	158	(4.8)	313	(4.8)	0.1
Preeclampsia - no. (%)	89	(2.71)	163	(2.49)	89	(2.7)	181	(2.8)	-0.4
Diabetes - no. (%)	134	(4.08)	232	(3.55)	133	(4.1)	273	(4.2)	-0.6
Dyslipidemia - no. (%)	16	(0.49)	26	(0.4)	16	(0.5)	32	(0.5)	-0.1
Obesity or overweight - no. (%)	77	(2.35)	128	(1.96)	77	(2.3)	149	(2.3)	0.4
Migraine/ headache - no. (%)	88	(2.68)	182	(2.78)	88	(2.7)	179	(2.7)	-0.4
Bipolar disorder - no. (%)	51	(1.55)	49	(0.75)	51	(1.6)	85	(1.3)	2.1
Personality disorder - no. (%)	147	(4.48)	142	(2.17)	146	(4.4)	253	(3.9)	2.9
Adjustment disorder - no. (%)	14	(0.43)	17	(0.26)	14	(0.4)	27	(0.4)	0.2
Neuropathic, Non-neuropathic, and Other Pain - no. (%)	300	(9.14)	609	(9.31)	300	(9.1)	583	(8.9)	0.8
Other Psychiatric disorders - no. (%)	147	(4.48)	169	(2.58)	147	(4.5)	263	(4)	2.3
Psychosis or Schizophrenia - no. (%)	42	(1.28)	51	(0.78)	42	(1.3)	78	(1.2)	0.8
Sleep disorder or Anxiety - no. (%)	180	(5.48)	172	(2.63)	179	(5.5)	316	(4.8)	2.8
Substance dependence - no. (%)	92	(2.8)	104	(1.59)	91	(2.8)	156	(2.4)	2.4
C - section - no. (%)∫	1219	(37.13)	2300	(35.16)	1219	(37.1)	2408	(36.8)	0.7
Preterm Birth - no. (%)∫	230	(7.01)	397	(6.07)	230	(7)	457	(7)	0.1
Concomitant medication∏									
Benzos - no. (%)	45	(1.37)	44	(0.67)	45	(1.4)	71	(1.1)	2.6
Triptans - no. (%)	292	(8.89)	731	(11.17)	292	(8.9)	610	(9.3)	-1.5
NSAIDs - no. (%)	1160	(35.33)	2343	(35.81)	1159	(35.3)	2311	(35.3)	-0.1

†To account for propensity score, the exposed before pregnancy only observations were weighted using the distribution of the exposed during pregnancy among propensity score strata.
‡Data related to the current pregnancy
∫Maternal covariates measured from any time before LMP through the end of the first trimester
∏Concomitant medication use measured during ant time pre-LMP

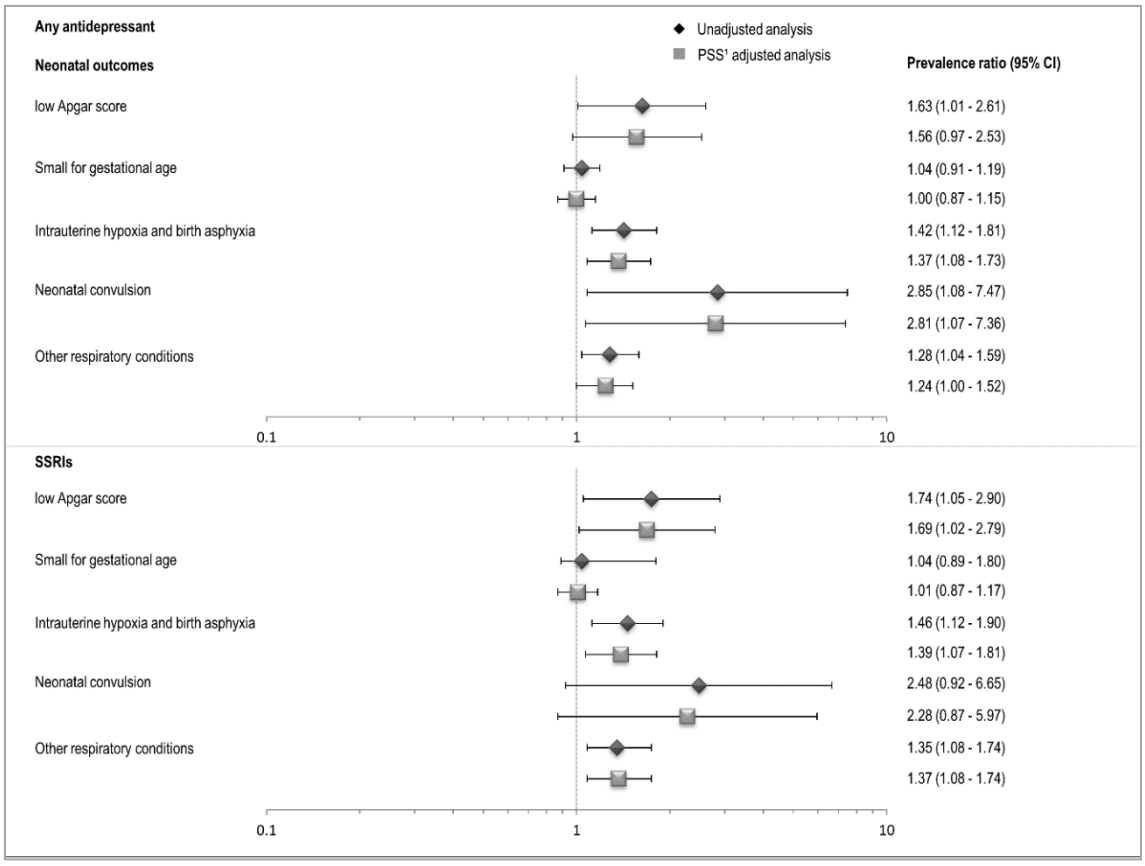

Figure 1. Prevalence ratios (and 95% confidence intervals) of selected neonatal outcomes according to maternal exposure to antidepressants during pregnancy, with respect to use only before pregnancy.
CI, confidence interval; PSS[1], propensity score stratification; SSRI, selective serotonin reuptake inhibitor.
Prevalence ratio, and 95% confidence interval, estimated with log-binomial regression. Estimates are adjusted for the covariates listed in Table 1.

SVILUPPO E VALIDAZIONE DEL FRAILTY PROGNOSTIC SCORE. UN INDICE RICAVATO DAI DATABASE AMMINISTRATIVI CAPACE DI PREDIRE ESITI LEGATI ALLA COMPLESSITÀ CLINICA

Rea Federico[1,2], Di Martino Mirko[3], De Palma Rossana[4], Scondotto Salvatore[2,5], Fusco Danilo[3], Lallo Adele[3], Belotti Laura[4], Ferrante Mauro[6], Pollina Addario Sebastiano[5], Merlino Luca[2,7], Mancia Giuseppe[8], Carle Flavia[2,9], Corrao Giovanni[1,2]

1. *Dipartimento di Statistica e Metodi Quantitativi, Università di Milano-Bicocca, Milano, Italia;*
2. *Centro Interuniversitario Healthcare Research & Pharmacoepidemiology;*
3. *Dipartimento di Epidemiologia, Regione Lazio, Roma, Italia;*
4. *Direzione Generale Cura della Persona, Salute e Welfare, Regione Emilia-Romagna, Bologna, Italia;*
5. *Osservatorio epidemiologico, Regione Sicilia, Palermo, Italia;*
6. *Dipartimento di Culture e Società, Università di Palermo, Palermo, Italia;*
7. *Osservatorio epidemiologico, Regione Lombardia, Milano, Italia;*
8. *Università di Milano-Bicocca, Milano, Italia;*
9. *Centro interdipartimentale di Epidemiologia Biostatistica e Informatica medica, Università Politecnica delle Marche, Ancona, Italia*

Introduzione

Riuscire a quantificare la complessità clinica degli individui è di fondamentale importanza perché (i) permette al decisore di politica sanitaria di stratificare la popolazione in accordo alla stato di fragilità, e (ii) consente un miglior controllo del confondimento negli studi epidemiologici.
L'obiettivo del lavoro consiste nel sviluppare e validare un indice basato sugli archivi sanitari presenti nelle diverse Regioni, capace di predire la mortalità per tutte le cause, ed altri esiti legati alla complessità clinica (numero di ricoveri e costi ospedalieri), tra i beneficiari del Servizio Sanitario Nazionale.

Metodi

Partendo dalle condizioni incluse nel Charlson Comorbidity Index (CCI) [***1***], Elixhauser Index (EI) [***2***] e Chronic Disease Score (CDS) [***3***], abbiamo stilato una lista di 46 condizioni. Queste condizioni erano tracciabili dalle diagnosi riportate nelle schede di dimissione ospedaliera e dai farmaci rimborsabili dal Servizio Sanitario.
Un training set di 500,000 residenti in Lombardia è stato causalmente selezionato tra coloro che nel 2008 avevano almeno 50 anni d'età. La relazione tra le 46 patologie ed il tempo di sopravvivenza a un anno è stato investigato per mezzo di un modello di sopravvivenza parametrico basato sulla distribuzione di Weibull. Il metodo LASSO è stato applicato per selezionare le condizioni abili a predire l'esito in studio. Per assegnare un punteggio ad ogni condizione, i coefficienti stimati dal modello sono stati moltiplicati per dieci e arrotondati all'intero più vicino. Infine, i punteggi sono stati sommati per ottenere una misura aggregata; questa è stata poi categorizzata secondo la suddivisione 0-4, 5-9, 10-14, 15-19 e ≥ 20 assegnando i valori 0, 1, 2, 3 e 4. L'indice così ottenuto è stato definito Frailty Prognostic Score (FPS).
Le performance del FPS sono state investigate rispetto altri indici prognostici presi in considerazione (CCI, EI e CDS) in un validation set interno di altri 500,000 lombardi e tre validation set esterni di 500,000 individui residenti in Emilia-Romagna (anno di reclutamento: 2010), Lazio (2010) e Sicilia (2013). La potenza discriminatoria è stata valutata calcolando l'area sotto la curva (Area Under the Curve, AUC), mentre il Net Reclassification Improvement (NRI) [***4***] è stato determinato per quantificare il miglioramento nella classificazione del rischio.

Analisi secondarie hanno indagato se il FPS sia in grado di predire altri esiti di interesse, ovvero: mortalità per tutte le cause a cinque anni, tasso di ospedalizzazione per tutte le cause a uno e cinque anni, e i costi ospedalieri a due anni.

Risultati

Il FPS ha una migliore capacità discriminatoria rispetto a CCI, EI e CDS: i valori dell'AUC (e gli intervalli di confidenza al 95%) ottenuti dai quattro indici sono rispettivamente 0.78 (0.77, 0.79), 0.69 (0.68, 0.70), 0.65 (0.64, 0.66) e 0.69 (0.68, 0.70). Basandoci sul calcolo del NRI, l'indice ha mostrato un significativo miglioramento nella classificazione rispetto al CCI (39%), EI (69%) e CDS (27%).
Inoltre, è stata osservata una straordinaria consistenza nella capacità discriminatoria del FPS nelle diverse Regioni, riportando valori dell'AUC praticamente identici: 0.78 (0.77, 0.79), 0.78 (0.77, 0.79), 0.77 (0.76, 0.78), e 0.78 (0.77, 0.79) in Lombardia, Emilia-Romagna, Lazio e Sicilia. Le curve di sopravvivenza Kaplan-Meier di ogni Regione, mostrando una progressiva riduzione della sopravvivenza all'incremento del FPS, confermano l'abilità del FPS di predire il tempo alla morte.
L'indice ha dimostrato inoltre buone capacità di previsione anche per tutti gli altri esiti indagati. Infatti, gli individui con il valore più elevato (FPS=4) hanno riportato un tasso 9, 8, 6 e 8 volte maggiore rispetto a coloro con il punteggio minore (FPS=0) per gli esiti indagati: mortalità a cinque anni, tassi di ricoveri a uno e cinque anni, e costi ospedalieri a due anni.

Conclusioni

Le caratteristiche della costruzione del FPS che lo differenziano rispetto ad altri indici presenti in letteratura sono principalmente due: (i) è basata su ampio campione di popolazione non selezionata; (ii) parte da dati già fruibili da ogni Regione. Il FPS può quindi rappresentare sia un utile strumento per la pianificazione di politiche sanitarie, identificando i pazienti più fragili che richiedono maggiore assistenza, che un criterio di aggiustamento in studi epidemiologici.

Bibliografia

[1] Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40(5):373-83.
[2] Elixhauser A, Steiner C, Harris DR, et al., Comorbidity measures for use with administrative data. Med Care 1998;36(1):8-27.
[3] Von Korff M, Wagner EH, Saunders K. A chronic disease score from automated pharmacy data. J Clin Epidemiol 1992;45(2):197-203.
[4] Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr. et al. Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond. Stat Med 2008 Jan 30;27(2):157-72.

LE LINEE DI RICERCA STRATEGICHE DI UN'AZIENDA OSPEDALIERO-UNIVERSITARIA: PROPOSTA DI UN METODO QUALI-QUANTITATIVO CONDIVISO TRA CLINICI, RICERCATORI E DIREZIONE.

Iezzi Elisa[1], Ardissino Diego[2], Ferrari Carlo[3], Vitale Marco[4], Caminiti Caterina[1]

1 *UO Ricerca e Innovazione – Azienda Ospedaliero-Universitaria di Parma*
2 *UO Cardiologia – Azienda Ospedaliero-Universitaria di Parma*
3 *UO Malattie Infettive ed Epatologia – Università degli studi di Parma*
4 *Dipartimento di Medicina e Chirurgia – Università degli studi di Parma*

Introduzione

Nelle aziende sanitarie pubbliche e private si evince sempre più la necessità di identificare le principali linee di ricerca strategiche per poter favorire il reperimento e conseguente utilizzo delle risorse economiche, risorse strumentali/core–facilities, ma soprattutto per valorizzare i professionisti che operano al loro interno e la loro expertise. Per fare questo è necessario adottare ed utilizzare un metodo quali-quantitativo che utilizzi sia un sistema di indicatori di performance [***1***] in campo bibliometrico e di attrattività economica, che l'assegnazione di punteggi [***2***] da un Gruppo di Lavoro di esperti multidisciplinari clinici, universitari e metodologi.
Individuare con modalità oggettive e condivise le linee di ricerca strategiche.

Metodi

1) E' stato richiesto a tutti i professionisti dell'azienda di completare, via web, una scheda per proporre una propria proposta progettuale che contenesse più attività progettuali indicando all'interno, la descrizione delle attività, gli obiettivi 2016-2018, le pubblicazioni attinenti, i team leader ed il gruppo di ricerca, i laboratori e gli strumenti principali/core facilities utilizzati.
2) Dalle schede sono stati estratti i team leader per i quali sono stati misurati degli indicatori di performance comprendenti sia quelli bibliometrici che quelli di capacità attrattiva di risorse finanziare, facilmente reperibili.
3) E' stato costituito un Gruppo di Lavoro formato da 5 componenti (clinici, universitari e metodologi) allo scopo di assegnare i punteggi ai valori degli indicatori. Per ogni team leader è stato cosi calcolato un punteggio complessivo ed effettuata una graduatoria dalla quale sono stati individuati i team leader con maggiore punteggio.
4) L'analisi ha consentito al Gruppo di Lavoro di individuare le linee di ricerca, in ognuno delle quali confluisce almeno un team leader con elevato punteggio, e fatto confluire i restanti team leader in una delle linee individuate, in base al loro expertise.
5) Le linee di ricerca sono state presentate al Board per la Ricerca che ha confermato la valenza strategica e ritenuto rigoroso il metodo utilizzato.

Risultati

La realizzazione del Piano della Ricerca, anni 2016-2018, ha coinvolto tutti i professionisti ed ha richiesto sei mesi per un totale di 63 schede. Il reperimento delle informazioni da ISI e dall'anagrafe regionale della ricerca dell'Emilia Romagna (ARER) ha richiesto circa 2 settimane. Il Gruppo di Lavoro che si è incontrato 5 volte in un arco di tempo di 4 mesi, ha assegnato i punteggi ed in base all'analisi ha individuato 5 linee strategiche ognuna con un coordinatore ed un gruppo di team leader (range del numero di componenti per le aree 7 – 14). Il punteggio totale di carriera/bibliometrico è 337 e quello relativo alla capacità attrattiva è di

204 per un totale di 541. Il punteggio medio per ogni team leader è risultato pari a 9 (range 6-12) (***Tabella 1***).

Tabella 1. Piano della Ricerca per gli anni 2016-2018

Titolo linea strategica	# Team leader	Punteggio totale PUBBLICAZIONI	Punteggio totale GRANT	Punteggio medio per team leader
La personalizzazione del trattamento oncoematologico	10	72	45	*12*
Meccanismi di cronicizzazione delle malattie immunomediate	14	92	57	*11*
Vecchi e nuovi fattori di rischio per le malattie cardiovascolari	7	40	31	*10*
Disordini nutrizionali, malattie metaboliche e cronico degenerative	14	88	35	*9*
Marcatori molecolari diagnostici e predittivi	14	46	37	*6*
totale	**59**	**337**	**204**	***9***

Conclusioni

Si è cercato, attraverso indicatori di performance di sviluppare un metodo semplice e di facile reperibilità per poter individuare in modo oggettivo le linee di ricerca strategiche.
Il medesimo modello potrebbe essere esteso ad altre Aziende Ospedaliero-Universitarie / IRCCS sia in territorio regionale che nazionale per permettere alle direzioni di individuare con criteri oggettivi le linee strategiche della propria realtà.

Bibliografia

[**1**] Wootton R. A simple, generalizable method for measuring individual research productivity and its use in the long-term analysis of departmental performance,including between-country comparisons. Health Res Policy Syst 2013;14:11-2.

[**2**] Caminiti C, Iezzi E, Ghetti C, et al. A method for measuring individual research productivity in hospitals: development and feasibility. BMC Health Serv Res 2015;15:468.

CANCER MORTALITY PREDICTIONS FOR 2017 IN LATIN AMERICA

Carioli Greta[1], La Vecchia Carlo[1], Bertuccio Paola[1], Rodriguez Teresa[2], Levi Fabio[3], Boffetta Paolo[4], Negri Eva[5], Malvezzi Matteo[1]

1. *Department of Clinical Sciences and Community Health, Universitá degli Studi di Milano, Milan, Italy.*
2. *Navarra Health Service, Pamplona, Navarra, Spain.*
3. *Institute of Social and Preventive Medicine (IUMSP), Lausanne University Hospital, Lausanne, Switzerland.*
4. *Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, USA.*
5. *Department of Biomedical and Clinical Sciences, Universitá degli Studi di Milano, Milan, Italy.*

Introduction

Since death certification figures are only available with a few years lag, cancer mortality predictions for the current or future years are essential to evaluate cancer prevention and management strategies, and to plan public health resource allocation, albeit with some inherent prediction uncertainty [***1***].
Over several years, for the USA [***2***] and for the European Union (EU) [***3***] predictions for cancer mortality have been published and resulted reasonably valid [***2, 4***].
We predicted the number of deaths and mortality rates for all cancers and selected major cancer sites for 2017 in seven selected Latin American countries [***5***]. Moreover, in order to select the most appropriate prediction method for these data, we compared the method used in previous publications with others, including a "hybrid" regression (a mixture of linear, log-linear, power five, and square root regression).

Materials and Methods

We retrieved official death certification data from the WHO database (WHOSIS) [***6***] for cancer of the stomach, colorectum, pancreas, lung, breast, uterus (cervix and corpus), prostate, leukaemias and total neoplasms (malignant and benign). We obtained data for the seven Latin American countries with over 85% death certification coverage and over 10 million inhabitants (Argentina, Brazil, Chile, Colombia, Cuba, Mexico, and Venezuela) [***7***], for the 1980-2014 calendar period (for Venezuela and Colombia up to 2013). We obtained resident population estimates, based on official censuses from the Pan American Health Organization (PAHO) database [***8***].
Using the matrices of certified deaths and resident population, we calculated age-specific death rates for each 5-year age-group (from 0-4 to 80+ years), sex and calendar year. We computed age-standardized rates per 100,000 person-years at all ages, with the direct method based on the world standard population. For lung cancer, we also calculated rates for the 25-44, 45-64, 65-74 and over 75 years age groups.
We based mortality data projections on an age-period joinpoint model [***9***]. A Poisson count data joinpoint regression model, allowing for up to five joinpoints, is fit to the logarithm of the number of age-specific certified deaths for each 5-year age-group to identify the most recent trend segments. We applied a linear regression model to the mortality data for each age-group over the period identified by the last segment of the joinpoint model in order to estimate the regression coefficients. We used this model to calculate the number of expected age-specific deaths for 2017 and the corresponding 95% prediction intervals (PIs). These are calculated with a standard error that takes the variability of the new observation into account [***4, 10, 11***]. Predicted age-standardized mortality rates, with corresponding 95% PIs, are calculated using the number of expected age-specific deaths and the projected population data for the period of interest from the PAHO database.
Before applying the linear regression method, we compared projections resulting from two other regression methods: log-linear and a "hybrid" regression. The latter blends the linear and log-linear regression models

along with power five and square root regressions. With the hybrid regression, the number of expected age-specific deaths for each of the previous models is calculated; we implemented an algorithm that chooses the linear transformation to be used considering the R-squared statistic values for each age-group, sex and cancer site. Thus, the resulting total number of deaths and age-standardized rate were calculated from different underlying linear transformations. In order to evaluate the results of these different methods, we used observed EU data from 1980 to 2001 as a training dataset, to which we applied the projection methods in order to predict data for the 2002-2011 period. We used observed data from 2002 to 2011 as a validating dataset. To measure the accuracy of the predicted figures and to compare the performance of the different projection methods we computed the average absolute relative deviation (AARD), i.e. the ratio between the prediction error and the observed rate. In general a prediction is considered reliable when the AARD value is less than 5% [***12***].
Moreover, for the seven selected Latin American countries, we estimated the numbers of avoided cancer deaths over the 1990-2017 period by comparing observed and expected deaths on the basis of 1990 age-specific rates.

Results
Overall, the AARDs from the three projection methods analyzed were quite similar, greater differences were only observed in few cases. For example, for stomach cancer, in men the AARD was 0.006 for the hybrid model, while for the linear and for the log-linear, it was respectively 0.032 and 0.014. Below the more noteworthy results: for stomach cancer, the predicted trend from the hybrid regression overlapped the observed rates perfectly. In the hybrid model, the log-linear projections seemed to be predominant. The log-linear figure in men was similar to the hybrid one, but the predicted trend overlapped the observed rates less precisely. Instead, the predictions produced by the linear regression model were less accurate. In women, the predicted trend of the hybrid model (AARD of 0.018) was similar to the linear one (AARD of 0.029), the worse projections were those of the log-linear model with wide confidence interval (AARD of 0.051). The prediction intervals tended to contain the true rates (except for the linear model in males, for which the PI was at the limit). In the case of colorectal tumour, the best predictive method in men was the log-linear model (AARD of 0.007), the predicted trend followed the observed rates closely. Similarly, the hybrid method worked well (AARD of 0.008). The linear model was less accurate (AARD of 0.02). In women the log-linear model had an AARD of 0.036; the hybrid method (AARD of 0.042) seemed influenced by the linear regression (AARD of 0.057), the worse method. Further, in women, the prediction intervals did not include the observed rates. For lung cancer, the more performing model was not the same between the sexes. In men the hybrid method provided an accurate predicted trend (AARD of 0.003), which overlapped the mortality rates perfectly and the prediction interval was narrow. However, the other two methods worked well, too (AARD of 0.015 for linear regression and of 0.007 for log-linear model). In women the best method was the log-linear (AARD of 0.009); but the prediction interval was wider than the other two methods' PIs. The hybrid method (AARD of 0.04) appeared influenced from the linear trend (AARD of 0.052), the worse one. Due to the lack of any real significant differences, we chose to continue using the linear method similarly to previous studies.
In all selected Latin American countries, total cancer mortality is predicted to decline. In men, the highest predicted rate is 132.3/100,000 for Cuba, with a percent difference of -5.4% since 2012. Argentina and Chile follow Cuba with a predicted rate of 119.8 (7% fall) and 111.2/100,000 men (6% fall), respectively. The lowest predicted total cancer rate is 64.7/100,000 men for Mexico with the largest fall in rates (-9.5%). In women, Cuba has the highest predicted rate 93.3/100,000 (-3.2%); Mexico has the lowest one, 60.6/100.000, with a decline of 8%. The greatest falls are predicted for Chile and Colombia. Women in Argentina show predicted rates for 2017 around 89/100,000, similar to those of 2012. Although favourable trends in rates, the number of deaths is predicted to increase in all countries and both sexes.

Total cancer mortality trends in men started to decline between 1990 and 2000, except for Cuba and Brazil. Overall, in women trends declined over the whole period considered, except for Cuba and Brazil, which showed unfavourable trends until the early-mid 2000's.
Regarding specific cancer sites (***Figure 1***), stomach cancer mortality has long been declining in both sexes; in contrast, colorectal cancer has been rising in most countries, with a tendency to level off over recent years. Pancreatic cancer rates are inconsistent. For most countries in men, lung cancer trends were moderately downwards in recent calendar periods. In women, lung cancer trends have been rising, except in Mexico. Breast cancer mortality was relatively low in most countries, except Argentina, and tended to decline over recent years. Conversely, cancer of the uterus mortality rates remain high in all Latin America, despite long term falls, particularly in Venezuela and Cuba, whose predicted rates remain around 10/100,000. Prostate cancer rates were particularly high in Cuba and Venezuela, but tended to decline moderately over the most recent years. Leukaemia mortality shows some falls over recent years.
Considering standardized mortality rates for lung cancer stratified by the age groups 25-44, 45-64, 65-74 and 75+ years, the patterns for men are favourable at all age groups and in all countries, generally larger in the young. Lung cancer rates are predicted to fall in women as well, albeit less than in men.
We estimated the number of avoided cancer deaths in men and women between 1990 and 2017, assuming the age-specific peak mortality rate in 1990 as constant. Over the 27-year period considered, a substantial amount of cancer deaths were avoided in Argentina (132,000 deaths, 88,000 in men and 44,000 in women), Chile (63,000 deaths, 16,000 in men and 47,000 in women), Colombia (83,000 deaths, 31,000 in men and 52,000 in women), Mexico (118,000 deaths, 39,000 in men and 79,000 in women) and Venezuela only for women (26,000 deaths). No appreciable reduction in cancer deaths was observed in Brazil, Cuba and Venezuelan men.

Conclusions
Regards the comparison of prediction methods, we found no measurable difference between the models under study. Most of AARDs were below the 5% threshold, thus the predictive trends for all the methods were quite satisfactory with negligible differences. However, the hybrid model projections, for any combination of cancer site and sex, were never the worst. Rather, it appeared as a compromise of the linear transformations considered, and in particular, it seemed sometimes greatly influenced by the linear or the log-linear method. Thus, when there are no strong cases for one model or another, using the hybrid model could be a good option, even if not the best. Since this hybrid method is still being refined and the results were only marginally different, we continued to use the linear regression for projection estimates also for comparability reasons.
Though there was appreciable variability across the seven Latin American countries considered, rates for all cancers and for most major cancer sites are predicted to decline to 2017.
Total cancer mortality rates in Argentina, Cuba and Chile were similar to those registered in Europe, North America and (for men) in Japan [***13***], but they were appreciably lower for both sexes in other Latin American countries considered. This reflects the historical low lung cancer (and probably other tobacco related cancers like pancreatic one) rates in these countries, due to less frequent cigarette use [***14***]. Stomach cancer predicted rates remain high in Latin America (particularly in Chile) [***13***]. This likely reflects the high prevalence of *Helicobacter Pylori* (HP) infection [***15***]. Colorectal cancer rates are much lower than in other areas of the world [***13***]; this is probably due to favourable aspects of local diet and physical activity [***16***], despite a high prevalence of overweight and obesity [***17***]. Breast cancer predicted rates decline, reflecting, as well as for the favourable predictions in prostate cancer and leukaemia rates, improved management and diagnosis [***4, 18, 19***]. Conversely, the rates for cancer of the uterus (cervix) are high, pointing the importance of prevention for this neoplasms [***20***].

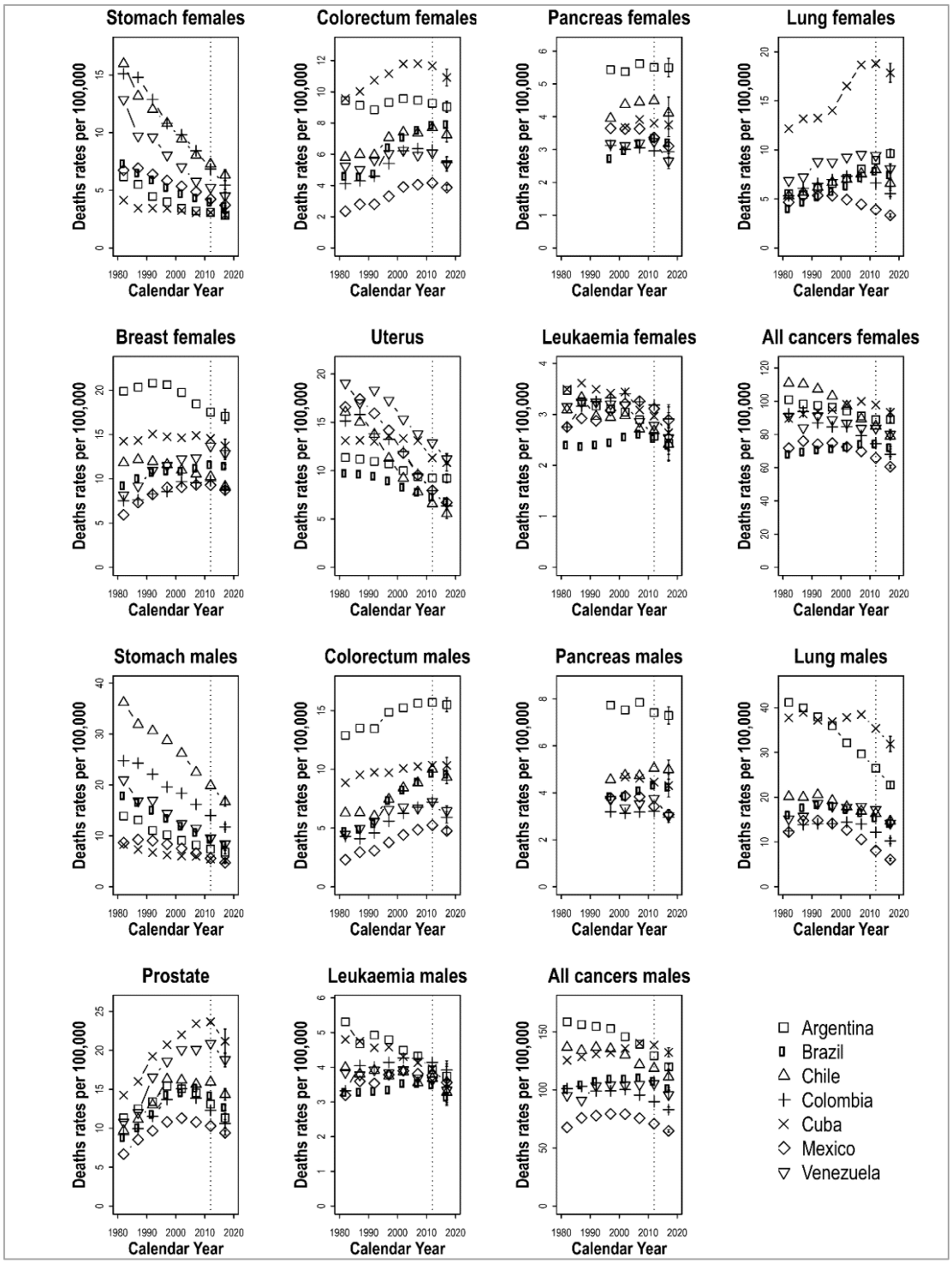

Figure 1. Age-standardized (world population) cancer mortality rate trends for men and women in quinquennia from 1980-84 to 2010-14 and predicted rates for 2017 with 95% PIs for stomach, colorectum, pancreas, lung, breast, uterus, prostate, leukaemias, and all cancers in the 7 selected Latin American countries.

The total number of avoided deaths in the selected Latin American countries was over 420,000, except Brazil, Cuba and Venezuelan men. This suggests the urgency of improving cancer prevention and management in these countries.
The number of avoided deaths may be underestimate, assuming that cancer deaths certification accuracy has improved over the last three decades. However, we considered only major cancer sites that are relatively easy to diagnose (except pancreas) and hence certify, and we included only countries with acceptable indicators of deaths certification validity in the WHO database.
An inherent limitation of predictions is their inability to model sudden changes or fluctuation in slope. Hence, caution in interpretation should be used and prediction estimates should be considered only as general indications for epidemiology and health planning.
These limitations taken into account, we predicted declines, though modest, in cancer mortality to 2017 in Latin America, with the exception of Cuba.

References

[1] Qiu Z, Hatcher J, Wang M. Review ancer projection methods for canadian partnership against cancer analytic network. Alberta Health Services for the Canadian Partnership Against Cancer.
[2] Siegel RL, Miller KD, Jemal A. Cancer Statistics 2017. CA Cancer J Clin 2017;67:7-30.
[3] Malvezzi M, Carioli G, Bertuccio P, et al. European cancer mortality predictions for the year 2017, with focus on lung cancer. Ann Oncol 2017;28:1117-23.
[4] Malvezzi M, Carioli G, Bertuccio P, et al. European cancer mortality predictions for the year 2016 with focus on leukaemias. Ann Oncol 2016; 27: 725-731 doi 710.1093/annonc/mdw1022.
[5] Carioli G, La Vecchia C, Bertuccio P, et al. Cancer mortality predictions for 2017 in Latin America [In press]. Ann Oncol 2017; doi: 10.1093/annonc/mdx1301.
[6] World Health Organization Statistical Information System. WHO mortality database Available from: http://www.who.int/healthinfo/statistics/mortality_rawdata/en/index.html (Last accessed October 2016).
[7] Mathers CD, Fat DM, Inoue M, et al. Counting the dead and what they died from: an assessment of the global status of cause of death data. Bull World Health Organ 2005;3:171-77.
[8] Pan American Health Organization (PAHO). Health Information Platform for the Americas. Available from: http://www.paho.org/data/index.php/en/indicators/demographics-core/308-poblacion-nac-en.html (Last accessed February 2017).
[9] Kim HJ, Fay MP, Feuer EJ, et al. Permutation tests for joinpoint regression with applications to cancer rates. (Erratum in: Stat Med 2001;20: 655). Stat Med 2000;19:335-51.
[10]. Faraway JJ. Linear Models with R. Texts in statistical science. vol. 63. Boca Raton:Chapman & Hall/CRC. 2005.
[11] Verzani J. Using R for Introductory Statistics. Chapman & Hall,2005.
[12] Lee TC, Dean CB, Semenciw R. Short-term cancer mortality projections: a comparative study of prediction methods. Stat Med 2011;30:3387-402.
[13] Hashim D, Boffetta P, La Vecchia C, et al. The global decrease in cancer mortality: trends and disparities. Ann Oncol 2016;27:926-33.
[14] Shafey O, Dolwick S, Guindon G. Tobacco control country profiles 2003. Atlanta, Georgia: American Cancer Society, World Health Organization, International Union Against Cancer, 2003.
[15] Peleteiro B, La Vecchia C, Lunet N. The role of Helicobacter pylori infection in the web of gastric cancer causation. Eur J Cancer Prev 2012;21:118-25.
[16] Ortiz-Hernandez L, Ramos-Ibanez N. Sociodemographic factors associated with physical activity in Mexican adults. Public Health Nutr 2010;13:1131-38.

[**17**] Garmendia ML, Ruiz P, Uauy R. [Obesity and cancer in Chile: estimation of population attributable fractions]. Rev Med Chil 2013;141:987-94.
[**18**] Amadou A, Torres-Mejia G, Hainaut P, et al. Breast cancer in Latin America: global burden, patterns, and risk factors. Salud Publica Mex 2014 56:547-54.
[**19**] Cuzick J, Thorat MA, Andriole G et al. Prevention and early detection of prostate cancer. Lancet Oncol 2014;15:484-92.
[**20**] Luciani S, Cabanes A, Prieto-Lara E, et al. Cervical and female breast cancers in the Americas: current situation and opportunities for action. Bull World Health Organ 2013;91:640-49.

LA QUALITA' DELL'ASSISTENZA DELLE CURE NEL FINE VITA

D'Arienzo Sara[1], Lastrucci Vieri [2], Collini Francesca[1], Zuppiroli Alfredo[1], Forni Silvia[1], Vannucci Andrea[1]

1. *Agenzia Regionale di Sanità della Toscana*
2. *Università degli studi di Firenze - Dipartimento di medicina sperimentale e clinica*

Introduzione

La qualità dell'assistenza fornita nel fine vita alle persone affette da patologie a prognosi infausta, è una delle grandi sfide di oggi, visto il crescente numero di persone che soffrono di una o più patologie croniche [***1, 2***]. Grazie ai progressi della Medicina e all'aumentata aspettativa di vita, il numero di persone con una o più patologie croniche che potenzialmente può evolvere in una condizione di terminabilità sta costantemente crescendo. Tale fenomeno è inoltre sostenuto anche dal cambiamento demografico, determinato dalle popolose coorti dei cosiddetti baby-boomer che sempre più si avvicinano e raggiungono l'età anziana [***3***].
L'Organizzazione mondiale della Sanità definisce la qualità delle cure come l'insieme di sei dimensioni [***4***]:
1) efficacia: assistenza basata sulle evidenze e che portino a miglioramenti dello stato di salute
2) efficienza: raggiungere risultati in termini di salute con il minor impegno di risorse possibile
3) accessibilità: assicurare cure appropriate a coloro che ne hanno bisogno
4) accettabilità: cure che siano centrate sul paziente e che quindi tengano conto delle preferenze e aspirazioni del paziente
5) equità: assenza di discriminazione
6) Sicurezza: cure che minimizzano il rischio e il danno al paziente.
Declinare questa definizione generale di qualità delle cure nell'assistenza al fine vita significa inevitabilmente confrontarsi con i concetti di proporzionalità e non proporzionalità delle cure erogate, di futilità e accanimento terapeutico, di sospensione delle cure, di rispetto delle preferenze e volontà del paziente e dei modelli e percorsi assistenziali delle Cure palliative.
Da queste peculiarità nasce la necessità di sviluppare strumenti di monitoraggio ad hoc per misurare la qualità delle cure e dell'assistenza erogata nel fine vita. Molti degli indicatori a oggi utilizzati sono stati derivati dal lavoro di ricerca di Earle e collaboratori pubblicato nel 2003 [***5***]. Tale lavoro ha definito un set di indicatori di qualità nell'assistenza nel fine vita nei pazienti oncologici, attraverso un approccio combinato di revisione della letteratura, focus group con pazienti e familiari e la costituzione di un panel di esperti. Proprio utilizzando tali indicatori è possibile notare quanto frequentemente i pazienti con malattie in fase terminale vadano incontro a ospedalizzazioni non programmate, spesso caratterizzate dall'impiego di interventi intensivi e invasivi [***5, 6, 7***]. Considerando che la maggior parte dei pazienti con malattia terminale esprime la preferenza di passare il maggior tempo possibile nella propria casa [***8***], il ricorso a un'alta intensità di cura spesso si configura come intervento clinicamente inappropriato ed eticamente reprensibile.
La maggioranza delle ricerche di valutazione della qualità dell'assistenza nel fine vita riguarda l'ambito oncologico per la maggiore prevedibilità del decorso clinico dei casi con prognosi infausta, per i quali è più facile capire il limite di efficacia delle cure e quando cominciare ad affrontare con chiarezza l'assistenza di fine vita [***9, 10***] Le malattie croniche non tumorali, invece, sono caratterizzate dal lento degrado della funzionalità a cui si associano picchi di riacutizzazioni della malattia, con un andamento che rende difficile individuare il momento della vicinanza alla terminalità, per forti resistenze di natura culturale, sia da parte dei professionisti sanitari sia da parte della famiglia [***11***]. Per queste ragioni l'assistenza nel fine vita per queste malattie si caratterizza per un ricorso minore alle cure palliative e sensibilmente più elevato alle cure intensive rispetto a quelle di natura oncologica [***12, 13***]. Tuttavia, anche in questi soggetti è possibile una pianificazione anticipata delle cure palliative e di fine vita, data l'esistenza di scale e modelli affidabili per

predire una ridotta sopravvivenza per molte malattie croniche non tumorali, come la bronco-pneumopatia cronica ostruttiva (BPCO) e lo scompenso cardiaco [***14-17***]. Non solo, le frequenti riacutizzazioni possono essere utilizzate come occasione per attivare in maniera precoce e tempestiva i corretti percorsi assistenziali [***18***]. Recentemente numerose linee guida e protocolli hanno iniziato a riconoscere il ruolo centrale delle cure palliative nelle malattie croniche non tumorali [***19-21***].
L'obiettivo è quello di sviluppare misure di monitoraggio ad hoc, ovvero strumenti adeguati di governance clinica, per valutare la qualità dell'assistenza nel fine vita, evidenziando, attraverso l'utilizzo di un set di specifici indicatori eventuali differenze nell'erogazione di servizi sanitari in pazienti affetti da tumore e/o patologie croniche.

Metodi
Le fonti informative amministrative della Regione Toscana utilizzate per lo studio sono anagrafe sanitaria, schede di morte informatizzate, schede di dimissione ospedaliera, pronto soccorso, hospice. I flussi sono stati interconnessi attraverso metodi di record linkage deterministico utilizzando come chiave primaria l'identificativo univoco del paziente.
Per selezionare la popolazione dei soggetti deceduti è stato utilizzato un algoritmo che, eseguendo un record linkage tra tutti i flussi amministrativi comprendenti informazioni sulla data di morte (DM), ne deframmenta la rendicontazione, rendendo disponibile uno strumento unico, snello e di facile accesso, in cui siano contenuti contemporaneamente tutti gli individui deceduti e la cui data di decesso, sia essa registrata in uno o più dei flussi distinti utilizzati per la sua creazione, viene specificata in maniera univoca.
La popolazione in studio comprende tutti i residenti in Toscana di età maggiore di 18 anni, deceduti nel periodo tra il 1 gennaio 2015 e il 31 dicembre 2015, con una storia clinica di tumore, malattia cronica od entrambe (tumore + malattia cronica) riscontrata nei 36 mesi precedenti il decesso. Dall'analisi sono stati esclusi i decessi per traumatismi avvenuti in pronto soccorso e ospedale.
Nell'analisi descrittiva della popolazione in studio, suddivisa per le tre patologie, sono state valutate le distribuzioni per sesso, classe di età, titolo di studio e Indice di Charlson che misura il grado di complessità assistenziale. Gli esiti in studio riguardano l'accesso a cure nell'ultimo mese di vita.

Risultati
Nel 2015 in Toscana 18.601 persone sono decedute con una storia clinica di tumore e/o patologie croniche (scompenso e BPCO), pari al 41% di tutti i decessi; di questi il 50% soffriva solo di patologie croniche, il 33% di tumore e il 17% di entrambi. Nel 73% dei casi si tratta di cittadini ultra 75enni, con un rapporto uomo/donna molto vicino a 1.
Da notare che tra i cronici la quota di ultra 75enni sale al 90%. Meno del 20% dei casi ha un titolo di studio alto. Nel complesso la popolazione in studio ha livelli di comorbilità elevati, con una quota non trascurabile di persone con Indice di Charlson superiore a 3. Tra i deceduti con storia di tumore, sola o associata a malattie croniche, le comorbilità sono maggiori che tra i deceduti con patologie croniche. Nell'ultimo mese di vita più di un terzo dei pazienti ha effettuato almeno un accesso al Pronto soccorso, e solo una minima percentuale (circa il 6,4%) degli accessi al Pronto soccorso nell'ultimo mese di vita non hanno dato esito a ricovero. La grande maggioranza dei soggetti (il 75%) ha effettuato almeno un ricovero nel mese precedente il decesso, con maggiore probabilità di ospedalizzazione per i pazienti affetti da malattie croniche non neoplastiche. Circa il 13% dei ricoveri è stato caratterizzato dal passaggio in reparto di Terapia intensiva, con una probabilità più che doppia nei pazienti con malattie croniche rispetto a quelli con patologie neoplastiche. Oltre un quarto dei ricoveri è stato caratterizzato da uno o più interventi di supporto vitale intensivo, senza differenze significative tra i diversi gruppi. Se invece si analizza soltanto il dato relativo alle trasfusioni, si osserva una maggiore probabilità di intervento nei casi con tumore, a conferma di una modalità di cura orientata a lenire le sofferenze per grave anemia; la differenza tra i gruppi di patologie, inoltre, risulta

statisticamente significativa. Il passaggio in Terapia intensiva e gli interventi invasivi si sono verificati più frequentemente, senza differenze di patologie, nel sesso maschile, nelle fasce di età più giovani, nelle persone con un titolo di studio più elevato e con un minor numero di comorbidità. Il decesso è avvenuto in ospedale nel 45,5% dei casi, con differenze sostanziali a seconda della patologia di base: 55% per le malattie croniche, 35% per le neoplasie. L'8% dei soggetti ha fatto ricorso all'Hospice nell'ultimo mese di vita, con percentuali altamente differenti a seconda della condizione clinica: 17% nel caso dei tumori, solo l'1% tra i pazienti con malattia cronica. Da sottolineare come tra i pazienti deceduti che hanno fatto ricorso all'Hospice, ben il 50% lo abbia fatto solo nell'ultima settimana di vita.

Conclusioni

Rispetto ad altre ricerche sulla valutazione della qualità dell'assistenza nel fine vita, questo studio si caratterizza per aver preso in considerazione non solo l'ambito oncologico ma anche quello delle malattie croniche degenerative non neoplastiche, che tra l'altro rappresentano le condizioni di maggior prevalenza tra le cause di morte. Un primo dato da sottolineare è che nel 2015 in Toscana i residenti deceduti per scompenso cardiaco o BPCO sono stati una volta e mezzo più numerosi rispetto a quelli deceduti per tumore, e che ben il 17% dei soggetti soffriva sia di una malattia cronica che di un tumore; questo significa che in futuro le ricerche sul fine vita dovranno tenere sempre più conto della crescente complessità dei pazienti e dell'aumento dell'impatto delle malattie croniche non neoplastiche. Va in questo senso anche il dato dell'età, in quanto i tre quarti dei deceduti ha più di 75 anni; nelle malattie croniche gli ultra 75enni rappresentano addirittura l'89% e gli ultra 85enni il 59%. La realtà dell'assistenza nel fine vita in Toscana nel 2015 è una realtà, come mostrano i risultati, ancora centrata sull'Ospedale e conseguentemente su cure ad alto livello di intensità e invasività. Sappiamo invece che, nei casi di malattie in fase terminale, i desideri dei pazienti e delle loro famiglie vanno nella direzione opposta. I riferimenti normativi in tema di Cure palliative, sia a livello nazionale che regionale, offrono la possibilità concreta di un'inversione di tendenza, che ci viene richiesta non solo dagli standard di appropriatezza clinica, ma anche e soprattutto da esigenze etiche che si muovono sia nella dimensione individuale e privata del rispetto delle persone, che in quella collettiva e pubblica dell'utilizzo delle risorse. Per tutto questo è necessario un grosso sforzo culturale, per formare i professionisti e anche i cittadini a una visione delle cure palliative come risorsa integrativa e non alternativa per i pazienti affetti da patologie croniche, neoplastiche e non, giunte ormai in prossimità della loro fase terminale.

Bibliografia

[1] Sarmento VP, Higginson IJ, Ferreira PL, Gomes B. Past trends and projections of hospital deaths to inform the integration of palliative care in one of the most ageing countries in the world. Palliative Medicine, 2016;30(4):363-73.

[2] Gomes B, Higginson IJ. Where people die (1974-2030): past trends, future projections and implications for care. Palliat Med 2008;22 (1):33-41.

[3] Eurostat EU. EurostatEUROPOP2013-population projection. Eurostat Eu, Luxembourg, 2014 Luxembourg

[4] Bengoa R, Key P, Leatherman S, Massoud R, Saturno P. Quality of Care A process for making strategic choices in health systems. ISBN 92 4 156324 9 (NLM classifi cation: W 84.1) ISBN 978 92 4 156324 6. World Health Organization (WHO), 2006

[5] Earle CC, Park ER, Lai B, Weeks JC, Ayanian JZ, Block S.Identifying potential indicators of the quality of end-oflife cancer care from administrative data. J Clin Oncol 2003;21:1133-38.

[6] Bekelman JE, Halpern SD, Blankart CR, et al, for the International Consortium for End-of-Life Research (ICELR). Comparison of site of death, Health care utilization, and hospital expenditures for patients dying with cancer in 7 developed countries. JAMA 2016;315(3):272-83.

[**7**] Morden NE, Chang CH, Jacobson JO, et al. End-of-life care for Medicare beneficiaries with cancer is highly intensive overall and varies widely. Health Aff 2012;31(4):786-96.

[**8**] Gomes B, Higginson IJ, Calanzani N, et al, on behalf of PRISMA. Preferences for place of death if faced with advanced cancer: a population survey in England, Flanders, Germany, Italy, the Netherlands, Portugal and Spain. Ann Oncology 2012;23:2006-15.

[**9**] Lunney JR, Lynn J, Foley DJ, Lipson S, Guralnik JM. Patterns of Functional Decline at the End of Life. JAMA 2003;289(18):2387-92.

[**10**] Murray SA, Kendall M, Boyd K, Sheikh A. Illness trajectories and palliative care. BMJ 2005;330:1007-11.

[**11**] Barnato AE, Cohen ED, Mistovich KA, Chang CH. Hospital End-of-Life Treatment Intensity Among Cancer and Non-Cancer Cohorts. Journal of Pain and Symptom Management 2015;49(3):521-9.

[**12**] Hyasat K. Sriram KB. Evaluation of the patterns of care provided to patients with copd compared to patients with lung cancer who died in hospital. American Journal of Hospice & Palliative Medicine 2016;33(8):717-22.

[**13**] Lyngaa T, Christiansen CF, Nielsen H, et al. Intensive care at the end of life in patients dying due to non-cancer chronic diseases versus cancer: a nationwide study in Denmark. Crit Care 2015;19:413.

[**14**] Divo M, Cote C, de Torres JP, et al. Comorbidities and risk of mortality in patients with chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2012;186(2):155-161.

[**15**] Celli BR, Cote CG, Marin JM, et al. The body-mass index, airflow obstruction, dyspnea, and exercise capacity index in chronic obstructive pulmonary disease. N Engl J Med. 2004;350(10):1005-12.

[**16**] Bouvy ML, Heerdink ER, Leufkens HGM, Hoes AW. Predicting mortality in patients with heart failure: a pragmatic approach. Heart 2003;89:605-09.

[**17**] Pocock SJ, Wang D, Pfeffer MA, et al. Granger CB. Predictors of mortality and morbidity in patients with chronic heart failure. Eur Heart J 2006;27(1):65-75.

[**18**] Hyasat K. Sriram KB. Evaluation of the patterns of care provided to patients with copd compared to patients with lung cancer who died in hospital. American Journal of Hospice & Palliative Medicine 2016;33(8):717-22.

[**19**] Siouta N, van Beek K, Preston N, et al. Towards integration of palliative care in patients with chronic heart failure and chronic obstructive pulmonary disease: a systematic literature review of European guidelines and pathways. BMC Palliative Care 2016;15:18.

[**20**] SIAARTI - Società Italiana di anestesia Analgesia Rianimazione e Terapia Intensiva. Grandi insufficienze d'organo "end stage": cure intensive o cure palliative? "documento condiviso"per una pianificazione delle scelte di cura. 2014

[**21**] SIAARTI - Società Italiana di anestesia Analgesia Rianimazione e Terapia Intensiva. Cure palliative dei pazienti con patologie respiratorie croniche avanzate non oncologiche, 2011

GLI ESITI DEI PAZIENTI CON DIMISSIONE CONTRO IL PARERE DEL MEDICO.

Bachini Letizia[1], Forni Silvia[1], Sergi Alessandro[1,2]

1. *Agenzia Regionale di Sanità della Toscana*
2. *Azienda USLToscana Centro*

Introduzione

Nel corso degli anni è stata posta molta attenzione allo studio del fenomeno delle dimissioni contro il parere del medico [***1***]. In particolare gli studi condotti si sono focalizzati su specifiche popolazioni di malati, come i pazienti traumatici o con particolari patologie infettive. Vari studi internazionali hanno dimostrato come questi pazienti sperimentino esiti negativi in termini di riammissioni e mortalità [***2***]. Inoltre recentemente sistemi di valutazione delle performance ospedaliera a livello nazionale e internazionale prendono in considerazione il fenomeno delle dimissioni contro il parere del medico. In particolare in Italia il Sistema di valutazione del Laboratorio MES considera questo indicatore come una proxy della soddisfazione del paziente [***3***]. Invece il sistema di valutazione adottato dalla Regione Lombardia [***4***] considera la dimissione volontaria come misura di efficacia ex-post (*effectiveness*) di un ospedale o reparto. Saia *et al.* [***5***] nel 2014 hanno pubblicato uno studio mirato a descrivere il trend di questo fenomeno in una regione italiana. Nel loro lavoro caratterizzano questi pazienti, ma non vengono valutati gli esiti.
L'obiettivo è quello di dare la dimensione del fenomeno e di valutarne gli esiti in termini di mortalità a 30 giorni, riaccessi in Pronto Soccorso (PS) entro 7 giorni e riammissioni a 30 giorni in ospedale dalla dimissione.

Metodi

Sono stati consultati i seguenti flussi informativi amministravi della Regione Toscana: anagrafe sanitaria, schede di dimissione ospedaliera (SDO), schede di morte informatizzate ed EMUR. Tali flussi sono stati interconnessi tramite record linkage deterministico, utilizzando come chiave primaria l'identificativo univoco del soggetto.
I ricoveri in studio sono costituiti da tutte le dimissioni in regime ordinario di pazienti residenti in Toscana nel periodo 2015-2016. Dall'analisi sono stati esclusi tutti i soggetti che nei suddetti flussi informativi presentavano un identificativo univoco nullo o errato per l'impossibilità di rintracciare il loro percorso sia in ospedale che in PS, i pazienti dimessi da reparti riabilitativi (codici specialità 28, 75, 56, 98, 00), i pazienti deceduti alla dimissione oppure entro due giorni dalla dimissione. La dimissione volontaria è definita sulla base della modalità di dimissione.
Gli esiti riguardano: mortalità a 30 giorni calcolata come differenza tra la data di dimissione del ricovero e la data di morte presente in anagrafe o nelle schede di morte informatizzate; riammissioni a 30 giorni per qualsiasi causa come differenza tra data di dimissione del ricovero e successiva ammissione in SDO; riaccessi in PS entro 7 giorni come differenza tra la data di dimissione del ricovero e la data di accettazione dell'accesso in PS.

Risultati

In Toscana nel biennio 2015-2016 sono state registrate 4.632 dimissioni volontarie, pari allo 0,68% dei ricoveri in studio. La ***tabella 1*** sotto riportata mostra che la dimissione volontaria è una prassi maggiormente diffusa tra le donne; il 60% circa dei dimessi ha tra i 20 e i 70 anni; l'11% è di nazionalità non italiana; oltre l'80% dei dimessi ha un titolo di studio medio-alto; un 30% circa dei dimessi volontari ha un basso indice di deprivazione e una stessa percentuale di pazienti ha un alto livello di deprivazione.

Tutti gli esiti esaminati risultano essere significativamente maggiori tra i dimessi volontari, in particolare le percentuali di riammissioni in ospedale e di riaccessi in PS sono molto superiori.

Tabella 1. Caratteristiche dei pazienti con (DV) o senza (Altro) dimissione contro il parere del medico

		DV		Altro	
		N	%	N	%
Totale		**6.432**	***0,68***	**936.842**	***99,32***
Sesso	M	3.151	*48,99*	436.745	*46,62*
	F	3.281	*51,01*	500.097	*53,38*
Età	0-19	882	*13,71*	140.205	*14,97*
	20-49	2.333	*36,27*	215.844	*23,04*
	50-69	1.546	*24,04*	231.417	*24,70*
	70+	1.671	*25,98*	349.376	*37,29*
Nazionalità	Italiana	5.688	*88,43*	864.143	*92,24*
	Altro	741	*11,52*	71.985	*7,68*
	--	3	*0,05*	714	*0,08*
Titolo di studio	Basso	920	*14,30*	141.269	*15,08*
	Medio	2.909	*45,23*	477.089	*50,93*
	Alto	2.603	*40,47*	318.484	*34,00*
Deprivazione	Bassa	1.712	*26,62*	269.261	*28,74*
	Media	877	*13,63*	138.875	*14,82*
	Alta	2.078	*32,31*	298.031	*31,81*
	--	1.765	*27,44*	230.675	*24,62*
Outcome					
Mortalità a 30 giorni		260	*4,04*	24.536	*2,62*
Riaccessi in PS entro 7 giorni		589	*9,16*	41.105	*4,39*
Riammissioni in H entro 30 giorni		792	*12,31*	40.271	*4,30*

Tra i pazienti in dimissione volontaria la mortalità a 30 giorni è superiore negli uomini di nazionalità italiana di età 70 e più. Anche riaccedere in PS dopo la dimissione è più frequente in soggetti di sesso maschile, in età adulta (20-69) con titolo di studio medio alto. Le riammissioni a 30 giorni in ospedale sono più frequenti oltre i 20 anni di età e tra coloro che hanno un maggior livello di istruzione. Riammissioni in ospedale e riaccessi al PS non sono influenzati da nazionalità e livello di deprivazione.

Conclusioni

Questo studio mostra come il fenomeno della dimissione contro il parere del medico riguardi una piccola percentuale dei dimessi. Inoltre questi pazienti mostrano percentuali più alte per tutti gli esiti in studio. Questi risultati sono in accordo con analoghi studi in letteratura internazionale. Tra i pazienti in dimissione volontaria gli esiti si differenziano a seconda di caratteristiche socio demografiche e mettono in luce la necessità di ulteriori approfondimenti per caratterizzare tali pazienti e sviluppare strategie di intervento appropriate.

Bibliografia

[1] Kraut A, Fransoo R, Olafson K, Ramsey CD, Yogendran M, Garland A. A population-based analysis of leaving the hospital against medical advice: incidence and associated variables. BMC health services research 2013;13:415-.

[2] Garland A, Ramsey CD, Fransoo R, et al. Rates of readmission and death associated with leaving hospital against medical advice: a population-based study. Canadian Medical Association Journal 2013;185(14):1207-14.

[3] Murante AM, Seghieri C, Brown A, et al. How do hospitalization experience and institutional characteristics influence inpatient satisfaction? A multilevel approach. The International journal of health planning and management 2014; 29(3):e247–e260.

[4] Centro di ricerca interuniversitario per I servizi di pubblica utilità. Manuale del Sistema di valutazione della performance degli ospedali Lombardi. Available from: http://www.crisp-org.it/vittadini/wp-content/papers/statistic/manualevalutazione.pdf

[5] Saia M, Buja A, Mantoan D, et al. Frequency and trends of hospital discharges against medical advice (DAMA) in alarge administrative database. Ann Ist Super Sanita 2014;50(4):357-62.

USE OF NON-INSULIN BLOOD GLUCOSE LOWERING DRUGS AND HEART FAILURE RISK: RESULTS FROM THE SAFEGUARD PROJECT

Scotti Lorenza[1], Romio Silvana[2], Leal Ingrid[2], Schmedt Niklas[3], De Berardis Giorgia[4], Gil Miguel[5], Bezemer Irene[6], Pecchioli Serena[7], Seeger John D[8], McGrogan Anita[9], Pladeval-Villa Manel[10], Smits Mark M[11], Rijnbeek Peter[2], Sturkenboom Miriam[2] and Corrao Giovanni[1].

1. *Department of Statistics and quantitative methods, University of Milano-Bicocca, Milan, Italy*
2. *Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, Netherlands*
3. *Leibniz-Institute for Prevention Research and Epidemiology – BIPS GmbH, Bremen, Germany*
4. *Center for Outcomes Research and clinical Epidemiology, Pescara, Italy*
5. *Spanish Agency for Drugs and Medical Devices, Madrid, Spain*
6. *PHARMO Institute, Utrecht, Netherlands*
7. *Health Search, Italian College of General Practitioners, Genomedics, Florence, Italy*
8. *The Brigham and Women's Hospital, Harvard Medical School, Boston, United States*
9. *University of Bath, Bath, United Kingdom*
10. *RTI Health Solutions, Barcelona, Spain*
11. *VU University Medical Center, Amsterdam, Netherlands*

Introduction

Type 2 Diabetes (T2DM) is a chronic and progressive disease characterized by hyperglycemia due to a defect in insulin signalling and insulin deficiency. These features leads to hyperglycemia and dyslipidaemia, which can impair insulin secretion and action [***1***] and further cause microvascular and macrovascular complications [***2***]. The treatment of T2DM is based on lifestyle changes and pharmacological treatment. Its main goal is to prevent and control hyperglycemia and then reduce its complications. For decades treatment options were based on lifestyle changes, metformin, sulfonylureas, and insulin. By the end of the 90's and during the last decade, new compounds with different mechanisms of action have been developed and authorized for marketing [***3***]: thiazolidinediones (TZD), meglitinides, Glucagon-like peptide 1 receptor agonists (GLP-1RA), Dipeptidyl-peptidase 4 (DPP-4) inhibitors and amilyn analogs. The progressive nature of T2DM and its associated glycemia tends to lead to increases in the dose and the use of combinations of non-insulin blood glucose lowering drugs (NIBGLDs) or the addition of insulin over time to meet the goals for metabolic control. Although safety issues associated with blood glucose lowering drugs are not new, recently, the safety of these treatments has been questioned and highly publicized. It has been reported that some of them increase the risk or modify the prognosis of diseases such as cancer, cardiovascular (CVD) or pancreatic diseases [***4***]. The Safety Evaluation of Adverse Reaction in Diabetes (SAFEGUARD) is a project funded by the FP7-HEALTH program aimed to assess the cardio/cerebrovascular and pancreatic safety of blood glucose lowering drugs in order to estimate the relationship between use of glucose lowering agents and risk of selected events, including heart failure (HF).
To estimate the association between use of NIBGLDs and risk of HF within the SAFEGUARD project.

Methods

The data used in the project were retrieved from 9 databases (DBs): BIFAP (Spain), GePaRD (Germany), Regional DBs of Puglia and Lombardy, Health Search (Italy), IPCI, PHARMO (The Netherlands), CPRD (UK) and Medicare (US). These DBs include different types of data sources (electronic medical records, healthcare utilization DBs and record linkage systems), different settings (out and in patient care, primary care), and the use of different terminology and coding systems for the registry of events and medications. Moreover, the data availability period differs among DBs covering overall a time span from 1999 to 2012

Therefore, an harmonization of outcome, exposure and covariates definition was performed to ensure the consistency of the definitions between databases. Within each database, all subject who received at least one prescription of NIBGLDs (ATC codes A10B) one year after database's enrolment were included in the initial cohort and the date of the first prescription was defined as cohort entry date. From these cohorts, we excluded those subjects who, in the year before cohort entry date i) were not enrolled in the databases ii) were treated with any antidiabetic drug (both NIBGLD and insulin ATC code A10) and those with a malignant cancer any time before cohort entry. Patients were followed from the cohort entry date until the date of the earliest event among hospitalization/diagnosis of heart failure (depending on the type of database), migration, death and end of data availability (depending on the DB).

Cases ascertainment

Cases were all patients who experienced a hospitalization/a diagnosis of heart failure during follow-up (ICD-9CM codes 398.91, 402.01, 402.11, 402.91, 404.01, 404.03, 404.11, 404.13, 404.91, 404.93, 428.xx for Regional DBs of Lombardy and Puglia, PHARMO and MEDICARE, same codes with the addition of free text search for Health search, ICD-10 codes: I50.x, I11.0, I13.0 and I13.2 for GePaRD, ICPC code K77 and free text search for IPCI and BIFAP and several Read Codes including G58..0 for CPRD). The index date was defined as the date of the first HF occurrence. Within each database, up to ten controls were randomly selected and matched to each case by sex, age (±1 year), year of cohort entry and index date. The index date of the case was assigned to the corresponding matched controls to ensure an equal length of follow-up between cases and controls.

Exposure assessment

All prescriptions of NIBGLDs (Metformin (ATC code A10BA02), Sulfonylureas (ATC codes A10BB) Thiazolidinediones (ATC codes A10BG) Meglitinides (ATC codes A10BX02 and A10BX03) α-glucosidase inhibitors (ATC codes A10BF), GLP-1RA (ATC codes A10BX04, A10BX07), DPP-4 inhibitors (ATC codes A10BH), Amilyn analog (ATC code A10BX05), Others (A10BX06) dispensed to cases and controls between cohort entry and index dates were selected. Both monotherapy and drug combinations (fixed or extemporaneous) were considered. Exposure to NIBGLD was categorized according to the coverage of the last prescription in: current use if the 30 days before the index date were covered by a prescription, recent use (between 31 and 62 days before index date) and past use (more than 62 days before index date). Current use was further classified according to the specific drug used as monotherapy or dual combination with metformin and any other NIBGLD including dual combination with other drugs, triple combinations and monotherapies of drugs with very low prevalence of use identified as less than 5 exposed subjects.

Identification of the covariates

Several potential confounders were identified in the year before cohort entry such as smoking habit, obesity, hypertension and hyperlipidemia. Other variables were detected 30 days before cohort entry, in particular, use of anticoagulants, lipid lowering drugs, ace-inhibitors, ATII antagonists diuretics, calcium channel blockers, beta-blockers, antihypertensives, vasodilators, systemic combined contraceptives, hormone replacement therapy and oestrogens. Aspirin use, antiplatelet, glucocorticoids, antiarrhythmics were identified 30 days before index date. Finally, several diseases were evaluated both before cohort entry and before index date, namely, myocardial infarction, atrial fibrillation, ventricular arrhythmia, pulmonary hypertension, cardiomyopathies, congestive heart disease, valve disorders, ischemic heart diseases, chronic liver disease, chronic kidney disease, peripheral artery disease, COPD, any cerebrovascular disease, myocarditis, endocarditis and renal failure.

Statistical analysis

Within each database, a multivariable conditional logistic regression model was applied to estimate adjusted Odds Ratios (OR) and corresponding 95% confidence intervals (95%CI) of current use of each considered NIBGLD (monotherapy or metformin combinations), recent use, and past use of any NIBGLD and risk of HF considering as reference current use of metformin in combination with sulfonylureas. To provide an overall

estimate of the association measure, the results obtained in each database were pooled using a meta-analytic approach. Heterogeneity was assessed using the I^2 index. An I^2 index greater than 60% imply high heterogeneity. In presence of heterogeneity, random effects estimates were considered.

Results

Overall, 29,668 HF cases were identified. The adjusted OR and the corresponding 95%CI for current use of specific NIBGLD and metformin combination compared to current use of metformin in combination with sulfonylureas obtained after the meta analytic pooling are reported in ***figure 1***.
Compared to current use of combination therapy of metformin and sulfonylureas, current use of rosiglitazone (pooled OR 1.75, 95% CI 1.50-2.03), pioglitazone (1.40, 1.24-1.58) and glipizide (1.24, 1.11-1.38) increased HF risk. No statistically significant association was found for current use of incretin-based treatment, alpha glucosidase inhibitors and glinides. Current use of metformin and glibenclamide seems, conversely to decrease HF risk.

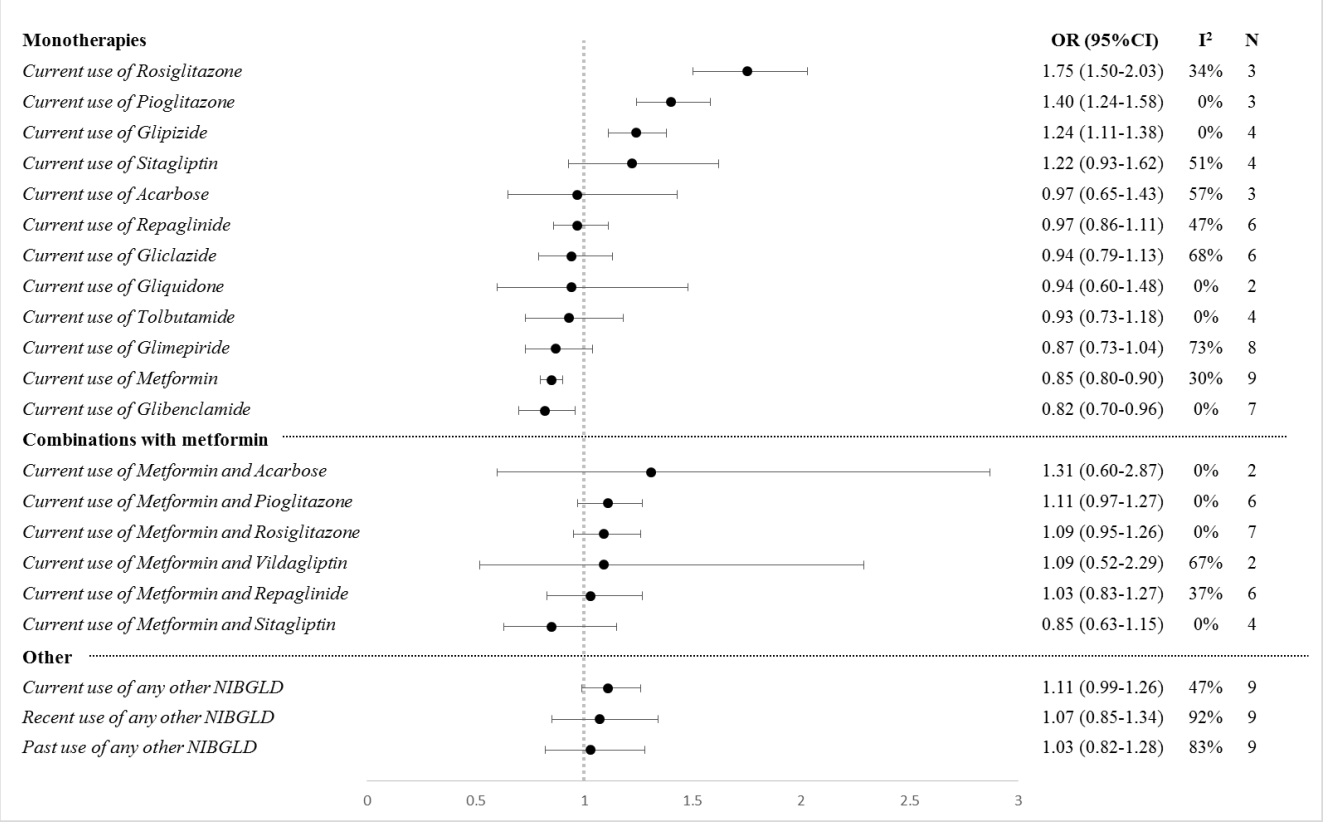

Figure 1. Meta-analytic pooled estimate of the association between current use of specific NIBGLD and metformin combination
I^2 value of the I^2 index used to assess between study heterogeneity
N number of databases considered in the pooled estimate

Conclusions

The results of the large international multicentre case-control study based on secondary data confirm the data available in the scientific literature regarding the increased risk of HF risk among current users of thiazolidinediones and glipizide as monotherapy, but not in combination with metformin. The lack of association between current use of incretin-based treatment, alpha glucosidase inhibitors and glinides, in particular sitagliptin may be due to the low exposure prevalence leading to a lack of power to detect the potential association.

References

[**1**] Muoio DM, Newgard CB. Mechanisms of disease:Molecular and metabolic mechanisms of insulin resistance and beta-cell failure in type 2 diabetes. Nat Rev Mol Cell Biol 2008;9(3):193-205.

[**2**] Fowler MJ. Microvascular and Macrovascular Complications of DiabetesClinical Diabetes 2008;26(2):77-82.

[**3**] EMA, European public assessment reports. Diabetes Mellitus Type 2, 2011.

[**4**] Bergenstal RM, Bailey CJ, Kendall DM. Type 2 diabetes: assessing the relative risks and benefits of glucose-lowering medications. Am J Med 2010;123(4):374.e 9 -18.

ANTIDEPRESSIVI E RISCHIO DI ARITMIA IN PAZIENTI ANZIANI CARDIOPATICI: RISULTATI DI UNO STUDIO MULTICENTRICO ITALIANO

Biffi Annalisa[1], Rea Federico[1], Scotti Lorenza[1], Lucenteforte Ersilia[2], Mugelli Alessandro[2], Bettiol Alessandra[2,3], Chinellato Alessandro[3], Roberto Giuseppe[4], Agabiti Nera[5], Onder Graziano[6], Trifirò Gianluca[7], Vitale Cristiana[8], Corrao Giovanni[1], on the behalf of the the Italian Group for Appropriate Drug prescription in the Elderly (I-GrADE)[9]

1. *Università degli studi di Milano-Bicocca*
2. *Università degli studi di Firenze*
3. *ULSS 9 Treviso*
4. *ARS Toscana, Firenze*
5. *ASL 1 Roma*
6. *Università cattolica Sacro Cuore, Roma*
7. *Università degli studi di Messina*
8. *IRCCS San Raffaele Pisana, Roma*
9. *I-GrADE (progetto finanziato da AIFA-FARM9LBBBL): Nera Agabiti[5], Claudia Bartolini[4], Roberto Bernabei[6], Alessandra Bettiol[3], Stefano Bonassi[8], Achille Patrizio Caputi[7], Silvia Cascini[5], Alessandro Chinellato[3], Francesco Cipriani[4], Giovanni Corrao[1], Marina Davoli[5], Massimo Fini[8], Rosa Gini[4], Francesco Giorgianni[7], Ursula Kirchmayer[5], Francesco Lapi[4], Niccolò Lombardi[2], Ersilia Lucenteforte[2], Alessandro Mugelli[2], Graziano Onder[6], Federico Rea[1], Giuseppe Roberto[4], Chiara Sorge[5], Michele Tari[7], Gianluca Trifirò[7], Alfredo Vannacci[2], Davide Liborio Vetrano[6], Cristiana Vitale[8]*

Introduzione

La depressione rappresenta un serio problema di salute pubblica dato il suo alto livello di prevalenza, in particolare nella popolazione anziana. Infatti, attualmente, circa un quarto della popolazione europea soffre di sindrome depressiva [***1***]. L'adeguato trattamento della depressione con farmaci antidepressivi, è necessario per ridurre l'onere della morbidità, della disabilità e della mortalità. Dal 1950, diversi tipi di farmaci antidepressivi (AD) sono stati sviluppati per trattare i sintomi depressivi. Al giorno d'oggi sono disponibili diversi trattamenti, quali i farmaci triciclici (TCA), gli inibitori selettivi della ricaptazione della serotonina (SSRI) e gli antidepressivi atipici (AA).

L'uso di farmaci antidepressivi è stato associato ad un incremento di rischio cardiovascolare (CV) [***2, 3***] e in particolare di aritmia [***4***]. Difatti TCA, SSRI e AA, in particolare venlafaxina, possono provocare un prolungamento dell'intervallo QT [***5-7***] che può portare ad aritmie tra cui la torsione di punta, potenzialmente fatale [***8***]. La relazione tra trattamento con AD e aritmia rappresenta, quindi, una questione di grande rilevanza clinica [***9***], in particolare in una popolazione di soggetti anziani con precedente ricovero CV.

Valutare l'associazione tra trattamento antidepressivo (SSRI, TCA e AA) e rischio di sviluppare aritmia in una popolazione di anziani cardiopatici utilizzando i dati provenienti dai database sanitari amministrativi delle unità partecipanti al progetto I-Grade. In particolare, sono stati confrontati i risultati ottenuti da due disegni di studio (caso-controllo innestato in una coorte e case-crossover) caratterizzati da una differente capacità di controllo del confondimento.

Metodi

I dati utilizzati per il presente studio sono stati ottenuti dai database sanitari di cinque unità territoriali italiane che partecipano al programma I-Grade (Italian Group for Appropriate Drug Prescription in the Elderly) finanziato dall'Agenzia Italiana del Farmaco (AIFA), al fine di valutare l'appropriatezza prescrittiva in pazienti anziani soggetti a precedente ricovero CV.

Le unità territoriali partecipanti riguardano tre regioni (Lazio, Lombardia, Toscana) e due unità sanitarie locali (Caserta, Treviso). I dati di circa 21 milioni di beneficiari del Servizio Sanitario Nazionale (SSN) residenti in queste aree (che rappresentano quasi il 35% della popolazione italiana). Al fine di raccogliere le informazioni relative alle prestazione sanitarie erogate, il SSN ha istituito un sistema di banche dati che include tra gli altri:(1) un archivio dei soggetti residenti, che riporta dati demografici e amministrativi (ad esempio età, genere), oltre che la data di inizio e termine in qualità di beneficiario dei servizi sanitari (data di nascita/immigrazione e morte/emigrazione); (2) un database riguardante i ricoveri ospedalieri, comprese le informazioni sulla diagnosi primaria e fino a cinque condizioni e procedure codificate secondo la Classificazione Internazionale delle Patologie (ICD-9); (3) un database comprendente tutte le prescrizioni ambulatoriali rimborsate dal SSN e codificate secondo il sistema di Classificazione Anatomico Terapeutico e Chimico (ATC).
Le informazioni relative a ogni assistito contenute in queste banche dati sono tra loro collegabili tramite un codice di identificazione univoco che consente il *record linkage* tra tutti i database a disposizione. Al fine di preservare la privacy, il codice di identificazione originale è stato sostituito con un codice criptato.
I beneficiari del SSN (i) residenti nelle unità territoriali nominate sopra, (ii) con almeno 65 anni di età, (iii) con precedente ricovero per patologie CV (scompenso cardiaco, malattie cerebrovascolari o cardiopatia ischemica) durante tra l'1/1/2008 e il 31/12/2010, sono stati inclusi nella coorte. La data di dimissione del primo ricovero CV verificato durante questo periodo è stata definita come data di ingresso nella coorte.
I soggetti sono stati esclusi se non registrati presso il SSN nei due anni antecedenti l'ingresso in coorte (per effettuare una valutazione delle covariate e applicare gli appropriati criteri di esclusione) e con almeno sei mesi di follow-up (per considerare un adeguato tempo di esposizione ai farmaci di interesse). Da questa coorte iniziale sono stati esclusi i soggetti che, durante i due anni antecedenti l'ingresso, avevano ricevuto almeno una prescrizione di (i) antineoplastico o con segni di cancro, così da escludere pazienti con più elevata fragilità, (ii) farmaco antiaritmico o ricovero per aritmia, per assicurare l'osservazione dei soli casi incidenti di aritmia durante il periodo di follow-up, (iii) antidepressivo, per includere solo i pazienti che hanno iniziato la terapia antidepressiva durante il follow-up, per includere i soli nuovi utilizzatori di AD [***10***].
I soggetti rimanenti sono stati seguiti dalla data di ingresso nella coorte fino al verificarsi di uno dei seguenti eventi: evento in studio (ricovero per aritmia), morte, emigrazione, segni di cancro o termine del periodo di studio, definito dal termine della disponibilità dei dati di ciascuna delle unità territoriali partecipanti al progetto (31 Dicembre 2011 per Lazio, 31 Dicembre 2012 per Lombardia, Toscana e Caserta e 31 Dicembre 2014 per Treviso).
Definizione dell'evento
L'esito principale dello studio è stato il primo ricovero per aritmia cardiaca durante il follow-up. In più, il ricovero per aritmia ventricolare è stato considerato outcome secondario, escludendo i soggetti che avevano ricevuto una prescrizione di antiaritmico o ricovero per aritmia durante il follow-up. La data del primo ricovero per aritmia è stata definita come *data indice.*

Disegno caso-controllo innestato in una coorte
Selezione dei controlli
Uno studio caso-controllo innestato è stato implementato per la coorte considerata. Ad ogni caso sono stati appaiati fino a cinque controlli selezionati casualmente e appaiati al corrispondente caso per unità territoriale di assistenza sanitaria, genere, età (±3 anni), e data di ingresso in coorte (±7 giorni).
Definizione dell'esposizione
Sono state identificate tutte le prescrizioni di AD dispensate durante il periodo di follow-up. L'esposizione è stata categorizzata in base alla data dell'ultima prescrizione di AD in: uso corrente (se la data cadeva nei 15 giorni antecedenti la data indice) e non uso. Il periodo corrente è stato ulteriormente categorizzato a seconda della classe di antidepressivo: SSRI, TCA e antidepressivi atipici.

Identificazione delle covariate
Per ogni caso e controllo diversi potenziali fattori confondenti sono stati identificati. Queste includono le caratteristiche dei soggetti misurate nei due anni antecedenti la data di ingresso in coorte o entrambi, nei due anni precedenti e durante il follow-up. Le caratteristiche misurate nei due anni antecedenti l'ingresso includono la diagnosi di ricovero CV considerata per il reclutamento (scompenso cardiaco, malattia cerebrovascolare o cardiopatia ischemica), i precedenti ricoveri CV e l'indice di comorbidità di Charlson (classificato come 0, 1 o ≥2) [***11***]. Le caratteristiche misurate nei due anni antecedenti e durante il periodo di follow-up includono co-trattamenti con farmaci CV (come antiaritmici, digossina e nitrati), antidiabetici, antiipertensivi e farmaci ipolipemizzanti, e ricoveri ospedalieri con diagnosi di diabete mellito, ipertensione e iperlipidemia. Durante il follow-up è stato inoltre valutato l'uso di farmaci psicotropi (antiepilettici, antiparkinson, psicolettici, psicoanalettici).

Disegno case-crossover
Definizione dell'esposizione
L'esposizione ad antidepressivo è stata misurata nei 15 giorni prima della data indice (periodo di rischio) e in 5 finestre di controllo, ciascuna di ampiezza 15 giorni. Un periodo di washout di 15 giorni è stato considerato tra il periodo a rischio e il primo dei periodi di controllo. Un individuo era identificato come esposto in un periodo se la data di prescrizione cadeva all'interno della finestra stessa.
Identificazione delle covariate
Alcuni potenziali fattori confondenti tempo dipendenti sono stati identificati durante le rispettive finestre (rischio e controllo) quali utilizzo di farmaci cardiovascolari (antiaritmici, digossina e nitrati), antidiabetici, antipertensivi, ipolipemizzanti, altre classi di AD, farmaci psicotropici (antiepilettici, antiparkinson, psicolettici, psicoanalettici); ricovero per diabete, ipertensione e iperlipidemia.
Analisi statistica
Sono stati utilizzati due approcci di analisi dei dati. In primo luogo, è stato utilizzato un modello di regressione logistica condizionata per il caso-controllo innestato 1:5 utilizzato per stimare gli Odds Ratio (OR), aggiustati per diversi confondenti e i relativi Intervalli di Confidenza (IC) al 95%, di aritmia associati con l'utilizzo di AD, avendo come riferimento la categoria nessun uso. In secondo luogo, è stato applicato l'approccio case-crossover [***12***]. Le stime sono state aggiustate per diverse covariate e per l'indice di Charlson. Nell'analisi secondaria è stato considerato come outcome il ricovero per aritmia ventricolare.

Risultati
È stata considerata una coorte di 199,569 soggetti tra le cinque unità territoriali, che ha generato 17,277 casi incidenti di aritmia, di cui 3,030 di aritmia ventricolare. Nell'analisi caso-controllo innestato i casi sono stati appaiati a 85,432 controlli.
L'età media dei pazienti all'ingresso in coorte era di circa 79 anni e il 55% erano uomini. Rispetto ai controlli, i casi facevano maggior uso di antidepressivi e sono stati più spesso diagnosticati per scompenso cardiaco e meno frequentemente per le malattie cerebrovascolari e cardiopatie ischemiche.
La ***Figura 1*** riporta gli OR e i corrispondenti 95%IC dell'associazione tra utilizzo corrente di AD e il rischio di sviluppare aritmia.
Nello studio caso controllo, gli utilizzatori correnti di SSRI e AA rispettivamente hanno mostrato incremento di rischio di aritmia, OR 1.43 (95% IC, 1.24-1.65) e 1.45 (95% IC, 1.20-1.76) rispetto ai non utilizzatori mentre utilizzando il disegno case-crossover, si osserva un incremento pari a OR 1.54 (95% IC, 1.25-1.89) e 1.81 (95% IC, 1.38-2.37). Valutando l'associazione tra il trattamento antidepressivo nel suo complesso è stato rilevato un incremento di rischio di aritmia, OR 1.44 (95% IC 1.29-1.62) e 1.62 (95% IC 1.37-1.92) rispettivamente per il disegno caso-controllo innestato e case-crossover. Un significativo incremento di rischio di aritmia ventricolare è stato anche rilevato tra gli utilizzatori correnti di AD nel loro complesso, e

nello specifico SSRI, sia nello studio caso-controllo innestato (AD: OR 1.61 (95% IC 1.25-2.07); SSRI: 1.61 (95% IC 1.17-2.20) che nello studio case-crossover (AD: OR 2.00 (95% IC 1.35-2.98) e SSRI: 2.30 (95% IC 1.41-3.75).

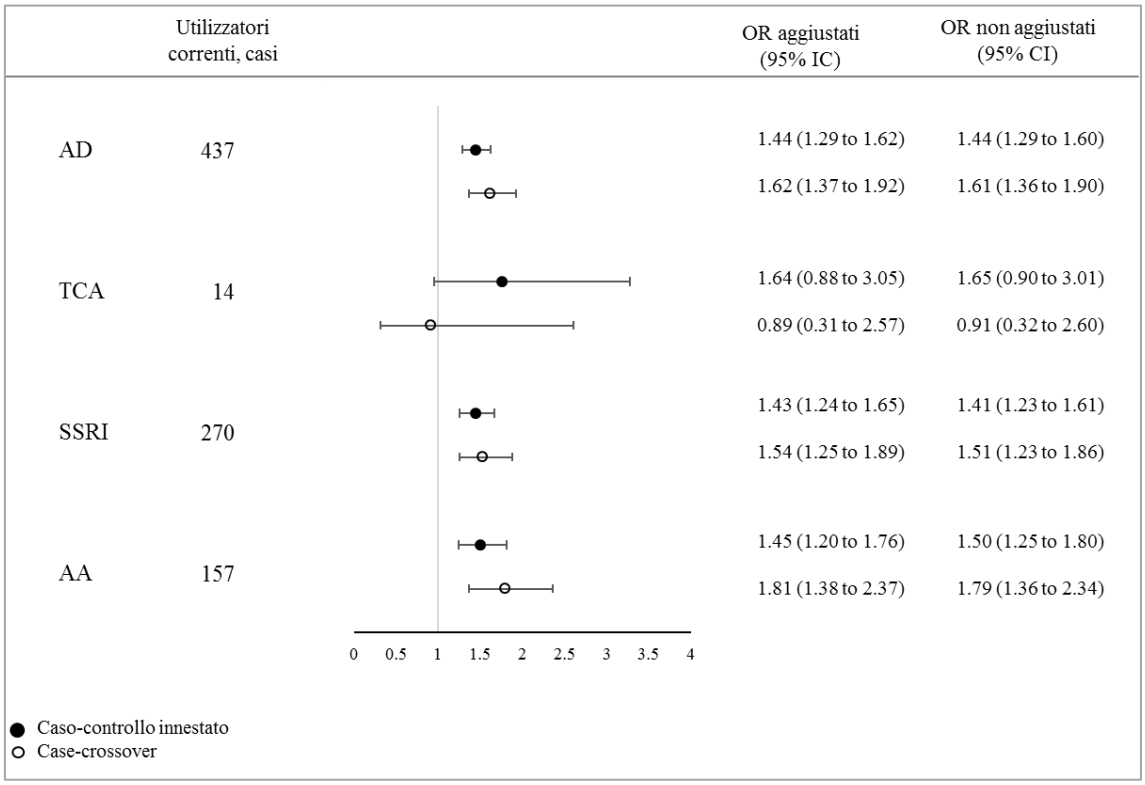

Figura 1. OR e corrispondenti 95%IC dell'associazione tra utilizzo corrente di AD e il rischio di sviluppare aritmia

Conclusioni

In questo ampio studio di popolazione, in cui sono stati considerati pazienti anziani con precedente ricovero CV, è stato riscontrato un incrementato rischio di sviluppare aritmia tra gli utilizzatori correnti di SSRI e AA. L'incremento del rischio di aritmia associato al trattamento con antidepressivi evidenziato nello studio basato sul disegno caso-controllo innestato potrebbe essere spiegato dalla presenza di confondimento residuo, dato da una porzione non misurabile circa la severità della malattia, la comorbidità, lo stato socio-economico e vari fattori legati allo stile di vita; caratteristiche non registrate dalle banche dati amministrative. Per tener conto del potenziale confondimento residuo è stato effettuato uno studio applicando il disegno case-crossover. I risultati ottenuti da questo secondo studio erano paragonabili a quelli ottenuti dall'applicazione del disegno caso-controllo innestato e case-crossover, suggerendo che in questa analisi il confondimento non sembra avere un forte impatto sulle stime.

Bibliografia

[1] Depression. A global Public Health Concern. Reperibile su: http://www.who.int/mental_health/ (last accessed 15 March 2017).

[2] Biffi A, Scotti L, Corrao G. Use of antidepressants and the risk of cardiovascular and cerebrovascular disease: a meta-analysis of observational studies. Eur J Clin Pharmacol 2017 Apr;73(4):487-97.

[**3**] Shin D, Oh YH, Eom CS, et al. Use of selective serotonin reuptake inhibitors and risk of stroke: a systematic review and meta-analysis. J Neurol 2014;261(4):686-95.
[**4**] Waring WS. Clinical use of antidepressant therapy and associated cardiovascular risk. Drug Healthc Patient Saf 2012;4:93-101.
[**5**] Wenzel-Seifert K., Wittmann M. and Haen E. QTc prolongation by psychotropic drugs and the risk of Torsade de Pointes. Dtsch Arztebl Int 2011;108(41):687-93.
[**6**] Bavle A. Venlafaxine induced QTc interval prolongation in a therapeutic dose. Asian J Psychiatr 2015;16:63-4.
[**7**] Isbister GK. Electrocardiogram changes and arrhythmias in venlafaxine overdose. Br J Clin Pharmacol 2009;67(5):572-76.
[**8**] El-Sherif N. and Turitto G. Torsade de pointes. Curr Opin Cardiol 2003;18(1):6-13.
[**9**] Howland RH. A critical evaluation of the cardiac toxicity of citalopram: part 2. J Psychosoc Nurs Ment Health Serv 2011;49(12):13-6.
[**10**] Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. Am J Epidemiol 2003;158(9):915-20.
[**11**] Charlson ME, Charlson RE, Peterson JC, et al. The Charlson comorbidity index is adapted to predict costs of chronic disease in primary care patients. J Clin Epidemiol 2008;61(12):1234-40.
[**12**] Delaney JA. and Suissa S. The case-crossover study design in pharmacoepidemiology. Stat Methods Med Res 2009;18(1):53-65.

VALIDAZIONE DELL'UTILIZZO DI BANCHE DATI AMMINISTRATIVE PER STUDI DI FARMACO UTILIZZAZIONE IN TERAPIE ONCOLOGICHE

Bartolini Claudia[1], Roberto Giuseppe[1], Moscatelli Valentino[2], Barchielli Alessandro[3], Paoletti Davide[4], Giorgi Silvano[4], Donnini Sandra[2], Ziche Marina[2], Monti Maria Cristina[5], Gini Rosa[1]

1. *Agenzia Regionale di Sanità della Toscana, Firenze;*
2. *Università degli studi di Siena, Siena;*
3. *Istituto per lo studio e la prevenzione oncologica, Firenze;*
4. *Azienda Ospedaliera Universitaria Senese, Siena*
5. *Università di Pavia, Dipartimento di Sanità pubblica, Medicina sperimentale e Forense, Pavia*

Introduzione

Il rituximab è un anticorpo monoclonale autorizzato per il trattamento di neoplasie ematologiche (Linfoma non-Hodgkin [***1***], leucemia linfatica cronica [***2***]), malattie autoimmuni (artrite reumatoide [***3***]) e alcuni tipi di vasculiti. Le evidenze disponibili circa l'utilizzo di rituximab nella pratica clinica reale sono ad oggi limitate, particolarmente per quanto riguarda l'uso oncoematologico. Ciò è in buona parte da ricondursi all'uso esclusivo di rituximab in ambiente ospedaliero o ad esso assimilabile dove generalmente le grandi banche dati elettroniche sanitarie non sono in grado di tracciare l'uso del farmaco a livello di paziente e/o d'indicazione d'uso [***4***].
Validare la qualità dei flussi di dati amministrativi della regione Toscana (DART) come fonte per la conduzione di studi di farmaco utilizzazione di farmaci oncologici, utilizzando i dati della Farmacia Ospedaliera di Siena (FOS) e il rituximab come caso di studio.

Metodi

La coorte di studio è stata identificata a partire dai pazienti registrati nella banca dati della FOS di età superiore ai 18 anni e con almeno una somministrazione di Rituximab ricevuta nel reparto di ematologia o oncologia nel periodo compreso tra l'1 gennaio 2011 e il 31 dicembre 2014. Le informazioni contenute in FOS sono state collegate attraverso un record linkage deterministico al DART. Questo è stato possibile perché entrambe le banche dati contenevano il codice d'identificazione anonimo regionale, che è stato la chiave primaria per tale processo di matching. La prima dispensazione di rituximab registrata in FOS o in DART costituiva la data d'ingresso nella coorte.
Abbiamo analizzato le dispensazioni nei primi 365 giorni successivi alla data d'ingresso nella coorte e calcolato la percentuale di pazienti della coorte che risultavano presenti anche in DART. Tra questi pazienti, per calcolare la percentuale di episodi di trattamento con rituximab identificabili in entrambe le fonti di dati, abbiamo assunto che le dispensazioni registrate in ciascuna banca dati con meno di 3 giorni differenza corrispondessero allo stesso episodio di trattamento (i.e. somministrazione). È stato calcolato il contributo di ciascuna fonte di dati rispetto all'identificazione dei pazienti nella coorte di studio. Successivamente, abbiamo confrontato questi risultati con quelli ottenuti utilizzando la data esatta come criterio di matching delle dispensazioni registrate nelle due fonti di dati. È stata osservata la percentuale di somministrazioni presenti solo in FOS che erano avvenute in regime di ricovero. Per le prescrizioni registrate in entrambe le fonti, è stato confrontato il dosaggio medio registrato. Per ciascun paziente identificabile in entrambe le fonti, è stato confrontato il numero medio di dispensazioni totali identificate utilizzando FOS o DART, con il numero medio di dispensazioni identificabili in DART.

Risultati

In FOS sono stati ritrovati 307 pazienti con un codice regionale d'identificazione anonimo valido: tra questi, in DART è stata identificata almeno una dispensazione di rituximab per 295 pazienti (96%).
Per questi 295 pazienti, considerando come un unico episodio di trattamento le dispensazioni registrate con differenza di data inferiore a 3 giorni, sono state identificate 1951 somministrazioni di rituximab in almeno una delle due fonti: 1757 (90,1%) sono state registrate in FOS, 1576 (80.7%) in DART, e 1382 (70.8%) in entrambe. Utilizzando la data esatta di dispensazione, invece, il numero totale di episodi di trattamento sarebbe stato 2093, con un aumento di 142 (+6,7%?) somministrazioni. Tra le somministrazioni identificate in FOS e non in DART, quelle avvenute durante un ricovero ordinario sono state il 3%. La dose media registrata in FOS è risultata minore del 3% rispetto a quella registrata in DART.
Tra i pazienti identificabili in entrambe le fonti, il numero medio di dispensazioni registrate in DART era significativamente minore rispetto al numero medio presente in almeno una delle due, 20% (5,13 vs. 6,35).

Conclusioni

I risultati preliminari di questo studio di validazione dimostrano che DART identifica in modo affidabile i pazienti in trattamento con rituximab. Anche gli episodi di trattamento possono essere identificati con una buona sensibilità (4 somministrazioni su 5 si trovano in DART). Il dosaggio delle dispensazioni è registrato in modo sostanzialmente coerente tra le due fonti. Inoltre, i risultati dello studio suggeriscono che il numero medio di somministrazioni per paziente potrebbe essere sottostimato utilizzando solo DART.
Le somministrazioni avvenute durante la degenza, è verosimile che non siano registrate in DART. Tuttavia tra le somministrazioni identificate solo in FOS, la percentuale di somministrazioni avvenute durante un ricovero ospedaliero, è risultata essere trascurabile. Pertanto, questa non appare essere una spiegazione plausibile per la mancata presenza di tali somministrazioni.
Questo studio si presta a numerosi sviluppi: verranno investigati il tempo di calendario, genere, età, tipo di indicazione e durata di terapia come possibili determinanti delle discrepanze osservate tra DART e FOS. Inoltre, saranno validati le indicazioni di utilizzo derivate in DART attraverso algoritmi basati sulle diagnosi informazioni contenute nelle Schede di Dimissione Ospedaliera (SDO), confrontandoli con quelle riportate FOS.
Ulteriori approfondimenti saranno necessari per comprendere quali sono gli episodi di trattamento da cui scaturisce la registrazione in una fonte e non nell'altra.

Bibliografia

[1] Dotan E, Aggar CW, Smith MR. Impact of Rituximab (Rituxan) on the Treatment of B-Cell Non-Hodgkin's Lymphoma, Pharmacology & Therapeutics 2010;35(3):148-57.
[2] Gentile M, Vigna E, Mazzone C, et al., Rituximab for the treatment of patients with chronic lymphocytic leukemia, Cancer Management and Research 2010;2:71–81.
[3] Randall KL. Rituximab in autoimmune diseases, Australian Prescriber 2016;39(4):131-34.
[4] Kavcic M, Fisher BT, Seif AE, Li Y, Huang YS, Walker D, Aplenc R. Leveraging Administrative Data to Monitor Rituximab Use in 2875 Patients at 42 Freestanding Children's Hospitals across the United States, J Pediatr. 2013 162(6):1252–8.

CLINICAL SIGNIFICANCE OF NATIVE VS. STATIN-ASSOCIATED DIABETES: EVIDENCE FROM A LARGE POPULATION-BASED COHORT

Monzio Compagnoni Matteo[1,2], Rea Federico[1,2], Merlino Luca[3], Catapano Alberico L[4,5], Mancia Giuseppe[6], Corrao Giovanni[1,2]

1. *Department of Statistics and Quantitative Methods, Division of Biostatistics, Epidemiology and Public Health, University of Milano-Bicocca, Milan, Italy*
2. *Interuniversity Centre of Healthcare Research & Pharmacoepidemiology, located at the Laboratory of Healthcare Research & Pharmacoepidemiology, University of Milano-Bicocca, Milan, Italy*
3. *Operative Unit of Territorial Health Services, Lombardy Region, Milan, Italy*
4. *Department of Pharmacological and Biomolecular Sciences, Centre of Epidemiology and Preventive Pharmacology (SEFAP), University of Milano, Milan, Italy*
5. *IRCSS Multimedica, Sesto San Giovanni, Milan, Italy*
6. *Professor Emeritus of Medicine, University of Milano-Bicocca, Milan, Italy*

Introduction

A large number of studies has shown that use of statins is accompanied by an increased risk of developing type 2 diabetes [***1-5***], which is thus currently listed as an inconvenience of these drugs that may attenuate in some patients their protective effect. However, several aspects of the statin-induced diabetes have not been adequately clarified. For example, albeit several hypotheses have been advanced, the mechanisms through which statins favour the alteration of glucose metabolism that leads to the appearance of hyperglycemia and diabetes remain unclear [***4, 6***]. Furthermore, although statin-induced diabetes is generally believed not to offset the protective lipid-lowering effect of statins on the cardiovascular (CV) system [***7-9***], limited information exists on whether statin-induced diabetes has the same adverse prognostic significance of native diabetes, i.e., whether it is associated with a similar increasing risk of diabetes-related macrovascular complications. This information is of fundamental importance to reliably quantify the impact of statin-induced diabetes on the role played by statins on primary and secondary CV prevention [***10, 11***]. Since we have previously shown that at the population level an increasing adherence with statin treatment is accompanied by a clear-cut progressive increase in the risk of new onset type 2 diabetes [***12***], thus the purpose of the present study was to provide information on the extent to which type 2 diabetes more likely induced by statins affects the risk of macrovascular complications to a similar or different degree compared to diabetes of a more likely native nature.

Methods

The data used for this study were retrieved from the healthcare utilization databases of Lombardy, a Region of Italy which accounts for about 16% (almost ten million) of its population, by means of a record-linkage procedure among databases. In Italy, the whole population is covered by the National Health Service and in Lombardy this has been associated since 1997 with an automated system of databases to collect a variety of information.

All the residents in Lombardy, with an age between 40 and 80 years and at least one prescription of statins between 2003 and 2005, were identified. The date of the first dispensation was considered as the step-1 index date. Four patient categories were excluded: (i) patients who received one or more statin prescriptions within three years prior the step-1 index date; (ii) patients who received at least one antidiabetic agent, or were hospitalized with a diagnosis of diabetes, within the three years before the step-1 index date; (iii) patients who were hospitalized for CV disease or received prescriptions of CV drugs such as nitrates or

digitalis within the three years before the step-1 index date; and (iv) patients who did not renew the initial prescription of statins and/or did not reach at least one year of follow-up.
The remaining 84,828 patients newly treated with statins between 2003 and 2005 were enrolled in the study and represented the step-1 cohort, each of its members accumulating person-years of follow-up from the step-1 index date (i.e., index statin prescription) until the earliest among the dates of starting antidiabetic drug therapy (step-1 outcome, i.e., the appearance of diabetes as diagnosed by the prescription of antidiabetic drugs) or censoring, e.g. death from any cause, emigration or step-1 phase stopping (i.e., December 31th 2009).
The 4,391 step-1 cohort members who exhibited new onset diabetes and the 77,893 statin-treated patients who did not have any antidiabetic drug dispensation were considered eligible for inclusion in the step-2 cohort. Cohort members who experienced diabetes were 1:1 randomly matched with those who did not developed diabetes for gender, age at cohort entry (±1 year), step-1 index date (±30 days) and previous adherence with statin therapy. A patient without diabetes had also to be at risk of developing diabetes when this was identified in the matched patient with diabetes.
The 1,070 couples (formed by a patient with diabetes and his/her match without diabetes) who experienced a hospitalization for CV disease during the step-1 follow-up were excluded. The remaining 3,321 diabetic and non-diabetic matched pairs represented the step-2 cohort, each of its members accumulating person-years of follow-up from the date of the first antidiabetic drug prescription (i.e., the step-2 index date), until the earliest among the dates of step-2 outcome (first admission to public or private hospitals for macrovascular complications), death from any cause, emigration, or December 31th 2012. Macrovascular complications included myocardial infarction, peripheral vascular disease, myocardial revascularization, heart failure and cerebrovascular disease, as reported by the diagnosis at discharge from hospital.
All prescriptions of statins dispensed to the cohort members during the step-1 follow-up were identified. Adherence with therapy was quantified as the cumulative number of days during which the medication was available divided by the number of days of follow-up, i.e., the "proportion of days covered" (PDC) by treatment.
A two-stage data analysis was carried out. In the first stage, we looked for replication of our previous findings [***12***] that increasing the level of adherence with statins increases the risk of developing diabetes (the step-1 outcome), using the step-1 cohort. Cox proportional hazard regression was used to estimate the hazard ratio (HR), and the 95% CI, for patients to develop diabetes in relation to the exposure to statin treatment. The exposure to statin treatment was quantified by four PDC categories, i.e. a very low (PDC < 10%), low (PDC from 11% to 50%), intermediate (PDC from 51% to 89%) and high (PDC ≥ 90%) adherence with statins. Because drug exposure may vary over time, adherence categories were included in the model as time-dependent variables, thereby accounting for their cumulative and varying nature. Data were also adjusted for 1) the type of the index prescribed statin; 2) previous use of antihypertensive, antiarrhythmic, antithrombotic and antidepressant agents, as well as of no-steroidal anti-inflammatory drugs and drugs used for chronic obstructive lung disease; and 3) the Charlson comorbidity index score, calculated in the three years prior the step-1 index date.
The second-stage data analysis focused on whether the greater or lesser chance for diabetes to be induced by statins translated into a different risk of macrovascular complications. To this aim, the 3,321 matched pairs forming the step-2 cohort were stratified according to the four categories of adherence with statins used for the step-1 follow-up. The overall risk of macrovascular complications associated with the diabetic condition was calculated for the entire cohort, as well as separately in each category of adherence with statins, using the HR, and 95% CI, derived from the Cox proportional hazard model. The risk of each specific macrovascular complication was also. Data were adjusted for the aforementioned covariates (excluding matching variables) as time-fixed. At the step-2 index date, covariates also included adherence with statins during the step-1 follow-up as well as with antidiabetic agents during the step-2 follow-up.

Two ancillary analyses were performed to check whether interpretation of the risk of macrovascular complications in diabetic patients with a different adherence to statin might be affected by (i) a delayed protective effect of statin dispensed during the step-1 follow-up and/or (ii) the exclusion of patients with CV events during the step-1 follow-up or before the inclusion in the study, with thus a selection of those less susceptible to the protective effect of the drug.
The Statistical Analysis System Software (version 9.4; SAS Institute, Cary, North Carolina, USA) was used for the analyses. For all hypotheses tested, two-tailed p-values less than 0.05 were considered to be significant.

Results
The 84,828 patients belonging to the step-1 cohort accumulated 467,317 person-years of follow-up, on average 5.5 years per patient. During this period, 6,935 patients started antidiabetic therapy (the step-1 outcome) with an incidence of 14.8 cases every 1,000 person-years. The risk of developing diabetes raised progressively and significantly as adherence with statin increased. Compared to patients with very low adherence, the increase was 24% (95% CI: 12% to 37%), 72% (95% CI: 56% to 90%), and 95% (95% CI: 60% to 139%) for patients with low, intermediate and high adherence, respectively.
The 3,321 matched pairs of diabetic and non-diabetic patients belonging to the step-2 cohort accumulated 33,623 person-years of follow-up, on average 5.0 years per patient. During this period, 376 and 272 diabetic and non-diabetic patients respectively experienced the step-2 outcome, the corresponding incidence being 22.9 and 16.1 hospitalizations every 1,000 person-years. The diabetes - macrovascular disease association experienced during the step-2 follow-up from the entire cohort of diabetic and non-diabetic patients, as well as within each category of step-1 adherence with statins, is shown in ***Figure 1***. In the entire cohort, patients with diabetes had a risk excess of macrovascular complications 35% (15% to 58%) higher than that of patients without diabetes. As shown in the top panel of ***Figure 1***, the potential of diabetes for generating macrovascular complications decreased with the increasing adherence to statins, the risk excess being 88% (95% CI: 17% to 204%), 51% (18% to 94%), 16% (-8% to 46%) and 4% (-64% to 197%) for very low, low, intermediate and high adherence, respectively. P-value for heterogeneity (p-value=0.0217) raises the significance. This trend was similar for the specific events separately considered, for all of which the diabetes-related risk was much less in patients with a high as compared to low adherence: 111% vs 9% for myocardial infarction, 112% vs 16% for myocardial revascularization and 59% vs 8% for cerebrovascular disease (Figure 1, bottom panel).
The results of the ancillary analysis further confirmed the results obtained.

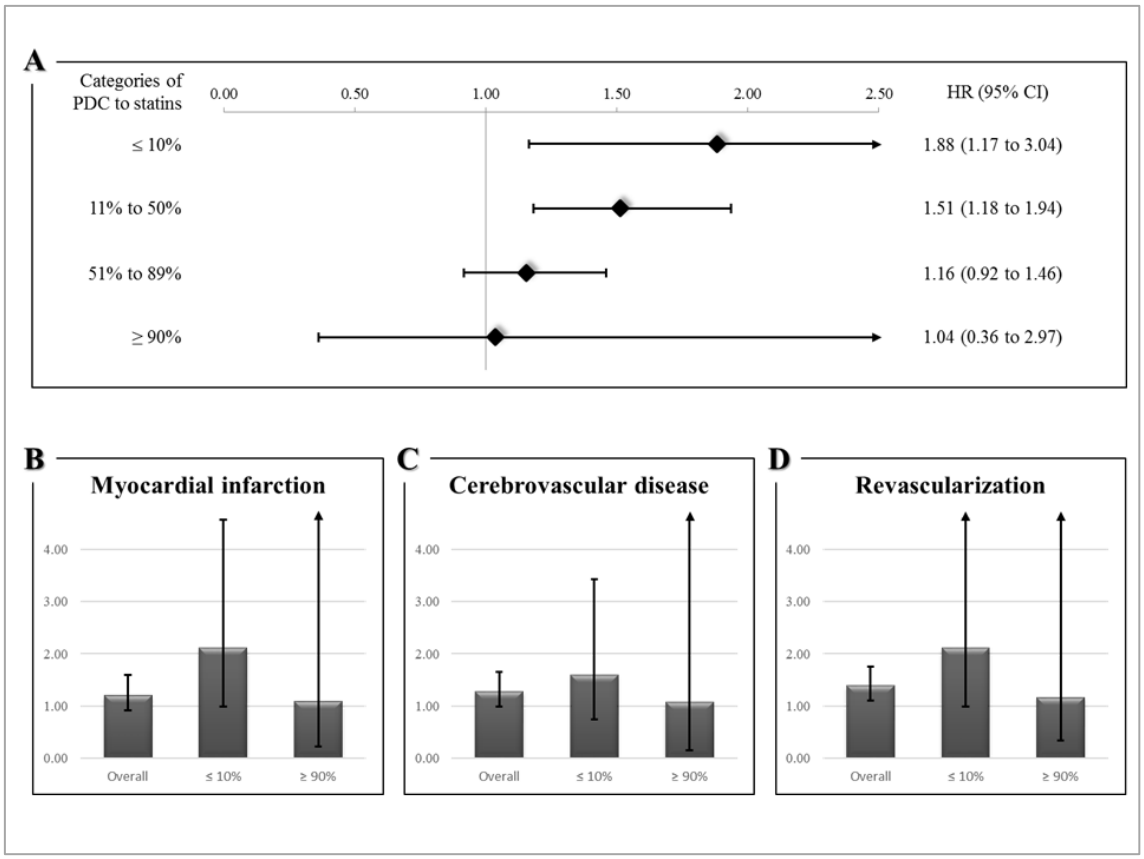

Figure 1. Effect of diabetes on the hazard ratio (HR) of hospitalization for macrovascular complications according to step-1 adherence with statin therapy. Macrovascular complications on the whole and specific macrovascular outcomes (i.e., myocardial infarction, cerebrovascular disease and myocardial revascularization) are shown in top and bottom panel respectively.

Conclusions

In summary, our data confirm that there is a definite increase in the development of diabetes with statin therapy (as observed during the step-1 follow-up). Type 2 diabetes lost its association with increasing macrovascular risk when previous adherence with statin treatment was very high, and thus the chance of its induction by the drug greater. Statin-dependent type 2 diabetes might be prognostically less adverse than native diabetes.Trials reflecting the clinical relevance of treatment-induced diabetes mellitus compared to naïve diabetes mellitus regarding macrovascular complications are required to confirm this finding.

References

[1] Dormuth CR, Filion KB, Paterson JM, et al, Canadian Network for Observational Drug Effect Studies Investigators. Higher potency statins and the risk of new diabetes: multicentre, observational study of administrative databases. BMJ 2015;348:g3244.

[2] Navarese EP, Szczesniak A, Kolodziejczak M, et al, Statins and risk of new-onset diabetes mellitus: is there a rationale for individualized statin therapy? Am J Cardiovasc Drugs 2014;14:79-87.

[3] Beckett RD, Schepers SM, Gordon SK. Risk of new-onset diabetes associated with statin use. SAGE Open Med, 2015;3:2050312115605518.

[**4**] Shah RV, Goldfine AB. Statins and risk of new-onset diabetes mellitus. Circulation 2012;126:e282-4.
[**5**] Bang CN, Okin PM. Statin treatment, new-onset diabetes, and other adverse effects: a systematic review. Curr Cardiol Rep 2014;16:461-465.
[**6**] Banach M, Malodobra-Mazur M, Gluba A, et al, Statin therapy and new-onset diabetes: molecular mechanisms and clinical relevance. Curr Pharm Des 2013;19:4904-12.
[**7**] Ridker PM, Danielson E, Fonseca FA, et al, JUPITER Study Group. Rosuvastatin to prevent vascular events in men and women with elevated C-reactive protein. N Engl J Med 2008;359:2195-207.
[**8**] Bell DS, DiNicolantonio JJ, O'Keefe JH. Is statin-induced diabetes clinically relevant? A comprehensive review of the literature. Diabetes Obes Metab 2014;16:689-94.
[**9**] Whayne TF Jr. Problems and possible solutions for therapy with statins. Int J Angiol 2013;22:75-82.
[**10**] Wang KL, Liu CJ, Chao TF, et al, Risk of new-onset diabetes mellitus versus reduction in cardiovascular events with statin therapy. Am J Cardiol 2014;113:631-6.
[**11**] Belalcazar LM, Raghavan VA, Ballantyne CM. Statin-induced diabetes: will it change clinical practice? Diabetes Care 2009;32:1941-3.
[**12**] Corrao G, Ibrahim B, Nicotra F, et al, Statins and the risk of diabetes: evidence from a large population-based cohort study. Diabetes Care 2014;37:2225-32.

DRUG SWITCHABILITY: A SYSTEMATIC REVIEW OF OBSERVATIONAL STUDIES.

Belleudi Valeria, Trotta Francesco, Vecchi Simona, Amato Laura, Addis Antonio, Davoli Marina

Department of Epidemiology, ASL Roma 1, Lazio Region

Introduction

The continuous patent expiration of important and costly medicines, such as biologics, may represent a great opportunity for the sustainability of the National Health Services (NHSs). However, the acceptance of off-patent drugs (generics or biosimilars) in the medical community as well as by patients continues to be limited. In the naïve population the evidence based on RCTs and observational studies, in particular new-users cohort [***1, 2***] approach is growing. Results from these studies are reducing the skepticism in the prescription of the off-patent medicines while the interchangeability between originators and off-patent drugs still remains an open issue. Recent EMA-guidance [***3***] specifies that interchangeability, switching and substitution are under the jurisdiction of member states. Evaluating switchability between drugs in experimental context requires specific and well-design RCTs studies [***4***]. Moreover clinical trials usually involve limited sample which may impact on transferability of results and the conduction of these studies requires a long time span [***5***]. Thus, an important challenge is to evaluate how switchability in observational contextutilizing administrative data is analyzed and the characteristics of the different study designs available.

To develop a systematic review of observational studies evaluating drug switchability to identify both methodological issue and study designs to answer this research question.

Methods

PubMed, EMBASE, and Web of Science were systematically searched to identify studies published up February 2017. No language limits was applied. We included switchability observational studies assessing effectiveness/safety outcomes and compliance in patients with any disease; a comparator group was required. Studies with a sample size less or equal 500 patients were excluded. Two reviewers independently screened the title and abstract of studies retrieved from the search. Study characteristics such as: year of publication, source of data, study design, disease, sample size, intervention/comparison groups and adjustment technique were extracted. Strategies to control the confounding, in particular, previous drug use and switching reasons were considered. All findings were summarized in a descriptive analyses.

Results

32 studies out of 2563 identified, met the inclusion criteria. All studies were published in the last 10 years, 67% in the last 5 years. Rheumatology, epilepsy and cardiovascular diseases were the most frequently represented. The sample size ranged from 616 to 81,356. The 79% of the studies reported data on effectiveness or safety outcomes, while only 21% on compliance outcomes. The most frequent study design was cohort (67%) followed by case-crossover (18%) and case-control/nested case-control (15%). In particular, in the cohort studies switcher had two different comparators: no switcher (64%) or another switching category (32%). When no switcher comparator was present, the index date (of no switcher) was defined using a random dispensing or considering an exposure time similar to switcher.

Among cohort studies the adjustment technique most frequent utilized was multivariate regression (59%), followed by propensity score matching (27%). All studies considered the previous drug use as a key factor for the analyses; 59% tested this as a confounding factor in terms of number of prescriptions, duration of treatment or adherence/persistence (e.g. medical possession rate). In the remaining 41% of studies, the

confounding associated with previous drug use was solved selecting only adherent patients. Three main reasons for switching were considered: lack of efficacy, adverse event or low adherence. As indirect reason, clinical parameters or previous occurrence of outcomes were measured to identify switching connected with lack of efficacy or adverse events. In this context, it was adopted usually a time-windows very close to the index date. However, around 30% of cohort studies did not report any information that allow direct or indirect reasons for switching.

Conclusions

The evidence on the observational studies evaluating switchability of drugs is recent and relative to specific clinical areas. The most used study design is the cohort. The heterogeneity in terms of design and methodology adopted in the included studies, highlight the lack of consensus on a study design capable to address all clinical questions associated with the drug switching. Previous drug use and switching reasons are the key factors to be considered in planning these observational studies. This systematic review allows to identify specific strategies which could be taking into account to remove confounding and bias in observational study on drug switchability.

References

[1] Schneeweiss S. A basic study design for expedited safety signal evaluation based on electronic healthcare data. Pharmacoepidemiol Drug Saf 2010;19(8):858-68.

[2] Suissa S, Moodie EE, Dell'Aniello S. Prevalent new-user cohort designs for comparative drug effect studies by time-conditional propensity scores. Pharmacoepidemiol Drug Saf 2017;26(4):459-68.

[3] European Medicines Agency. Biosimilars in the EU Information guide for healthcare professionals. 2017. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Leaflet/2017/05/WC500226648.pdf

[4] Faccin F, Tebbey P, Alexander E, et al. The design of clinical trials to support the switching and alternation of biosimilars. Expert Opinion on Biological Therapy 2016;16(12):1445-53.

[5] Jørgensen KK, Olsen IC, Goll GL, et al. Switching from originator infliximab to biosimilar CT-P13 compared with maintained treatment with originator infliximab (NOR-SWITCH): a 52-week, randomised, double-blind, non-inferiority trial. Lancet 2017;389(10086):2304-16.

CONFRONTO TRA CASO CONTROLLO INNESTATO, CASE CROSSOVER E SELF CONTROLLED CASE SERIES PER VALUTARE L'ASSOCIAZIONE TRA ANTIDEPRESSIVI E FRATTURE.

Airoldi Chiara, Scotti Lorenza, Corrao Giovanni

Healthcare Research & Pharmacoepidemiology, Università Milano Bicocca

Introduzione

La depressione è una malattia comune caratterizzata da tristezza persistente, repulsione alle attività che influenza pensieri e comportamenti [***1***]. La malattia rappresenta un importante problema di salute pubblica a causa della sua alta prevalenza (2-15% durante la vita) e alla disabilità ad essa associata [***2***].
I farmaci antidepressivi (AD) sono frequentemente utilizzati per alleviare i sintomi della patologia e tra i più comuni vi sono i triciclici (TCA) e gli inibitori selettivi della ricaptazione della seratonina (SSRI). In passato i TCA rappresentavano lo standard nella cura della depressione mentre attualmente hanno una maggiore diffusione gli SSRI in quanto più tollerati e caratterizzati da un miglior profilo di sicurezza. I farmaci AD hanno infatti una provata efficacia ma alcuni studi hanno osservato un incremento di rischio di eventi cardiovascolari, cadute e conseguenti fratture tra gli utilizzatori di questi farmaci [***3-5***]. Per quanto riguarda il rischio di cadute, l'utilizzo di AD, infatti, può causare ipotensione ortostatica, aritmia e confusione a causa di effetti anticolinergici e dopaminergici che possono compromettere l'equilibrio e il controllo motorio [***6***].
La maggior parte degli studi osservazionali che hanno stimato l'associazione tra utilizzo di antidepressivi e fratture sono basati su disegni classici quali coorte e caso-controllo innestato in una coorte (CCI) [***5, 7***]. Le stime ottenute utilizzando questo tipo di disegni potrebbero essere distorte sia a causa del confondimento residuo dovuto all'impossibilità di aggiustare le stime di associazione per fattori non misurati o non misurabili e sia per la potenziale scarsa comparabilità tra i gruppi a confronto (e.g. esposti e non esposti ad un farmaco). Dal 1990 alcuni disegni (case-only), quali il case crossover (CC) e il self controlled case series (SCCS), sono stati introdotti con l'obiettivo di minimizzare l'impatto dei confondenti non tempo-dipendenti non misurati e/o non misurabili [***8, 9***]. La peculiarità di questi approcci è l'inclusione dei soli soggetti che hanno sperimentato l'evento di interesse e l'esposizione viene quindi confrontata entro pazienti e non tra pazienti. Il problema del confondimento non misurato potrebbe essere particolarmente rilevante negli studi basati sui database sanitari amministrativi in quanto consentono di effettuare studi su popolazioni ampie e non selezionate ma mancano di alcune informazioni riguardanti potenziali fattori confondenti legati agli stili di vita.
Valutare la relazione tra l'utilizzo di antidepressivi (SSRI e TCA) e il rischio di fratture confrontando i risultati di uno studio caso-controllo innestato in una coorte (CCI) con le stime ottenute applicando il CC e il SCCS disegni di studio caratterizzati da una diversa capacità di controllo del confondimento utilizzando i dati provenienti dai database amministrativi della regione Lombardia.

Metodi

La popolazione target è rappresentata da tutti i soggetti di età maggiore di 18 anni residenti in Lombardia. Per il disegno CCI sono stati selezionati tutti coloro che avevano ricevuto una prima prescrizione di antidepressivi (ATC N06A) tra il 1/1/2005 e il 31/12/2007. La data della prima prescrizione è definita come la data di ingresso nella coorte.
Sono stati esclusi i soggetti che nei 5 anni precedenti la data di ingresso avevano avuto almeno una prescrizione di AD, erano stati ricoverati per fratture ossee o patologiche, avevano avuto almeno una prescrizione di bifosfonati, terapia ormonale sostitutiva o altri farmaci ossei così da poter includere solo i nuovi utilizzatori di AD, le sole fratture incidenti e i soggetti a basso rischio di frattura. I pazienti che l'anno

precedente la data di ingresso avevano avuto un ricovero con diagnosi di depressione, evento traumatico, malattia di Paget, osteomalcia, sindrome di Cushing, ipertiroidismo, iperparatiroidismo, celiachia sono stati esclusi per poter analizzare gli individui con depressione incidente ma liberi da patologie associate alle fratture. Sono stati, inoltre, esclusi i pazienti con diagnosi di tumore maligno nei 5 anni precedenti la data di ingresso in quanto i pazienti oncologici hanno un rischio superiore di frattura rispetto alla popolazione generale. Infine, sono stati esclusi i soggetti con meno di 90 giorni di follow-up.
I membri della coorte sono stati seguiti dalla prima prescrizione di antidepressivo fino al più precoce dei seguenti eventi: primo ricovero per frattura (ICD-9 CM: 800-829), morte, migrazione o termine del follow-up (31/12/2010). I casi sono rappresentati da tutti i soggetti che hanno sperimentato un ricovero per frattura durante il follow-up e la data della prima frattura è stata indicata come data indice. Ad ogni caso sono stati appaiati casualmente fino a 10 controlli per sesso, età, data di ingresso scelti casualmente dalla popolazione ancora a rischio di sviluppare l'evento. Ogni soggetto è stato categorizzato in base all'esposizione in: utilizzatore corrente di ogni specifica classe di AD (TCA: ATC N06AA, SSRI: ATC N06AB o altri farmaci antidepressivi: ATC N06AF, N06AG, N06AX) se almeno uno dei 30 giorni precedenti la data indice era coperto dal trattamento, utilizzatore recente (30-60 giorni prima della data indice) e utilizzatore passato (>60 giorni prima della data indice). I potenziali confondenti analizzati riguardano l'utilizzo dei farmaci anti-infiammatori, anti-psicotici, bifosfonati, corticosteroidei (orali), terapia ormonale sostitutiva, inibitori della pompa protonica, sedativi e narcotici e la presenza di artrite, malattie ossee, malattie cardiovascolari, cataratta e glaucoma, diabete, epilessia, ipertensione, ipotensione ortostatica, disturbi mentali, malattie polmonari ostruttive, Parkinson, malattie tiroidee e indice di Charlson.
Per i disegni case-only sono stati selezionati i soli soggetti che tra il 01/01/2005 e il 31/12/2010 erano stati ricoverati con diagnosi di frattura. Sono stati applicati gli stessi criteri di esclusione utilizzati nel CCI considerando periodi di 1 o di 5 anni precedenti. Nel CC sono stati esclusi anche i soggetti che avevano avuto la frattura nei primi 90 giorni di follow up e che non avevano neanche una prescrizione di AD dall'01/01/2005. Nel SCCS sono stati invece esclusi i pazienti con un periodo di follow up inferiore a 90 giorni, inoltre se la distanza tra le date di ricovero di due fratture successive era inferiore a 14 giorni l'evento di frattura è stato considerato come unico.
Nel CC l'esposizione per ogni individuo è stata valutata in un periodo immediatamente precedente la frattura (30 giorni prima) rispetto a un periodo più indietro nel tempo (60-90 giorni) mantenendo la distinzione tra utilizzo di SSRI, TCA o altro. Un periodo di washout è stato incluso per eliminare l'effetto carryover. I potenziali confondenti analizzati riguardano l'utilizzo dei farmaci anti-infiammatori, anti-psicotici, corticosteroidei (orali), inibitori della pompa protonica, sedativi e narcotici e la presenza di epilessia.
Nel Self Controlled Case Series (SCCS) il follow up è stato diviso in differenti finestre: ad alto rischio nei primi 30 giorni dopo l'assunzione di antidepressivo, a rischio nel restante tempo in cui un soggetto è coperto e baseline quando non è esposto. Nel periodo ad alto rischio si mantiene la stratificazione per i differenti farmaci assunti.
Sono stati implementati un modello di regressione logistica condizionata e un modello di Poisson per stimare rispettivamente gli Odds Ratio aggiustati (OR), i Rapporti tra Tassi (RT) e i corrispondenti intervalli di confidenza (IC) al 95% dell'associazione tra uso di antidepressivi e rischio di frattura.
L'efficienza relativa degli stimatori nei diversi disegni è stata comparata tramite il rapporto degli errori quadratici medi.

Risultati

Nel CC sono stati inclusi 9975 casi e 99497 controlli, nel CC rispettavano i criteri di inclusione 24067 casi mentre nel SCCS 24920 soggetti per un totale di 28836 fratture.
Gli OR, RT e corrispondenti 95%Ci ottenuti dalle analisi relative ai 3 studi per i vari livelli di esposizione sono riportati in ***figura 1***.

I risultati nel CCI mostrano un incremento di rischio di fratture tra gli utilizzatori correnti di SSRI (ORa: 1.75; IC95% 1.67-1.84) e TCA (ORa: 1.63; IC95% 1.29-2.05) rispetto agli utilizzatori passati (riferimento). Risultati analoghi si osservano nel CC: gli ORa sono 1.58 (IC95% 1.45-1.73) e 1.22 (IC95% 0.91-1.64) per utilizzatori di TCA e SSRI rispettivamente. Questi risultati sono poi confermati anche dal SCCS: gli RT stimati sono 1.60 (IC95% 1.51-1.69) e 1.57 (IC95% 1.25-1.98) per SSRI e TCA nel periodo ad alto rischio rispetto al baseline (riferimento). Nel CCI e nel SCCS è stato possibile anche valutare l'esposizione recente e nel periodo a rischio di antidepressivi: le stime sono positive e pari a 1.25 (IC95% 1.13-1.39) e 1.70 (IC95% 1.63-1.78) rispettivamente. Il SCCS ha un'efficienza comparabile a quella del CCI con rapporti delle varianze pari a 1 mentre le stime ottenute dal CC presentano una precisione minore con IC molto ampi tanto da non raggiungere la significatività statistica durante l'esposizione a TCA.

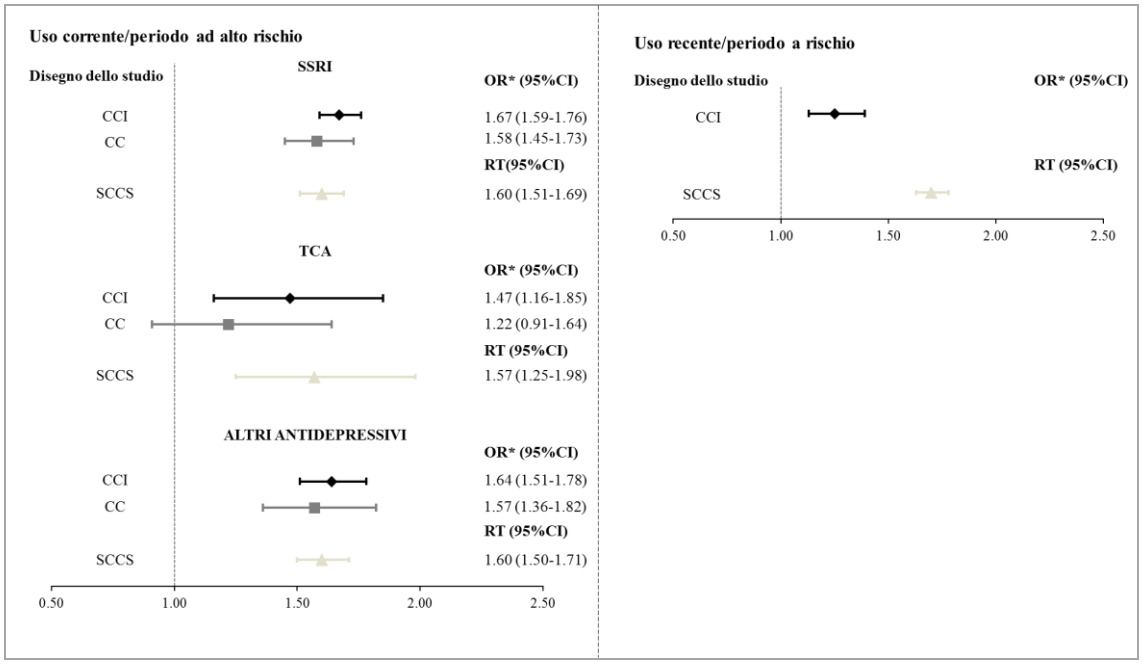

Figura 1. Forest plot dell'associazione tra utilizzo di antidepressivi e rischio di fratture distinto. Nella parte di sinistra sono riportate le stime dell'uso corrente o nel periodo ad alto rischio distinguendo per esposizione a SSRi, TCA o altro. Nella parte di destra sono riportate le stime dell'uso recente o nel periodo a rischio per la generica esposizione ad antidepressivo. Nel CCI gli OR sono aggiustati per uso di anti-infiammatori, anti-psicotici, bifosfonati, corticosteroidei (orali), terapia ormonale sostitutiva, inibitori della pompa protonica, sedativi e narcotici e per presenza di artrite, malattie ossee, malattie cardiovascolari, cataratta e glaucoma, diabete, epilessia, ipertensione, ipotensione ortostatica, disturbi mentali, malattie polmonari ostruttive, Parkinson, malattie tiroidee e indice di Charlson. Nel CC gli OR sono aggiustati per uso di anti-infiammatori, anti-psicotici, corticosteroidei (orali), inibitori della pompa protonica, sedativi e narcotici e per presenza di epilessia.

Conclusioni

I risultati ottenuti dal confronto dei tre diversi disegni mostrano un incremento di rischio di fratture negli utilizzatori di SSRI e TCA.

Le stime puntuali ottenute nelle due classi di antidepressivi non sono significativamente diverse tra di loro ma, contrariamente a quanto proposto dalla letteratura, i TCA sembrano avere una miglior sicurezza e un numero inferiore di eventi avversi. Tale risultato potrebbe essere giustificato dalla normale pratica clinica: i medici prescrivono con maggior frequenza gli SSRI e soprattutto prediligono il loro utilizzo nei soggetti più gravi e con una fragilità superiore.

Le associazione trovate utilizzando gli studi case only sono simili ai risultati ottenuti nel classico CCI e la perdita di efficienza del SCCS è trascurabile. Tali risultati suggeriscono come i disegni entro soggetti siano strumenti validi soprattutto quando si hanno a disposizione database amministrativi. Le banche dati utilizzate (ricoveri ospedalieri, delle prescrizioni farmaceutiche territoriali e dell'anagrafe) permettono infatti di studiare ampie popolazioni descrivendo la reale pratica clinica ma in esse sono assenti potenziali confondenti che con i disegni case only siamo in grado di controllare.

Bibliografia

[1] Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5). American Psychiatric Association, 2013

[2] Moussavi S, Chatterji S, Verdes E, et al. Depression, chronic diseases, and decrements in health: results from the World Health Surveys. Lancet 2007;370(9590):851-8.

[3] Arroll B, Macgillivray S, Ogston S, et al. Efficacy and tolerability of tricyclic antidepressants and SSRIs compared with placebo for treatment of depression in primary care: a meta-analysis. Ann Fam Med 2005;3(5):449-56.

[4] Biffi A, Scotti L, Corrao G. Use of antidepressants and the risk of cardiovascular and cerebrovascular disease: a meta-analysis of observational studies. Eur J Clin Pharmacol 2017;73(4):487-97.

[5] Rabenda V, Nicolet D, Beaudart C, et al. Relationship between use of antidepressants and risk of fractures: a meta-analysis. Osteoporos Int 2013;24(1):121–37.

[6] Liu B, Anderson G, Mittmann N, et al. Use of selective serotonin-reuptake inhibitors or tricyclic antidepressants and risk of hip fractures in elderly people. Lancet 1998;351(9112):1303-7.

[7] Souverein P, Abbing-Karahagopian V, Martin E, et al. Understanding inconsistency in the results from observational pharmacoepidemiological studies: the case of antidepressant use and risk of hip/femur fractures. Pharmacoepidemiol Drug Saf 2016;25(Suppl. 1):88–102.

[8] Maclure M, The case-crossover design: a method for studying transient effects on the risk of acute events. Am J Epidemiol 1991;133(2):144-53.

[9] Farrington P, Pugh S, Colville A, et al. A new method for active surveillance of adverse events from diphtheria/ tetanus/pertussis and measles/mumps/rubella vaccines. Lancet 1995 Mar;345(8949):567-9.

RANDOMIZED CONTROLLED TRIAL ON A NOVEL HUMAN MILK FORTIFIER FROM DONKEY MILK, IN FEEDING PRETERM INFANTS

Ansaldi Giulia[1], Coscia Alessandra[1], Cresi Francesco[1], Spada Elena[1,2], Milani Silvano[2], Moro Guido[3], Bertino Enrico[1]

1. *Neonatology and Neonatal Intensive Care Unit, Department of Public Health and Pediatrics, University of Turin, Turin, Italy*
2. *Laboratory of Medical Statistics, Biometry and Epidemiology "G.A. Maccacaro", Department of Clinical Sciences and Community Health – Università degli Studi di Milano, Milan, Italy*
3. *Italian Association of Human Milk Banks (AIBLUD)*

Introduction

Very preterm newborns (gestational age lower than 32 weeks) and Very Low Birthweight (VLBW) infants (birthweight lower than 1500 grams) currently represent the majority of patients cared in Neonatal Intensive Care Units (NICU). The increase of the survival rate for these newborns, due to improvements in perinatal care, has opened new perspectives regarding their outcome and has a significant impact on their health status in adulthood.

In these groups of infants, nutrition represents a fundamental factor not only for neonatal survival and short-term outcome, but also for long-term consequences and quality of life. The main issue is to ensure an adequate qualitative and quantitative nutrition, particularly in terms of protein intake, which is the main cause of post-natal growth deficit [***1***].

Human milk is the recommended food for all neonates including preterm infants [***2***]. Breast milk alone, however, does not meet the recommended nutritional needs in preterm infants [***3***]. The most common strategy is to cope with potential nutrient deficits by supplementing breast milk with additional nutrients (mainly proteins and minerals) to satisfy the special nutritional requirements of these infants [***4***]. Concerning the composition of fortifiers, fortification of human milk requires simple, high-quality and well-tolerated nutritional supplements. At present commercially available fortifiers are based on bovine milk, with a protein composition very different from that of human milk. Bovine milk (BM) protein intake has raised concerns because of its association with allergies [***5***] and a possible role as a trigger of intestinal inflammation in preterm neonates [***6***]. Recently human milk (HM) based fortifiers have been proposed, but their utilization is limited by high costs and ethical issues.

Consistently, milk from monogastric animals, rather than from ruminants, has been suggested during recent years to be more suitable for human nutrition based on their physiochemical properties [***7***]. In previous studies, we found that protein and lipid fraction of donkey milk are very similar to those of human milk [***8, 9***]. Besides we observed that donkey milk (DM) was well tolerated in a group of highly problematic cow's milk allergic children [***10***].

Our hypothesis is that feeding preterm infants with HM supplemented with fortifiers derived from DM will improve the feeding tolerance.

The purpose of the present randomized clinical trial is to compare the use of DM-derived fortifiers with commercial BM-based fortifiers in infants with birthweight ≤1500 g or gestational age <32 weeks in terms of nutritional tolerance.

Methods

The study was performed in the NICU of the University, City of Health and Science of Turin, and was approved by Ethics Committee. Informed written consent was obtained from parents before enrollment. The

inclusion criteria were: gestational age <32 weeks or birthweight ≤ 1500 grams, exclusive feeding with human milk (fresh own mother's or donor milk) and enteral feeding ≥80ml/kg/day reached within the first 4 weeks of life. The neonates affected by severe gastrointestinal pathologies (necrotizing enterocolitis, colostomy, intestinal obstruction, peritonitis, blood in the stools), chromosomal abnormalities or major malformations, metabolic diseases, intravascular disseminated coagulopathy (IDC), shock, patent ductus arteriosus (PDA) requiring medical care or surgery at time of randomization, and severe renal failure were excluded.

Study design. After informed written parental consent was obtained, infants were randomized 1:1 by a software-generated list in two arms: BF-arm (Bovine Fortifier) and DF-arm (Donkey Fortifier). Observation period was defined as 21 days since the beginning of fortification. The same Adjustable Fortification regimen was utilized in the two arms. In BF-arm, the human milk was fortified with commercial multi-component fortifier (FM85 Nestlè) and protein concentrate (Protifar Nutricia), derived from bovine milk, while in DF-arm with multi-component fortifier and protein concentrate derived from donkey milk (Fortilat), produced by an ultrafiltration process of pasteurized donkey milk in a pilot stainless steel.

The same nutritional protocol was followed for all study participants. Because the nurses in charge of feeds preparation and administration were also in charge of evaluating signs of feeding tolerance, it was not possible to achieve a double blindness of the intervention.

Advancing of enteral feeds was strictly regulated according to the feeding protocol adopted in our NICU, based on the evaluation of signs of feeding intolerance.

Data on necrotizing enterocolitis (NEC) occurred after randomization, sepsis, mortality, hospital stay duration, intraventricular haemorrhage (IVH), retinopathy of prematurity (ROP), were collected.

The criteria for hospital discharge were satisfactory weight gain while receiving full oral feeding, maintenance of adequate thermal stability and resolution of acute medical conditions.

Primary and secondary endpoints. The primary endpoint was the occurrence of at least one episode of feeding intolerance, defined as interruption of enteral feeding for at least 8 consecutive hours during the observation period. We also evaluated the total number of feeding intolerance episodes, the number of feeding interruption episodes of any duration, the total hours of enteral feeding interruption, the time required to reach full enteral feeding (150 ml/kg/day), and the short-term weight gain.

Study size. Before the beginning of this study, approximately 45% VLBWI or preterm infants admitted to our NICU had at least one episode of feeding intolerance. A 25% reduction in the frequency of the primary endpoint was regarded as the minimum clinically important difference (MCID): 62 infants per arm (planned study) had to be recruited to ensure an 80% study power, if the risk of type I error is set to the usual level of 5%. Since the occurrence of primary endpoint in our study population resulted to be much lower than that assumed in the protocol, and no adverse effect was observed, when the planned study size was achieved, it was decided to continue the enrollment until the end of the stock of FortiLat (extended study).

Statistical analysis. In the *intention-to-treat* population (ITT, all randomized infants) failure included the occurrence of at least one episode of feeding intolerance, NEC, death, and transferred before 21 days of observation. Subjects discharged at home before 21 days were instead considered successes, assuming that these babies had maintained good tolerance also at home. Time-depending endpoints were evaluated only on subjects observed for 21 days from enrollment (*per protocol* population). The difference in the occurrence of failure (primary endpoint) between the two arms of the trials was tested with Fisher's exact test. A Poisson linear model was used to analyse the number of feeding intolerance episodes and the number of feeding interruption episodes (of any duration); a general linear model was used to analyse, after rank transformation, the total hours of enteral feeding interruption; median time required to reach full enteral feeding was estimated according to Kaplan and Meyer. Weight was expressed as SDS (standard deviation score) on the basis of INeS charts [***11***]. The short term auxological outcome was expressed as weight-ΔSDS/day, i.e. the mean daily weight-SDS variation between weight-SDS at discharge and

weight-SDS on the first day of exclusive enteral feeding, and anlysed with a general linear model. Primary and secondary endpoints were evaluated separately for the planned and for the extended study.

Results

The initial preparation time for this study took 8 months. In this period, the study protocol was designed, the donkey derived fortifiers were developed, and the new products were tested to assess solubility and to certify safety. The study protocol was submitted to the Ethics Committee of our hospital and the health care staff of the NICU was trained.

A total of 168 infants were assessed for eligibility, 124 of which were enrolled (Planned study: BF-arm n= 62; DF-arm n=62). Further 34 infants were assessed for eligibility and, after the exclusion of 2 of them, a total of 156 infants was enrolled (Extended study: BF-arm n= 79, DF-arm n=77). Infants small-for-gestational-age (SGA: birthweight<10th centile) were slightly more numerous in DF than in BF-arm (BF *vs* DF: 25.8% *vs* 32.3% in *planned sample*, 24.4% *vs* 35.1% in *extended sample*), as well as Infants with Intra Uterine Growth Restriction diagnosis (IUGR) (BF *vs* DF: 27.9% *vs* 37.1% in *planned sample*, 25.6% *vs* 44.2% in *extended sample)*. One baby from each arm died following NEC. Patent Ductus Arteriosus (PDA) before randomization was more frequent in BF-arm (BF *vs* DF: 32.2% *vs* 17.7% in *planned sample*, 32.9% *vs* 14.3% in *extended sample* even if at the time of randomization it was a clinical condition already solved.

Primary endpoint. The primary endpoint was estimated in both the ITT population and the *per-protocol* population. In both populations, risk of failure in the *planned sample* has resulted lower in DF-arm than in BF-arm with a Relative Risk Reduction (RRR) [CI(95%)] of 0.40 [-0.27; +0.72] in the ITT and 0.63 [-0.29; +0.90] in the *per-protocol* population. The trend of these results was similar in the *extended sample*: 0.46 [-0.09; +0.73] in the ITT and 0.58 [-0.27; +0.86] in the *per-protocol* population.

Secondary endpoints. Secondary endpoints are time depending and, for this reason, were estimated on the *per protocol* population. The mean [IC(95%)] of the total episodes of feeding intolerance and the mean [IC(95%)] of enteral feeding interruptions of any duration was lower in DF-arm (0.11 [0.01; 0.21] and 0.27 [0.12; 0.42] respectively) than in BF-arm (0.20 [0.07; 0.34] and 0.39 [0.20; 0.57] respectively). The time (days) to reach full enteral feeding was overlapping in BF and DF arms. Indeed the median days [IC(95%] resulted 19 [15;22] in BF- and 19 [15;23] in DF-arm in *planned sample*, and 19 [15;22] in both arms in *extended sample*. No differences were observed between BF-arm and DF-arm about total number of hours of feeding interruptions, neither in the planned sample (p=0.456) nor in the extended sample (p=0.451). Despite the non-significant results, the trend for these secondary endpoints was further improving in favor of the donkey milk based fortifier.

The short-term auxological outcome (weight ΔSDS/day) did not differ between the two arms, mean [IC(95%)] being: BF -0.013 [-0.018; -0.009] and DF -0.012 [-0.016; -0.008]. The results were similar in the *extended sample*. Mean [IC(95%)]:: BF -0.012 [-0.016; -0.008], DF -0.013 [-0.018; -0.008].

Conclusions

The aim of the study is to assess the effects of a new human milk fortifier derived from donkey milk on VLBWI feeding tolerance. To the best of our knowledge our trial is unique, as there are no other studies investigating the use of a fortifier derived from animal species other than bovine for the nutrition of preterm and VLBWI. In our trial all the infants (both in the BF-arm and in the DF-arm) received exclusively human milk (fresh own mother's milk or pasteurized donor milk) without any preterm formula supplementation. For instance the trial of Sullivan [***12***], comparing a human milk based fortifier with a bovine milk-based fortifier, included in the group supplemented with the bovine fortifier also subjects fed with preterm formula. This represents a confounding variable masking the effects of the two different fortifiers as a sole supplement of human milk. Moreover, in that study the incidence of NEC was very high (15%) in the bovine group while in the human milk based fortifier- group it was similar to reports from units using bovine's milk based fortifier

[***13***]. Our first results about donkey milk based fortifiers are promising. Compared to the bovine fortifiers, the donkey milk based fortifier seems to have a better feeding tolerance even if these results are not significant. At baseline, the incidence of the total episodes of feeding intolerance resulted surprisingly lower than that assumed in the protocol. A possible explanation is that the introduction of a RCT may itself influence and improve the behavior of clinical staffs. This effect is generally referred to as "Hawthorne effect", widely discussed in the literature [***14, 15***]. Nutrition is managed closely and consistently in our NICU, and our current nutrition practices, mostly since the introduction of a specific feeding-surveillance protocol during our RCT, has first demonstrated excellent improvements: all these practices could probably have led to the lower incidence of the total episodes of feeding intolerance, as previously reported. When we extended the study including more patients (extended sample), we observed a further slight improvement in feeding tolerance in favor of the DF-arm. We can therefore hypothesize that increasing the sample size we would observe a significant better feeding tolerance in DF-arm.
To confirm our data it will be useful to perform a multicenter study in order to increase the sample size. It will be interesting to analyze metabolic, auxological and neurodevelopmental outcomes in order to determine if the use of donkey milk fortifiers will have also long-term benefits.

References

[**1**] Dinerstein A, Nieto RM, Solana CL, *et al*. Early and aggressive nutritional strategy (parenteral and enteral) decreases postnatal growth failure in very low birth weight infants. J Perinatol. 2006;26(7):436-442.
[**2**] Moro GE, Arslanoglu S, Bertino E, *et al*. XII. Human Milk in Feeding Premature Infants: Consensus Statement. J Pediatr Gastroenterol Nutr. 2015;61 Suppl 1:S16-19.
[**3**] Agostoni C, Buonocore G, Carnielli VP, *et al*. Enteral nutrient supply for preterm infants: commentary from the European Society of Paediatric Gastroenterology, Hepatology and Nutrition Committee on Nutrition. J Pediatr Gastroenterol Nutr. 2010;50(1):85-91.
[**4**] Dutta S, Singh B, Chessell L, et *al*. Guidelines for feeding very low birth weight infants. Nutrients. 2015;7(1):423-442.
[**5**] Sampson HA, Aceves S, Bock SA, *et al*. Food allergy: a practice parameter update-2014. J Allergy Clin Immunol. 2014;134(5):1016-1025.e1043.
[**6**] Abdelhamid AE, Chuang SL, Hayes P, Fell JM. Evolution of in vitro cow's milk protein-specific inflammatory and regulatory cytokine responses in preterm infants with necrotising enterocolitis. J Pediatr Gastroenterol Nutr. 2013;56(1):5-11.
[**7**] Barłowska J., Szwajkowska M., Litwińczuk Z., Król J. Nutritional Value and Technological Suitability of Milk from Various Animal Species Used for Dairy Production. Comprehensive Reviews in Food Science and Food Safety. 2011;10(6):291-302.
[**8**] Bertino E, Gastaldi D, Monti G, *et al.* Detailed proteomic analysis on DM: insight into its hypoallergenicity. Front Biosci (Elite Ed). 2010;2:526-536.
[**9**] Gastaldi D, Bertino E, Monti G, et al. Donkey's milk detailed lipid composition. Front Biosci (Elite Ed). 2010;2:537-546.
[**10**] Monti G, Bertino E, Muratore MC, *et al.* Efficacy of donkey's milk in treating highly problematic cow's milk allergic children: an in vivo and in vitro study. Pediatr Allergy Immunol. 2007;18(3):258-264.
[**11**] Bertino E, Spada E, Occhi L, et al.Neonatal anthropometric charts: the Italian neonatal study compared with other European studies. J Pediatr Gastroenterol Nutr. 2010;51(3):353-61
[**12**] Schanler RJ, Kim JH, *et al.* An exclusively human milk-based diet is associated with a lower rate of necrotizing enterocolitis than a diet of human milk and bovine milk-based products. J Pediatr. 2010;156(4):562-567.e561.

[**13**] Luoto R, Matomäki J, Isolauri E, Lehtonen L. Incidence of necrotizing enterocolitis in very-low-birth-weight infants related to the use of Lactobacillus GG. Acta Paediatr. 2010;99(8):1135-1138.

[**14**] McCambridge J, Witton J, Elbourne DR. Systematic review of the Hawthorne effect: new concepts are needed to study research participation effects. J Clin Epidemiol. 2014;67(3):267-277.

[**15**] McCambridge J, Kypri K, Elbourne D. Research participation effects: a skeleton in the methodological cupboard. J Clin Epidemiol. 2014;67(8):845-849.

RELAZIONI TRA EVENTI CARDIORESPIRATORI E REFLUSSO IN NEONATI CON SINTOMI DI GERD

Maggiora Elena[1], Martinelli Domenico[2], Dell'Anna Vito Andrea [1], Peila Chiara[1], Baricco Marta[1], Coscia Alessandra[1], Bertino Enrico[1], Cresi Francesco[1]

1. *S.C.D.U. di Neonatologia e Terapia Intensiva Neonatale, Dipartimento di Scienze della Sanità Pubblica e Pediatriche, Università di Torino, Torino, Italia*
2. *U.O. di Neonatologia e Terapia Intensiva Neonatale, Dipartimento di Scienze Biomediche e Oncologia Umana, Università di Bari "Aldo Moro", Bari, Italia*

Introduzione

I neonati ricoverati nelle unità di Terapia Intensiva Neonatale presentano spesso eventi cardiorespiratori (CR) quali apnea, bradicardia e desaturazione. Tali eventi rappresentano un problema di difficile gestione clinica che frequentemente ritarda il momento della dimissione.[***1, 2***] Il reflusso gastroesofageo (GER), cioè il passaggio del contenuto gastrico nell'esofago, è un fenomeno fisiologico nei neonati e generalmente si risolve spontaneamente nel primo anno di vita.[***3, 4***] La maggioranza degli episodi di GER sono brevi e asintomatici [***5***]. Tuttavia, sono spesso associati a rigurgito, vomito, e ruminazione, e possono essere correlati con eventi CR.[***1, 6***] La percezione di una relazione causale tra gli eventi CR e i sintomi GER è rafforzata dal fatto che entrambi i tipi di sintomi spesso coesistono nei neonati, soprattutto nei neonati prematuri.[***7, 8***] La presunta associazione causale tra eventi CR e GER è causa di un'eccessiva prescrizione di farmaci antiacidi che possono incrementare il rischio di sepsi e enterocolite necrotizzante.[***9–11***] In realtà, non è stata ancora chiarita la natura della relazione tra questi eventi. Sebbene alcuni autori abbiano dimostrato l'esistenza di una correlazione causale tra eventi CR e GER, [***6, 12, 13***] risultati di studi recenti rigettano tale ipotesi.[***8, 14, 15***] Sebbene l'associazione tra eventi GER e CR nei neonati sia generalmente accidentale, [***16***] la nostra ipotesi è che una associazione significativa e causale tra questi due tipi di eventi possa esistere in alcuni pazienti sintomatici con specifiche caratteristiche di GER.
Obiettivi Scopo dello studio è valutare l'associazione temporale tra eventi GER e CR in un grande numero di neonati sintomatici e individuare le caratteristiche del GER associati agli eventi CR.

Metodi

Abbiamo rivalutato i tracciati di tutti i neonati che sono stati sottoposti a monitoraggio CR sincronizzato con pH-impedenziometria multicanale intraluminale (MII/pH) ricoverati presso la S.C.D.U di Neonatologia e Terapia Intensiva Neonatale del presidio Sant'Anna-Regina Margherita da gennaio 2006 a dicembre 2016. Sono stati considerati eligibili i neonati monitorizzati per la presenza concomitante di eventi CR (apnea, desaturazione d'ossigeno e/o bradicardia) e sintomi di malattia da GER (criteri NASPGHAN). Sono stati inclusi nello studio i neonati che rispettavano i seguenti criteri d'inclusione: 1. Monitoraggio CR e MII/pH sincronizzato per più di 12 ore, escluso il tempo dei pasti; 2. Assenza di malattie infettive, genetiche, metaboliche e neurologiche; 3. Assenza di supporto ventilatorio e/o supplementazione d'ossigeno al momento del monitoraggio CR e MII/pH sincronizzato; 4. Assenza di terapia farmacologiche con effetti sul GER per almeno 1 settimana prima del monitoraggio CR e MII/pH sincronizzato. 5. Alimentazione orale esclusiva per suzione al seno o al biberon con latte materno o formulato; Per realizzare il monitoraggio CR è stato utilizzato il sistema VitaGuard VG3100 (Getemed Medizin und Information stechnik AG, Teltow, Germania), equipaggiato con Signal Extraction Technology (SET) (Masimo Corp., Irvine, CA, USA). Frequenza cardiaca, saturazione d'ossigeno transcutanea e frequenza respiratoria sono state misurate attraverso un sensore pulso-ossimetrico posizionato al polso o al piede destro e tre elettrodi cardiaci posizionati sul torace. I tracciati CR sono state analizzati visivamente da un operatore esperto, in cieco

rispetto alle registrazioni MII/pH, utilizzando il software VitaWin3®, versione 3.3. Gli eventi CR sono stati definiti come episodi di apnea della durata di più di 20 secondi, o più di 5 secondi se seguiti da desaturazione o bradicardia, episodi di desaturazione con saturazioni di O2 inferiori all'80%, ed episodi di bradicardia con frequenza cardiaca inferiore a 80 battiti per minuto. La valutazione degli eventi GER è stata eseguita attraverso il monitoraggio MII/pH (Sleuth monitoring system, the Canadian Scientific, Highlands Ranch, CO, USA). Il metodo utilizzato per l'identificazione e la caratterizzazione degli eventi GER è già stato descritto in un precedente lavoro del nostro gruppo.[4] Sono state valutate le seguenti caratteristiche: – Frequenza dei MII-GER:episodi di reflusso/ora; – Livello di risalita del bolo (BRE): livello prossimale raggiunto dal reflusso, indicato dal numero di canali coinvolti consecutivamente; – Tempo di clearance del bolo (BCT): durata del reflusso, espressa in secondi; – Indice di esposizione al bolo (BEI): totale del tempo di reflusso rispetto alla durata del tracciato, espresso in percentuale; – Frequenza dei pH-GER: episodi di reflusso/ora; – Indice di reflusso (RI): totale del tempo con pH <4 rispetto alla durata del tracciato, espresso in percentuale; Ogni evento MII-GER è stato definito come debolmente acido, acido o debolmente alcalino a seconda del valore minimo di pH raggiunto durante l'evento. I monitor CR e MII/pH sono stati sincronizzati utilizzando un riferimento esterno prima di ogni sessione di registrazione. Immediatamente dopo l'inizio e la fine di ogni sessione, appariva un segnale simultaneo sui due tracciati per convalidare la corretta sincronizzazione durante l'analisi dei dati. È stato utilizzato l'indice symptom association probability (SAP) per individuare i pazienti con un numero significativo di eventi CR e GER temporalmente associati [***17***]. Utilizzando queste informazioni, i pazienti sono stati suddivisi in due gruppi: gruppo SAP positivo (indice SAP >95%) and gruppo SAP negativo (indice SAP ≤95%). Abbiamo poi analizzato eventuali differenze nelle caratteristiche dei GER tra i due gruppi. Abbiamo anche analizzato la sequenza cronologica degli eventi CR e GER con tempo di insorgenza minore di 30 secondi l'uno dall'altro, per chiarire la natura di queste associazioni. Le analisi statistiche sono state effettuate con il pacchetto software STAT per Windows (StatSoft, Inc., Tulsa, Oklahoma, USA). I risultati sono stati espressi come mediana e range interquantile (IQR) se non altrimenti specificato, mentre il p value è stato posto a 0.05. La distribuzione delle variabili è stata analizzata con il test di Shapiro-Wilk. È stato utilizzato il t-test di Student per analizzare le differenze tra le variabili con distribuzione normale, e il test di Mann-Whitney per le variabili con distribuzione non normale.

Risultati

Sono stati inclusi nello studio 72 neonati. Sei non avevano un valido tracciato CR e MII/pH e sono stati esclusi. Sessantasei neonati soddisfacevano tutti i criteri di inclusione e sono stati considerati per l'analisi. Tali pazienti presentavano un'età gestazionale di 38 (32-40) settimane e un peso alla nascita di 2890 (1720-3190) grammi. Al momento del monitoraggio sincronizzato CR e MII/pH i neonati eligibili avevano 29 (18-45) giorni di vita, con un'età post-mestruale di 41.5 (39-43) settimane e un peso di 3310 (2700-3800) grammi. Complessivamente sono state analizzate 1377 ore di monitoraggio sincronizzato CR e MII/pH, 1316 (1139-1366) minuti per paziente. Otto (12.2%) pazienti non hanno avuto eventi CR durante il monitoraggio e sono stati esclusi da ulteriori analisi, con un campione finale di 58 (87.8%) pazienti nello studio. Questi pazienti hanno mostrato un totale di 1331 eventi CR, 6 (3-23) eventi per paziente: 114 (8.6%) apnee>20 secondi, 237 (17.8%) apnee>5 secondi associate con desaturazioni e/o bradicardie, 874 (65.7%) desaturazioni isolate, e 106 (7.9%) bradicardie isolate. Sono stati registrati un totale di 5239 eventi GER. I 4071 (78%) eventi MII-GER hanno avuto una frequenza di 2.99 (2.15-3.81) GER/ora per paziente, BCT di 22.87 (20.48-29.35) secondi, e un BRE di 4.5(4.22-4.89) canali. Sono stati registrati 973(24%) eventi GER acidi. Un totale di 1168 (22%) reflussi sono stati individuati attraverso la sola pH-metria. Cinquantuno (88%) pazienti appartenevano al gruppo SAP negativo e 7 (12%) a quello SAP positivo. Non sono state osservate differenze nella frequenza degli eventi CR tra i gruppi SAP positivo e negativo (0.61 vs 0.26 eventi/ora; p=0.516). I pazienti nel gruppo SAP positivo hanno presentato una maggiore frequenza di MII-GER, BCT,

BRE e BEI rispetto al gruppo SAP negativo. Gli eventi GER debolmente acidi e acidi nel gruppo SAP positivo hanno dimostrato maggiori frequenza e BRE. La frequenza dei pH-GER era maggiore nel gruppo SAP positivo (***Tabella 1***).

Tabella 1. Differenze nelle caratteristiche del reflusso secondo l'indice SAP. I dati mostrati sono da intendere come mediane (range interquartile). Legenda: SAP: symptom association probability; MII: impedenzometria intraluminale multicanale; GER: reflusso gastroesofageo; BRE: livello di risalita del bolo; BCT: tempo di clearance del bolo; BEI: indice di esposizione al bolo; WA: debolmente acido (weakly acidic); A: acido (acidic); WALK: debolmente alcalino (weakly alkaline); RI: indice di reflusso.

		Gruppo SAP positivo (SAP index >95%)	Gruppo SAP negativo (SAP index ≤95%)	test t di Student
		n=7	n=51	p
MII-ALL-GER	Frequenza (GER/ora)	4.00 (3.35- 4.88)	2.97 (2.15- 3.67)	**0.018**
	BRE (canali)	4.63 (4.31- 4.92)	4.14 (3.54- 4.44)	**0.029**
	BCT (secondi)	38.27 (24.15-41.92)	22.41 (20.60-26.86)	**0.048**
	BEI (%)	3.53 (3.24- 3.65)	1.76 (1.35- 2.74)	**0.001**
MII-WA-GER	Frequenza (GER/ora))	3.07 (2.32- 4.02)	2.15 (1.64- 2.76	**0.047**
	BRE (canali)	4.67 (4.25- 5.00)	3.96 (3.64- 4.22)	**0.044**
	BCT (secondi)	31.60 (26.49-39.80)	21.99 (19.15-26.58)	**0.023**
MII-A-GER	Frequenza (GER/ora)	1.01 (0.71- 1.87)	0.65 (0.32- 1.00)	**0.010**
	BRE (canali)	4.67 (4.25- 5.00)	3.96 (3.64- 4.22)	**0.019**
	BCT (secondi)	35.83 (14.77-42.80)	22.12 (17.73-30.80)	0.542
MII-WALK-GER	Frequenza (GER/ora)	-	0.00 (0.00- 0.00)	-
	BRE (canali)	-	4.50 (4.00- 5.00)	-
	BCT (secondi)	-	17.95 (12.88-30.00)	-
pH-GER	Frequenza (GER/ora)	3.14 (1.89- 3.54)	1.30 (0.51- 1.82)	**0.012**
	RI (%)	9.63 (4.71-15.64)	4.26 (1.70- 7.44)	0.316

Sono state identificate 145 associazioni tra eventi GER and CR con un'insorgenza entro 30 secondi l'uno dall'altro; in 68 (46.90%) associazioni gli eventi CR erano seguiti da quelli GER, mentre in 77 (53.10%) gli eventi CR erano preceduti da quelli GER (p>0.05). Queste associazioni sono state identificate in 7(100%) pazienti del gruppo SAP positivo (56 associazioni) e in 27 (52.94%) pazienti del gruppo SAP negativo (89 associazioni). Il numero di associazioni per paziente era maggiore nel gruppo SAP positivo 10.00 (4.5011.50) rispetto a quello SAP negativo 1.00 (0.00-2.00) (p<0.001). Nel gruppo SAP positivo, il 30.77% (21.54-75.00) degli eventi CR era associato ad eventi GER, in quello SAP negativo questa percentuale era del 2.70% (0.00-15.30) (p<0.001). Nel gruppo SAP positivo gli eventi CR preceduti da eventi GER erano l'83.33% di tutti gli eventi CR associati a GER, mentre nel gruppo SAP negativo questa percentuale era del 45.00% (0.00-66.67) (p=0.04). Non è stata individuata alcuna differenza tra gli eventi GER acidi preceduti (36.3%) versus seguiti da eventi CR (30.4%) (p=0.65). Gli eventi GER preceduti e seguiti da eventi CR hanno avuto un BCT rispettivamente di 23.10 (10.40-78.20) e 19.20 (5.80-35.45) secondi (p=0.122) e un BRE rispettivamente di 4.00 (4.00-5.00) e 3.00 (2.00-4.50) canali (p=0.004).

Conclusioni

La teoria secondo la quale gli eventi GER possano giocare un ruolo scatenante gli eventi CR è basata sull'esistenza di un chemiorecettore nelle vie aeree superiori in grado di indurre l'apnea a scopo protettivo delle vie aeree stesse.[***18***] Sono stati pubblicati numerosi studi sulla relazione tra eventi CR e GER nei neonati, e sulla natura di tale relazione, ma questi studi hanno riportato risultati contrastanti.[***6, 8, 12, 13,***

19–21] Tali studi sono difficilmente confrontabili a causa delle differenti tecniche utilizzate per identificare gli eventi CR e GER, delle differenze nell'età e caratteristiche delle popolazioni studiate, delle diverse finestre temporali utilizzate per definire un'associazione tra eventi, e delle diverse definizioni degli eventi CR [***22***]. In aggiunta, gli eventi CR e GER potrebbero risultare temporalmente associati più per la frequenza di entrambi gli eventi nei neonati che per un vero rapporto di causalità [***16***]. La maggior parte degli studi ha analizzato la relazione temporale tra CR e GER considerando tutti gli eventi dello studio come un unico gruppo di dati, senza analizzare la significatività delle associazioni in ogni singolo paziente. Questo metodo non permette di individuare la presenza di significative associazioni tra eventi CR e GER se questi sono presenti solo in un piccolo numero di pazienti [20]. Nel nostro studio, abbiamo calcolato l'indice SAP per identificare i pazienti con significative associazioni temporali tra gli eventi CR e GER. L'indice SAP utilizza il test di Fisher a 2 code per determinare la probabilità che le associazioni tra i due gruppi di eventi non siano casuali. L'indice SAP è già stato utilizzato in altri precedenti studi su neonati, [***12, 20, 21, 23***] tuttavia, questo è il primo lavoro nel quale viene utilizzato per studiare neonati sintomatici sottoposti a monitoraggio sincronizzato CR e MII/pH e nel quale viene utilizzato come discriminante per lo studio delle caratteristiche del GER. Nel nostro studio, il gruppo SAP positivo rappresentava il 12% del campione finale, suggerendo che, almeno in alcuni neonati, ci sia davvero una relazione tra eventi CR e GER. I pazienti con indice SAP positivo mostravano una maggiore frequenza di eventi GER, eventi GER più lunghi, e un più alto BRE. Questi dati suggeriscono che le associazioni tra eventi CR e GER sono molto più frequenti in soggetti con GER più severo, ma non forniscono informazioni sulla natura causale di questa relazione. Per comprendere meglio tale relazione, abbiamo analizzato le associazioni tra eventi con insorgenza entro 30 secondi l'uno dall'altro, in accordo con altri studi recenti che hanno incluso popolazioni simili e hanno utilizzato tecniche analoghe.[***14, 24–26***] Sembra sensato utilizzare una finestra temporale di 30 secondi per due ragioni: permette di identificare la maggior parte delle associazioni significative (è improbabile che eventi con una latenza maggiore possano influenzarsi reciprocamente) e rende improbabile la presenza di multipli eventi CR e GER nella stessa finestra temporale, considerando che la durata media di tali eventi è di circa 15-30 secondi. Abbiamo scoperto che queste associazioni erano molto rare se paragonate all'alto numero di eventi CR e GER, coinvolgendo solo il 4% di tutti gli eventi GER e l'11% di tutti gli eventi CR; tuttavia, risultano circa 10 volte più frequenti nel gruppo SAP positivo, rappresentando circa un terzo di tutte le associazioni in questi pazienti. Questi risultati confermano la nostra ipotesi iniziale che, in alcuni soggetti con specifiche caratteristiche di reflusso, ci sia una significativa associazione tra gli eventi CR e GER. Focalizzando l'analisi sulle associazioni di tali eventi nel gruppo SAP positivo, abbiamo notato che circa l'83% degli eventi CR erano preceduti da GER, e in questi casi, i GER avevano un più alto BRE rispetto a quelli seguiti dai CR; non sono state osservate differenze in BCT o valore di pH. Questi dati confermano che il reflusso prossimale possa rappresentare uno stimolo per gli eventi CR indipendentemente dall'acidità del reflusso, [***13, 25***] come precedentemente osservato in modelli animali [***27***]. L'apnea e la conseguente desaturazione d'ossigeno potrebbero essere il risultato di un meccanismo protettivo delle vie aeree contro l'inalazione stimolato sia da meccanocettori che da chemiocettori.[***18, 28, 29***] Da un punto di vista fisiopatologico, sarebbe interessante chiarire il meccanismo attraverso il quale alcuni reflussi possano indurre eventi CR, ma, da un punto di vista clinico, la rarità di questo fenomeno suggerisce che l'impatto degli eventi GER su neonati con eventi CR sia limitato ad un ridotto numero di soggetti sintomatici e sia indipendente dall'acidità dei reflussi. In conclusione, nel nostro studio su neonati con eventi CR e sintomi di malattia da GER, il 12% dei soggetti hanno avuto una significativa relazione temporale tra eventi CR e GER. Nell'83% di queste associazioni, gli eventi CR erano preceduti da quelli GER. Questi risultati suggeriscono che, in una popolazione selezionata, la valutazione simultanea degli eventi CR e GER possa essere utile ad identificare i neonati con eventi CR nei quali sia indicato il trattamento del GER. Tuttavia, il fatto che gli eventi GER registrati siano per lo più non acidi suggerisce che il trattamento farmacologico con antiacidi sia, in molti casi, inappropriato.

Bibliografia

[1] Dhillon AS, Ewer AK. Diagnosis and management of gastro-oesophageal reflux in preterm infants in neonatal intensive care units. Acta Paediatr Oslo Nor 1992 2004;93:88–93.

[2] Malcolm WF, Gantz M, Martin RJ, Goldstein RF, Goldberg RN, Cotten CM. Use of medications for gastroesophageal reflux at discharge among extremely low birth weight infants. Pediatrics 2008;121:22–7.

[3] Poets CF. Gastroesophageal reflux: a critical review of its role in preterm infants. Pediatrics 2004;113:e128-132.

[4] Cresi F, Locatelli E, Marinaccio C, Grasso G, Coscia A, Bertino E. Prognostic Values of Multichannel Intraluminal Impedance and pH Monitoring in Newborns with Symptoms of Gastroesophageal Reflux Disease. J Pediatr 2013;162:770–5.

[5] Rudolph CD, Mazur LJ, Liptak GS, Baker RD, Boyle JT, Colletti RB, et al. Guidelines for evaluation and treatment of gastroesophageal reflux in infants and children: recommendations of the North American Society for Pediatric Gastroenterology and Nutrition. J Pediatr Gastroenterol Nutr 2001;32 Suppl 2:S1-31.

[6] Wenzl TG, Schenke S, Peschgens T, Silny J, Heimann G, Skopnik H. Association of apnea and nonacid gastroesophageal reflux in infants: Investigations with the intraluminal impedance technique. Pediatr Pulmonol 2001;31:144–9.

[7] Abu Jawdeh EG, Martin RJ. Neonatal apnea and gastroesophageal reflux (GER): is there a problem? Early Hum Dev 2013;89:S14–6.

[8] Slocum C, Arko M, Di Fiore J, Martin RJ, Hibbs AM. Apnea, bradycardia and desaturation in preterm infants before and after feeding. J Perinatol Off J Calif Perinat Assoc 2009;29:209–12.

[9] Terrin G, Passariello A, De Curtis M, Manguso F, Salvia G, Lega L, et al. Ranitidine is associated with infections, necrotizing enterocolitis, and fatal outcome in newborns. Pediatrics 2012;129:e4045.

[10] Guillet R, Stoll BJ, Cotten CM, Gantz M, McDonald S, Poole WK, et al. Association of H2-blocker therapy and higher incidence of necrotizing enterocolitis in very low birth weight infants. Pediatrics 2006;117:e137-142.

[11] More K, Athalye-Jape G, Rao S, Patole S. Association of inhibitors of gastric acid secretion and higher incidence of necrotizing enterocolitis in preterm very low-birth-weight infants. Am J Perinatol 2013;30:849–56.

[12] Nunez J, Cristofalo E, McGinley B, Katz R, Glen DR, Gauda E. Temporal Association of Polysomnographic Cardiorespiratory Events With GER Detected by MII-pH Probe in the Premature Infant at Term. J Pediatr Gastroenterol Nutr 2011;52:523–31.

[13] Corvaglia L, Zama D, Gualdi S, Ferlini M, Aceti A, Faldella G. Gastro-oesophageal reflux increases the number of apnoeas in very preterm infants. Arch Dis Child Fetal Neonatal Ed 2009;94:F188192.

[14] Di Fiore J, Arko M, Herynk B, Martin R, Hibbs AM. Characterization of cardiorespiratory events following gastroesophageal reflux in preterm infants. J Perinatol Off J Calif Perinat Assoc 2010;30:683–7.

[15] Peter CS, Sprodowski N, Bohnhorst B, Silny J, Poets CF. Gastroesophageal reflux and apnea of prematurity: no temporal relationship. Pediatrics 2002;109:8–11.

[16] Poets CF. Gastroesophageal Reflux and Apnea of Prematurity - Coincidence, Not Causation. Commentary on L. Corvaglia et al.: A Thickened Formula Does Not Reduce Apneas Related to Gastroesophageal Reflux in Preterm Infants (Neonatology 2013;103;98-102). Neonatology 2012;103:103–4.

[17] Bredenoord AJ, Weusten BL a. M, Smout AJPM. Symptom association analysis in ambulatory gastro-oesophageal reflux monitoring. Gut 2005;54:1810–7.

[**18**] Davies AM, Koenig JS, Thach BT. Upper airway chemoreflex responses to saline and water in preterm infants. J Appl Physiol Bethesda Md 1985 1988;64:1412–20.

[**19**] Di Fiore JM, Arko M, Whitehouse M, Kimball A, Martin RJ. Apnea is not prolonged by acid gastroesophageal reflux in preterm infants. Pediatrics 2005;116:1059–63. doi:10.1542/peds.20042757

[**20**] Mousa H, Woodley FW, Metheney M, Hayes J. Testing the association between gastroesophageal reflux and apnea in infants. J Pediatr Gastroenterol Nutr 2005;41:169–77.

[**21**] Funderburk A, Nawab U, Abraham S, DiPalma J, Epstein M, Aldridge H, et al. Temporal Association Between Reflux-like Behaviors and Gastroesophageal Reflux in Preterm and Term Infants: J Pediatr Gastroenterol Nutr 2016;62:556–61.

[**22**] Smits MJ, van Wijk MP, Langendam MW, Benninga MA, Tabbers MM. Association between gastroesophageal reflux and pathologic apneas in infants: a systematic review. Neurogastroenterol Motil Off J Eur Gastrointest Motil Soc 2014;26:1527–38.

[**23**] Cresi F, Castagno E, Storm H, Silvestro L, Miniero R, Savino F. Combined esophageal intraluminal impedance, pH and skin conductance monitoring to detect discomfort in GERD infants. PloS One 2012;7:e43476.

[**24**] Corvaglia L, Spizzichino M, Zama D, Aceti A, Mariani E, Legnani E, et al. Sodium Alginate (Gaviscon®) does not reduce apnoeas related to gastro-oesophageal reflux in preterm infants. Early Hum Dev 2011;87:775–8.

[**25**] Corvaglia l, Zama D, Spizzichino M, Aceti A, Mariani E, Capretti MG, et al. The frequency of apneas in very preterm infants is increased after non-acid gastro-esophageal reflux: Apneas in preterm are increased by non-acid GER. Neurogastroenterol Motil 2011;23:303-e152.

[**26**] Apnea Is Not Prolonged by Acid Gastroesophageal Reflux in Preterm Infants - Di Fiore.pdf n.d.

[**27**] Bauman NM, Sandler AD, Schmidt C, Maher JW, Smith RJ. Reflex laryngospasm induced by stimulation of distal esophageal afferents. The Laryngoscope 1994;104:209–14.

[**28**] Jadcherla SR. Upstream effect of esophageal distention: effect on airway. Curr Gastroenterol Rep 2006;8:190–194.

[**29**] Jadcherla SR, Duong HQ, Hoffmann RG, Shaker R. Esophageal body and upper esophageal sphincter motor responses to esophageal provocation during maturation in preterm newborns. J Pediatr 2003;143:31–8.

COMPREHENSIVE LANDSCAPE OF SUBTYPE-SPECIFIC CODING AND NONCODING RNA TRANSCRIPTS IN BREAST CANCER

Calza Stefano[1,2], Vu Trung Nghia[2], Pramana Setia[2], Suo Chen[2], Lee Donghwan[2,3], Pawitan Yudi[2]

1. *Department of Molecular and Translational Medicine, Univerity of Brescia, Brescia, Italy*
2. *Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, SE 17177 Stockholm, Sweden*
3. *Department of Statistics, Ewha Womans University, Seodaemun-gu, Seoul 120-750, South Kore*

Molecular classification of breast cancer into clinically relevant subtypes helps improve prognosis and adjuvant-treatment decisions. The aim of this study is to provide a better characterization of the molecular subtypes by providing a comprehensive landscape of subtype-specific isoforms including coding, long noncoding RNA and microRNA transcripts. Isoform-level expression of all coding and non-coding RNAs is estimated from RNA-sequence data of 1168 breast samples obtained from The Cancer Genome Atlas (TCGA) project. We then search the whole transcriptome systematically for subtype-specific isoforms using a novel algorithm based on a robust quasi-Poisson model. We discover 5451 isoforms specific to single subtypes. A total of 27% of the subtype-specific isoforms have better accuracy in classifying the intrinsic subtypes than that of their corresponding genes. We find three subtype-specific miRNA and 707 subtype-specific long non-coding RNAs. The isoforms from long non-coding RNAs also show high performance for separation between Luminal A and Luminal B subtypes with an AUC of 0.97 in the discovery set and 0.90 in the validation set. In addition, we discover 1500 isoforms preferentially co-expressed in two subtypes, including 369 isoforms co-expressed in both Normallike and Basal subtypes, which are commonly considered to have distinct ER-receptor status. Finally, analyses at protein level reveal four subtype-specific proteins and two subtype co-expression proteins that successfully validate results from the isoform level.

IL SISTEMA DI SORVEGLIANZA OKkio ALLA SALUTE: UNO STRUMENTO PER IL MONITORAGGIO E LA VALUTAZIONE DEI PROGRAMMI DI PROMOZIONE DELLA SALUTE E PREVENZIONE

Puci Mariangela Valentina[1], Coppola Liliana[2], Marella Margherita[2], Pirrone Lucia[2], Celata Corrado[2], Montomoli Cristina[1], Gruppo OKkio alla Salute 2016[3]

1. *Dipartimento di Sanità Pubblica, Medicina Sperimentale e Forense, U.O. di Biostatistica e Epidemiologia Clinica, Università degli Studi di Pavia*
2. *Direzione Generale Welfare, Regione Lombardia*
3 *Agenzia di Tutela della Salute (ATS) della Lombardia*

Introduzione

Attualmente il sovrappeso e l'obesità rappresentano, con sedentarietà, tabagismo e scorretto consumo di alcool, i principali fattori di rischio comportamentali di diverse patologie croniche [***1***]. In età infantile, tali condizioni hanno delle implicazioni dirette sulla salute del bambino poiché sono fattori di rischio per l'insorgenza di gravi patologie in età adulta. Secondo l'Organizzazione Mondiale della Sanità (OMS), nel 2014 la quota di bambini sotto i cinque anni in condizioni di sovrappeso e obesità, nel mondo, era pari a 42 milioni [***2***]. L'obesità è imputabile, nella maggior parte dei casi, a stili di vita scorretti quali una dieta ipercalorica o la sedentarietà. Pertanto, nel tempo sono stati implementati in Europa dei sistemi di sorveglianza atti a comprendere la dimensione del fenomeno e i comportamenti associati [***3***]. A partire dal 2007, in Italia il Ministero della Salute/CCM ha promosso e finanziato lo sviluppo e l'implementazione nel tempo del sistema di sorveglianza OKkio alla SALUTE, che aderisce al progetto "*Childhood Obesity Surveillance Initiative"* (COSI) della Regione Europea dell'Organizzazione Mondiale della Sanità, coordinato dal Centro Nazionale di Epidemiologia, Sorveglianza e Promozione della Salute dell'Istituto Superiore di Sanità e condotto in collaborazione con le Regioni e il Ministero dell'Istruzione, dell'Università e della Ricerca [***4***]. La definizione di priorità sulla base dell'analisi dei profili di salute della popolazione, e dei relativi stili di vita, rappresenta un processo chiave nello sviluppo delle strategie di promozione della salute e prevenzione declinate nel Piano Regionale di Prevenzione di Regione Lombardia (PRP 2014 – 2018) [***5***] in coerenza con il Programma "Guadagnare salute" [***6***] e il Piano Nazionale della Prevenzione. In un simile scenario OKkio alla SALUTE rappresenta uno strumento importante grazie al quale è possibile descrivere la variabilità geografica e l'evoluzione temporale dello stato ponderale, degli stili alimentari e dell'abitudine all'esercizio fisico dei bambini della classe primaria. I dati qui presentati sono il risultato preliminare di uno studio attualmente in corso che sarà concluso entro fine anno.
Descrivere le abitudini alimentari e l'attività fisica dei bambini afferenti alla scuola primaria per orientare e valutare le strategie di promozione della salute e prevenzione.

Metodi

Studio descrittivo trasversale condotto sui bambini di 8 - 9 anni afferenti alle scuole primarie della regione Lombardia nell'anno scolastico 2015/2016. Sono stati inclusi i bambini afferenti alle classi campionate nell'ambito del sistema di sorveglianza OKkio alla SALUTE, tramite campionamento a grappolo (cluster survey design). Le variabili in studio sono state raccolte attraverso 4 strumenti: scheda antropometrica compilata da operatori ASL, questionario indirizzato alla scuola, ai bambini e ai genitori. L'analisi ha riguardato il calcolo delle principali statistiche descrittive sulle variabili raccolte. Per verificare la normalità delle distribuzioni è stato utilizzato il test di Shapiro Wilk. La relazione tra le caratteristiche antropometriche dei bambini e i genitori è stata indagata tramite l'utilizzo del coefficiente di correlazione di Spearman. È stato utilizzato il test chi-quadro per valutare l'associazione tra variabili categoriche e il test t di Student o il

corrispettivo non parametrico test di Mann - Whitney per quelle quantitative. Il livello di significatività considerato è stato p<0.05.

Risultati

I bambini in studio sono 2244 (53% maschi) per un totale di 118 classi. L'età mediana è di 8.83 anni sia nel campione (25°-75° percentile: 8.58 - 9.08) che tra i maschi e le femmine. Per quanto riguarda le variabili antropometriche il peso mediano del campione è di 29.45 kg (26.05 - 34.57), per i maschi è di 29.76 Kg (26.33 - 34.94) e per le femmine è di 29.05 Kg (25.68 - 34.21), mentre l'altezza mediana è pari a 133.20 (129.05 - 137.40) cm nel campione, 133.80 cm (129.70 - 138.00) nei maschi e 132.50 cm (128.30 - 136.50) nelle femmine. Sia per il peso che per l'altezza le differenze tra maschi e femmine sono statisticamente significative (rispettivamente p=0.0066 e p < 0.0001). L'indice di massa corporea mediano del campione è pari à 16.63 kg/m2 (15.21 - 18.84), mentre nei maschi è di 16.53 kg/m2 (15.25 - 18.89) e nelle femmine è 16.70 kg/m2 (15.18 - 18.80; p = 0.5151). Per quanto concerne le caratteristiche dei genitori, la maggioranza è di nazionalità italiana (padri 83% e madri 80%), e possiede un titolo di studio che corrisponde al diploma di scuola superiore o alla laurea (padri 65%, madri 76%). Il 50% dei padri e il 76% delle madri sono in una condizione di sottopeso o normopeso. In merito alla condizione economica, la maggioranza dei genitori (55%) dichiara di arrivare facilmente a fine mese. La ***tabella 1*** riporta i risultati inerenti alle categorie di BMI, alle abitudini alimentari e del tempo libero dei bambini in base al sesso del bambino.

Tabella 1. Abitudini alimentari e attività fisica

	Totale (n=2244)	Maschi n=1195 (53.25%)	Femmine n=1049 (46.75%)	P value
Classi BMI				
Sottopeso	32 (1.43)	19 (1.59)	13 (1.24)	0.910
Normopeso	1672 (74.51)	889 (74.39)	783 (74.64)	
Sovrappeso	416 (18.54)	220 (18.4)	196 (18.68)	
Obeso	124 (5.53)	67 (5.61)	57 (5.43)	
Tipo di colazione				
Non fa colazione	96 (4.28)	50 (4.19)	46 (4.39)	0.566
Non adeguata	739 (32.98)	406 (34.00)	333 (31.81)	
Adeguata	1406 (62.74)	738 (61.81)	668 (63.80)	
Tipo di merenda				
No fa merenda	172 (7.67)	112 (9.38)	60 (5.72)	0.001
Non adeguata	1043 (46.50)	566 (47.40)	477 (45.47)	
Adeguata	1028 (45.83)	516 (43.22)	512 (48.81)	
Consumo di frutta e verdura				
No	1819 (83.21)	984 (84.90)	831 (81.39)	0.029
SI	367 (16.79)	175 (15.10)	190 (18.61)	
Ore giornaliere trascorse a giocare ai videogiochi, al computer o guardando la tv				
0-2 ore	1438 (65.45)	704 (60.38)	734 (71.19)	< 0.001
3-4 ore	636 (28.95)	383 (32.85)	253 (24.54)	
5 o più ore	123 (5.60)	79 (6.78)	44 (4.27)	

C'è un'associazione significativa tra il genere e il consumo della merenda, il consumo di almeno un alimento che sia frutta o verdura e le ore giornaliere trascorse a giocare ai videogiochi, al computer o davanti alla tv (rispettivamente p=0.001, p=0.029 e < 0.001). Le bambine tendono ad avere delle abitudini più sane rispetto

ai bambini per quanto riguarda la merenda, il consumo di frutta e verdura e lo stile sedentario. Nel primo caso solo il 6% delle bambine non fa la merenda rispetto al 9% dei bambini, le bambine inoltre consumano di più una merenda adeguata (49%) rispetto ai bambini (43%). Il consumo di frutta e verdura rimane limitato sia per quanto riguarda le bambine che i bambini (rispettivamente solo il 19% delle bambine e il 15% dei bambini consuma questi alimenti).
Anche per quanto concerne la sedentarietà, le bambine manifestano uno stile di vita più sano, poiché la maggioranza (71%) non trascorre più di due ore a giocare ai videogiochi, computer o a guardare la tv (contro il 60% dei bambini). Dai risultati si evince una correlazione debole tra il peso e il BMI dei bambini e quello dei genitori.

Conclusioni

Questi risultati suggeriscono la necessità di implementare programmi di promozione della salute inerenti a stili alimentari corretti e allo svolgimento di una adeguata attività fisica sia in ambito scolastico che nella comunità.

Bibliografia

[1] WHO Regional Office for Europe. Action Plan for implementation of the European Strategy for the Prevention and Control of Non communicable Diseases 2012−2016 Reperibile su: http://www.euro.who.int/data/assets/pdf.

[2] Global Status Report on Non-Communicable Diseases, 2010. Reperibile su: http://www.who.int/nmh/publications/ncd_report_full_en.pdf.

[3] WHO European Childhood Obesity Surveillance Initiative (COSI). Reperibile su: http://www.euro.who.int/en/healthtopics/diseaseprevention/nutrition/activities/monitoring-and-surveillance/who-europeanchildhood-obesity-surveillanceinitiative-cosi.

[4] OKkio alla Salute: sistema di sorveglianza su alimentazione e attività fisica nei bambini della scuola primaria. Risultati 2014. Reperibile su: http://www. epicentro.iss.it/okkioallasalute.

[5] Regione Lombardia, Delibera Consiglio Regionale n. 1497 dell'11 aprile 2017 - "Piano regionale di prevenzione 2014 – 2018, ai sensi dell'Intesa Stato Regioni del 13 novembre 2014 e della legge regionale 30 dicembre 2009, n. 33 (Testo unico delle leggi regionali in materia di Sanità) e s.m.i". Reperibile su: http://www.regione.lombardia.it/wps/portal/istituzionale/HP/DettaglioRedazionale/istituzione/direzioni-generali/direzione-generale-welfare/red-piano-regionale-prevenzione-ser.

[6] Decreto del Presidente del Consiglio dei Ministri, 4 maggio 2007. Guadagnare salute. Rendere facili le scelte salutari. Gazzetta Ufficiale n. 117 del 22 maggio 2007. Reperibile su: http://www.salute.gov.it/imgs/C_17_pubblicazioni_605_allegato.pdf.

POSTER

ANALYSIS OF CASUAL RELATIONSHIPS BETWEEN MEDITERRANEAN DIET AND ATHEROSCLEROSIS. PATH ANALYSIS AND STRUCTURAL EQUATION MODELING.

Barbieri Simone[1], Bonomi Alice[1], Veglia Fabrizio[1], Tedesco Calogero[1], Ecca Eleonora [1], Amato Mauro[2], Baldassarre Damiano[2], Tremoli Elena[3].
On behalf of the IMPROVE Study Group

1. *Unit of Biostatistics, Centro Cardiologico Monzino;*
2. *Unit for the Study of Arterial Morphology and Function, Centro Cardiologico Monzino;*
3. *Scientific Director, Centro Cardiologico Monzino.*

Introduction

The associations between Mediterranean Diet and cardiovascular disease have been mainly studied by using regression analysis, which involves a single dependent variable and multiple predictors. However, this approach is unable to discover complex relationships between variables. 'Path analysis' allows to study and to visualize the most meaningful relations among the whole set of selected variables, and Structural Equation Modeling (SEM) allows to detect relationships between the so called 'latent variables', i.e. variables that are not directly observed but are rather inferred.
The aim of this project is to identify causal pathways in atherosclerosis, as indexed by measures of carotid intima-media thickness (IMT) and to analyze the effect of Mediterranean Diet.

Methods

The IMPROVE [***1***] database, including 3703 subjects from five different European countries, was employed. We used partial correlation network analysis, a technique that allows investigating the relationships between variables, free from the influence of all the other covariates in the model. The links between the variables in the 'path diagram' represent partial R^2; only coefficients > 0.05 are depicted; sizes refer to the strength of the relations and colours to the sign (red=positive, green=negative).
We inferred a latent variable from multiple observed items typical of Mediterranean Diet, and its relationships with other variables were imputed by SEM (PROC CALIS of the software SAS 9.4).

Results

In the Path diagram (***Figure 1***), baseline IMT and IMT progression (PROG IMT) were negatively associated with the latent variable DIET (β=-0.1 and β=-0.04, respectively). Moreover, DIET was also significantly associated with BMI (β=-0.25), physical activity (PHYSACT, β=0.1) and latitude (LAT, β=-0.53), and acted as a mediator between these factors and atherosclerosis.

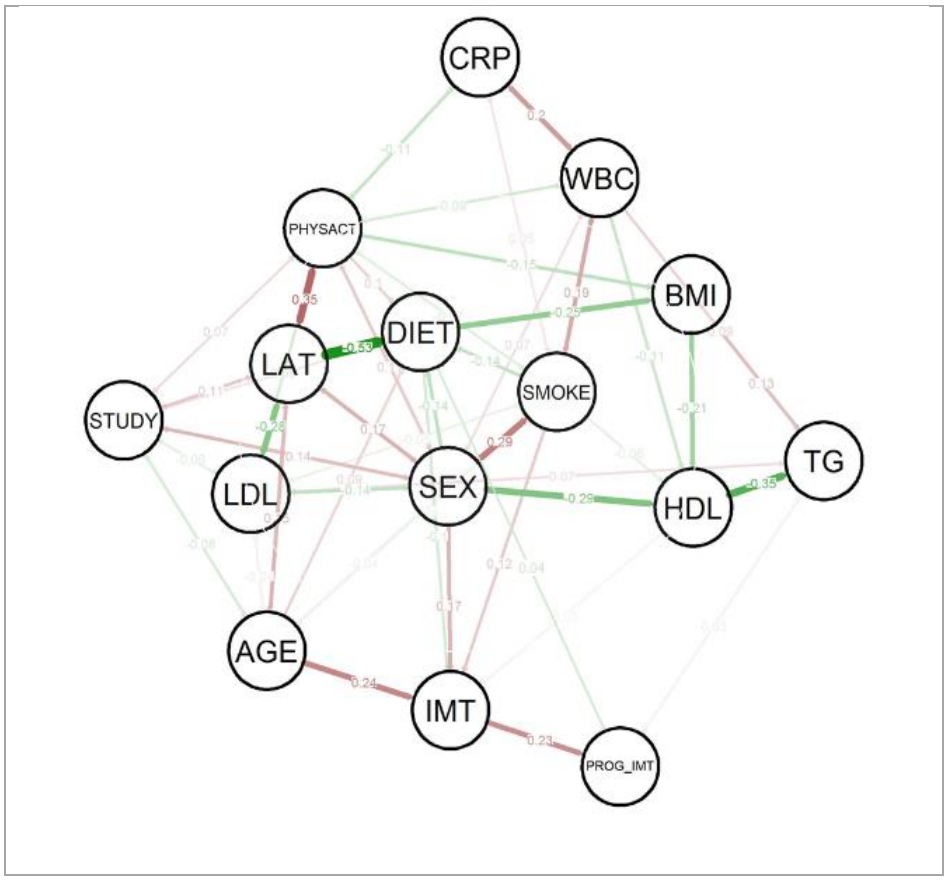

Figure 1. Path Diagram

Conclusions

Mediterranean Diet appears to be an important determinant of subclinical atherosclerosis and atherosclerosis progression, and exhibits complex interrelations with cardiovascular risk factors.

References

[1] Baldassarre D, Nyyssonen K, Rauramaa R, et al. Cross-sectional analysis of baseline data to identify the major determinants of carotid intima-media thickness in a European population: the IMPROVE study. Eur Heart J 2010;31:614 –22.

RESPIRATORY HEALTH IN ASSOCIATION WITH DIETARY PATTERNS IN ADULTS FROM EUROPEAN COMMUNITY RESPIRATORY HEALTH SURVEY III

Chamitava Liliya[1], Bakolis Ioannis[2], Burney Peter G.J[3], Zanolin Maria Elisabetta[1], Jarvis Deborah[3], Garcia-Larsen Vanessa[3,4], on behalf of the ECRHS III Nutrition Working Group

1. *Unit of Epidemiology and Medical Statistics (SESM), Department of Public Health and Community Health, University of Verona, Italy*
2. *Department of Biostatistics, Department of Health Services and Population Research, Institute of Psychiatry, Psychology & Neuroscience, King's College London, UK*
3. *Population Health and Occupational Disease, National Heart and Lung Institute, Imperial College London, UK*
4. *Department of International Health, Johns Hopkins Bloomberg School of Public Health, US*

Introduction

Recent developments of nutritional epidemiology consider dietary patterns rather than single food or nutrient. Dietary pattern analysis considers foods in a synergic way that aims to recognize potential interactions and collinearity among nutrients and foods and might address better the issue of confounding than single food or nutrient analyses [***1***]. Studying dietary patterns could provide a comprehensive approach in preventive medicine [***2***]. In nutritional epidemiology two statistical approaches are commonly used to derive dietary patterns from empirical data: factor or cluster analysis. The first one breaks down the data onto dietary patterns considering intercorrelations between food frequency items, while the cluster analysis is based on individual differences in mean intakes. Principal component analysis (PCA) is most-used factor analysis to derive dietary patterns [***3***]. In case of orthogonal rotation after PCA, independent factors can be obtained and, thus, can be included together in regression models without affecting each other and regression analysis outcomes [***2, 3***]. Meta-analysis is used to address differences across studies or countries and to combine results computing their mean size of effect [***4***]. Food frequency questionnaire (FFQ) is one of primary methods to assemble individual dietary information into dietary analysis, especially dietary pattern analysis.

In this work we aimed at investigating the cross-sectional association of dietary patterns with respiratory outcomes in adults participating in the European Community Respiratory Health Survey (ECRHS) III, using meta-analysis of Principal Component Scores (MPCA) and a standardized food frequency questionnaire (FFQ).

Methods

Study design. Cross-sectional data from 3627 adults from the European Community Respiratory Health Survey III (ECRHS III) were included in the current analysis. Participants from eight countries (Iceland, Norway, Sweden, Denmark, UK, France, Spain and Australia) answered a standardised general questionnaire on respiratory health and related factors, and underwent objective measures of lung function (forced expiratory volume in 1 second (FEV_1) and forced vital capacity (FVC). An asthma score (range 0 to 5) was built based on answering "Yes" to any of the following questions on symptoms (in the last 12 months): 1) breathless while wheezing; 2) woken up with a feeling of chest tightness; 3) attack of shortness of breath at rest; 4) attack of shortness of breath after exercise; 5) woken by attack of shortness of breath [***5***]. Dietary intake was ascertained using the internationally validated and standardized GA²LEN Food frequency Questionnaire (FFQ), which enquired on frequency of intake of 250 foods over the past 12 months [***6***]. Using standard portion sizes (Food Standards Agency, 2002), daily grams of intake were derived, and nutrient intake estimated based on the McCance and Widdowson´s British Food Composition Table.

Dietary patterns. The Rosenthal's method was used to prepare data to the meta-analysis. Firstly, the r-values of each correlation matrix were converted to Fisher's z-scores [**7**]. Secondly, the r-to-z transformed matrices were weighted by the degrees of freedom for each country and combined to yield a summed matrix of z-scores. This matrix was then divided by the summed degrees of freedom and then converted back to correlation coefficients to obtain a pooled matrix of mean r-values. PCA (with varimax rotation) was applied to the matrix of pooled correlation coefficients, giving dietary pattern scores appropriate for all countries. Dietary patterns were labelled according to the food items with an absolute correlation coefficient >0.3.
Statistical analyses. The unconditional chi^2 test and the Kruskal–Wallis test by ranks were used to determine differences within countries in terms of outcomes (FVC, FEV_1, FEV_1/FVC and asthma score) and potential confounders (age, weight, height or BMI (body mass index), gender, smoking habits (never-, ex-, current smoker), total energy intake, employment status (employed, unemployed) and regular intake of nutritional supplements) with statistically significant level $p<0.05$. Adjusted multivariable linear regression models were used to examine the association between dietary patterns and spirometry (FEV_1, FVC, FEV_1/FVC). Adjusted negative binomial regression analysis was used to evaluate association between dietary patterns and asthma score. Trend of association of respiratory outcomes and asthma score with dietary patterns were evaluated assigning consecutive integers (1 to 3) to the dietary pattern tertiles. Regression results where pooled across countries using meta-analysis. Heterogeneity between countries was assessed with the use of the I-squared statistics. Some countries were excluded from the meta-analyses because of the lack of observations. Such outcomes as asthma, COPD and chronic bronchitis were not analyzed because of the lack of subjects with these diseases.

Results
Country samples differed in terms of outcomes: FVC (ml) – highest Median(IQR)= 4144(3539-4966) in Norway and lowest Median(IQR)=3709(3155-4428) in UK, $p<0.001$; FEV_1 (ml) - highest Median(IQR)= 3294(2751-3856) in Norway and lowest Median(IQR)=2900(2489-3422) in UK, $p<0.001$; FEV_1/FVC - highest Median(IQR)=0.80(0.75-0.83) in Spain and lowest Median(IQR)=0.78(0.75-0.82) in Australia, $p<0.05$; asthma score – bigger 1.8% of subjects with highest asthma score 5 in Australia and bigger 83.6% of subjects with zero asthma score in Denmark, $p<0.001$) and of potential confounders (age, weight, height, gender, and others) with statistically significant levels $p<0.05$ and $p<0.001$.
The scree plot from the PCA analysis showed a clear break in the curve after two components with eigenvalue >5. The "Prudent" pattern was associated with fruits, vegetables, olive oil and fish consumption in all countries. The "Western" pattern was characterized by higher consumption of meat, potatoes and sweets.
The meta-analysed effect estimates (adjusted ratio of mean scores [AR]) of all countries showed a positive association between asthma score and having a 'Western pattern' (AR 1.16 per trend of dietary pattern score; 95% CI 1.00, 1.36), with no evidence of heterogeneity (I^2=0.0%). FVC was unrelated to Prudent or Western patterns (β-coefficient 6ml; 95% CI -49, 61; and β-coefficient 6ml; 95% CI -58, 46, respectively). However if stratified by gender and smoking habit, i.e. in the group of never-smoking men among 7 European countries, the association of FVC with Prudent dietary pattern was negative at statistically significant level (b.coef -176 per trend of dietary pattern score; 95% CI -292, -61), with heterogeneity (I^2=29.4%). In the sample of currently smoking men among 5 European countries (Iceland, Norway, Denmark, France and Spain this association was positive (b.coef 183 per trend of dietary pattern score; 95% CI 0, 366), with no evidence of heterogeneity (I^2=0.0%).

Conclusions
In this multi-country, population-based cross-sectional study we demonstrated that a meta-analytic approach can be used to examine the association between dietary patterns derived from PCA and health

outcomes. Further use of meta-analytic approach to PCA in nutritional epidemiology could contribute to explore underlying differences and heterogeneity issues in multi-centric studies.
The results of the conducted analyses showed that a higher (worse) asthma score was associated with having a 'Western' dietary pattern in European adults; and that adult smoking male population consuming more fruits, vegetables, olive oil and fish, worth on nutrients with antioxidant properties can improve their ventilatory function (FVC) modulating lung function decline in adults exposed to cigarette smoke.

References

[1] Bakolis I, Burney PGJ, Hooper R. Principal components analysis of diet and alternatives for identifying the combination of foods that are associated with the risk of disease: a simulation study. Br J Nutr 2014;112(1):61-9.
[2] Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. Current opinion in lipidology 2002;13(1):3-9.
[3] Newby PK, Tucker KL. Empirically derived eating patterns using factor or cluster analysis: a review. Nutrition reviews 2004;62(5):177-203.
[4] Grube BS, Bilder RM, Goldman RS. Meta-analysis of symptom factors in schizophrenia. Schizophrenia research 1998;31(2-3):113-20.
[5] Sunyer J, Pekkanen J, Garcia-Esteban R, et al. Asthma score: predictive ability and risk factors. Allergy 2007;62(2):142-8.
[6] Arcia-Larsen V, Luczynska M, Kowalski ML, et al. Use of a common food frequency questionnaire (FFQ) to assess dietary patterns and their relation to allergy and asthma in Europe: pilot study of the GA2LEN FFQ. Eur J Clin Nutr 2011;65(6):750-6.
[7] Grube BS, Bilder RM, Goldman RS. Meta-analysis of symptom factors in schizophrenia. Schizophr Res 1998;31(2-3):113-20.

ABITUDINI ALIMENTARI NEGLI SPORT DI SQUADRA:VALIDAZIONE DI UN QUESTIONARIO

Borrelli Paola[1], Ferraris Cinzia[2], Tagliabue Anna[2], Vandoni Matteo[3], Montomoli Cristina[1]

1. *Unità di Biostatistica e Epidemiologia Clinica, Dipartimento Sanità Pubblica, Medicina Sperimentale e Forense, Università di Pavia, Pavia*
2. *Centro Interdipartimentale di Studi e Ricerche sulla Nutrizione Umana e i Disturbi del comportamento Alimentare, Dipartimento Sanità Pubblica, Medicina Sperimentale e Forense, Università di Pavia, Pavia*
3. *Laboratorio di Attività Motoria Adattata (LAMA), Dipartimento Sanità Pubblica, Medicina Sperimentale e Forense, Università di Pavia, Pavia*

Introduzione

Gli sport di squadra condividono la caratteristica comune di uno schema di attività intermittente ad alta intensità, ma presentano tuttavia una variabilità delle caratteristiche di gioco tra sport, ruoli e stili di gioco all'interno dello sport stesso e tra una competizione e l'altra. Questo crea una diversità a livello dei cambiamenti fisiologici e delle esigenze nutrizionali per gli atleti di sport di squadra [***1***]. L'apporto dietetico adeguato è un fattore importante che incide sulla capacità degli atleti di mantenere una condizione fisica ottimale e di allenarsi in modo efficace. Il Comitato Olimpico Internazionale ha pubblicato nel 2010 una dichiarazione di consenso sull'alimentazione sportiva, mentre nel 2016 l'American College of Sports Medicine (ACSM) ha promulgato una position su nutrizione e performance sportiva [***2***]. Gli sportivi usano l'alimentazione e la supplementazione nutrizionale come strumento per migliorare le loro prestazioni. Appare, quindi, evidente la necessità di orientare la ricerca scientifica sulla nutrizione sportiva verso l'indagine delle conoscenze alimentari degli atleti per verificare l'aderenza alle indicazioni a loro dedicate. Generalmente i metodi utilizzati per l'indagine delle abitudini alimentari possono essere diversi e comprendono: la storia dietetica, il diario alimentare, il recall dietetico delle 24 ore, il questionario sulle conoscenze alimentari e quello sulle frequenze di consumo degli alimenti (FFQ). Tutti questi metodi possono essere utilizzati su diversi gruppi d'indagine e possono determinare la relazione fra abitudini alimentari e rischio di malattia. Gli studi fino ad ora condotti per valutare le abitudini alimentari sui giocatori di sport di squadra hanno riportato un'assunzione calorica media abituale inferiore rispetto a quanto raccomandato e, di conseguenza, abitudini alimentari inadeguate [***3-7***]. Tuttavia questi studi hanno adottato metodologie d'indagine diverse e non validate sul target di popolazione indagato. Risulta, pertanto, fondamentale creare uno strumento di indagine sulle abitudini alimentari degli sportivi.
Messa a punto di un questionario sulle abitudini alimentari di sportivi dilettanti che praticano sport di squadra, valutazione della sua affidabilità e confronto tra tipologie di sport.

Metodi

Lo studio sulle abitudini alimentari di sportivi dilettanti praticanti sport dell'Università di Pavia nell'anno 2017 si è svolto inizialmente mediante la somministrazione del questionario a 15 soggetti con un età media tra i 22-25 anni per verificare la comprensione degli item inseriti. Successivamente a questa verifica il questionario, composto da una prima sezione riguardante i dati socio-anagrafici e da 3 sezioni specifiche (frequenze di consumo degli alimenti (B), abitudini alimentari (C), utilizzo di integratori alimentari (K) per un totale di 59 item), è stato somministrato due volte (a distanza di una settimana per metodo del test-retest) a tutti i partecipanti ed auto compilato insieme al questionario standardizzato PREDIMED volto ad indagare l'aderenza alla Dieta Mediterranea (DM) [***8***]. L'analisi statistica ha riguardato l'utilizzo degli usuali indici descrittivi per la sintesi dei dati e delle tecniche volte a misurare l'affidabilità e il grado di stabilità nel tempo dello strumento e di generalizzabilità dei risultati (alpha di Cronbach e coefficiente di correlazione di

Pearson) delle sezioni C e K. Per la sezione B, non essendo una vera e propria scala, è stato calcolato solo il grado di stabilità per ogni item tra le due somministrazioni tramite il coefficiente di correlazione di Pearson. Per la sezione C è stato possibile anche validare esternamente lo score utilizzando i risultati del questionario PREDIMED verificando la relazione tra i due questionari (Test t di Student). Infine, tramite un'analisi inferenziale si è studiata la relazione tra gli item e il campione raggruppato in due categorie in base alla tipologia di sport (sport di terra: calcio, pallavolo, rugby, pallacanestro, pallamano e pattinaggio e sport in acqua: canoa e pallanuoto). Il software statistico utilizzato è STATA 14.2, College Station, Texas 77845 USA.

Risultati

101 sportivi dilettanti che praticano sport di squadra hanno compilato e completato il questionario in entrambe le somministrazioni con un 'età media di 22±2 anni (range 19-26) di cui 58 femmine e 43 maschi. La quasi totalità degli sportivi analizzati (90%) riporta abitudini alimentari adeguate mentre nella quota restante sono parzialmente adeguate. L'analisi di affidabilità ha evidenziato per la sezione C un alpha di Cronbach di 0.71 e un coefficiente di correlazione tra le due somministrazioni del test pari a 0.90 ($p<0.001$), mentre per la sezione K un alpha di 0.75 e un coefficiente di 0.87 ($p<0.001$). Questi risultati indicano una buona consistenza interna ed un ottimo grado di stabilità nel tempo del questionario. Per la sezione B il coefficiente va da un minimo di 0.35 ad un massimo 0.87. Tutti i coefficienti sono risultati essere statisticamente significativi a dimostrazione del fatto che per la maggior parte degli item esiste un buon grado di stabilità nella ripetibilità del questionario nei due tempi di somministrazione. L'analisi di validazione esterna ha mostrato che gli sportivi con un alto indice di aderenza alla DM (PREDIMED), presentano uno score della sezione C staticamente più alto rispetto a quelli con un indice inferiore (39±4 *vs* 35±6, $p=0.015$), evidenziando la validità della scala C nel misurare lo stesso costrutto.
Per quanto riguarda il confronto tra le due tipologie di sport (83% praticante sport di terra e 17% sport in acqua), si è riscontrato una differenza statisticamente significativa nello score della sezione C e K tra gli sportivi che praticano sport in acqua rispetto a quelli che praticano sport di terra, rispettivamente 39±4 *vs* 35±6 ($p=0.041$) e 7±4 *vs* 4.2±2 ($p<0.001$). Nello specifico, tra gli sportivi che svolgono sport di terra il 12% mostra delle abitudini parzialmente adeguate, l'88% adeguate, mentre in chi pratica sport in acqua il 100% ha delle abitudini alimentari adeguate. Per l'utilizzo degli integratori si riscontra una percentuale del 58% degli sportivi di acqua *vs* il 24% degli sportivi di terra ($p=0.004$).

Conclusioni

Il questionario adottato mostra un buona affidabilità nell'indagare le abitudini alimentari (C) e l'utilizzo degli integratori (K) negli sportivi dilettanti che svolgono attività di squadra. Quanto riscontrato nella sezione C risulta essere in linea con i risultati del PREDIMED per quanto riguarda l'aderenza alla dieta mediterranea. Quanto emerso invece dalla sezione relativa all'uso di integratori merita una riflessione. I professionisti della nutrizione sportiva convengono che gli atleti non dovrebbero usare supplementi dietetici per compensare una dieta sbilanciata, ma piuttosto consumare cibi "interi" per soddisfare le loro esigenze nutrizionali. Solo quando questo non è possibile gli atleti si possono affidare agli integratori alimentari [***9***]. I risultati ottenuti suggeriscono, quindi, un'analisi più approfondita sul tipo di dieta seguita dagli sportivi dilettanti.

Bibliografia

[**1**] Mujika I, Burke LM. Nutrition in team sports. Ann Nutr Metab 2010;57(2):26-35.
[**2**] American College of Sports Medicine Joint Position Statement, Nutrition and Athletic Performance. J Am Diet Assoc 2016;48(3):543-68.
[**3**] Mutsumi O, Eileen K, Sue R, et al. Nutrition and culture in professional football. A mixed method approach. Appetite 2012;58(1):98-104.

[4] Zapolska J, Witczak K, Mańczuk A, et al. Assessment of nutrition, supplementation and body composition parameters on the example of professional volleyball players. Rocz Panstw Zakl Hig 2014;65(3):235-242.

[5] Mielgo-Ayuso J, Zourdos MC, Calleja-Gonzalez J, et al. Dietary intake habits and controlled training on body composition and strength in elite female volleyball players during the season. Appl Physiol Nutr Metab 2015;40:827–34.

[6] Petri C, Mascherini G, Bini Vi, et al. Evaluation of physical activity and dietary behaviours in young athletes: pilot study. Minerva Pediatr 2017;69(6):463-9.

[7] Parnell JA, Wiens KP, Erdman KA. Dietary Intakes and Supplement Use in Pre-Adolescent and Adolescent Canadian Athletes. Nutrients 2016;8(9):526

[8] Martinez-Gonzalez MA, Fernandez-Jarne E, Serrano-Martinez M, et al. Development of a short dietary intake questionnaire for the quantitative estimation of adherence to a cardioprotective Mediterranean diet. Eur J Clin Nutr 2004;58:1550-2.

[9] Desbrow B, McCormack J, Burke LM, et al. Sports Dietitians Australia Position Statement: Sports nutrition for the adolescent athlete. Int J Sport Nutr Exerc Metab 2014;24:570–84.

L'ABUSO DI ALCOLICI FRA GLI ADOLESCENTI ITALIANI: L'ANDAMENTO DEGLI ULTIMI 15 ANNI ATTRAVERSO I DATI HBSC

Charrier Lorena[1], Berchialla Paola[2], Dalmasso Paola[1], Cavallo Franco[1]

1. *Department of Public Health and Paediatrics, University of Torino (Italy)*
2. *Department of Clinical and Biological Sciences, University of Torino (Italy)*

Introduzione

Il consumo e l'abuso di alcolici, così come altri comportamenti a rischio (fumo, *gambling*, uso di cannabis e altre sostanze illecite), tendono ad instaurarsi durante l'adolescenza, e quanto più precocemente ciò avviene, tanto più difficile diventa poi riuscire ad abbandonarli. [***1***] La sorveglianza di questi fenomeni, che sovente tendono a clusterizzare negli stessi soggetti, e dei contesti (familiare, scolastico, cerchia di amici) in cui vivono gli adolescenti risulta fondamentale non solo per monitorarne gli andamenti nel tempo ma anche per ricercarne i fattori ad essi associati. [***2***] Nello specifico, l'alcol è una delle sostanze più ampiamente diffusa e più comunemente utilizzata dagli adolescenti che possono ricorrervi per rispondere a bisogni sociali o personali, instaurare nuove relazioni con i pari, sentirsi da questi accettati. [***3, 4***] L'uso e l'abuso di alcolici in giovane età costituiscono d'altro canto uno dei principali problemi di sanità pubblica nei paesi europei e del Nord America: l'inizio precoce, il consumo frequente e l'abitudine ad eccedere sono associati a conseguenze in ambito psichico, sociale e di salute che vanno dal ritardo scolastico, agli atti di violenza, agli incidenti, all'uso e abuso di altre sostanze [***5***].
Valutare, attraverso i dati HBSC delle 4 indagini condotte fra il 2002 e il 2014, eventuali cambiamenti nell'abuso di alcolici (esperienze di ubriachezza e di *binge drinking,* quest'ultimo fenomeno indagato solo nelle due rilevazioni più recenti, 2010 e 2014) fra gli adolescenti italiani, appartenenti storicamente ad un paese a cultura "bagnata" che sembra progressivamente uniformarsi a quella del Nord Europa, caratterizzata da consumi globali minori ma più frequenti episodi di eccesso.

Metodi

Lo studio HBSC è una sorveglianza internazionale attiva sin dal 1982, condotta in ambito scolastico, che raccoglie dati sulla salute e il benessere degli adolescenti, l'ambiente in cui vivono e i comportamenti correlati alla salute. L'Italia è coinvolta sin dal 2001 e ha condotto sino ad ora 4 indagini, le prime 2 su campioni rappresentativi a livello nazionale, le ultime 2 su campioni rappresentativi anche a livello regionale. Il disegno dello studio è trasversale, ripetuto ogni 4 anni, al fine di simulare uno studio longitudinale. La procedura di campionamento segue le linee guida tracciate a livello internazionale con l'obiettivo di selezionare campioni rappresentativi della popolazione di 11, 13 e 15 anni. I soggetti sono selezionati mediante procedure di campionamento a cluster in cui l'unità di campionamento primaria è la classe, selezionata mediante procedimento sistematico dalla lista completa delle scuole, pubbliche e private, fornita dal Ministero dell'Istruzione. La dimensione campionaria raccomandata è di 1500 soggetti per ciascuna fascia di età. Il questionario, standardizzato, è auto-compilato in classe dai ragazzi, in forma anonima; la partecipazione allo studio è su base volontaria [***6***].
Le analisi sono condotte sul database relativo alle indagini HBSC 2002, 2006, 2010 e 2014, e tengono conto dell'effetto del disegno complesso (cluster, stratificazione e peso campionario) sulla precisione delle stime. Le analisi dei trend temporali sono state effettuate mediante modelli di regressione logistica, stratificando per genere ed età. OR e relativi intervalli di confidenza ($IC_{95\%}$) sono calcolati per ciascuna indagine e complessivamente, utilizzando, ad esempio, come variabile dipendente l'essersi ubriacati almeno 2 volte nella vita (dicotomizzando la variabile in sì *versus* no), e l'anno dell'indagine e il Family

Affluence Scale (FAS) - un indicatore del benessere socio-economico utilizzato in HBSC, come variabili indipendenti. Le analisi sono state condotte stratificando per genere e per fascia di età.

Risultati

Dal confronto fra i dati delle due rilevazioni più recenti sembra emergere, coerentemente con quanto è stato rilevato in molti paesi, un calo del consumo di alcolici, soprattutto di quello frequente (almeno una volta alla settimana o ogni giorno) in entrambi i generi, principalmente fra i maschi di 15 anni che passano, nel nostro paese, da una prevalenza del 40% nel 2010 al 33% nel 2014. Una riduzione di quattro punti percentuali è stata rilevata anche fra i maschi di 13 anni e di 3 punti fra quelli di 11. Anche fra le femmine di 13 e 15 anni è emersa una riduzione, seppur meno evidente rispetto ai coetanei maschi. Parallelamente alla riduzione del consumo settimanale/quotidiano, i dati del 2014 sembrano tuttavia registrare un aumento rispetto al passato della quota di adolescenti che dichiarano forme di abuso di bevande alcoliche: sia fra i maschi che fra le femmine di 15 anni emerge un aumento significativo della prevalenza di ragazzi che dichiarano di essersi ubriacati almeno 2 volte nella vita (***Figura 1***). Il fenomeno del *binge drinking*, rimane sostanzialmente stabile rispetto al 2010 fra i maschi di 15 anni (38-39%), mentre passa dal 24 al 30% la quota di ragazze della stessa età che riferiscono di aver fatto questa esperienza almeno una volta nell'ultimo anno.

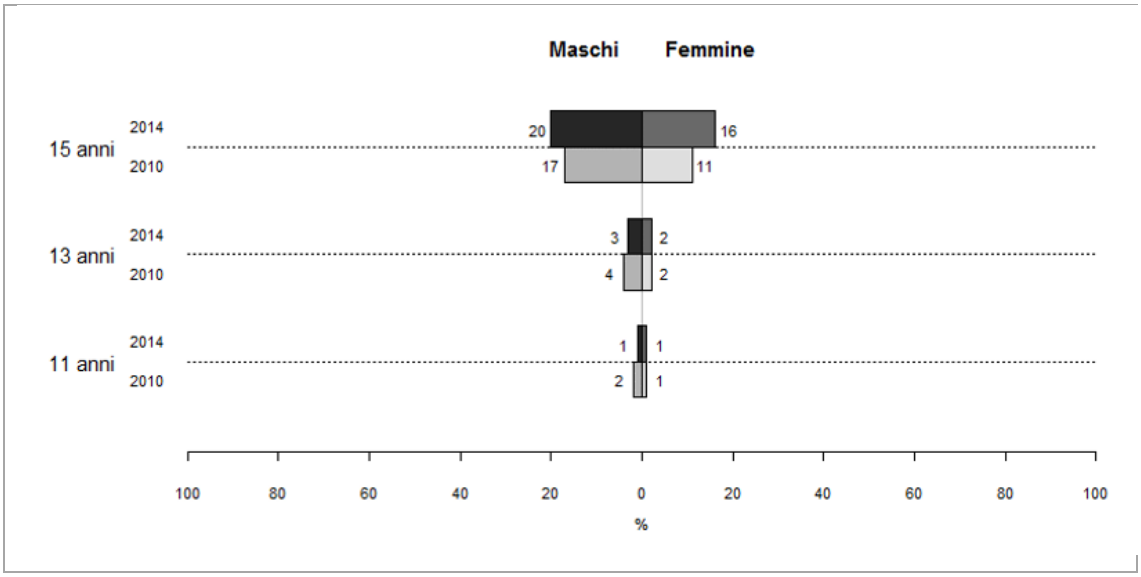

Figura 1. Percentuale di ragazzi che si sono ubriacati almeno 2 volte nella vita – per genere, età e anno di indagine

Conclusioni

Il quadro che emerge sembra fotografare un cambiamento del comportamento degli adolescenti italiani rispetto al consumo e all'abuso di alcolici, con un progressivo avvicinamento alla cultura del bere tipica dei Paesi nord-europei, caratterizzata da consumi globali minori e, al contempo, da più frequenti episodi di eccesso.

Bibliografia

[**1**] Dick B, Ferguson BJ. Health for the World's Adolescents: A Second Chance in the Second Decade. J Adolesc Health 2015;56(1):3-6

[**2**] Spring B, Moller AC, Coons MJ. Multiple health behaviours: overview and implications. J Public Health 2012;34(1):i3-i10

[**3**] Anderson P, Baumberg B. Alcohol in Europe. A public health perspective. Brussels: European Commission. 2006 Available from: (http://ec.europa.eu/health/archive/ph_determinants/life_ style/alcohol/documents/alcohol_ europe_en.pdf, [accessed 13 June 2017]

[**4**] Engels RCME, ter Bogt T. Influences of risk behaviours on the quality of peer relations in adolescence. J Youth Adolesc 2001;30(6):675–95

[**5**] Boden JM, Fergusson DM. The short and long term consequences of adolescent alcohol use. In: Saunders J, Rey JM, editors. Young people and alcohol: impact, policy, prevention and treatment. Chichester: Wiley-Blackwell; 2011:32–46

[**6**] 4° Rapporto sui dati italiani dello studio internazionale HBSC. A cura di: Cavallo F, Lemma P, Dalmasso P, et al. Stampatre s.r.l. – Torino; 2016

THE PROPORTION OF RA CAUSED BY SMOKING INCREASE WITH NUMBER OF PACK YEARS SMOKED

Carlucci Claudia[1], Berglund Daniel[2], Bellocco Rino[1], Santacatterina Michele[3], Klareskog Lars[4], Alfredsson Lars[3*], Westerlind Helga[3*]

1. *University of Milano-Bicocca, Italy*
2. *Department of Mathematics, Royal Institute of Technology (KTH), Sweden*
3. *Institute of Environmental Medicine (IMM), Karolinska Institutet, Sweden*
4. *Department of Medicine (Solna), Karolinska Institutet, Sweden*

* *These authors contributed equally*

Introduction

Rheumatoid arthritis (RA) is an autoimmune disease resulting in symmetric inflammation and destruction of the joints. Like other complex diseases, the risk factors are both genetic and environmental and two of the strongest risk factors is the genetic factor shared epitope (SE) and smoking. Although factors associated with the risk for the disease are continuously being discovered, the causality of RA has not been fully clarified yet. When working with risk for complex diseases, a statistical model is used to investigate the underlying susceptibility. One model for complex diseases is the sufficient cause model proposed by Rothman [***1***]. A complex disease can be caused by several different mechanisms. By understanding what risk factors are within the same mechanism, we can learn more about the etiology of the disease. For this additive (also referred to as biological or sufficient cause) interaction can be used. An additive interaction is present if the risk of contracting the disease for people exposed to more than one factor is greater than the sum of the risk of the factors separately. The above definition of additive interaction is based on dichotomous risk factors, but in real life risk factors are often continuous.

An example of a validated additive interaction in RA would be smoking and SE. In a recent study [***2***] it was observed that individuals with two copies of SE alleles that had smoked over 20 pack years of cigarettes were almost 20 times more likely to develop RA than individuals with no copies of SE smoking the same amount of cigarettes. However, it is likely that dichotomizing of continuous factors causes loss of valuable information.

The aim of this study was to apply the method for assessing interaction with continuous variables suggested by Knol et al on the RA risk factors SE and smoking.

Methods

Data was obtained from the Swedish Epidemiological Investigation of Rheumatoid Arthritis (EIRA) [***3***], an ongoing prospective observational study in Sweden aimed at investigating the causal factors, both genetic and environmental, for RA. This case-control study begun in 1996, and across the years it has so far involved 4000 patients diagnosed with RA and 7000 controls randomly selected from a population register and matched for age, sex and residential area.

Each case was diagnosed according to the 1987 ACR criteria for RA by rheumatologists at the corresponding clinic. All participants completed an extensive questionnaire covering a wide spectrum of issues such as smoking, occupational history, exposures of working environment, alcohol, psychosocial environment and history of other diseases. A blood sample was also taken for genetic analyses.

Data on smoking frequency and duration was obtained from the questionnaire. Smokers were defined as individuals who reported that they smoked or had previously smoked cigarettes and non-smokers were defined as those who reported that they had never smoked cigarettes. Those who smoked pipes or cigars

were excluded, thus ever smokers were only cigarette smokers. Pack-years (PYs) was used as the measurement for quantifying the amount of smoking where one pack year is equivalent to smoking one pack of 20 cigarettes per day for one full year. Therefore, an individual smoking two pack-years means that he or she either smoked one pack of cigarettes every day for two years, two packs of cigarettes per day for one year, or an equivalent amount of cigarettes during any time period.
The method for estimation of interaction between continuous determinants suggested by Knol and Geerlings [***4***] was used. We fitted a logistic regression model to investigate the role of smoking, SE and their interaction of the two factors on the risk of RA occurrence. In this model, smoking was entered as a continuous independent variable. This way we were able to assess odds ratios (OR) and corresponding 95% confidence intervals (CI). Further- more, we estimated the three measures of additive interaction introduced by Rothman and Hosmer and Lemeshow: the relative excess risk due to interaction (RERI), the attributable proportion due to interaction (AP) and the synergy index (S) [***5***]. To calculate RERI, AP and S, the following three ORs are needed: e^{β_1} , e^{β_2} , $e^{\beta_3+\beta_2+\beta_1}$ .The formulas to asses presence of interaction on an additive scale and to estimate the RERI are: $RERI = e^{\beta_3+\beta_2+\beta_1} - e^{\beta_1} - e^{\beta_2} + 1$. The formula of the other two measures are AP=$\frac{RERI}{e^{\beta 3+\beta 2+\beta 1}}$ and S=$\frac{e^{\beta 3+\beta 2+\beta 1}-1}{(e^{\beta 1}-1)+(e^{\beta 2}-1)}$. Delta Method was used to calculate the corresponding 95% CIs. Finally, to verify the robustness of the estimates derived from delta method, CIs were reassessed using bootstrap techniques.

Results

Through EIRA, we were able to identify 2103 ACPA-positive RA patients and 2409 controls with smoking data and information on SE alleles. ***Table 1*** shows the results of the additive interaction test between SE and different amounts of PYs smoked. A significant additive interaction between each category of cigarette smoking and the presence of any SE allele was noticeable through the values of the three measures previously described. For five PYs smoked and one copy of the allele, RERI is 0.5 (95% CI 0.36 to 0.65), AP is 0.12 (95% CI 0.09 to 0.15) and S is 1.19 (95% CI 1.06 to 1.33). A stronger additive interaction was observed for the same amount of PY and two copies of SE alleles, with a RERI of 1.61 (95% CI 0.92 to 2.31), an AP of 0.13 (95% CI 0.09 to 0.19) and an S of 1.17 (95% CI 1.07 to 1.27).
We could see not only that RERI increased with the number of cigarettes smoked, but also, more importantly, that the proportion of cases attributable to this interaction increased. In the group that smoked five pack years and had two copies of SE alleles, 13% of the cases were attributed to the interaction whereas, in the group that smoked 25 pack years, a stunning 50% of the cases were attributed to the interaction. This difference was statistically significant.

Table 1. Odds ratios and the measures of additive interaction for different doses of smoking and SE alleles, regarding the risk to develop ACPA positive RA.

HLA-SE	Pack-years smoking	OR	*[95% CI]*	RERI	*[95% CI]*	AP	*[95% CI]*	S	*[95% CI]*
Single	5	4.16	*[3.50;4.83]*	0.50	*[0.36;0.65]*	*0.12*	*[0.09;0.15*	*1.19*	*[1.06;1.33]*
	10	4.92	*[4.26;5.86]*	1.10	*[0.77;1.44]*	*0.22*	*[0.17;0.28]*	*1.39*	*[1.21;1.60]*
Shared	15	5.82	*[5.15;6.49]*	1.82	*[1.23;2.40]*	*0.31*	*[0.24;0.38]*	*1.61*	*[1.40;1.88]*
	20	6.89	*[6.20;7.58]*	2.67	*[1.76;3.59]*	*0.39*	*[0.30;0.47]*	*1.83*	*[1.53;2.19]*
Epitope	25	8.15	*[7.42;8.88]*	3.70	*[2.36;5.04]*	*0.45*	*[0.36;0.55]*	*2.07*	*[1.70;2.52]*
	P value for trend						***0.1858***		
Double	5	12.20	*[9.65;14.76]*	1.61	*[0.92; 2.31]*	0.13	*[0.09;0.19]*	1.17	*[1.07;1.27]*
	10	14.26	*[11.73;16.80]*	3.50	*[1.89; 5.12]*	0.25	*[0.15;0.34]*	1.36	*[1.21;1.53]*
Shared	15	16.67	*[14.15;19.19]*	5.71	*[2.88; 8.54]*	0.34	*[0.22;0.47]*	1.57	*[1.35;1.84]*
	20	19.48	*[16.96;22.01]*	8.30	*[3.89;12.71]*	0.43	*[0.28;0.57]*	1.82	*[1.50;2.20]*
Epitope	25	22.77	*[20.23,25.31]*	11.33	*[4.88;17.78]*	0.50	*[0.34;0.65]*	2.09	*[1.66;2.63]*
	P value for trend						***0.0160***		

Conclusions

Results from this large cohort case-control study confirms that the coexistence between smoking and SE genes highly increases the risk of incident RA among Swedish people through a significant additive interaction. Furthermore, it has been confirmed that gene-environment interactions for both single and double SE alleles in ACPA-positive disease are additive, basing upon three independent measures of interaction: RERI, AP, and S. Patients with both risk factors (SE and smoking more than 20 PYs) were eight to thirty times more likely to have ACPA-positive RA than patients lacking both risk factors. However, the risk was higher for double SE alleles carriers who had smoked 25 PYs (OR: 22.7 [95% CI: 20.23-25.31] AP: 0.50 [95% CI: 0.34 - 0.65]), meaning that 50% of the "excess" risk in this group is directly attributable to the gene-environment interaction. Although the different number of SE alleles are associated with different magnitudes of increased risk of ACPA-positive RA, their interaction with smoking seems to be similar according to the magnitude of the AP values.

This results highlight not only the importance of prevention of RA through smoking cessation, but also casts further light on the etiology of RA through the impact of smoking and SE and risk for developing disease.

References

[1] Kenneth J Rothman, Sander Greenland, and Timothy L Lash. Modern epidemiology. Lippincott Williams & Wilkins, 2008.

[2] Ka¨ llberg Hk, Ding B, Padyukov L, Bengtsson C, Ro¨ nnelid J, Klareskog L, Alfredsson L, EIRA Study Group, et al. Smoking is a major preventable risk factor for rheumatoid arthritis: estimations of risks after various exposures to cigarette smoke. Annals of the rheumatic diseases 2011;70(3):508–11.

[3] Stolt P, Bengtsson C, Nordmark B, Lindblad S, Lundberg I, Klareskog L, Al- fredsson L. Quantification of the influence of cigarette smoking on rheumatoid arthri- tis: results from a population based case-control study, using incident cases. An- nals of the rheumatic diseases 2003;62(9):835–41.

[4] Knol MJ, van der Tweel I, Grobbee DE, Numans ME, and Geerlings MI. Estimating interaction on an additive scale between con- tinuous determinants in a logistic regression model. International journal of epidemiology 2007;36(5):1111–8.

[5] Knol MJ, VanderWeele TJ, Groenwold RHH, Klungel OH, Rovers MM, and Grobbee DE. Estimating measures of interaction on an additive scale for preventive exposures. European journal of epidemiology 2011;26(6):433–8.

PREVALENZA DI OSAS NELLA PROVINCIA DI PESARO-URBINO: SCREENING SU TITOLARI DI PATENTE SUPERIORE E/O PROFESSIONALE DELL'AREA VASTA 1.

Rocchi Marco Bruno Luigi[1], Vittoria Emanuela[2], Carlotti Eugenio[3], Amatori Stefano[1], Grossi Paola[4], Sisti Davide[1]

1. *Dipartimento di Scienze Biomolecolari Unità di statistica medica e biometria - Università di Urbino "Carlo Bo"*
2. *Dipartimento di Giurisprudenza - Università di Urbino "Carlo Bo"*
3. *ASUR Marche - Area Vasta 1 - Dipartimento di Prevenzione – UOC SPSAL*
4. *Medico Competente Specialista in Medicina del Lavoro*

Introduzione

La sindrome delle apnee ostruttive del sonno (*Obstructive Sleep Apnea Syndrome*, OSAS) è una patologia determinata da episodi ripetuti di ostruzione parziale (ipopnea) o completa (apnea) delle vie aeree superiori, che determinano desaturazioni ossiemoglobiniche e frammentazione del sonno, sintomatologia diurna conseguente alla privazione di sonno, spesso associata a comorbidità cardiovascolari e metaboliche. La maggior parte dei pazienti al risveglio si sente stanca e non riposata, a prescindere dalla durata del tempo trascorso a letto. Uno dei sintomi diurni caratteristici è rappresentato dalla sonnolenza eccessiva. È dimostrato che la sonnolenza diurna conseguente alla deprivazione di sonno riscontrabile in circa il 50% dei pazienti OSAS ha un impatto negativo sulla loro capacità di condurre veicoli [***1***], determinando un aumento del rischio di incidenti stradali, gravi o mortali e infortuni lavorativi.Il sospetto clinico di OSAS deve essere preso in considerazione in presenza di alcuni sintomi notturni (russamento abituale e persistente, ovvero presente tutte le notti da almeno 6 mesi; pause respiratorie nel sonno riferite dal "bed-partner"; risvegli con sensazione di soffocamento o "*chocking*", nicturia in assenza di altre cause evidenti) e diurni (sonnolenza valutata con questionari ad hoc o con metodiche oggettive, astenia, iporeattività, cefalea mattutina, irritabilità, diminuzione delle prestazioni sessuali). L'esame obiettivo è volto a ricercare alcuni segni predittivi di OSAS, quali: Indice di Massa Corporea (*Body Mass Index*, BMI) >29 Kg/m², Circonferenza del Collo >43 cm nei maschi e >41 cm nelle femmine, eventuali dismorfismi cranio-facciali (ad esempio ipoplasia della mandibola) o anomalie oro-faringee che determinano una riduzione del calibro delle prime vie respiratorie comportando aumento della resistenza al passaggio del flusso aereo. [***2***]. La diagnosi di OSAS può essere formulata tramite Polisonnografia (gold standard) o Monitoraggio Cardiorespiratorio Notturno (MCRN). Gli studi epidemiologici sull'OSAS hanno portato a risultati variabili in relazione a differenze nelle caratteristiche delle popolazioni esaminate, nelle metodologie di analisi, nelle strumentazioni e negli standard diagnostici impiegati. Secondo recenti studi internazionali, la sindrome è più frequente nei soggetti di età compresa fra 40 e 65 anni, con una prevalenza del 2-4% nei maschi e dell'1-2% nelle femmine [***3***].
Lo scopo del seguente lavoro è studiare a prevalenza di OSAS nella popolazione da cui è stato selezionato il campione; inoltre al fine di ottimizzare la capacità di diagnosi in funzione e razionalizzare l'utilizzo del MCRN, si è studiata l'associazione tra la presenza di OSAS ed alcuni variabili antropometriche e cliniche.

Metodi

Il lavoro, in collaborazione con il Servizio Prevenzione e Sicurezza Ambienti di Lavoro dell'ASUR Marche Area Vasta 1 e dalla Commissione Medica Locale Patenti di Guida (CML), è finalizzato a rilevare la possibile presenza di OSAS nei soggetti possessori di patenti di categoria superiore alla B, di categoria B ad uso professionale e nei soggetti segnalati dalla Motorizzazione Civile per revisione della patente conseguente a "colpo di sonno" alla guida. I soggetti sottoposti a visita medica di screening per OSAS quindi sono stati reclutati non a partire dalla popolazione generale ma da un particolare sottogruppo selezionato dalla

popolazione generale. Nella valutazione clinica conclusiva è segnalata l'indicazione per MCRN (nel caso di sospetta OSAS) o la non indicazione al MCRN (nel caso di screening ad esito negativo). Sono state rilevate per ogni soggetto in screening le seguenti variabili: età, sesso, BMI, circonferenza del collo, angolo collo-mento; classe Mallampati; punteggio ESS; familiarità per patologie cardiovascolari, ipertensione, cardiopatia ischemica, fibrillazione striale, ictus, patologie metaboliche. Per studiare la capacità di indentificare correttamente un soggetto con OSAS in funzione di alcune variabili anamnestiche, si è utilizzata un modello di regressione logistica multipla, con backward stepwise elimination.

Risultati

Su un totale di 553 (248 a Urbino e 305 a Pesaro) soggetti sottoposti a visita di screening per OSAS possessori di patente di guida superiore e/o professionale, i maschi erano il 97,2% ad Urbino e il 97,4 % a Pesaro. I soggetti apparivano mediamente in sovrappeso, con molteplici casi di conclamata obesità. L'età media (59.7 anni ad Urbino, 57.8 anni a Pesaro) era del tutto simile nei campioni provenienti dai due diversi centri, così come la circonferenza del collo (41.6 vs 41,7). Nei campioni circa il 60% dei soggetti erano fumatori od ex fumatori, mentre il 73.7% ad Urbino ed il 50.2% a Pesaro ha dichiarato di assumere abitualmente bevande alcoliche. Tra le patologie associate, prevalgono l'ipertensione arteriosa (36.2% e 31.0%), le cardiopatie (15.0% e 11.3%), le malattie metaboliche (16.3% e 11.3%). I soggetti, probabilmente anche per l'età, assumevano farmaci nel 56.3% e 48.5% dei casi. Considerando solo il campione risultati positivi allo screening (74 a Urbino, 84 a Pesaro), il 32.3% dei soggetti sottoposti alle prove per la diagnosi di OSAS non ne erano affetti, mentre il 67.7% dei soggetti mostravano almeno sintomi da apnee ostruttive notturne; da sottolineare come il 21.5% dei soggetti mostrasse una forte gravità dei sintomi riconducibili all'OSAS. (***Figura 1***) In generale, nei campioni provenienti dalle due città le percentuali di OSAS erano molto simili, indicando una sostanziale sovrapposizione ($p\ \chi^2>0.05$)

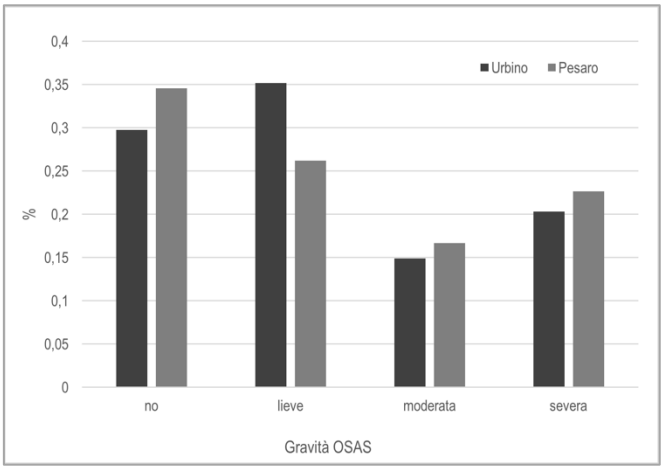

Figura 1. Percentuali di OSAS negativi e di OSAS a differente gravità (lieve, moderata, severa) nei centri di Urbino e Pesaro.

La probabilità che un soggetto sia soggetto ad OSAS (almeno lieve) è associato solamente al BMI (O.R=1.091; IC 1.007-1.182; p<0.033) e all'età, (O.R=1.057; IC 1.022-1.093; p<0.001), con sensibilità=96.3%, ma con specificità insolitamente bassa (16.7%), con una accuratezza complessiva del 76.4%

Discussione

La maggior frequenza di OSAS ottenuta entro la fascia di età compresa tra i 45 e i 65 anni è coerente con i dati epidemiologici attualmente disponibili in Letteratura [***3***]. I dati relativi alla prevalenza, confermano che la OSAS è una sindrome a prevalenza sottostimata. I dati relativi al BMI e all'età, confermano che la OSAS tende a manifestarsi maggiormente (e nelle forme di maggior gravità) in soggetti con obesità; tuttavia poco meno della metà della popolazione analizzata affetta da OSAS non presenta una concomitante obesità. L'indagine conferma che la OSAS è una patologia ampiamente diffusa e la cui prevalenza risulta sottostimata nella popolazione generale e lavorativa. I soggetti che ne sono maggiormente affetti (e che presentano la malattia nella sua forma a maggiore gravità) sono obesi. Tuttavia, considerato che quasi la metà della popolazione oggetto dello studio aveva un BMI<30 Kg/m², i segni e i sintomi a supporto del sospetto della patologia vanno ricercati anche nei soggetti che non presentano obesità. Una temuta complicanza in corso di OSAS è data dall'eccessiva sonnolenza diurna, che può comportare deficit attentivi fino all'addormentamento in circostanze in cui è richiesto uno stato di vigilanza integro, determinando condizioni che predispongono a rischi per la sicurezza (soprattutto sulle strade e negli ambienti di lavoro). Ulteriori gravi complicanze sono date dalle frequenti comorbidità (cardiache, vascolari e metaboliche), che inseriscono la OSAS in un quadro di coinvolgimento multisistemico complesso. La ricerca di segni e sintomi a supporto del sospetto diagnostico di OSAS consente infatti di tutelare la salute e la sicurezza dei lavoratori rispetto ad una patologia ampiamente sottostimata e sottodiagnosticata, caratterizzata da possibili complicanze potenzialmente gravi e mortali. Sono tutt'ora in corso studi ulteriori su altre province della regione Marche.

Bibliografia

[**1**] Pizza F, Contardi S, Mondini S, Cirignotta F.G. Impact of sleep deprivation and obstructive sleep apnea syndrome on daytime vigilance and driving performance: a laboratory perspective. Ital Med Lav Ergon 2012; 34(3):375-7.

[**2**] Mallampati SR, Gatt SP, Gugino LD, Desai SP, Waraksa B, Freiberger D, Liu PL.A clinical sign to predict difficult tracheal intubation: a prospective study. Can Anaesth Soc J 1985;32(4):429-34.

[**3**] Maspero C, Giannini L, Galbiati G, Rosso G, Farronato G. Obstructive sleep apnea syndrome: a literature review. Minerva Stomatol 2015;64(2):97-109.

ANGULAR SMOOTHING AREA: A NEW APPROACH REVEALED BALANCE FEATURES OF RHYTHMIC GYMNASTICS IN A SIMPLE BIPODAL STEADINESS.

Sisti Davide[1], Calavalle Anna Rita[2], Amatori Stefano[2], Gervasi Marco[2], Rocchi Marco Bruno Luigi[1]

1. *Dipartimento di Scienze Biomolecolari, Unità di statistica medica e biometria, Università degli Studi di Urbino.*
2. *Dipartimento di Scienze Biomolecolari, Istituto di ricerca sull'attività motoria, Università degli Studi di Urbino.*

Introduzione

I sistemi visivo, vestibolare e somato-sensoriale inviano informazioni al sistema senso-motorio per il mantenimento dell'equilibrio. Anche il controllo posturale dipende da questa complessa interazione. La capacità di equilibrio e il controllo posturale sono collegati all'età e a fattori genetici, che possono però essere influenzati dall'allenamento [***1, 2***]. La force platform è utilizzata ormai da qualche decennio per l'evidenziazione di particolari strategie utilizzate da atleti e non nel controllo della postura, particolarmente importante in sport come la ginnastica sia artistica che ritmica [***3, 4***]. L'interesse verso questo studio deriva dal fatto che il controllo posturale è alla base di ogni tecnica sportiva, per migliorare la performance e prevenire infortuni. Inoltre la capacità di controllo della postura nella vita quotidiana è fondamentale come prevenzione a incidenti da caduta [***5***]. La misurazione della stabilità posturale viene di solito eseguita tramite la pedana di forza, che visualizza lo spostamento del centro di pressione (COP) nel tempo. La traiettoria risultante, visualizzata nel piano orientato lungo l'asse antero-posteriore (AP) e medio-laterale (ML) viene denominata stabilogramma (sway). Sono stati proposti, da diversi autori [***6***], svariati indici correlati allo sway; in particolare si è cercato di associare una superficie allo sway, in modo da quantificare l'area associata alla distribuzione dei COP. In letteratura si è evidenziato che la maggiore abilità di controllo dell'equilibrio negli atleti si riscontra tra i ginnasti, ma si esplica solo aumentando la difficoltà della posizione, con limitazione di input propriocettivi o con aggiunta di movimenti degli arti liberi e del busto. Anche nelle prove ad occhi chiusi i ginnasti si confermano migliori, rispetto a atleti di altri sport e non atleti, ma sempre in situazioni di aumentata difficoltà. Inoltre anche i ginnasti si dimostrano meno controllati senza il riferimento ottico rispetto ai propri risultati ad occhi aperti, in tutti gli indici posturali calcolati a partire dallo sway [***2, 3, 7***]. In questo lavoro si utilizzano i dati del campione di ginnaste testate in una ricerca precedente [***4***] nella quale lo sway era stato analizzato con gli indici più usati in letteratura, nel dominio del tempo (Discrete Fourier Trasform) e dello spazio (assi antero-posteriore e medio-laterale, l'Area del Cerchio di Confidenza, L'Area dell'Ellisse Standard). I risultati evidenziavano che le ginnaste esperte a differenza degli altri atleti si dimostravano migliori di atlete amatoriali di altri sport, in una posizione bipodalica di base, non propria della loro expertise, in particolare sulla direzione ML nei tempi da 0.5 a 10 secondi erano sempre migliori delle Non Atlete. Inoltre, seppur dipendenti dalla visione non si dimostravano peggiori del controllo. I risultati però non mostravano differenze relative area dell'ellisse di confidenza tra atlete e controlli.

Nel presente lavoro è stato utilizzato un nuovo indice da noi proposto riferito all'area dello sway, l'ASmA (Angular Smoothing Area). Questo indice consente di evidenziare lo spostamento del COP del soggetto in relazione al proprio baricentro descrivendo la "forma" dello sway. È quindi possibile calcolare lungo quale direzione si sono ottenute le maggiori oscillazioni, e valutare la percentuale di scostamento lungo qualsiasi direzione, relativamente alla distanza media. L'indice è stato utilizzato per verificare la presenza di eventuali differenze tra un gruppo di atlete di ginnastica ritmica e un campione di soggetti omogenei per età, sesso e BMI.

Metodi

Lo sway è stato misurato in un gruppo di 15 atlete di ginnastica ritmica (RG) e un gruppo di 20 studentesse della Scuola di Scienze Motorie (OS). Le condizioni di equilibrio sono state testate chiedendo ai soggetti di rimanere in posizione eretta in appoggio bipodalico, scalzi, su una pedana di forza cercando di rimanere immobili per 60 secondi. Due condizioni sperimentali sono state testate: occhi aperti (EO) e occhi chiusi (EC). I soggetti posizionavano i piedi al centro di gravità della pedana di forza (CG), formando un angolo di circa 30° e con i talloni distanziati di circa 3 cm. I COP dei soggetti, quindi, corrispondevano al CG della pedana. Nella condizione EO, ai soggetti è stato chiesto di fissare un obiettivo all'altezza degli occhi, distante circa 2 metri dalla pedana. È stata effettuata una prova per ogni condizione sperimentale. Le forze di reazione verticali sono state misurate utilizzando una pedana di forza campionata a 100 Hz e trasferite su un PC utilizzando un convertitore A/D a 10 bit. L'ASmA media è legata alla distribuzione locale dei COP; si ottiene calcolando il baricentro dei punti interni ad un angolo di ampiezza data b. Tale angolo viene fatto ruotare con il principio della media mobile, ovvero l'angol b ruota in senso antiorario di un angolo p minore di b; se p = b i settori sono contigui senza sovrapposizione, mentre se p < b i settori si sovrappongono. In seguito, unendo i baricentri di ogni settore, si definisce il tracciato che delimita l'ASmA. Inoltre è stata ottenuto un ulteriore indice di forma (ASmA max) in cui il peso di ciascun punto è il punteggio z derivante da una gaussiana il cui massimo coincide con il punto con il modulo più grande rispetto al centro. Con tale procedura si ottiene un poligono che definisce la forma globale, racchiudendo al proprio interno la quasi totalità dei punti. L'area dell'ASmA (media e max) e le relative variabili (dmax, dmin, dmedia) sono state soggette a un'analisi della varianza a due vie (ANOVA) a misure ripetute. Il fattore tra soggetti era il gruppo di appartenenza (ginnaste e controlli; RG-OS) mentre il fattore entro soggetti era lo stato della vista (occhi aperti chiusi; EO-EC). L'errore family-wise è stato fissato a p = 0.05. Le analisi statistiche sono state effettuate utilizzando lo Statistical Package for Social Sciences (SPSS 20.0) e Excel 2007 (Microsoft software) con opportune programmazioni di Macro in VBA.

Risultati

ASmA Media

Occhi Aperti: RG e OS sono significativamente differenti per quanto riguarda la distanza minima (-91.2%) e la distanza media (-20.5%), mentre non ci sono differenze significative per l'area e la distanza massima. Occhi Chiusi: sono state trovate differenze significative tra RG e OS solo per quanto riguarda la distanza minima (-93.4%), mentre non ci sono differenze significative per l'area, la distanza media e distanza massima.

ASmA Max

Occhi Aperti: sono state trovate differenze significative tra RG e OS per quanto riguarda l'area (-38.0%), la distanza minima (-92.8%) e la distanza media (-28.1%), mentre non ci sono differenze significative per la distanza massima. Occhi Chiusi: sono state trovate differenze significative tra RG e OS per quanto riguarda l'area (-29.9%), la distanza minima (-95.1%) e la distanza media (-22.8%), mentre non ci sono differenze significative per la distanza massima. I risultati del confronto tra le aree dell'ASmA media e ASmA max, tra RG e OS, in condizioni di EO e EC, sono mostrati in ***figura 1***, dove si riporta a titolo esemplificativo la rappresentazione grafica dell'ASmA media e ASmA max relativi ad uno specifico sway.

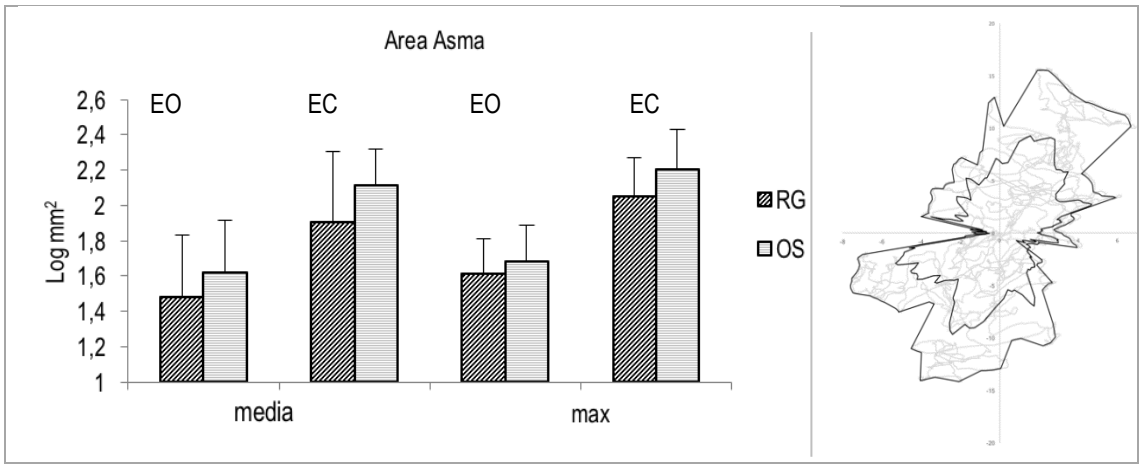

Figura 1. A sinistra: le medie dei valori dell'ASmA rispetto alla distanza media e alla distanza massima, sui due gruppi (RG - OS) nelle due condizioni sperimentali testate (EO - EC). A destra: grafico di ASmA media e max (in nero) relative ad uno sway (in grigio).

Conclusioni

L'ASmA consente di ottenere delle differenze tra RG e OS basandosi su un'area associata allo sway, laddove l'ellissi di confidenza risultava non significativa. Le differenze tra RG ed OS si evidenziano soprattutto considerando l'ASmA max; da ciò si può dedurre che le atlete non si comportino in maniera differente dai soggetti di controllo nell'area definita dai baricentri dei settori angolari e il centro dello sway, cioè considerando AsmA media. In quest'area si ha il massimo di equilibrio in quanto ogni spostamento è vicino al baricentro. Interessante è notare come le oscillazioni che fuoriescono dall'area delimitata dall'ASmA media, quantificate dall'ASmA max siano significativamente minori nelle atlete. Tali risultati validano, almeno nello specifico contesto sportivo, l'utilizzo di questo indice forma-dipendente e suggerisce l'applicazione di tale indice anche in contesti clinici, soprattutto nell'ambito della prevenzione delle cadute e dei disturbi dell'equilibrio nei soggetti anziani. È quindi possibile sintetizzare gli aspetti positivi dell'AsmA, che ci consentono di:

- Associare un poligono allo sway
- ottenere una informazione completa, poiché tiene conto di tutti i dati a disposizione a differenza di altri indici forma dipendenti
- calcolare lungo quale direzione si sono ottenute le maggiori oscillazioni;
- calcolare l'area del poligono;
- definire, attraverso la dipendenza dalla forma, lo stato di anisotropia dello sway;

Bibliografia

[1] Golomer E, Dupui P, Sereni P, et al. The contribution of vision in dynamic spontaneous sways of male classical dancers according to student or professional level. J Physiol 1999;93:233-7.

[2] Vuillerme N, Danion F, Marin L, et al. The effect of expertise in gymnastics on postural control. Neurosci Lett 2001;303:83-86.

[3] Asseman FB, Caron O, Cremieux J, Are there specific conditions for which expertise in gymnastics could have an effect on postural control and performance? Gait Posture 2008;27:76-81.

[4] Calavalle AR, Sisti D, Rocchi MBL, et al. Postural trials: expertise in rhythmic gymnastics increases control in lateral directions. Eur J Appl Physiol 2008;104:643-9.

[**5**] Brouwer B, Musselman K, Culham E, Physical function and health status among seniors with and without fear of falling. Gerontology 2004;50:135-41.

[**6**] Prieto TE, Hoffman RG, Lovett EG, et al. Measures of postural steadiness: differences between healthy young and elderly adults. IEEE, 1996;43(9):956-66.

[**7**] Carrick FR, Oggero E, Pagnacco G, et al. Posturographic testing and motor learning predictability in gymnasts Disability and Rehabilitation 2007;29(24):1881-9.

HELMET USAGE PATTERNS AMONG NON-COMPETITIVE ITALIAN CYCLISTS

Ferraro Ottavia Eleonora, Popa Ioana, Orsi Chiara, Morandi Anna, Montomoli Cristina

Department of Public Health, Experimental and Forensic Medicine, Biostatistics and Clinical Epidemiology Unit, Centre of Study and Research on Road Safety, University of Pavia, Italy

Introduction

Bicycle is playing an important role in the future transportation scenario [***1***] and consequently it is interesting to understand more about helmet and its use. Although a large number of studies proved helmet's protective role [***2-3***] there are contrasting opinion on its real usefulness [***4***].
The aim of this study is to analyse which factors are associated with using or not using of the helmet among non-competitive Italian cyclists.

Methods

Data about helmet use and potential factors associated with its use were collected from an ad hoc questionnaire through a web survey with the collaboration of the Italian Federation of Friends of the Bicycle, an association of non-competitive Italian cyclists spread over the national territory. The respondents are member of this cycling organisation in Italy and were selected, on a voluntary basis, to answer questions on their self-reported helmet use.
A multinomial logistic regression was implemented to catch which factors are relevant among five different categories of helmet users: from those who are enthusiastic-users, to those who use it often but not always, to those who never wears it.

Results

Overall, 2,502 Italian cyclists took part in the survey, only 2,072 subjects answered the question about helmet use. Most bicycle riders (63.1%) were between 45 and 64 years of age and 13.7% were over 65; while only 0.8% were younger than 25. Among bicycle riders, 62.3% were males. More than two-thirds (69.8%) live in the Northern Italy, one-fifth in Central Italy and 10.7% in Southern Italy. In the sample, 71.5% of riders use a city bike and 28.5% a sport bike. More than half (53.2%) circulate almost every day by bicycle and 37.3% cycle more than 60 km per week. More than three-quarters have got a helmet; however, not all bicyclists who have got one use it. The sample seems perfectly divided regarding helmet use: one-quarter (24.2%) of the riders reported never wearing a helmet, one-quarter (25.3%) reported always wearing a helmet. While, half of the riders (50.5%) wearing a helmet sometimes, rarely or most of the time. Nearly one cyclist in five had been involved in any accidents in the last year.
The results of the multinomial logistic regression are shown in ***Table1***. Using "always users" as a reference category, females are more likely to be a "never and rarely users" than males. The not use of the helmet decreases with age. Cyclists from North of Italy are less likely to be "always users" than those coming from Centre\South.
A lower use of the helmet is related to short weekly distance. Cyclists not involved in a bicycle accident are more likely to be "rarely or never users" versus "always users" than those involved in at least one accident. Different situation is for the daily use of bicycle, in fact those who cycling almost daily are more likely to be a "rarely users". Cyclists riding city bike are less likely to use helmet than those riding sport bikes.

Table1. Multinomial logistic regression for identifying factors associated with bicycle helmet use

	"Often users" RRR (95% CI)	p-value	"Sometime users" RRR (95% CI)	p-value	"Rarely users" RRR (95% CI)	p-value	"Never users" RRR (95% CI)	p-value
Gender (Female vs Male)	1.39 (0.99-1.95)	0.057	1.23 (0.84-1.81)	0.278	1.79 (1.18-2.72)	0.006	2.24 (1.59-3.16)	<0.001
Age	1 (0.99-1.02)	0.158	0.99 (0.98-1.01)	0.895	0.97 (0.95-0.99)	0.002	0.97 (0.95-0.98)	<0.001
Geographical area (North vs Centre-South)	1.93 (1.41-2.65)	<0.001	2.34 (1.59-3.45)	<0.001	2.16 (1.4-3.32)	<0.001	1.68 (1.2-2.35)	0.002
Weekly cycling distance (<60km vs >=61)	1.15 (0.83-1.58)	0.392	2.12 (1.46-3.08)	<0.001	2.15 (1.41-3.27)	<0.001	3.16 (2.23-4.48)	<0.001
Involvement in an accident (no vs yes)	1.3 (0.91-1.83)	0.137	1.15 (0.76-1.74)	0.483	1.7 (1.04-2.78)	0.034	2.55 (1.65-3.92)	<0.001
Cycling almost daily (yes vs no)	1.73 (1.26-2.37)	0.001	2.61 (1.8-3.78)	<0.001	3.39 (2.22-5.17)	<0.001	2.21 (1.58-3.1)	<0.001
Bike type (City bike vs Sport bike)	1.56 (1.13-2.14)	0.006	2.91 (1.92-4.4)	<0.001	3.52 (2.14-5.79)	<0.001	4.08 (2.76-6.05)	<0.001

° Reference category: "Always users" CI, confidence interval; RRR, relative risk ratio;

Conclusions

From our result it is interesting to underline that young people, female, short distance travelled weekly but almost daily and not to be involved in a road accident in the last year are related to a lower use of the helmet. From these findings, it is possible to understand which are the factors influencing more the use or not use of the helmet. Our evidence generates new information for effective enforcement, education and promotion measure to increase helmet use by bicyclist.

References

[1] European Commission, 2014. Special Eurobarometer 422a "Quality of Transport" 137.
[2] Orsi C, Ferraro OE, Montomoli C, et al. Alcohol consumption, helmet use and head trauma in cycling collisions in Germany. Accid Anal Prev 2014;65: 97–104.
[3] Harada MY, Gangi A, Ko A, et al. Bicycle trauma and alcohol intoxication. Int J Surg 2015;24:14–9.
[4] Uibel S, Müller D, Klingelhoefer D, et al. Bicycle helmet use and non-use-recently published research. J Occup Med Toxicol 2012;7(1):7-9.

PIANIFICAZIONE DI UNO STUDIO DI INCIDENZA UTILIZZANDO MODELLI CATTURA-RICATTURA CON ETEROGENEITÀ: LA MORTE CARDIACA IMPROVVISA NELLO SPORT (MCIS)

Di Rocco Arianna[1], Farcomeni Alessio[1], Alunni Fegatelli Danilo[1], Castelli Vincenzo[2], Vestri Annarita[1]

1. *Dipartimento di Sanità Pubblica e Malattie Infettive, Sapienza – Università di Roma*
2. *Ospedale Madre Giuseppina Vannini - Roma*

Introduzione

La Società Europea di Cardiologia definisce la Morte Cardiaca Improvvisa (MCI) come una "morte naturale dovuta a cause cardiache, preceduta dalla rapida perdita di coscienza, entro un'ora dall'insorgere della sintomatologia acuta, in soggetti con o senza cardiopatia nota preesistente, ma per i quali tempi e modalità del decesso sono inaspettati"[1].

La MCI durante l'attività sportiva (MCIS) è un evento raro che colpisce spesso atleti giovani che incarnano il concetto comune di buona salute e che secondo le norme di legge dovrebbero essere sottoposti a periodici screening medico sportivi. La morte in campo del pallavolista Vigor Bovolenta (Macerata, 24 marzo 2012) e del calciatore Piermario Morosini (Pescara, 14 aprile 2012), a distanza di poco tempo l'uno dall'altro, ha portato alla pubblicazione sulla GU del 20-7-2013 del "decreto Balduzzi" da parte del Ministero della Salute. Il decreto prevede oltre alla certificazione medica per l'attività sportiva, agonistica e non, anche l'obbligo per tutte le Associazioni Sportive Dilettantistiche (A.S.D) di dotarsi di un Defibrillatore Automatico Esterno (DAE) e di formare adeguatamente una o più persone sul corretto utilizzo di tali dispositivi.

Nonostante siano passati quattro anni dalla sua pubblicazione, non ci sono ancora i decreti attuativi e l'entrata in vigore è prevista al 30 luglio 2017.

La reale incidenza, le caratteristiche demografiche, le cause e le circostanze in cui avvengono gli eventi di MCIS sono attualmente poco conosciuti; sulla base di studi retrospettivi è risultato che negli Stati Uniti il tasso di incidenza è di 0,6 MCIS ogni 100.000 [***2***], in Francia di 1/100.000 [***3***] e in Israele di 2,6/100.000 [***4***]. Le informazioni più attendibili relative ad incidenza e cause della MCIS negli atleti derivano dall'Italia, precisamente dalla regione Veneto (2,1/100.000) [***5***] che nel 1987 ha istituito un registro regionale delle morti per evento cardio vascolare under/over 35 anni.

L'estrema variabilità del dato è legata alla modalità di acquisizione delle informazioni riguardanti l'evento; queste, in mancanza di registri regionali e nazionali, sono ottenute attraverso le notizie di MCIS riportate da quotidiani e siti internet.

L'inadeguatezza metodologica, la non uniformità delle modalità di raccolta dati e le analisi statistiche *naive* eseguite finora hanno portato sistematicamente a sottostimare l'incidenza dell'evento MCIS.

I modelli di cattura-ricattura possono aiutare a correggere le stime in presenza di casi non registrati. Sono basati su liste separate di eventi, che verranno ottenute dal web, permettendo di stimare in maniera non-distorta l'incidenza dell'evento anche se qualcuno di questi non sia stato incluso in alcuna delle liste. Il razionale è molto semplice ed è basato sulla sovrapposizione tra le liste stesse. Le stime risultanti sono non-distorte (ovvero, non sovrastimano o sottostimano sistematicamente l'incidenza).

Lo studio si pone lo scopo di mettere in evidenza la reale incidenza della MCIS nel nostro territorio, sottolineando l'importanza dell'immediata attuazione del decreto Balduzzi. Questo, oltre a mantenere in vigore le attuali regole riguardo lo screening pre- partecipativo all'attività fisica (che ha già ridotto di molto l'incidenza degli eventi cardiovascolari negli agonisti), obbliga le società a disporre di un DAE che, riducendo l'incidenza dei decessi se utilizzato nelle giuste tempistiche.

Metodi

La logica sulla quale si basano i modelli di cattura-ricattura è la seguente: supponiamo che ci siano due fonti giornalistiche che riportano informazioni su eventi MCIS. Un certo numero di eventi, chiamato *n10*, apparirà solo nella prima lista. Un certo numero, chiamato *n01*, apparirà solo nella seconda lista. Infine, *n11* apparirà in entrambi gli elenchi. Chiamate anche il numero totale di pazienti che appaiono almeno in una lista. Il numero totale di eventi è ovviamente dato da $N = n + n00$, dove *n00* è il numero di eventi MCIS che non appaiono in nessuna delle liste. Il metodo di cattura- ricattura consente di stimare *n00*, e di conseguenza la dimensione totale della popolazione. Un semplice stimatore di *N* per due elenchi, in assoluta omogeneità e assunzioni di indipendenza, è dato da $(n10 + n11)(n01 + n11) / n11$. Questo è il noto estimatore Lincoln-Peterson e segue da un semplice assunto di indipendenza nella tabella 2x2. In casi più generali e con più di due elenchi non esiste una forma chiusa per lo stimatore di *N*. Questo deve essere ottenuto massimizzando numericamente la verosimiglianza. D'altra parte, quando vengono raccolti più elenchi, la precisione aumenta. Inoltre, le ipotesi di omogeneità e indipendenza che stanno alla base della stima Lincon-Petersen sono raramente credibili nella pratica. Per quanto riguarda la stima del numero di casi (o di nuovi casi) di MCIS, è evidente che dovranno essere presi in considerazione varie fonti di eterogeneità e di interazione tra cui, in primo luogo, eventi che riguardano atleti ad alto livello più probabilmente compariranno in elenchi diversi, vista la risonanza che in questi casi ha la notizia; in secondo luogo, eventi avvenuti in piccoli centri avranno maggior probabilità di essere riportati sulla stampa locale rispetto a quelli delle grandi città in cui, per questo motivo, ci aspettiamo un'incidenza minore. Anche l'eterogeneità non osservata può essere presa in considerazione, almeno in parte. Tutti questi problemi, se ignorati, sono una fonte di pregiudizio o di perdita di informazioni che portano ad aumentare le lunghezze degli intervalli di confidenza. Le stime verranno inoltre stratificate per sesso e classi di età.

Risultati

I risultati al momento riguardano la pianificazione dello studio. Per la pianificazione del numero di liste necessarie a ottenere una stima precisa del numero di morti improvvise per arresto cardiaco nel periodo 2006-2017 utilizziamo un modello cattura-ricattura che tiene conto della possibile diversa probabilità di inclusione di ciascun soggetto in ciascuna lista (sorgenti d'informazione). Con tale modello è possibile specificare una diversa probabilità di cattura per ciascuna lista (associata o ad un giornale o ad un sito internet). In particolare, nel nostro caso, saranno considerate due diverse probabilità di cattura, una associata alle liste relative ai giornali (*pg*) e l'altra associata alle liste relative ai siti internet (*psi*). Con tale assunzione la probabilità che l'evento sia riportato su ciascun giornale sarà la stessa per tutti i giornali e pari a *pg*; in maniera analoga la probabilità che l'evento sia riportato su ciascun sito internet sarà uguale per tutti i siti e pari a *psi*.

Per quanto riguarda la pianificazione dello studio si utilizza la procedura di Alunni Fegatelli e Farcomeni [***6***]. Considerando un numero totale di 40 liste (20 giornali e 20 siti internet), un valore del parametro di interesse pari a N=1500 ed assumendo pg= 0.1 e psi = 0.01 si attende un'ampiezza relativa dell'intervallo di circa 0.45 (da cui un intervallo di confidenza di ampiezza poco inferiore a 70). Si noti che il modello finale utilizzato per la stima sarà comunque più complesso, in quanto saranno considerati anche altri fattori di eterogeneità per le probabilità di inclusione (ad esempio, se l'evento è accaduto in una grande città o meno, l'età del soggetto, etc.).

Bibliografia

[1] Felicani C, Moccia E, Naccarella, et al. La morte improvvisa da sport. Due database prospettici 1990-2004 e 2005-2016: aspetti epidemiologici, preventivi e assistenziali. Giornale Italiano di Aritmologia e Cardiostimolazione 2008;11(2): 69-82.

[2] Maron BJ, Haas TS, Doerer JJ, et al. Comparison of U.S. and Italian experiences with sudden cardiac deaths in young competitive athletes and implications for preparticipation screening strategies. Am J Cardiol 2009;104(2):276-80.

[3] Marijon E, Tafflet M, Celermajer DS, et al, Sports-related sudden death in the general population. Circulation 2011;124(6):672-81.

[4] Steinvil A, Chundadze T, Zeltser D, et al. Mandatory electrocardiographic screening of athletes to reduce their risk for sudden death proven fact or wishful thinking? J Am Coll Cardiol 2011;57(11):1291-6.

[5] Corrado D. Migliore F, Zorzi A, et al. Preparticipation electrocardiographic screening for the prevention of sudden death in sports medicine. G Ital Cardiol 2011;12(11):697-706.

[6] Alunni Fegatelli D, Farcomeni A. On the design of closed recapture experiments, Biometrical Journal 2006;58:1273-94.

IL SUPPORTO SOCIALE NEGLI ADOLESCENTI MIGRATI IN ITALIA

Dalmasso Paola[1], Borraccino Alberto[1], Charrier Lorena[1], Berchialla Paola[2], Lemma Patrizia[1], Cavallo Franco[1]

1. *Dipartimento di Scienze della Sanità Pubblica e Pediatriche, Università di Torino*
2. *Dipartimento di Scienze Cliniche e Biologiche, Università di Torino*

Introduzione

Recentemente l'Italia è stata oggetto di un importante flusso migratorio e, come in altri Paesi europei, si registra una crescente percentuale di adolescenti immigrati, che rappresentano il 9,2% della popolazione scolastica nazionale [***1***]. Come i loro pari italiani, i giovani immigrati devono affrontare una serie di problemi, quali i rapporti sociali con insegnanti, familiari e coetanei, cui si aggiunge la sfida del processo di acculturazione.
Il supporto sociale è stato riconosciuto come uno dei più importanti fattori per il successo del processo di acculturazione [***2, 3***] ed è considerato un moderatore degli effetti negativi dello stress ambientale e sociale sul benessere individuale sia tra le popolazioni ospitanti che su quelle immigrate, in particolare nell'adolescenza [***4***].
In questo studio si è valutato il livello di sostegno sociale percepito tra gli adolescenti immigrati in Italia di prima e seconda generazione, confrontandolo con quello dei loro pari della popolazione ospitante.

Metodi

Utilizzando i dati dell'indagine HBSC italiana raccolti nel 2014, gli intervistati di età compresa tra 11, 13 e 15 anni sono stati classificati come "popolazione ospitante" se entrambi i genitori sono nati in Italia, come "immigrati di prima generazione" se sono nati all'estero e almeno uno dei loro genitori è nato all'estero, e come "immigrati di seconda generazione" se sono nati in Italia e almeno uno dei loro genitori è nato all'estero [***5***]. Inoltre, in base al Paese di nascita della madre, i giovani immigrati sono stati classificati come provenienti da Paesi dell'Europa occidentale, dell'Europa dell'Est, o da Paesi extra europei. Il supporto sociale è stato analizzato attraverso le relazioni con la famiglia, gli insegnanti, i pari e i compagni di classe, valutati con scale multidimensionali, standardizzate e validate, quale la Scala Multidimensionale del sostegno sociale percepito [**6, 7**].
Sono stati utilizzati modelli di regressione logistica multivariata per valutare l'associazione tra ciascun tipo di supporto esaminato ed il background etnico, aggiustando per caratteristiche sociodemografiche quali l'età, il genere e lo status socioeconomico.

Risultati

L'analisi è stata condotta su 47399 intervistati, di cui il 13,4% immigrati di prima o seconda generazione (4,0% dai paesi occidentali, 4,1% dai paesi dell'Est europeo e 5,3% da paesi extra-europei).
Nel complesso, gli immigrati adolescenti provenienti dai Paesi dell'Europa dell'Est e dai Paesi Extraeuropei hanno riportato un sostegno sociale sostanzialmente inferiore rispetto alle loro controparti della popolazione ospitante in tutti gli ambiti esplorati. Per gli adolescenti provenienti dai paesi occidentali, invece, i risultati non sono significativi, ad eccezione del sostegno degli insegnanti.

Discussione

Nel nostro studio sono emersi due diversi modelli di immigrazione: il modello occidentale, dai paesi più ricchi, ed il modello dell'est. Tra questi ultimi, gli immigrati di seconda generazione hanno mostrato il livello più basso di sostegno in tutti gli ambiti esaminati.

Bibliografia

[1] MIUR. Gli alunni stranieri nel sistema scolastico italiano: A.S. 2014/2015. Roma: MIUR - Ufficio di Statistica; 2015: Reperibile su: http://www.istruzione.it/allegati/2015/Notiziario_Alunni_Stranieri_1415.pdf. [Accessed 24/03/2017].

[2] Cauce AM. Social networks and social competence: Exploring the effects of early adolescent friendships. Am J Community Psychol 1986;14(6):607-28.

[3] Cauce AM, Felner RD, Primavera J. Social support in high-risk adolescents: Structural components and adaptive impact. Am. J. Community Psychol 1982;10(4):417-28.

[4] Oppedal B, Røysamb E, Sam DL. The effect of acculturation and social support on change in mental health among young immigrants. International Journal of Behavioral Development 2004;28(6):481-94.

[5] Cavallo F, Lemma P, Dalmasso P, Vieno A, Lazzeri G, Galeone D. Report Nazionale dati HBSC Italia 2014: 4° Rapporto sui dati HBSC Italia 2014. Torino: Stampatre s.r.l.; 2016

[6] Zimet GD, Dahlem NW, Zimet SG, Farley GK. The multidimensional scale of perceived social support. J Pers Assess 1988;52(1):30-41.

[7] Zimet GD, Powell SS, Farley GK, Werkman S, Berkoff KA. Psychometric characteristics of the multidimensional scale of perceived social support. J Pers Assess 1990;55(3-4):610-7.

FATTORI MATERNI, OSTETRICI E PERINATALI ASSOCIATI AL RISCHIO DI SVILUPPARE VIZI DI REFRAZIONE E/O AMBLIOPIA NEI BAMBINI DI 4 ANNI IN TRENTINO

Pertile Riccardo[1], Trettel Cristina[2], Bombarda Lucia[1], Piffer Silvano[1]

1. *Servizio di Epidemiologia Clinica e Valutativa, Azienda Provinciale per i Servizi Sanitari di Trento*
2. *Servizio Coordinamento Screening, Azienda Provinciale per i Servizi Sanitari di Trento*

Introduzione

In provincia di Trento è attivo uno screening ortottico prescolare (secondo anno di scuola materna) già dalla seconda metà degli anni '80 e attualmente copre tutta la provincia.
Il presente studio valuta il livello di associazione tra prevalenza di vizi di refrazione/ambliopia con i fattori materni/ostetrici e perinatali corrispondenti dei bambini delle coorti di nascita 2008-2010 e sottoposti a screening ortottico nel periodo 2012-2014.

Metodi

Si è proceduto ad un record linkage tra database dello screening ortottico relativo alle coorti di nascita 2008-2010 e archivio provinciale della Natalità, alimentato dal flusso CedAP (Certificato di Assistenza al Parto). Tale flusso raccoglie tutti i dati inerenti la gravidanza, il parto ed il neonato, relativamente alle gestanti assistite presso i punti nascita della provincia di Trento. Per il solo campione di bambini linkati al corrispettivo certificato di assistenza al parto, è stato eseguito un confronto tra il gruppo di casi con almeno un vizio di refrazione (e/o con ambliopia) ed il gruppo di controllo (bambini sani) in termini di peso neonatale, età gestazionale, l'essere stato o meno ricoverato o rianimato alla nascita e presentare condizioni malformative note. Le analisi multivariate sono consistite in due modelli di regressione logistica per valutare la probabilità di insorgenza rispettivamente di almeno un vizio di refrazione e di ambliopia.

Risultati

Su un totale di 15.782 bambini iscritti al 2° anno delle scuole materne, ne è stato valutato il 91% negli anni 2012-2014. Il rimanente 9% non è stato visitato o perché già in cura per patologie oculari (n=764), o perché non frequentante o ritirato dalla scuola materna (n=94), o perché i genitori non hanno dato il consenso alla visita ortottica (n=121), oppure perché il bambino era assente nel giorno della prima visita e non si è presentato alla seconda visita prenotata in ambulatorio (n=457). Dei 14.346 bambini visitati nel periodo 2012-2014, 911 erano affetti da almeno un vizio di refrazione con una prevalenza pari al 6,4%. Stratificando per ciascun vizio di refrazione, si riscontra una prevalenza pari al 2,7% di ipermetropia, 0,5% di miopia e 5,1% di astigmatismo. L'ambliopia riguarda l'1,8% della casistica di bambini visitati. La tabella 1 riporta i medesimi dati distinti per ciascuna coorte di nascita.
Il linkage con il corrispondente archivio CedAP è riuscito per 12.012 casi, pari all'83,7%. L'analisi di regressione logistica multivariata che valuti la probabilità di sviluppare almeno un vizio di refrazione in relazione ai fattori neonatali, materni ed ostetrici considerati evidenzia che (***Tab. 1***): i bambini nati da madri con età ≥45 anni presentano un rischio di sviluppare un vizio di refrazione 3,5 volte maggiore rispetto alle donne under 45 anni (OR aggiustato=3,52, I.C. al 95% 1,18-10,55). A parità delle altre variabili inserite nel modello d'analisi, le madri con cittadinanza straniera hanno un rischio del 27% più alto rispetto alle madri italiane che al proprio figlio venga diagnosticato un vizio di refrazione. Anche il parto con ventosa e il presentare un peso alla nascita molto basso (<1500 g) rappresentano significativi fattori di rischio nello sviluppare almeno un vizio di refrazione (OR aggiustati rispettivamente pari a 1,43 e 3,81). L'analisi

multivariata condotta per un confronto tra gruppo di ambliopi e gruppo di controllo non ha portato ad alcuna differenza statisticamente significativa.

Tabella 1. Analisi multivariata per verificare quali o quale fattore pesi di più sulla probabilità di avere un vizio di refrazione a 4 anni

Variabile analizzata	**OR**	**95% I.C.**	***p*-value**
Età materna (30-44 anni vs. <30 anni)	1,00	0,85 – 1,19	n.s.
Età materna (≥45 anni vs. <30 anni)	3,52	1,18 – 10,55	**0,02**
Fumo in gravidanza (Sì vs. No)	1,13	0,84 – 1,51	n.s.
Cittadinanza materna (Straniera vs. Italiana)	1,27	1,05 – 1,54	**0,01**
Modalità del parto (taglio cesareo vs. spontaneo)	0,91	0,76 – 1,10	n.s.
Modalità del parto (ventosa vs. spontaneo)	1,43	1,03 – 2,00	**0,03**
Peso alla nascita (<1500 g vs. ≥2500 g)	3,81	1,14 – 12,68	**<0,01**
Peso alla nascita (1500-2499 g vs. ≥2500 g)	0,79	0,51 – 1,23	n.s.
Settimane gestazionali (<37 vs. ≥37)	0,86	0,57 – 1,31	n.s.
Punteggio Apgar a 5 min. (<7 vs. ≥7)	0,45	0,10 – 1,98	n.s.
Ricovero alla nascita (Sì vs. No)	1,12	0,82 – 1,52	n.s.
Rianimazione alla nascita (Sì vs. No)	1,09	0,68 – 1,75	n.s.
Presenza di anomalie congenite note alla nascita (Sì vs. No)	0,95	0,49 – 1,86	n.s.

n.s.=non significativo

Conclusioni

Ogni programma pubblico di prevenzione per cui è disponibile un'evidenza di efficacia dovrebbe essere implementato nella pratica operativa, verificandone l'efficacia nella pratica (*effectiveness*) rispetto all'efficacia teorica (*efficacy*). La valutazione organica dello screening ortottico prescolare in atto in provincia di Trento consente di delineare una serie di aspetti positivi tra cui un'attività consolidata nel tempo che consente il suo sviluppo sull'intero territorio di riferimento, la disponibilità di procedure operative standardizzate e condivise, un'elevata adesione allo screening. Il record linkage tra il flusso informativo dello screening ortottico e quello del CedAP ha permesso l'individuazione di alcuni potenziali fattori di rischio materni, ostetrici e neonatali. I sottogruppi di bambini con maggior rischio di sviluppare un vizio di refrazione dovrebbero essere identificati e monitorati per consentire una diagnosi precedente al periodo di screening. Le informazioni su caratteristiche perinatali e ostetriche potrebbero orientare il clinico ed il pediatra di base in fase di pianificazione delle visite ortottiche ed oculistiche, sia precoci che da screening.

ACCESS TO HEALTH CARE FOR WOMEN MIGRANTS IN LOMBARDY REGION, ITALY

Popa Ioana[1], Ferraro Ottavia Eleonora[1], Collivasone Luigi[2] , Sisto Patrizio[3], Montomoli Cristina[1]

1. *Biostatistics and Clinical* Epidemiology *Unit, Department of Public Health, Experimental and Forensic Medicine, University of Pavia, Italy*
2. *Il Sole Family Counseling Service in Vigevano and Gambolò, Pavia, Italy*
3. *AMPRA Associazione di Medicina e Psicologia per la Ricerca-Azione, Pavia , Italy*

Introduction

The Italian National Health Service (SSN-Servizio Sanitario Nazionale) is a public system aiming to grant universal access to a uniform level of health care throughout the country. Local health authorities (ATS) are responsible for the delivery of health care services at local level. Access to health care refers to the ease with which an individual can obtain needed medical services. The social, cultural, economic, and geographic factors influence health care access worldwide. Women migrants present in Italy, may be with or without documents. Undocumented women migrants do not have the right to register in the Italian National Health Service [***1***]. Lombardy is the region with the largest number of migrants from countries with a strong migratory pressure, 1,149,000 estimated at 2016. [***2***]. One of the point of entry into Italian Health System, for migrant women, is Family Counseling Services [***3***].
Objectives are (i) to present the foreign population distribution according to age and sex, in Lombardy Region, Italy, (ii) to describe the access to health care in Lombardy Region, Italy.

Methods

Demographic analysis: the population growth and pyramid of ages were used to present the foreign population distribution a ccording to age and sex, in Lombardy Region, Italy. Frequency distribution of main causes of hospitalisation of women migrants in Lombardy Region. Study population: women migrants in fertility age [15 years, 49 years] in Lombardy. To study the main causes of hospitalization, we used the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) [***4***]. Source of data: Istat and National Database SDO elaborated by the Italian Ministry of Health [***5, 6***].

Results

Foreign Population in Lombardy Region (***Figure 1***). The total foreigner residents in Italy, in 2016, were 5.026153 and represent 8.3% of the resident population. 52.6% are females, 25% of foreigners live in the Lombardy Region. The foreign population is represented mainly by people from Eastern Europe (36%) with a huge prevalence of Romanians and Albanians (23.2%).The recent migration flows is named “feminization of migration”.

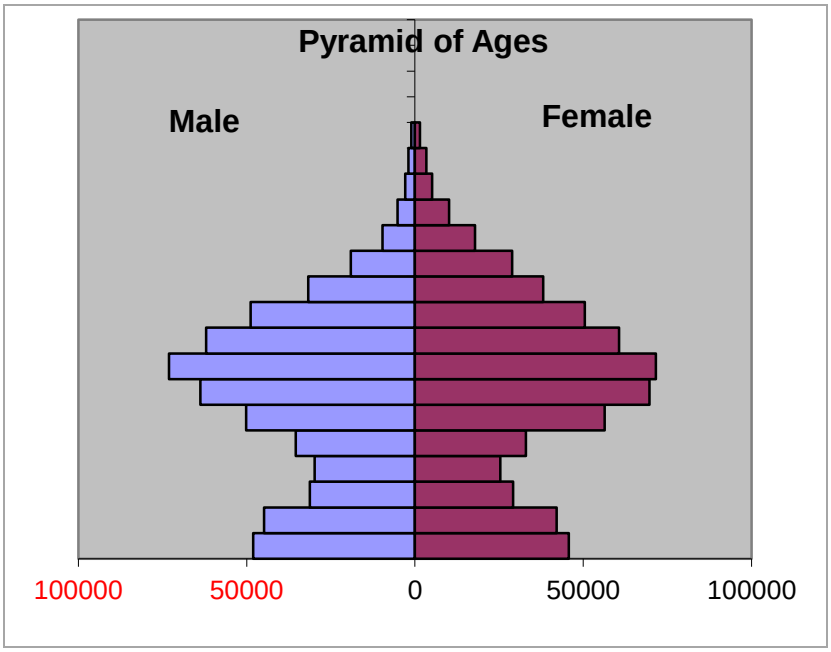

Figure 1. Foreign Population in Lombardy Region

Access to health care in Lombardy region All foreign citizens are entitled to healthcare in Italy, a service provided by the Italian National Health Service (SSN). It is possible to obtain healthcare and medical services through: (i) compulsory registration, (ii) voluntary registration or (iii) entering country for medical treatment. Having registered with the SSN, foreign citizens receive the TEAM card, i.e. European Health Insurance Card, entitling them to receive medical care in the same way as Italian citizens.
Foreign citizens not registered with NHS and not having sufficient financial means are in any case guaranteed casualty/outpatient services, emergency hospital care or continuous care for diseases or injuries and preventive medicine programs (e.g. vaccinations) in accredited public or private facilities.
For the women migrant *health services guaranteed are: i)* protection of pregnancy and maternity*, ii)* protection of children's health*, iii)* vaccinations*, according to the law authorized by the Region, iv)* international prophylaxis and v) prophylaxis, diagnosis and treatment of infectious diseases. In the National Health Program 1998-2000, the Italian Ministry of Health introduced a project to improve maternal and child health.
The analysis of gender hospitalizations of migrants living in Lombardy Region, in 2014, shows 83,486 female hospital admissions. The number of female hospital admissions in Lombardy is due mainly to Complications of Pregnancy, Childbirth, and the Puerperium (ICD-9 Chapter 11). Hospital admission for these reasons account for 29% of all admissions of foreign citizens and 44.5% of all admission of foreign females. Hospital admission of migrant women, in Lombardy Region, for childbirth represent 20,3% of all births occurring in Lombardy and 40.3% of all abortions.
In Lombardy, the major causes of hospitalization of women migrant in fertility age [15 years, 49 years] , are: Complications of Pregnancy, Childbirth, and the Puerperium (52.3%), Diseases of the genitourinary system (9.5%), Diseases of the digestive system (7.6%), Neoplasm (4.6%), Diseases of the respiratory system (4.1%).
Access of women migrant to health care can be *limited by*: language, health culture and education, age, functional handicap, religious perceptions. *Access facilitation*: cultural and linguistic mediators, local and

community religious confessions, local authorities, NGOs. For the women migrants, at local level, Family Counseling Service is a primary care unit and a point of entry into Italian Health System.

Conclusions

1. The Italian health care system offers universal health coverage to all citizens and access to a wide range of services. Citizens have the free choice of GP, who acts as gatekeeper.
2. For the women migrant health services guaranteed are: *i)* protection of pregnancy and maternity, *ii)* protection of children's health, *iii)* vaccinations, according to the law authorized by the Region *iv)* international prophylaxis and v) prophylaxis, diagnosis and treatment of infectious diseases.
3. In Lombardy, the major cause of hospitalization of women migrant in fertility age is Complications of Pregnancy, Childbirth, and the Puerperium (52.3%).

References

[**1**] Access to Health Care for Undocumented Migrants in Europe. Available from: http://picum.org/picum.org/uploads/file_/Access_to_Health_Care_for_Undocumented_Migrants.pdf.
[**2**] Italian Demographic Statistics. Available from: http://www.tuttitalia.it
[**3**] Ministero della Sanità. Decreto ministeriale del 24/4/2000. Progetto Obiettivo Materno Infantile. Gazzetta Ufficiale n. 131 Suppl. Ord. n. 89 del 7/6/2000.
[**4**] International Classification of Diseases – 9th revision – Clinical Modification (ICD-9-CM) Available from: https://www.cdc.gov/nchs/icd/icd9cm.htm
[**5**] Istat, Populazione Straniera residente per citadinanza e anno, 2016. Available from: https://www.istat.it/it/archivio/stranieri
[**6**] Elaborazione Banca dati nazionale SDO, Ministero della Salute, Direzione Generale della Programmazione Sanitaria.

EVALUATION OF HEART FAILURE PATTERNS OF CARE VIA MULTI STATE MODELS: A POPULATION BASED APPROACH BASED ON ADMINISTRATIVE DATA

Gasperoni Francesca[1], Ieva Francesca[1], Barbati Giulia[2,3], Scagnetto Arjuna[2], Iorio Annamaria[3,4], Sinagra Gianfranco[5], Di Lenarda Andrea[3]

1. *MOX-Modelling and Scientific Computing, Department of Mathematics, Politecnico di Milano, Milano, Italy*
2. *Department of Medical Sciences, Università di Trieste, Trieste, Italy*
3. *Cardiovascular Center, Trieste, Italy*
4. *Cardiology Unit, Papa Giovanni XXIII Hospital, Bergamo, Italy*
5. *Cardiovascular Department, Azienda Sanitaria-Universitaria Integrata Trieste 'ASUITS', Trieste, Italy*

Introduction

How different risk profiles of heart failure (HF) patients can influence multiple readmissions and outpatient management is largely unknown. Methods that are able to jointly model both endpoints of interest and the process dynamics in a flexible way might ease a deeper comprehension of the clinical evolution of patients' patterns of care, expecially in chronic diseases like HF [***1, 2, 3***].
We propose the application of two multi-state models in real world setting to jointly evaluate the impact of different risk factors on multiple hospital admissions, Integrated Home Care (IHC) activations, Intermediate Care Unit (ICU) admissions and death [***4, 5***].

Methods

The first model (model 1, ***fig.1***) concerns only hospitalizations as possible events and aims at detecting the determinants of repeated hospitalizations. The second model (model 2) considers both hospitalizations and ICU/IHC events and aims at evaluating which profiles are associated with transitions in intermediate care with respect to repeated hospitalizations or death. Both are characterized by transition specific covariates, adjusting for risk factors.

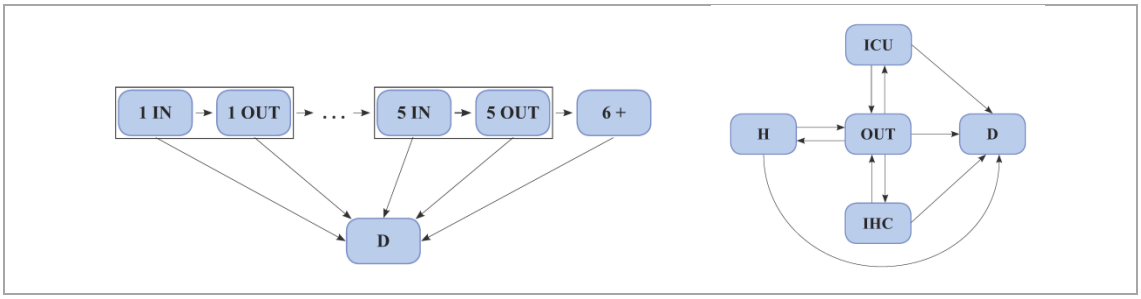

Figure 1. Multi state models 1 and 2

Results

We identified 4,904 patients (4,129 de novo and 775 worsening heart failure, WHF) hospitalized for HF from 2009 to 2014. 2,714 (55%) patients died. Advanced age and higher morbidity load increased the rate of dying and of being rehospitalized (model 1), decreased the rate of being discharged from hospital (models 1 and 2) and increased the rate of inactivation of IHC (model 2). WHF was an important risk factor associated with hospital readmission.

Conclusions: Multi-state models enable a better identification of two patterns of HF patients. Once adjusted for age and comorbidity load, the WHF condition identifies patients who are more likely to be readmitted to hospital, but does not represent an increasing risk factor for activating ICU/IHC. This highlights different ways to manage specific patients' patterns of care. These results provide useful healthcare support to patients' management in real world context. Our study suggests that the epidemiology of the considered clinical characteristics is more nuanced than traditionally presented through a single event.

References

[**1**] Corrao G, Ghirardi A, Ibrahim B et al., Short-and long-term mortality and hospital readmissions among patients with new hospitalization for heart failure: A population-based investigation from Italy. Int J Cardiol 2015,181:81-7.

[**2**] Tavazzi L, Ventura C. 'Observational medicine': registries and Electronic Health Recording for science and health systems governance, Eur J Heart Fail 2016,18 (9):1093-5.

[**3**] Bakal JA, McAlister FA, Liu W, et al. Heart failure re-admission: measuring the ever shortening gap between repeat heart failure hospitalizations, PLoS One 2014;9(9):e106494.

[**4**] Postmus D, van Veldhuisen DJ, Jaarsma T, et al., The COACH risk engine: a multistate model for predicting survival and hospitalization in patients with heart failure, Eur J Heart Fail 2012;14(2):168-75.

[**5**] Ieva F, Jackson CH, Sharples LD. Multi-State modelling of repeated hospitalisation and death in patients with Heart Failure: the use of large administrative databases in clinical epidemiology. Stat Methods Med Res 2017;26(3):1350-72.

PERSISTENCE AND DISCONTINUATION OF APIXABAN AMONG PATIENTS WITH ATRIAL FIBRILLATION: A RETROSPECTIVE ANALYSIS IN FIVE ITALIAN DISTRICTS

Verdecchia Paolo[1], Lanati Elena Paola[2], Verdini Valeria[2], Corrao Giovanni[3], Scotti Lorenza[3], Airoldi Chiara[3], Iorio Arianna[2].

1. *Department of Medicine, Hospital of Assisi, Italy*
2. *MA Provider, Milan, Italy*
3. *Department of Statistics and Quantitative Methods, University of Milan-Bicocca*

Introduction

Novel oral anticoagulants (NOACs), including apixaban, have demonstrated at least similar efficacy and safety in stroke prevention compared with warfin. However the lack of monitoring, may affect treatment persistence which is important in ensuring optimal stroke prevention.
Despite this, little is known about the rate and reasons for apixaban discontinuations in daily care.
This study aimed to collect data from Italian hospitals on patients with non-valvular atrial fibrillation (NVAF) in treatment with apixaban, in order to examine the characteristics and persistence in the real world setting.

Methods

In this multi-center, retrospective, observational study we collected data from clinical database on consecutive patients with a diagnosis of NVAF who had initiated apixaban from 1 January 2014 up to March 31, 2016.
Peristence, defined as proportion of patients still on initial drug therapy after a fixed time, was estimated considering both permanent discontinuations, when patients had a treatment gap of more than 30 day, and temporary if patients had a treatment gap of less than 30 days.

Results

The analyzed sample consisted of 766 patients affected by NVAF from five Italian Centers. The mean age was 74.2 years, 53.5% of patients were women and the median $CHADS_2$ and CHA_2DS_2VASc scores were respectively 2.0 and 4.0. The most frequent co-morbidities were cardiovascular diseases (hypertension – affecting 84.1% of patients; previous vascular disease [34.1%]; heart failure [22.1%]), renal impairment (34.4%), diabetes mellitus (22.5%) and anemia (12.5%). At baseline (V0), 15.7% of patients had recorded at least one event among stroke, transient ischemic attack or systemic embolism. In the whole cohort, the half of patients (50.7%) were naïve to oral anticoagulants, while 219 patients had been previously treated with warfarin, heparin (66 patients), acetylsalicylic acid (52 patients), clopidogrel (1 patient) and a novel oral anticoagulant (NOAC, 40 patients).
The follow up duration was up to 3 years with different timespan: 750 patients completed the first follow up visit (V1), at median time from the baseline 339.5 days, 253 completed the second follow up visit (V2) at median time from V1 of 110.9 days and only for 84 patients the third visit was recorded (V3) at median time from V2 of 241.5 days.During the follow up period, 123 cases of therapy discontinuation were recorded, including also patients with more than one discontinuation. 58 were permanent discontinuations, of whom 41 occurred at V1, 14 at V2 and 3 at V3. The persistence of apixaban treatment was 83.5% (75.5-89.2) considering the absence of permanent discontinuation of treatment and 37.2% (10.0-65.4) considering the absence of any type of discontinuation.Regarding the reasons of 58 permanent discontinuations: 10 occurred for worsening of renal function, 5 for minor bleeding, 5 for itching or rush, 3 for cancer, 4 for surgery, 4 for major bleeding, 3 for stroke, 3 for patient no-compliance, 3 for nausea, 2 for clinician decision, 1 for

dyspepsia, 1 for valve disease, 1 for worsening of hepatic function, 1 for anemia, 1 for heart failure, 1 for TIA, 1 for hematoma and 8 for other causes

Conclusions

Apixaban demonstrated better persistence than warfarin data available in literature. Further studies are needed to identify the consequences of discontinuation and evaluate the impact on outcomes.

PATTERN OF NEW ORAL ANTICOAGULANTS USE IN NON-VALVULAR ATRIAL FIBRILLATION: A POPULATION-BASED STUDY FROM SOUTHERN ITALY

Ingrasciotta Ylenia[1], Giorgianni Francesco[2], Aliquò Giancristoforo[2], Ientile Valentina[2], Pizzimenti Valeria[2], Scondotto Giulia[2], Cananzi Pasquale[3], Pollina Addario Walter[4], Pastorello Maurizio[5], Tari Michele[6], Alibrandi Angela[7], Trifirò Gianluca[1]

1. *Department of Biomedical and Dental Sciences and Morphofunctional Imaging, University of Messina, Italy;*
2. *Unit of Clinical Pharmacology A.O.U. Policlinico "G. Martino", Messina, Italy;*
3. *Sicilian Regional Centre of Pharmacovigilance, Servizio 7-Farmaceutica, Health Department of Sicily, Palermo, Italy;*
4. *Department of Epidemiologic Observatory, Health Department of Sicily, Palermo, Italy;*
5. *Department of Pharmacy, Palermo Local Health Unit, Palermo, Italy;*
6. *Local Health Unit of Caserta, Caserta, Italy;*
7. *Department of Economics, University of Messina, Messina, Italy*

Introduction

New oral anticoagulants (NOAs) are a valid alternative to Vitamin K antagonists for the treatment of non-valvular atrial fibrillation (NVAF), because they have fewer disadvantages than traditional oral anticoagulants, such as a narrow therapeutic window, pharmacological interactions, risk of severe bleeding, and the need to routinely monitor the International Normalized Ratio (INR) values leading to poor compliance with warfarin therapy, especially in older patients [***1, 2***]. Three clinical trials of dabigatran (RE-LY study) [***3***], rivaroxaban (ROCKET study) [***4***] and apixaban (ARISTOTLE study) [***5***] versus warfarin, demonstrated the non inferiority of these drugs in terms of efficacy and a lower risk of developing intracranial bleeding. However, a higher risk of major gastrointestinal bleeding was reported [***6***]. Adherence and persistence to NOA therapy in clinical practice are essential for preventing NVAF-related thromboembolic complications and for reducing the risk of stroke. In clinical practice, few and heterogeneous data on adherence and persistence to NOA and on the impact of the clinical effects exist [***7-9***].
The study aims to explore NOA treatment persistence, adherence and switching pattern in NVAF patients using two Italian population-based healthcare databases.

Methods

A retrospective cohort study was conducted during 2012–2015 using Caserta and Palermo Local Health Unit (LHU) claims databases covering around 2.4 million inhabitants. Incident NOA users (no dispensing within one year prior to treatment start) were characterized at baseline. Rate of discontinuation (>60 days treatment gap) over time and adherence (as medication possession ratio-MPR: <40%; 40-80%; >80%) and switching pattern of different NOAs during first year of treatment were explored. Moreover, changes in NOA treatment after major bleeding events occurrence were evaluated.

Results

Overall, 7,953 NVAF patients started a NOA treatment during the study years (rivaroxaban: 3,716, 47%; dabigatran: 2,364, 30%; apixaban: 1,873, 23%). Of these, almost ¾ had high thromboembolic risk (CHA2DS2-VASc score≥3) and 55% were previously treated with warfarin. NOA treatment discontinuation rate was equal to 23% during a median follow-up of 6 months with highest rates reported for dabigatran (37% vs. 19% for rivaroxaban vs. 15% for apixaban) who stopped more frequently therapy during the first month. Dabigatran users showed also lowest level of treatment adherence during first year of follow-up (MPR<40%: N=305, 20%) than rivaroxaban (N=175, 10%) or apixaban (N= 55, 9%) incident users.

Moreover, 12.5% of incident NOA users switched to another oral anticoagulant, mostly to warfarin (58% of total switchers). During NOA therapy, bleeding requiring hospitalization occurred in 198 patients (2.5%, almost half of the cases being gastrointestinal bleeding). Of these, 23% discontinued any oral anticoagulant treatment thereafter, while 8% switched to warfarin during the following six months.

Conclusions

High proportion of NVAF patients had a high risk of thromboembolic events, in line with phase 3 clinical trials. Differences in persistence and adherence to therapy across NOAs, especially after occurrence of bleeding event, were found. Strategies to improve adherence and persistence to NAO treatment while minimizing risk of bleeding are needed.

References

[1] Bungard TJ, Ghali WA, Teo KK, McAlister FA, Tsuyuki RT. Why do patients with atrial fibrillation not receive warfarin? ArchIntern Med 2000;160:41-6.

[2] Whittle J, Wickenheiser L, Venditti LN. Is warfarin underused in the treatment of elderly persons with atrial fibrillation? Arch Intern Med 1997;157:441-5.

[3] Connolly Stuart J. et al. Dabigatran versus Warfarin in Patients with Atrial Fibrillation. N Engl J Med 2009;361:1139-51.

[4] Patel MR, Mahaffey KW, Garg J, et al. Rivaroxaban versus warfarin in non valvular atrial fibrillation. N Engl J Med 2011;365:883–91.

[5] Granger CB. (), Apixaban versus warfarin in patients with atrial fibrillation. N Engl J Med 2011;365(11):981-92.

[6] Hernandez I, Baik SH, Piñera A, Zhang Y. Risk of bleeding with dabigatran in atrial fibrillation, JAMA Int. Med 2015;175:18–24.

[7] Sanfélix-Gimeno G, Rodríguez-Bernal CL, Hurtado I, Baixáuli-Pérez C, Librero J, Peiró S. Adherence to oral anticoagulants in patients with atrial fibrillation—a population-based retrospective cohort study linking health information systems in the Valencia region, Spain: a study protocol. BMJ Open 2015;5:e007613.

[8] Schulman S, Shortt B, Robinson M, et al. Adherence to anticoagulant treatment with dabigatran in a real-world setting. J Thromb Haemost 2013;11:1295–9.

[9] Shore S, Carey EP, Turakhia MP, et al. Adherence to dabigatran therapy and longitudinal patient outcomes: insights from the veterans health administration. Am Heart J; 2014;167: 810–7.

ANALYSIS OF THE GENETIC VARIANTS ASSOCIATED WITH SGP130 IN THE IMPROVE STUDY

Bonomi Alice[1], Veglia Fabrizio[1], Baldassarre Damiano[1,2], Tremoli Elena[1,2], de Faire Ulf[3], Gigante Bruna[3,4] on behalf of IMPROVE Study Group

1. *Monzino Cardiology Center,IRCCS, Milan, Italy*
2. *Dipartimento di Scienze Farmacologiche e Biomolecolari, Università di Milano, Italy.*
3. *Unit of Cardiovascular Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stochkolm, Sweden*
4. *Division of Cardiovascular Medicine, Department of Clinical Sciences, Danderyd University Hospital, Stockholm, Sweden.*

Introduction

Genetic variants regulating sgp130, a natural antagonist of the pro-inflammatory/pro-atherogenic IL6/sIL6R complex, are largely unknown. Recently, a single nucleotide polymorphism (SNP) in the gene coding gp130, the SNP rs3729960, has been shown to be associated with a decreased risk of myocardial infarction and circulating levels of sgp130 [***1, 2***].

Objective is to identify novel SNPs associated with sgp130 circulating levels and to explore the potential association between SNPs associated with sgp130 and markers of sub-clinical atherosclerosis.

Methods

We performed the analysis in IMPROVE study (n=3703), a large multicentre study designed to investigate the determinants of carotid intima media thickness (c-IMT) as a measure of subclinical atherosclerosis [***3***]. Genomic DNA from IMPROVE study participants was genotyped with two combined arrays, the CardioMetaboChip and ImmunoChip, containing approximately 200.000 SNPs.

We tested the association of all SNPs with log-transformed sgp130 by linear regression adjusted for age, gender and population stratification. Principal component analysis (PCA) was applied to summarize the whole set of SNPs associated with sgp130 a reduced number of factors. The association between the factors and c-IMT was assessed by multinomial logistic regression.

Results

We identified 41 SNPs associated with sgp130 levels after adjustments. The PCA identified 10 factors explaining 75% of the variance. Factor 1 was strongly correlated with SNPs located in region of *CUX2* gene previously associated with the risk of coronary artery disease. The minor alleles at these SNPs were associated with lower levels of sgp130 (Beta=-0.373, p=8.19E-05).

Regarding the association with c-IMT, only factor 6, correlated with SNPs mapping at an unknown genetic location, was associated with lower levels of sgp130 (β=-0.011, SE=0.002, p<.0001) and lower level of c-IMT (β=-0.071, SE=0.033, p=0.031) after adjustments (***Table 1***).

Table 1. Association between Factors and sgp130 and between Factors and Max-IMT adjusted for age, gender and Center.

	Sgp130			Max-IMT		
Factors	**β**	**SE**	**p-value**	**β**	**SE**	**p-value**
Factor1	-0.009	0.002	0.0002	0.0266	0.0326	0.4144
Factor2	-0.009	0.002	<.0001	-0.0402	0.0329	0.2218
Factor3	0.008	0.002	0.0008	0.0534	0.0332	0.1074
Factor4	0.009	0.002	0.0001	0.0141	0.0331	0.6694
Factor5	0.008	0.002	0.0006	0.0191	0.0331	0.5652
Factor6	-0.011	0.002	<.0001	-0.0715	0.0332	0.0310
Factor7	0.008	0.002	0.001	-0.0017	0.0335	0.9604
Factor8	-0.011	0.002	<.0001	-0.0185	0.0328	0.5732
Factor9	-0.009	0.002	0.0002	-0.0061	0.0343	0.8580
Factor10	0.022	0.002	<.0001	0.0151	0.0332	0.6494

Conclusions

Circulating sgp130 levels are explained by the combined effect of 10 factors identified by PCA. Our results also suggest that common genetic variants might affect circulating levels of sgp130 and markers of subclinical atherosclerosis.

References

[**1**] Benrick A, Jirholt P, Wernstedt I, et al. A non-conservative polymorphism in the IL-6 signal transducer (IL6ST)/gp130 is associated with myocardial infarction in a hypertensive population. Regul Pept 2008;146 (1-3):189–96.

[**2**] Wonnerth A, Katsaros KM, Krychtiuk KA, et al. Glycoprotein 130 polymorphism predicts soluble glycoprotein 130 levels.Metabolism 2014;63(5):647-53.

[**3**] Baldassarre D,Nyyssönen K, Rauramaa R, et al. Cross-selection analysis of baseline data to identify the major determinants of carotis intima-media thickness in a European population: the IMPROVE study. Eur Heart J 2012;31(5):614-22.

A STUDY OF THE CLINICAL PHENOTYPE OF PATIENTS WITH MYOTONIC DYSTROPHY TYPE 2 (DM2) IN GERMANY

Mondello Stefania[1], Montagnese Federica[2], Wenninger Stephan[2], Wolfram Kress[3], Schoser Benedikt[2], Vestri Annarita[4]

1. *Department of Biomedical and Dental Sciences and Morphofunctional Imaging , University of Messina, Messina, Italy*
2. *Friedrich-Baur-Institute, Department of Neurology, Ludwig-Maximilians-University of Munich, Germany*
3. *Institute of Human Genetics, Würzburg, Germany*
4. *Department of Public Health and Infectious Disease, Sapienza University of Rome, Rome, Italy*

Introduction

Myotonic dystrophy type 2 (DM2) is a far younger and less well studied disease entity in comparison to myotonic dystrophy type 1 and the description of its phenotype mainly relies on studies performed on small cohorts [***1, 2***].

The purpose of this study was to analyze the association between clinical phenotype and patient's characteristic using the largest observational cohort study of patients with a genetic diagnosis of DM2 (n=307) collected in Germany between 2001 and 2016.

Methods

Detailed biological, neurological and electrophysiological data were collected on all patients alongside correlating clinical features and multisystem involvement. Multinomial regression analyses were applied.

Results

Our cohort comprised 186 females (61%). Proximal muscle weakness was the leading first symptom (55.4%), followed by pain (35.5%) and myotonia (25.4%). Proximal muscle weakness was more common in women than men (64.9% vs. 43.8%, p=0.0006). Patients with muscle weakness at onset were older than those with pain and myotonia (median 49 vs. 39 and 30yrs, p<0.0001). A multinomial regression model revealed that age at onset and sex are significantly and independently associated with specific types of symptoms. Specifically, being male was associated with significantly higher odds of developing pain (OR=2.94 [95%CI 1.53-5.67]; p=0.0012), while each additional disease year was associated with 10% lower odds of developing myotonia (OR=0.9 [95%CI 0.87-0.93], p<0.0001) and a 6% decrease in the odds of developing pain (OR=0.94 [95%CI 0.91-0.97], p<0.0001). Frequent multisystem comorbidities were: cataract (49%), dyslipidaemia (41%), thyroid dysfunction (32%), and diabetes (30%). Cataract and thyroid diseases occurred more often in women (p=0.002 and p<0.001).

Conclusions

We provide here an update of the clinical description of DM2 based on a large cohort, demonstrating a profound gender and age influence on the phenotype. These age- and gender-specific differences should be considered in the diagnostics, management and future clinical studies of patients with DM2.

References

[**1**] Schoser B, Timchenko L. Myotonic dystrophies 1 and 2: complex diseases with complex mechanisms. Curr Genomics 2010;11(2):77-90.

[**2**] Schoser BG, Schneider-Gold C, Kress W, et al. Muscle pathology in 57 patients with myotonic dystrophy type 2. Muscle Nerve 2004;29(2):275-81.

SVILUPPO DI UN METODO DI ANALISI PER L'IDENTIFICAZIONE DELLE ALTERAZIONI EPIGENETICHE STOCASTICHE ED APPLICAZIONE IN UNA CASISTICA DI BAMBINI NATI DA FECONDAZIONE IN VITRO.

Gentilini Davide[1], Somigliana Edgardo[2], Pagliardini Luca[3], Viganò Paola[3], Di Blasio Anna Maria [1], Oliveri Antonino [4], Bernardinelli Luisa [4]

1. *Istituto Auxologico Italiano IRCCS, Cusano Milano, Milan, Italy;*
2. *Infertility Unit, Fondazione Ca' Granda, Ospedale Maggiore Policlinico, Milan, Italy;*
3. *Reproductive Sciences Laboratory, Division of Genetics and Cell Biology, IRCCS Ospedale San Raffaele, Milan*
4. *Department of Brain and Behavioural Science, University of Pavia, Italy.*

Introduzione

Ad oggi le nascite con tecnologie di fecondazione assistita rappresentano circa l'1% del totale e l'affinamento tecnologico ha determinato un conseguente aumento dei tassi di successo. La procreazione medicalmente assistita (PMA) è ritenuta una pratica sicura ed è diffusa in tutto il mondo, tuttavia recentemente sono state sollevate perplessità circa possibili sequele a carico dei figli di genitori che si siano sottoposti a tali procedure. Alcuni studi hanno riportato una frequenza di difetti alla nascita in bambini nati con PMA aumentata di almeno il doppio rispetto a bambini concepiti naturalmente. Confrontando i bambini nati con tecnologie di PMA e concepiti naturalmente, si nota un aumento dell'incidenza di IUGR (intrauterine growth retardation) ed una maggiore incidenza di patologie da difetti di imprinting come sindrome di Beckwith Wiedemann o sindrome di Angelman. L'ipotesi più accreditata è che la metodica possa interferire a livello epigenetico agendo prevalentemente sullo stato di metilazione del genoma [***1***]. La metilazione del DNA rappresenta infatti un meccanismo di regolazione genetica molto importante ed è coinvolto in fenomeni cruciali come l'imprinting, lo sviluppo embrionale ed il controllo dell'espressione dei geni. Molti studi hanno cercato di far luce su questa spinosa questione cercando alterazioni nei livelli di metilazione a carico dei bambini nati da PMA, tuttavia ad oggi, i risultati rimangono contrastanti. Tutti gli studi hanno impiegato un disegno sperimentale guidato da un'ipotesi a priori puntando l'attenzione su un ristretto numero di geni potenzialmente coinvolti. Questa caratteristica che accomuna gli studi, potrebbe essere però considerata un limite se si è alla ricerca di alterazioni stocastiche che potrebbero cadere ovunque nel genoma. Inoltre la manipolazione embrionale potrebbe produrre a livello genomico danni stocastici e non condivisi tra i soggetti che difficilmente emergerebbero dalle analisi differenziali abitualmente condotte.
L'obiettivo dello studio è stato quello di verificare se le metodiche di PMA siano in grado di indurre alterazioni epigenetiche potenzialmente importanti o rischiose per la salute dell'individuo.

Metodi

Un gruppo di bambini nati da PMA (n=23) ed un gruppo di bambini nati naturalmente (n=61) sono stati reclutati nello studio a 20 settimane di gestazione. Una volta nati i livelli di metilazione di oltre 485.000 loci del loro genoma sono stati analizzati su DNA estratto dal sangue tramite metodica microarray (illumina 450K). Dati clinici e biometrici dei bambini e dei genitori sono stati raccolti contestualmente. L'analisi differenziale di metilazione è stata condotta utilizzando l'ambiente R ed il pacchetto RnBeads [***2***]. L' analisi delle componenti principali è stata parallelamente condotta sui dati fenotipici raccolti e le componenti principali risultate significativamente diverse tra i due gruppi in studio sono state impiegate come covariate nell'analisi differenziale. Un metodo alternativo è stato inoltre sviluppato utilizzando l'ambiente R allo scopo

di identificare le variazioni epigenetiche rare e stocastiche che non emergono dell'analisi differenziale. In ogni soggetto i valori di metilazione misurati sono stati confrontati verso una popolazione di riferimento (169 soggetti sani concepiti naturalmente). Ogni qualvolta il soggetto ha mostrato livelli di metilazione eccedenti oltre 3IQR rispetto alla popolazione di riferimento è stato classificato come "extreme outlier" o "epimutato". L'analisi è stata condotta per tutte le 485.000 sonde presenti sull'array identificando per ogni soggetto il numero totale delle mutazioni epigenetiche stocastiche e la loro posizione genomica. Successivamente è stata sviluppata una funzione che, utilizzando una finestra mobile capace di scorrere lungo il genoma ed impiegando un test che sfrutta la distribuzione ipergeometrica, ha identificato, in ogni soggetto, regioni dove i difetti epigenetici sono risultati superiori all'atteso in modo significativo. Il metodo è stato inoltre validato analizzando 10 campioni in duplicato e 48 soggetti con difetti epigenetici rari noti.

Risultati

Una volta corretti i risultati per test multipli l'analisi differenziale non ha identificato alcun locus differentemente metilato in modo significativo tra esposti e non esposti alla PMA. Non sono emersi risultati significativi nemmeno aggregando i dati di metilazione in base alla localizzazione genomica. L'analisi delle mutazioni epigenetiche stocastiche non ha evidenziato differenze significative tra i due gruppi in studio, il numero mediano delle SEM espresso in scala logaritmica è risultato avere il valore di 7.4 (Q1 = 7.2; Q2 = 7.6) negli esposti e 7.4 (Q1 = 7.2; Q2 = 7.6) nei non esposti alla PMA. Il numero delle mutazioni epigenetiche è risultato però associato ad altre variabili in esame come la modalità di parto ($p<0.0001$), il peso alla nascita $p<0.001$), il BMI della madre al momento del concepimento ($p<0.0001$) e l'età gestazionale ($p<0.0001$).L' analisi delle regioni in cui le mutazioni epigenetiche risultano superiori all'atteso ha infine identificato un soggetto portatore di un difetto di imprinting tra i bambini concepiti naturalmente.

Conclusioni

Lo studio ha valutato l'eventualità che le metodiche di PMA possano introdurre difetti epigenetici nei bambini concepiti in vitro. L'analisi differenziale non ha evidenziato nessun locus con differenze significative tra i due gruppi in studio ne considerando i loci analizzati indipendentemente ne aggregando i dati di metilazione in base alla localizzazione genomica. Inoltre lo studio, utilizzando un approccio inedito ed avvalendosi di una popolazione di riferimento, ha valutato in ogni soggetto il numero di mutazioni epigenetiche stocastiche. Questo approccio ha permesso di identificare le alterazioni epigenetiche rare e stocastiche che non sarebbero altresì emerse dall'analisi differenziale. Comparando le alterazioni stocastiche identificate nelle due popolazioni non sono emerse differenze significative ne nel loro numero ne nella distribuzione. Lo studio evidenzia che il numero di alterazioni epigenetiche non sembra aumentare in risposta ai trattamenti di manipolazione degli embrioni. Questa evidenza potrebbe suggerire che l'aumentata incidenza di difetti di imprinting nei bambini concepiti tramite PMA sia da ricondurre ad altri fattori e che tali metodiche siano da considerarsi sicure. Il numero di mutazioni epigenetiche stocastiche è risultato associato ad altre variabili come il BMI della madre prima del concepimento, l'età gestazionale, la modalità di parto ed il peso del bambino alla nascita suggerendo che tali fattori potrebbero giocare un ruolo a livello epigenetico. Il metodo sviluppato per l'identificazione delle mutazioni epigenetiche stocastiche ha permesso di identificare in un campione un difetto epigenetico che non sarebbe altrimenti stato evidenziato utilizzando altri approcci oggi in uso. Questo risultato supporta l'idea di poter applicare questo tipo di analisi come test di screening allo scopo di identificare precocemente difetti epigenetici congeniti.

Bibliografia

[**1**] Iliadou AN, Janson PC, Cnattingius S. Epigenetics and assisted reproductive technology. J Intern Med 2011;270(5):414-20**.**

[**2**] Assenov Y, Muller F, Lutsik P et al. Comprehensive analysis of DNA methylation data with RnBeads. Nat Methods 2014;11(11):1138-40.

AGREEMENT BETWEEN AN AUTOMATED VOLUME BREAST SCANNER AND HANDHELD ULTRASOUND FOR DIAGNOSTIC BREAST EXAMINATIONS

Barr Richard G[1,2], DeVita Robert[1], Destounis Stamatia[3], Manzoni Federica[4,5], De Silvestri Annalisa[5], Tinelli Carmine[5]

1. *Ohio Medical University, Rootstown, Ohio, USA*
2. *Southwoods Imaging,Youngstown, Ohio, USA*
3. *Elizabeth Wende Breast Care, Rochester, New York, USA*
4. *Biostatistics and Clinical Epidemiology Unit, Department of Public Health, Experimental and Forensic Medicine, University of Pavia, Pavia, Italy*
5. *Biometric and Clinical Epidemiology, Istituto di Ricovero e Cura a Carattere Scientifico Policlinico San Matteo, Pavia, Italy*

Introduction

Ultrasound (US) is often used in the workup of patients presenting with focal breast problems. It can be for a palpable mass, abnormal mammographic findings, abnormal magnetic resonance imaging findings, breast discharge, or pain. The examination is usually performed with a handheld US device to perform a limited examination over the area of concern. An automated breast volume scanner (ABVS) has been used for whole-breast supplemental screening. [***1–4***]

Previous studies have demonstrated that the ABVS has high sensitivity and fair interobserver concordance.[***5–7***] The model of physician-performed handheld US examinations is time-constraining for the physician. The development of ABVSs that provide 3-dimensional breast US volume scan replaces the need for a highly trained physician or sonographer to perform the examination, as the entire volume scanned is available for review and provides reconstructions in the coronal and perpendicular planes.

Objectives is to compare the agreement and interobserver variability of diagnostic handheld ultrasound (US) and a single volume on an automated breast volume scanner (ABVS) and to determine whether there was a significant difference if the ABVS was used by a sonographer or mammographic technologist.

Methods

Ninety patients scheduled for diagnostic US examinations were randomized to either handheld US or the ABVS first. The AVBS was randomized between a sonographer and a mammographic technologist performing the study. The studies were blinded, randomized, and read by 2 radiologists. The lesion with the highest Breast Imaging Reporting and Data System (BI-RADS) score was used in the analysis [***8***]. Final diagnoses were made by core biopsy or follow-up for 2 years. Lesions included 9 malignant and 81 benign.

Statistical Analysis

For sample size considerations, we powered the study as an agreement evaluation of malignancy between the techniques. In a test for agreement between the techniques using the k statistic, a sample size of 80 patients achieves 90% power to detect a true k value of 0.90 in a test of H0 (k = 0.60) versus H1 k < 0.60 or > 0.60) when there are 2 categories with frequencies equal to 0.80 and 0.20. This power calculation was based on a significance level of 0.05, and an extra 10% of patients were enrolled to account for loss to follow-up and missing data. Quantitative variables were described as mean and standard deviation or median and interquartile range, qualitative ones as counts and percentages. Agreement between readers and methods was evaluated by means of k statistics, presented along with their 95% confidence intervals (CIs). The k statistic measure of agreement is scaled to be 0 when the amount of agreement is what would be expected to be observed by chance and 1 when there is perfect agreement. For intermediate values,

Landis and Koch [***9***] suggest the following interpretations: below 0.0, poor; 0.00 to 0.20, slight; 0.21 to 0.40, fair; 0.41 to 0.60, moderate; 0.61 to 0.80, substantial; and 0.81 to 1.00, almost perfect. For independent comparisons, k values were compared by using z score statistics. Diagnostic performance assuming BI-RADS categories 1, 2, and 3 were benign and BI-RADS categories 4 and 5 were malignant and using biopsy or clinical stability for 2 years as the reference standard was assessed by receiver operating characteristic (ROC) curve analysis. The sensitivity, specificity, and area under the curves were presented along with 95% CIs. [***10***] Pair-wise comparisons of areas under the curves between readers and methods were also given by z statistics.[6] P< 0.05 was considered statistically significant. All tests were 2 sided. The data analysis was performed with the Stata statistical package (release 14.0; StataCorp, College Station, TX) and MedCalc (version 16.2; MedCalc Software bvba, Mariakerke, Belgium).

Results

The 90 patients had a mean age of 53.1 (sd 16.3) years. The k value for agreement between the ABVS and handheld US was 0.831 (95% confidence interval, 0.744–0.925), whereas the global agreement for a 7-point BI-RADS score was 0.488 (0.372–0.560). The agreement between the ABVS and handheld US was nearly the same when the ABVS was used by a mammographic technologist (k = 0.858 [0.723–0.963]) or sonographer (k = 0.803 [0.596–1.000]; p = 0.47).

The areas under the receiver operating curves for characterization by the ABVS were 0.91 (0.84–0.96) for reader 1 and 0.91 (0.83–0.96) for reader 2; those for handheld US were 0.91 (0.84–0.96) for reader 1 and 0.83 (0.74–0.90) for reader 2, with no statistical difference (***Figure 1***). The agreement based on pathologic images was k = 0.831 (0.718–0.944); for handheld US, k = 0.795 (0.623–0.967); and for the AVBS, k = 0.869 (0.725–1.000).

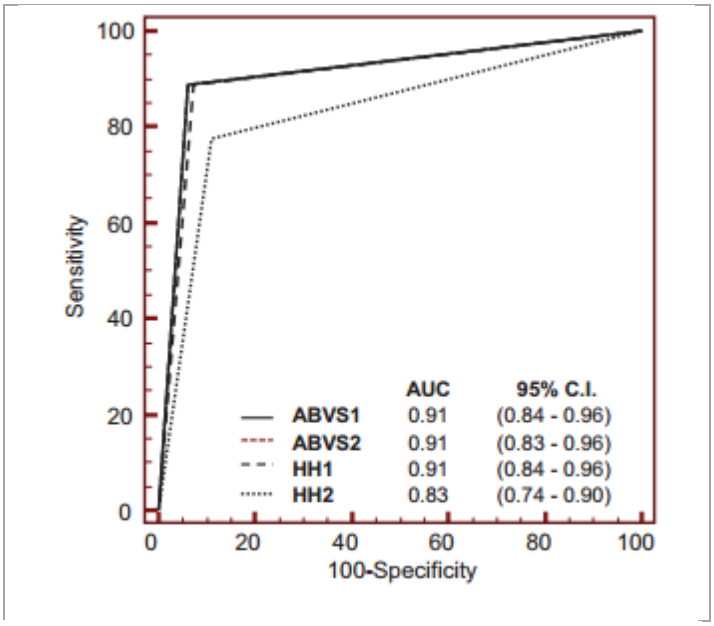

Figure 1. Areas under the ROC curves with 95% CIs concerning the ability to detect a pathologic breast image. ABVS1 indicates interpretation of ABVS images by the first reader; ABVS2, interpretation of ABVS images by the second reader; HH1, interpretation of handheld US images by the first reader; and HH2, interpretation of handheld US images by the second reader

Conclusions

Based on our results, a focal diagnostic US examination performed with an ABVS by either a sonographer or mammographic technologist has good agreement and the same diagnostic performance as a handheld US examination performed by a sonographer, although further studies are required to demonstrate that the diagnostic performance of the ABVS is the same as that of handheld US in the diagnostic setting.

References

[1] Wojcinski S, Farrokh A, Hille U, et al. The Automated Breast Volume Scanner (AVBS): initial experiences in lesion detection compared with conventional handheld B-mode ultrasound—a pilot study of 50 cases. Int J Womens Health 2011;3:337–46.

[2] Kelly KM, Dean J, Comulada WS, et al. Breast cancer detection using automated whole breast ultrasound and mammography in radiographically dense breasts. Eur Radiol 2010;20:734–42.

[3] Brem RF, Tabar L, Duffy SW, et al. Assessing improvement in detection of breast cancer with three-dimensional automated breast US in women with dense breast tissue: the SomoInsight Study. Radiology 2015;274:663–73.

[4] Shin HJ, Kim HH, Cha JH, Current status of automated breast ultrasound. Ultrasonography 2015;34:165–72.

[5] Wojcinski S, Gyapong S, Farrokh A, et al. Diagnostic performance and inter-observer concordance in lesion detection with the automated breast volume scanner (ABVS). BMC Med Imaging,2013;13:36.

[6] Shin HJ, Kim HH, Cha JH, et al. Automated ultrasound of the breast for diagnosis: interobserver agreement on lesion detection and characterization. AJR Am J Roentgenol 2011;197: 747–75.

[7] Lin X, Wang J, Han F, et al. Analysis of eighty-one cases with breast lesions using automated breast volume scanner and comparison with handheld ultrasound. Eur J Radiol 2012;81: 873–8.

[8] D'Orsi CJ, Sickles EA, Mendelson EB, et al. ACR BI-RADS Atlas: Breast Imaging Reporting and Data System. Reston, VA; American College of Radiology; 2013.

[9] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.

[10] Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology1983;148:839–43.

SINGLE AGENT TRASTUZUMAB OR LAPATINIB TO TREAT HER2-OVEREXPRESSING BREAST CANCER: COMBINING PAST AND CURRENT EVIDENCE IN A BAYESIAN REANALYSIS

De Vito Giuseppe[1], Baldi Ileana[2], Bottigliengo Daniele[3], Nuzzo Annamaria[4], Montemurro Filippo[4], Berchialla Paola[5]

1. *National Institute of Optics, National Research Council, Florence, Italy*
2. *Unit of Biostatistics, Epidemiology and Public Health, Department of Cardiac, Thoracic and Vascular Sciences, University of Padova, Italy*
3. *Department of Public Health and Pediatric Sciences, University of Torino, Italy*
4. *Department of Investigative Clinical Oncology, Fondazione del Piemonte per l'Oncologia, Candiolo Cancer Institute, Torino, Italy*
5. *Department of Clinical and Biological Sciences, University of Torino, Italy*

Introduction. Recent studies investigated the possible role of Human Epidermal Growth Factor Receptor 2 (HER2)-targeting compounds as first-line, single-agent therapy for HER2-over-expressing Breast Cancer (BC) with promising results [***1, 2***]. In particular, for a subgroup of patients the observed disease control duration was similar to that reported for the commonly-used anti-HER2 and chemotherapy combination treatment.
In order to gather further insights about the biomarkers that characterise the patients that can benefit from anti-HER2 single-agent therapy and to evaluate the efficacy of this therapy in patients not previously treated for HER2-positive metastatic BC, two clinical trials were initiated, HERLAP I and HERLAP II, both testing two anti-HER2 agents: trastuzumab and lapatinib [***3***]. However, HERLAP I was prematurely terminated, also due to the slow accrual of patients.
We devised to measure the Progression Free Survival (PFS) for patients in single-agent therapy from the HERLAP trial data, in order to compare it to the combination treatment. However, the small sample size makes it difficult to apply frequentist statistical approaches and calls for an integration of the information derived from the two trials. In this regard, the Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials, issued by the Food and Drug Administration, states the opportunity to use a Bayesian approach to combine prior information with new observations, suggesting basing this information on empirical evidences [***4***]. Using this approach, we generated prior distributions from the data of the early-stopped HERLAP I trial, devising to use them in the analysis of the HERLAP II trial results.

Methods.
We planned to employ a hierarchical Bayesian Weibull survival model to characterise both the biological PFS (i.e. taking in consideration only the period of exclusive administration of anti-HER2 agents) and the total PFS (regardless of protocol failures).
In particular, using non-informative prior distributions, we derived posterior distributions for the parameters of the Weibull model based on the HERLAP I data, and we have planned to use them in turns as prior distributions to derive the posterior distributions for the parameters based on the HERLAP II data, thus "borrowing strength" from the first trial to the second.

Results
After describing the statistical method in detail and presenting the data, in this contribution we shall discuss preliminary results that we obtained by deriving the posterior distributions from the HERLAP I data. In particular, we observed that the median survival times in days (and the extremes of their 95% credible

intervals) for the biological PFS and the total PSF are 190 (96; 355) and 333 (172; 672), respectively. If we take into account only the trastuzumab-treated patients, then these values become 335 (139; 893) and 442 (162; 1304); whereas considering only the lapatinib-treated patients they become 99 (46; 232) and 250 (102; 805). It is interesting to note that these survival times are similar to those reported for the combination treatment. The median PFS values, together with the PFS probability curves based on the posterior distributions of the Weibull distribution parameters, are graphically shown in ***Fig. 1***: curves based on the aggregated data are plotted on top row, and the treatment-specific curves are plotted on bottom row. It is worth noting that these survival times are similar to those reported in the literature for the combination treatment: in particular, the total PFS value (333 days) is perfectly in the range of the reported values: 225-356 days. In addition, the median PFS value for the trastuzumab treatment resulted larger with respect to the lapatinib treatment with 97% probability for the biological PFS and 81% probability for the total PFS.

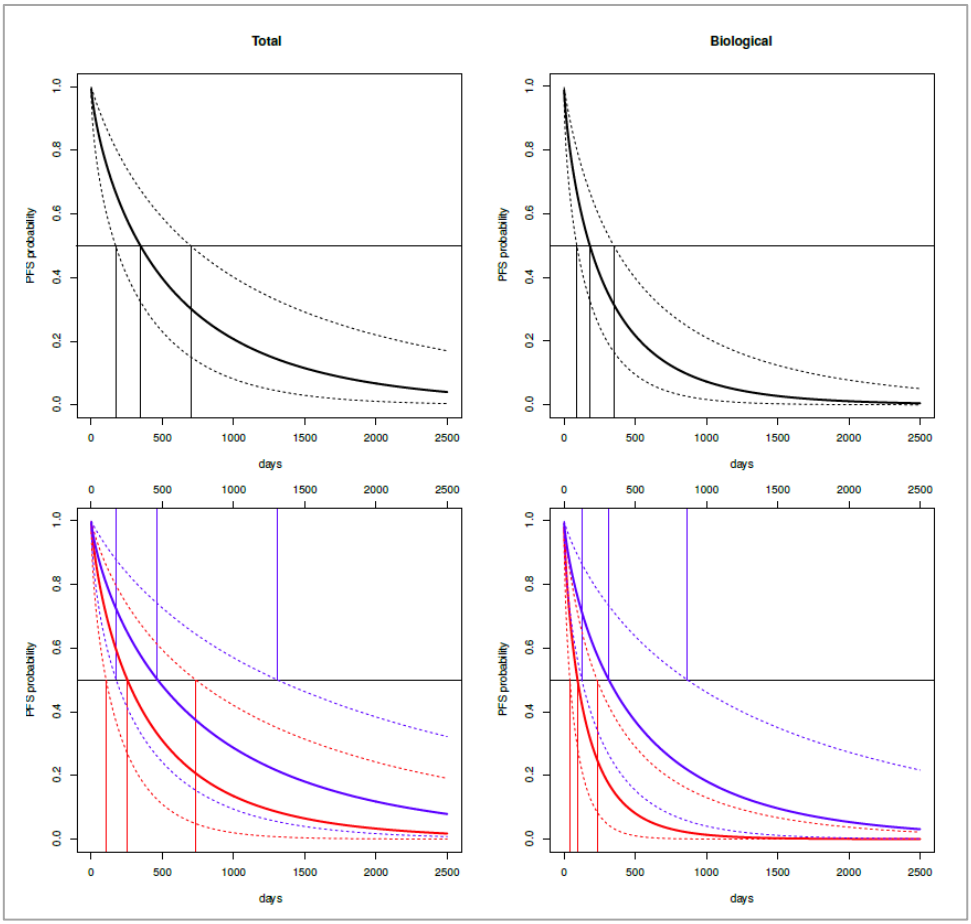

Figure 1. PFS probability curves based on the posterior distributions of the *r* and *m* parameters. Top row: curves drawn from the aggregated data. Bottom row: curves drawn considering only the trastuzumab-treated (blue) or the lapatinib-treated (red) patients. Left column: total PFS. Right column: biological PFS. The thick continuous curves represent the median values of the survival probability distributions, while the thin dashed curves represent the borders of its 95% credible interval. The vertical lines indicate the median PFS times (thick lines) and the extremes of its 95% credible interval (thin lines) for the respective populations indicated by their colour.

Conclusions

These data represent an additional suggestion for the efficacy of anti-HER2 single-agent therapy for HER2-positive metastatic BC. Therefore, they can be useful to improve the quality of life for therapy-responsive patients, sparing them the toxic collateral effects of chemotherapy at least for an initial period.
Moreover, the employed statistical technique is an exemplification of the FDA suggestions for the use of the Bayesian framework as a way to integrate prior empirical-based information in clinical trial data analysis, thereby enabling the future design of clinical trials with smaller sample sizes.

References

[1] Vogel CL, Cobleigh MA, Tripathy D, et al. Efficacy and safety of trastuzumab as a single agent in first-line treatment of HER2-overexpressing metastatic breast cancer. J Clin Oncol 2002;20;719-26.
[2] Gomez HL, Doval DC, Chavez MA, et al. Efficacy and safety of lapatinib as first-line therapy for ErbB2-amplified locally advanced or metastatic breast cancer. J Clin Oncol 2008;26,2999-3005.
[3] Montemurro F, Prat A, Rossi V, et al. Potential biomarkers of long-term benefit from single-agent trastuzumab or lapatinib in HER2-positive metastatic breast cancer. Mol Oncol 2014;8: 20-6.
[4] Food and Drug Administration, 2010. Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials.

RETROSPECTIVE EVALUATION OF LATE RADIATION DAMAGES AFTER FOCAL RADIOTHERAPY FOR CHILDHOOD BRAIN TUMORS

Lecchi Mara[1], Cavatorta Claudia[2] , Montin Eros[3], Oprandi Chiara[4], Meroni Silvia[2], Pecori Emilia[5], Spreafico Filippo[6], Diletto Barbara[5], Biassoni Veronica[6], Schiavello Elisabetta[6], Arrigoni Filippo[4], Poggi Geraldina[4], Massimino Maura[6], Mainardi Luca[3], Pignoli Emanuele[2], Gandola Lorenza[5], Verderio Paolo[1]

1. *Medical Statistics, Biometry and Bioinformatics, Fondazione IRCCS Istituto Nazionale Tumori, Milan Italy*
2. *Medical Physics, Fondazione IRCCS Istituto Nazionale Tumori, Milan Italy*
3. *Biosignal-Bioimage and Bioinformatics, Politecnico di Milano, Milan, Italy*
4. *Child Neuro-Oncology Rehabilitation, Pediatric Neuro-Radiology, Istituto Scientifico E. Medea, Bosisio Parini, Italy*
5. *Pediatric Radiotherapy, Fondazione IRCCS Istituto Nazionale Tumori, Milan Italy*
6. *Pediatric Oncology Unit, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy*

Introduction

Late neurocognitive sequelae in childhood brain tumors survivors have been extensively documented [***1, 2***] and correlated to several risk factors such as the tumor itself, young age, longer time since treatment and radiation treatments, in particular radiotherapy (RT) doses [***3***]. Magnetic Resonance Diffusion Tensor Imaging (DTI) and cognitive tests results could be used as indicators of brain damage [***4-6***].The lack of radiotherapy dose thresholds for different areas of child brain represents a significant gap in literature. Different dose levels related to the probability of neurocognitive damage for each specific brain areas could be used as dose constraints during RT treatment planning procedure.
To evaluate the relationship between brain tissue damage, by means of DTI metrics after focal RT for childhood brain cancer, neurocognitive outcome and RT dose levels and consequently identify tolerance doses of different sensitive areas of child brain to be used as dose constraints to optimize RT treatment planning.

Methods and Materials

45 retrospective patients (23 males, 22 female; median age at RT 6.2 years, range 1.1-22.5 years, median age at neurocognitive evaluations 11.1 years, range 4.4-25.9 years) treated with focal RT were recruited for DTI exams and neurocognitive tests. Patients' brains images were parceled in 116 Regions of Interest (ROIs) using the atlas segmented by Tzourio-Mazoyer et al. [***7***] on T1-weighted images provided by the Montreal Neurological Institute. By using a dedicated home-made non-rigid and multimodal registration framework, we collected mean RT doses and DTI metrics values for each ROI, selected DTI metrics were: Fractional Anisotropy (FA), Axial Diffusivity (AD), Radial Diffusivity (RD) and Trace of the tensor (TR). Neurocognitive and psychological assessments were based on standardized tests according to different patients' ages. Intellectual performances, attention, memory, executive functions and practical skills were assessed and scores of different tests were dichotomized according to standard cutoffs discriminating ordinary and not ordinary score. Pair-wise correlations (Spearman correlations) were estimated within each ROI to quantify the relationship between dose and DTI metrics. The pattern of association between dichotomous cognitive tests scores and RT dose as well as DTI values was assessed in each ROI by Wilcoxon Rank Sum test. Subsequently the value of RT dose discriminating patients with and without neurocognitive flaws (optimal cut-off value) was selected by using a Receiver Operating characteristic Curve (ROC).

Results

Dose resulted differently correlated with DTI parameters of each ROI. The statistical analysis allowed us to find at least one significant relationship between cognitive tests results and dose mean values of 99 ROIs. For the most significant relationships we obtained preliminary threshold dose values for specific ROIs at which could occur a cognitive damage. In 35 out of the 99 ROIs we also observed associations between cognitive tests and each of the DTI metrics.

Conclusions

Our results suggest the presence of a relationship between the RT dose and the considered damage variables. To better characterize the observed relationships, follow up data of prospectively collected patients will be considered with the final aim to provide a safe dose-threshold for each ROI. To this end, a prospective study in our Institution is already opened to children enrollment to assess the pre and post RT changes.

Acknowledgements

This work was supported by grant from Associazione Italiana per la Ricerca sul Cancro (AIRC, IG 2013).

References

[1] Muldhern R.K., Merchant T.E., Gajjar A. et al. Late neurocognitive sequelae in survivors of brain tumor in childhood. Lancet Oncol., 2004 Jul, 5(7):399-408

[2] Greene-Schloesser D., Robbins M.E., Peiffer A.N. et al. Radiation induced brain injury: a review. Frontiers in oncology, 2012 Jul, 2:73

[3] Qiu D.B.E., Kwong D.L.W., Chan G.C.F. et al. Diffusion tensor magnetic resonance imaging finding of discrepant fractional anisotropy between the frontal and parietal lobes after the whole-brain irradiation in childhood medulloblastoma survivors: reflection of regional white matter radiosensitivity?. Int. J. Radiat Oncol Biol Phys., 2007 Nov 69(3):846-851

[4] Hope T.R., Vardal J., Bjørnerud A. et al. Serial Diffusion Tensor Imaging for Early Detection of Radiation-Induced Injuries to Normal-Appearing White Matter in High-Grade Glioma Patients. Journal Of Magnetic Resonance Imaging, 2015 Jan, 41(2):414–423

[5] Palmer S.L., Glass J.O., Li Y. et al. White matter integrity is associated with cognitive processing in patients treated for a posterior fossa brain tumour. Neuro Oncol., 2012 Sep, 14(9):1185-93

[6] Khong P.K., Leung L.H.T., Fung A.S.M. et al. White Matter Anisotropy in Post-Treatment Childhood Cancer Survivors: Preliminary Evidence of Association With Neurocognitive Function. J Clin Oncol., 2006 Feb, 24(6):884-90

[7] Tzourio-Mazoyer N., Landeau B., Papathanassiou D. et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage, 2002 Jan, 15(1):273–289

RISK FACTORS FOR RELAPSE AND SURVIVAL IN A COHORT OF 1311 WOMEN WITH BOT

Breda Enrico[1], Ventriglia Mariacarla[2], Simonelli Ilaria[2] and Pasqualetti Patrizio[2]

1. *Medical Oncology Department, San Giovanni Calibita-Fatebenefratelli Hospital, Isola Tiberina, Rome, Italy;*
2. *Service of Medical Statistics and Information Technology, Fatebenefratelli Foundation for Health Research and Education, AFaR Division, Rome, Italy.*

Introduction

Borderline ovarian tumors (BOTs) represent a heterogeneous group of noninvasive tumors of uncertain malignant potential with characteristic histology. They have a more favorable outcome than the other ovarian cancers but current literature on this topic leads to a number of controversies about the optimal surgical management and staging of BOTs and information regarding prognostic factors is inconclusive. As suggested by Du Bois [***1***] large multicenter series may represent an appropriate alternative approach to better comprehend the natural course and prognostic profiles of BOT.
In order to evaluate the prognostic factor for women with BOT, the *multicentre Italian Trials in Ovarian cancer (MITO)* group started a retrospective–prospective cohort study of clinical characteristics, surgical management and surgical outcomes to identify variables affecting recurrence and survival in these cohort of women patients.

Methods

Consecutive patients with BOT were enrolled in 31 Italian hospitals between March 2000 and March 2015. Disease-free survival (DFS) was calculated as the months between the first surgery and the first relapse of cancer, while the overall survival (OS) was calculated as time to death. The DFS and OS probability were estimated using the Kaplan-Meier method. Survival in two or more groups of patients can be compared using Log-rank test. Multi-variable analysis was performed using the Cox regression model.

Results

Data were collected for 1311 women. Mean age was 45.7 (sd=15.6) years; 820 patients (62.5%) were below o equal to 50 years of age and 491 patients (37.5%) were above 50 years of age. 1138 patients were diagnosed in stage I (86.8%), 71 in stage II (5.4%), and 102 in stage III and IV (7.8%). 717 women had serous BOT (54.7 %), 459 mucinous BOT (35%), 54 seromucinous (4.1%) and 60 other types (4.6%), for 21 patients the information was missing. Overall 26 (2.0%) died within the observation period and 138 patients (10.5%) experienced a first relapse.
The overall estimated DFS was 44% (Standard Error, SE=12.9%). For 893 (20 events) patients we could calculate the OS time resulting equal to 92.4% (SE=2.9%). On the multivariable Cox regression model, age ≤50 years and FIGO(II or III STAGE vs I STAGE) and surgery (radical VS conservative) and serous histopathological subtypes were associated with significantly worse DFS. About OS, age ≤50 years was associated with significant improvement while FIGO(II or III STAGE vs I STAGE) was associated with significant worsening.

Conclusions

In summary, the age and the tumour or surgery related characteristics (i.e. FIGO stage, surgical treatment) were significant prognostic factors for recurrence and survival in women with BOT.

References

[1] du Bois A, Trillsch F, Mahner S, et al. Management of borderline ovarian tumors. Ann Oncol 2016;27(1):i20-i22.

MODELLO BAYESIANO CON STRATEGIA DI CAMPIONAMENTO LIMITATO PER LA STIMA DELL'AREA SOTTO LA CURVA CONCENTRAZIONE-TEMPO (AUC_{24H}) DELL'ANTIBIOTICO DAPTOMICINA NEL PAZIENTE ONCO-EMATOLOGICO

Giangreco Manuela[1], Cojutti Piergiorgio[1, 2], Pea Federico[1, 2], Isola Miriam[1]

1. *Dipartimento di area medica – DAME, Università degli Studi di Udine*
2. *Istituto di Farmacologia Clinica, Presidio Ospedaliero Universitario "Santa Maria della Misericordia", Azienda Sanitaria Universitaria Integrata di Udine*

Introduzione

Le batteriemie sostenute da patogeni Gram-positivi multiresistenti (MDR) sono una complicanza frequente delle terapie antiblastiche nel paziente neutropenico onco-ematologico. L'antibiotico daptomicina, grazie alla sua rapida attività battericida e alle sue caratteristiche fisico-chimiche di idrosolubilità, costituisce il farmaco di riferimento per queste infezioni. La variabilità interindividuale nell'esposizione plasmatica al farmaco può essere minimizzata attraverso la valutazione dell'area sotto la curva concentrazione-tempo delle 24h (AUC_{24h}), che consente l'aggiustamento posologico in corso di terapia nel singolo paziente [***1***]. Tuttavia, l'esecuzione di nove prelievi seriati per la stima dell'AUC_{24h} costituisce un disagio sia per il paziente che per il personale sanitario ed è un costo aggiuntivo per il reparto.[***1***]

L'obiettivo di questo studio è quello di sviluppare alcuni modelli statistici che consentano di stimare l'AUC_{24h} di daptomicina tramite l'esecuzione di soli 3 o 4 prelievi ematici.

Metodi

Sono state registrate le concentrazioni plasmatiche (n = 216) di daptomicina in 24 pazienti ai tempi 0, ½ ora, 1 h, 2 h, 3 h, 5 h, 7 h, 9 h, 12 h dalla somministrazione. L' AUC_{24h} (AUC_{ref}) per ciascun paziente è stata ottenuta tramite stima bayesiana ed è stata utilizzata come variabile dipendente all'interno dei modelli statistici bayesiani costruiti con strategia di campionamento limitato. Sono stati costruiti e testati modelli a quattro e tre punti variando anche la distribuzione a priori. Il modello migliore verrà scelto in base ai valori di MPE (median prediction error) e di RMSE (root median squared prediction error) oltre che per DIC (deviance information criterion) e test di convergenza dell'algoritmo bayesiano.[***2, 3***]

Risultati

Scegliendo una distribuzione a priori Normale sui coefficienti del modello e una distribuzione inversa-Gamma sul parametro di dispersione, quello migliore è risultato il modello lineare a tre punti $AUC_{24h} = 20.7951 + 7.44 \times C_0 + 4.80 \times C_5 + 12.90 \times C_9$. Il valore di MPE è pari a 1.71 mentre il valore di RSME risulta essere 12.37. L'algoritmo bayesiano è convergente e il valore DIC è pari a 228.56.

Conclusioni

Il modello sviluppato può essere applicato nella pratica clinica quotidiana nel contesto di programmi di monitoraggio dell'esposizione plasmatica di daptomicina al fine di personalizzare il regime posologico del farmaco in pazienti adulti con infezioni da patogeni Gram-positivi MDR.

Bibliografia

[**1**] Cojutti P, Candoni A, Ramos-Martin V, et al. Population pharmacokinetics and dosing consideration for the use of daptomycin in adult patients with haematological malignancies. J Antimicrob Chemother 2017;72(8):2342-50.

[**2**] Tsuruta H, Fukumoto M, Bax L, et al. Limited sampling strategies to estimate the area under the concentration–time curve. Biases and a proposed more accurate method. Methods Inf Med 2012;51:383–94.

[**3**] Lesaffre E, Lawson AB. Bayesian Biostatistics. John Wiley and Sons, 2012.

OCCURRENCE OF POTENTIAL BACTERIAL AND VIRAL PATHOGENS IN STABLE CHRONIC OBSTRUCTIVE PULMONARY DISEASE AND DURING ACUTE EXACERBATIONS OF THE DISEASE, IN ASIA PACIFIC [STUDY DESIGN]

Polito Letizia[1], Rosillon Dominique[2], Ashwani Kumar Arora[1], Rondini Simona[1], Taddei Laura[1]

1. *GlaxoSmithKline Vaccines, Siena, Italy*
2. *GlaxoSmithKline Vaccines, Wavre, Belgium*

Introduction

Chronic Obstructive Pulmonary Disease (COPD) is the third leading cause of mortality worldwide, accounting for 5% of all deaths globally [***1***]. COPD is a progressive life-threatening lung disease characterized by pulmonary airflow limitation that is not fully reversible [***2, 3***] and predisposes to exacerbations and serious illness. Acute exacerbation of COPD (AECOPD) accelerates pulmonary function decline, reduces quality of life and increases medical cost [***4***]. The aetiology of AECOPD remains not fully understood. Improved understanding of the role of respiratory pathogens in AECOPD is required and the use of molecular microbiological techniques may lead to insights into host–pathogen interactions and the development of more targeted therapeutic approaches [***5***].

The results of the EPIC Asia population-based survey suggest a high prevalence of COPD in the participating Asia-Pacific territories, and indicate a substantial socioeconomic burden of the disease in this region [***6***]. The infectious aetiology of AECOPD has been suggested to vary according to geographical regions. Few data are available in Asia Pacific.

This study (NCT03151395) aims to evaluate the occurrence of bacterial and viral pathogens in the sputum of stable COPD patients and at the time of AECOPD in several countries in Asia Pacific. This study will also evaluate the frequency, severity and duration of AECOPD, as well as the impact of AECOPD on health-related quality of life (HRQOL), healthcare utilization and lung function.

Methods

This is a prospective observational multi-country cohort study. This study targets enrolling approximately 200 stable moderate to very severe COPD patients, aged 40 years or older with at least 1 documented moderate or severe AECOPD in the year before enrolment. This study will be conducted in several sites among four countries in Asia Pacific [China (Hong Kong), Philippines, Korea, and Taiwan]. Patients will be screened to assess eligibility criteria. After enrollment, patients will be seen for three scheduled visits occurring at 6 months intervals. Patients' COPD symptoms will be daily monitored through the patient-completed electronic diary cards. Potential AECOPD will be detected as worsening of symptoms (two major symptoms or any major symptom together with any minor symptoms for at least two consecutive days) based on the Anthonisen criteria [***7***]. Unscheduled AECOPD visit will be performed, as well as follow-up phone call(s) to define end of AECOPD. In addition, a wide range of study procedures will be performed at study entry, scheduled visits and exacerbation visits (***table 1***). Sputum samples will be obtained by spontaneous or induced expectoration at each scheduled and unscheduled visit and assayed by culture and multiplex PCR for identification of respiratory pathogens (including, but not limited to, non-typeable *H. influenzae*, *M. catarrhalis*, *S. pneumoniae*, *P. aeruginosa*, *S. aureus*, respiratory syncytial virus, parainfluenza virus, rhinovirus, human metapneumovirus, influenza virus, adenovirus and coronavirus).

Table 1. List of study procedures for scheduled visits

Type of contact	Screening Visit	Visit 1	Visit 2	Visit 3	AECOPD visit
Time point	Pre-Month 0	Month 0	Month 6	Month 12	Unscheduled
Informed consent	●				
Check inclusion/exclusion criteria	●	O			
Record demographic data	●				
Record medical history	●				
Record history of AECOPD within the previous year	●				
Record intercurrent comorbidities	O	●	●	●	
Record history of pneumococcal and influenza vaccination	O	●	●	●	
Smoking exposure history (ATS-DLD-78A questionnaire)/ biomass exposure history	O				
Smoking status	●	●	●	●	
Physical examination including vital signs	●	O	O	O	●
Urine Pregnancy Test	●				●
Measure/ record height and weight	●			●	
Pre- and post-bronchodilator spirometry	●			●	
Chest X-ray	●				●
Record subject's COPD status (stable, recovered or not recovered)		●	●	●	
Record current medication for COPD		●	●	●	●
Record healthcare resource utilisation		●	●		●
Record additional COPD treatments prescribed by primary and secondary care physicians		●	●	●	●
Sputum sampling	●	●	●	●	●
Blood samples for biomarkers		●		●	
Record Adverse Events (AEs) related to study participation	●	●	●	●	●
Record serious adverse events (SAEs)	●	●	●	●	●
Record AECOPD severity					●
HRQOL questionnaires:					
COPD Assessment Tool (CAT)		O	O	O	O
St George's Respiratory Questionnaire for COPD (SGRQ-C)		O	O	O	

● is used to indicate a study procedure that requires documentation in the individual eCRF.
O is used to indicate a study procedure that does not require documentation in the individual eCRF.

Sample size calculation. The sample size calculation was based on the primary study endpoint of occurrence of potential bacterial and viral pathogens in sputum samples over the course of 1 year: Assuming that 10% of patients study discontinuation, on average 1 confirmed AECOPD per subject-year and 80% of a sputum sample at AECOPD. 200 patients will provide, for observed prevalence of 10%, 20%, 40%, the following precision (95% confidence interval): would be 6.0–15.3%, 14.4-26.6%, and 32.8-47.6%, respectively.
Data analysis. Prevalence of bacterial and viral pathogens in sputum sample will be compared between stable CODP and AECOPD visits using a Generalized Estimating Equations (GEE) model assuming a binomial distribution for the response variable with logit as link function and a compound symmetry correlation matrix (exchangeable structure) to account for the within-patient correlations [***8***]. The incidence

rate of all-cause AECOPD and of AECOPD having sputum containing bacteria\viral pathogens will be estimated using a model which accounts for repeated events, namely the generalized linear model assuming a negative binomial distribution for the response variable with logarithm as link function, and the logarithm of time for follow-up as an offset variable. Potential risk factors of exacerbations will be analysed using a multiple Cox regression models for recurrent events. The impact of AECOPD on health related quality of life, healthcare utilization and lung function (according to SGRQ-C/CAT score, total number of HCU, and post bronchodilator FEV_1%, respectively) will be assessed by the relationship between the changes from first visit to final scheduled visit versus the number of exacerbations. Correlation will be evaluated using Spearman's rank-order correlation coefficient.

Results
Not applicable

Conclusions
This epidemiological study will contribute to understand the causal relationship of bacterial and viral pathogen in exacerbation of COPD, in Asia Pacific. This knowledge will provide insights necessary to develop novel therapeutical approaches.

Funding: GlaxoSmithKline Biologicals SA

References
[1] World Health Organization (WHO). Chronic obstructive pulmonary disease (COPD) Fact Sheet. 2015. Available from: http://www.who.int/mediacentre/factsheets/fs315/en/. [Accessed: 13 February 2017].
[2] Soriano JB, Visick GT, Muellerova H, et al. Patterns of comorbidities in newly diagnosed COPD and asthma in primary care. Chest 2005;128(4):2099-107.
[3] Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global strategy for the diagnosis, management, and prevention of Chronic Obstructive Pulmonary Disease, updated 2017. Available from: http://goldcopd.org/gold-2017-global-strategy-diagnosis-management-prevention-copd/.
[4] Soler-Cataluña JJ, Martínez-García MA, Román Sánchez P, et al. Severe acute exacerbations and mortality in patients with chronic obstructive pulmonary disease. Thorax. 2005;60(11):925–31.
[5] Wilkinson TMA, Aris E, Bourne S, et al. AERIS Study Group. A prospective, observational cohort study of the seasonal dynamics of airway pathogens in the aetiology of exacerbations in COPD. Thorax. 2017;72(10):919-27.
[6] Lim S, Lam DC-L, Muttalif AR, et al. Impact of chronic obstructive pulmonary disease (COPD) in the Asia-Pacific region: the EPIC Asia population-based survey. Asia Pac Fam Med 2015;14(1):4.
[7] Anthonisen NR, Manfreda J, Warren CP, et al. Antibiotic therapy in exacerbations of chronic obstructive pulmonary disease. Ann Intern Med 1987;106:196-204.
[8] Liang KY, Zeger SL. Longitudinal Data Analysis Using Generalized Linear Models. Biometrika 1986;73:13–22.

CHALLENGES IN TEACHING CLINICAL TRIALS: THE EXPERIENCE OF TEACHING BIOSTATISTICS IN ONLINE POST-GRADUATE ACADEMIC COURSES TARGETED TO INDUSTRY BIOMETRICIANS

Cavaliere Laura[1], Perissinotto Egle[1], Baldi Ileana[1], Barbetta Beatrice[2], Gregori Dario[1]

1. *Unit of Biostatistics, Epidemiology and Public Health, Department of Cardiac, Thoracic and Vascular Sciences, University of Padova, Padova, Italy*
2. *Department of Clinical Research, Rottapharm Biotech, Monza, Italy*

Introduction

At today, new technologies and widespread web access are moving the traditional face-to-face teaching towards online and open digital teaching processes. New technologies provide widespread flexible and convenient learning tools overcoming any temporal and geographical constrains [**1**]. The instances of innovation gradually enhanced higher education institutions towards new models of e-learning [**2**]. The challenge is furtherly hard for teaching Biostatistics to health professionals and medical students, who are in general very little motivated to learn statistics in the traditional academic courses [**3**].

Material and methods

In the academic year 2015/2016, first in the teaching experience at the University of Padova (Italy), two fully online post-graduate courses of Biostatistics were established: the first one taught basic biostatistics and research methodology; the second one, advanced topics in biostatistics. In both cases, the course was organized into two phases: modules dealing with structured topics (25 weeks on the whole, each unit from 2 to 5 weeks) and a project-work (20 weeks). To perform statistical analysis, the R software was adopted. The student-centered model [**2**] was used instead of the traditional teacher-centered model. Each student independently manages her/his access to the web pages contents, without limits in number of accesses and time, within the timetable. She/he attends the didactic activities individually, preforming self-test and participating in the discussion online.

From the teaching point of view, the face-to-face lessons' components of contents, interaction and assessment were translated into digital and online contents and tools; from the logistic point of view, the publication time of the work tools was planned, the quality and accessibility of platform was conveniently tested and the students' access to the platform was monitored. The computerized portal for education was based on MOODLE platform. To realize the fully online courses, different tools were built: streaming videos (from 10 to 40 minutes) with in person teachers' explanation, slides highlighting central concepts, self-tests with multiple-choice questions and simulation-based tests with unlimited access, and homework with supervisor. Moodle platform allowed students for documents download and upload.

At the end of each module a questionnaire regarding the assessment of teaching has been reported, the questions were proposed in a Likert scale from 1 to 10.

Forum, chat, email and any other way not in presence were applied to contrast viewpoints and opinions and to summarize concepts in asynchronous discussion boards.

Results

Twelve students in the basic course and thirteen in the advanced course were able to access to web pages on Moodle platform. Most of the students were workers. Median number of individual accesses to Moodle platform was 458 for the basic course and 931 for the advanced one.

They mostly appreciated: self-administration of hour, number and time of access and stimulating discussion board online. They reported as worst limits: inadequate preliminary knowledge to understand new

biostatistics concepts (median score 6.3 for the basic course and 6.6 for the advanced one), too elevated study load in the advanced course (median score 6.7), time-spending searching concepts in videos during study and reviewing topics due to streaming modality.
The achievement of expected learning outcomes was tested by means of an in-person final examination.

Conclusions

Our experience supports feasibility and efficacy of online distance learning in teaching biostatistics. The experience suggests elaborating the following tools: videos length shorter than 20 min [***4***], lists of main concept and definition indicating the position (minute) in videos, widespread operative examples, timely matching of concepts and examples.

References

[**1**] Reis LO, Ikari O, Taha-Neto K A, et al. Delivery of a urology online course using moodle versus didactic lectures methods. Int J Med Inform 2015;84:149-54.
[**2**] Boyadzhieva E, Learner-centered Teaching and Learner Autonomy, Procedia – Social and Behavioral Sciences 2016;232:35-40.
[**3**] Sahai H, Ojeda M M, Problems and challenges of teaching biostatistics to medical students and professionals. Medical Teacher 1999;21(3):286-8.
[**4**] Syed M R. Strategic applications of distance learning technologies. IGI Global. (Ed.). (2008).

CREAZIONE DI UNA MACRO SAS PER EFFETTUARE META-ANALISI CON STIME INDIPENDENTI E DIPENDENTI

Soranna Davide[1,2], Zambon Antonella[1,2]

1. *Dipartimento di Statistica e Metodi Quantitativi, Università degli Studi di Milano – Bicocca, Milano, Italia.*
2. *Istituto Auxologico Italiano*

Introduzione

La meta-analisi di studi sperimentali combina i risultati indipendenti di più studi che hanno indagato l'effetto di uno specifico intervento. In molti casi però può essere necessario combinare risultati correlati tra loro. La correlazione può derivare dal fatto che: i) le stime di associazione considerano lo stesso gruppo di pazienti come gruppo di riferimento (es. nelle stime delle relazioni dose-risposta); ii) si studiano più eventi contemporaneamente (alcuni dei quali possono essere annidati tra loro); iii) più modalità di misurazione sugli stessi soggetti. La meta-regressione multivariata può essere utilizzata per modellare le correlazioni entro-studio migliorando così la precisione dell'analisi rispetto all' approccio univariato.
Lo scopo dello studio è di creare una macro SAS, facile da utilizzare e flessibile, che consenta di modellare i risultati di diversi studi al fine di ottenere sintesi meta-analitiche, mediante l'approccio a effetti fissi e causali, sia quando le stime sono indipendenti sia quando sono correlate. Questa macro è stata applicata ad una meta-analisi su studi clinici controllati e randomizzati condotta per stimare la riduzione della pressione arteriosa (R-PA) misurata sugli stessi pazienti con due tecniche differenti (Office e ABPM).

Metodi

Date N osservazioni ed m-1 covariate, il modello lineare misto applicato alla meta-analisi è il seguente: $y_i=X_i*\beta+Z_i*b_i+\varepsilon i$ dove: y_i è un vettore ni*1 delle stime d'effetto , X_i è una matrice ni*m di caratteristiche di interesse, β è un vettore degli effetti fissi m*1, ε_i è un vettore ni*1, b_i è un vettore degli effetti casuali k*1. Si assume che $b_i \sim N(0, G)$ e che $\varepsilon_i \sim N(0, R)$ con b_i indipendenti da ε_i. La matrice G rappresenta la matrice di varianze e covarianze a livello dello studio mentre la matrice R rappresenta la matrice delle varianze e covarianze entro studio. Quando si utilizza il modello misto per realizzare una meta-analisi, la matrice R dovrebbe riportare sulla sua diagonale principale le varianze campionarie osservate delle stime degli effetti (yi) in modo tale da pesare ciascuna stima per l'inverso della sua variabilità campionaria. La matrice G invece andrà utilizzata per stimare alcune componenti della variabilità legata a eventuali effetti casuali inseriti nel modello (ad es. ogni singolo studio incluso) [***1, 2***]. Nel caso in cui si volesse effettuare una meta-analisi a stime indipendenti con effetti fissi la matrice Z sarebbe assente mentre la matrice R andrebbe costruita considerando solo valori sulla diagonale principale. Nel caso in cui si volesse effettuare una meta-analisi ad effetti casuali con stime indipendenti allora la matrice R andrebbe costruita come precedentemente descritto mentre l'effetto casuale assunto per la variabilità tra studi verrebbe stimato dai dati assumendo una forma VC (variance components) per la matrice G a blocchi (blocco=studio). Nel caso in cui più stime siano presenti in uno studio allora si potrà tenere conto della correlazione tra le stime considerando queste covarianze nella matrice R.

Risultati

La macro in SAS consente i) di creare, da un generico dataset di partenza, la matrice di varianze e covarianze da utilizzare nel modello ad effetti misti; ii) di modellare le stime meta-analitiche scegliendo tra il modello ad effetti fissi e casuali sia in presenza che in assenza di correlazione tra stime; iii) di costruire il forest plot relativo al modello scelto. La macro SAS è stata applicata in una meta-analisi che aveva lo scopo

di stimare la riduzione della pressione arteriosa misurata sugli stessi pazienti, esposti ad un farmaco antiipertensivo, con due tecniche differenti Office e ABPM. La meta-analisi comprendeva 52 studi clinici controllati randomizzati in cui era misurata la pressione arteriosa sugli stessi pazienti con entrambe le tecniche. Inizialmente sono stati individuati 2415 articoli di cui 1800 sono stati esclusi perché non erano inerenti all'oggetto dello studio, 218 non erano studi randomizzati o i pazienti non erano trattati con antiipertensivi, 332 studi non riportavano le misurazioni per entrambe le tecniche ed infine 13 articoli non analizzavano dati originali. Nell'esempio esplicativo introdotto, si osserva che, per le stime meta-analitiche con dati indipendenti, la macro fornisce gli stessi risultati di quelli prodotti dalla libreria RMETA di R e dal comando MVMETA di STATA. Per quanto riguarda le stime di sintesi con dati dipendenti (R-PA con le tecniche Office e ABPM sugli stessi soggetti) si osserva una stima di sintesi di -17.87 (-19.07, 16.68) per la tecnica Office e -11.95 (-13.13, -10.77) per ABPM con la macro; di -17.95 (-19.25, -16.66) e di -11.88 (-12.84, -10.91) con STATA.

Conclusioni

Questa macro SAS consente, partendo da database dinamici e flessibili che possono essere costruiti in base agli obiettivi dello studio, di ottenere stime meta-analitiche con effetti fissi o casuali in presenza e in assenza di correlazione tra le stime incluse.

Bibliografia

[**1**] van Houwelingen HC, Arends LR, Stijnen T. Advanced methods in meta-analysis: multivariate approach and meta-regression. Stat Med 2002;28;21(4):589-624.

[**2**] Wei Y, Higgins JP. Estimating within-study covariances in multivariate meta-analysis with multiple outcomes. Stat Med 2013;32(7):1191-205

SAMPLE SIZE CALCULATION IN ANIMALS RESEARCH.

Tresoldi Claudio[1], Biganzoli Elia Mario[1], Milani Silvano[1], Cesana Bruno Mario[2].

1. *Università degli Studi di Milano, Dipartimento di Scienze Cliniche e di Comunità, Sezione di Statistica Medica e Biometria G.A. Maccacaro*
2. *Università degli Studi di Brescia, Dipartimento di Medicina Molecolare e Traslazionale, Sezione di Statistica Medica*

Introduzione

Limitiamo la nostra rassegna storica riguardante le raccomandazioni e le regole sull'uso degli animali nella ricerca biologica e biomedica al testo di Russell e Burch "The Principles of Humane Experimental Technique", pubblicato nel 1959 [***1***] e recentemente citato da Balls, nel 2010 [***2***], in cui sono state definite le Tre R per una ricerca con animali da esperimento che sia definibile "umana": Reduction, Refinement, and Replacement. Inoltre, vorremmo ricordare la "Guide to the Care and Use of Experimental Animals" (1993) [***3***], la "Guide for the Care and Use of Laboratory Animal" del 1996 [***4***], pubblicata per la prima volta nel 1963 sotto il titolo "Guide for Laboratory Animal Facilities and Care".

In tutti questi documenti è stata affrontata la problematica di un comportamento incurante del dolore e della sensibilità animale, ma è solo nella sezione dedicata alla revisione del protocollo (Protocol Review) dell'ottava edizione della "Guida for the Care and Use of Laboratory Animal" che sono stati presi in considerazione gli aspetti scientifici della ricerca. In particolare si è sottolineato che il numero di animali e le dimensioni dei gruppi sperimentali dovrebbero essere statisticamente giustificati; inoltre, sono stati riportati alcuni utili riferimenti statistici nell'Appendice A, nel paragrafo Disegno Sperimentale e Statistiche (Experimental Design and Statistics).

Secondo Allgoewer e Mayer (2017) [***5***], il calcolo della dimensione campionaria è obbligatorio quando si pianificano esperimenti su animali, anche se di solito è difficilmente effettuabile in modo appropriato. Conveniamo che tale calcolo non sia una procedura facile, a causa della necessità di stimare accuratamente i pertinenti parametri (differenza tra medie e variabilità, utilmente considerate assieme nella dimensione dell'effetto: "effect size"). Tuttavia riteniamo che il calcolo debba essere fatto per ragioni etiche e scientifiche. Infatti, le sperimentazioni con potenza insufficiente (dove l'animale è sacrificato e delle risorse sono utilizzate senza un'adeguata probabilità di ottenere risultati scientifici) o quelle sovradimensionate (dove troppi animali sono sacrificati inutilmente e troppe risorse sono sprecate) non sono scientificamente appropriate e, di conseguenza, non etiche. Inoltre, la presentazione adeguata del calcolo della dimensione del campione nel paragrafo relativo ai Metodi statistici del lavoro scientifico consente di comprendere meglio le analisi statistiche effettuate sull'obiettivo primario dello studio, rendendo anche più facile valutare la validità dello studio.

A seguito di un'esperienza (BMC) come biostatistico in un comitato etico per la ricerca sugli animali (ora: Organismo Preposto per il Benessere Animale), abbiamo deciso di svolgere un'indagine sulla modalità di riportare i metodi statistici (in particolare il calcolo della dimensione del campione) nei lavori sulla ricerca con animali di laboratorio pubblicati su riviste farmacologiche con elevato "impact factor", con l'obiettivo di confrontare i nostri risultati con quelli di Kilkenny et al. (2009) [***6***], riassunti in Significance (2010) [***7***] e di verificare l'impatto di quanto riportato da Charan e Kantharia [***8***]: "We want to suggest researchers to include a statement about method of calculation of sample size and justification of sample size in the manuscript they want to publish."

Lo scopo di questo lavoro è quello di riportare dei risultati preliminari su un campione casuale di articoli pubblicati sul Journal of Pharmacology and Experimental Therapeutic, dal vol. 358, numero 1, del luglio 2016 al vol. 359, numero 3, del dicembre 2016.

Materiali e metodi

Gli articoli sono stati recuperati in formato pdf da uno degli Autori (CT) e poi sono stati letti in accordo ad una griglia preparata dagli AA. In particolare, di ciascun articolo sono stati considerati i seguenti aspetti: (1) disegno dello studio; (2) scopo dello studio; (3) end-points; (4) cecità dei valutatori; (5) specie animale oggetto della sperimentazione; (6) numero dei gruppi e, se più di uno, (7) modalità di assegnazione dei trattamenti (randomizzazione o no); (8) numero degli animali in ciascun gruppo e numerosità totale, considerando anche (9) se era stata specificata nel paragrafo relativo all'analisi statistica o nel testo dell'articolo o, invece, solo nelle legende di grafici o tabelle; (10) numero di rilevazioni della/e variabile/i sperimentale/i (misure ripetute nel tempo); (11) calcolo della numerosità campionaria riportando: livello di significatività (errore di I tipo), potenza (complemento a 1 dell'errore di II tipo), "effect size" (entità della differenza divisa per la variabilità del fenomeno in studio) e test di significatività statistica; (12) analisi statistica riportata (sua coerenza con il modello sperimentale adottato). Sono stati inoltre considerati i seguenti due aspetti: il numero di esperimenti riportato in ciascun articolo e l'osservanza di un protocollo per la cura degli animali da laboratorio. Abbiamo estratto a caso 20 articoli dalla lista dei 79 articoli selezionati sulla base dei seguenti criteri. Presenza nel titolo dell'articolo o nell'abstract dei termini esperimento / sperimentale unitamente al riferimento di una specie animale con la successiva conferma nell'abstract che si trattava di uno studio sperimentale sull'animale da laboratorio. Gli articoli sono stati rassegnati da due AA (CT e BMC) e in caso di disaccordo il parere dirimente è stato dato da SM o EMB.

Risultati

Le specie utilizzate nelle sperimentazioni sono: topo (14 su 20; frequenza relativa – f.r. – 0.70, I.C. 0.457 - 0.881), ratto (4/20; f.r. = 0.20, I.C. 0.057 - 0.437). In 2 articoli sono riportate sperimentazioni su due specie: ratto e maiale (1/20; f.r. = 0.05; I.C. 0.001 - 0.249) in uno e gatto e uomo (1/20; f.r. = 0.05; I.C. 0.001 - 0.249) nell'altro. Il disegno dello studio sperimentale è controllato in 20/20 (f.r. = 1.00; I.C. 0.832 - 1.000), di cui 7 studi sono randomizzati (f.r. = 0.35; I.C. 0.154 - 0.592) e 13 non randomizzati (f.r. = 0.65; I.C. 0.408 - 0.846). Dei 7 studi randomizzati, uno è stato effettuato secondo un disegno entro soggetti a quadrato latino e un altro è in accordo a un disegno a blocchi incompleti e misure ripetute.

Dei 13 studi non randomizzati, 3 presentano la randomizzazione per alcuni aspetti procedurali; in particolare, uno ha solo la sequenza di training randomizzata e alcune associazioni di sostanze "according to a pseudo-Latin square design", uno ha la sequenza delle dosi randomizzata, uno riporta che gli animali "were pseudorandomized on each testing day".

La dimensione campionaria totale è dichiarata o calcolabile esattamente (considerando quanto riportato per la numerosità di ciascun gruppo) in 12 articoli su 20 (f.r. = 0.6; I.C. 0.361 - 0.809) e varia da 7 a 552 (mediana = 134.5) animali per esperimento. La numerosità dei gruppi è dichiarata solo per una parte dei gruppi in 5 casi su 20 (f.r. = 0.25; I.C. 0.087 - 0.491), come numero minimo e massimo per gruppo in 7 (f.r. = 0.35; I.C. 0.154 - 0.592) e con entrambe le suddette modalità in 2 (f.r. = 0.10; I.C. 0.012 - 0.317). La numerosità varia da 1 a 50 soggetti per gruppo con una mediana di 6 soggetti per gruppo. Le caratteristiche dei gruppi sono illustrate in ***Tabella 1***.

Il numero dei gruppi confrontati per articolo, compreso il controllo, presenta un minimo di 1, un massimo di 88 e una mediana di 11, come riportato in ***Tabella 1***, unitamente alla numerosità per gruppo e totale per articolo.

Rilevante è il fatto che la numerosità campionaria non sia giustificata in nessuno studio e, quindi, non vengano indicati tutti gli elementi necessari per il suo calcolo: il livello di significatività (tra l'altro riportato

espressamente in 16 articoli su 20: f.r. = 0.80, I.C. 0.563 - 0.943), la potenza, l'effect size (differenza biologicamente rilevante divisa per la variabilità del fenomeno indagato, costituita dalla deviazione standard), il test statistico da utilizzare per l'analisi. In conclusione, non c'è alcuna aderenza con quanto richiesto dalla linea guida ARRIVE (2010) [***9***] al paragrafo 10 relativo al "Sample size" specificamente ai punti 10.a: "Specify the total number of animals used in each experiment and the number of animals in each experimental group.", 10.b. "Explain how the number of animals was decided. Provide details of any sample size calculation used." e 10.c. "Indicate the number of independent replications of each experiment". Come ultimo punto, si deve sottolineare che l'appropriatezza dell'analisi statistica risulta discutibile in 3 lavori su 20 (f.r. = 0.15; I.C. 0.032 - 0.379): in particolare, in 1 caso su 17 (f.r. = 0.059, I.C. 0.001 - 0.287) ai confronti multipli effettuati mediante ANOVA non segue la dichiarazione dell'utilizzo di un post-hoc test; inoltre in 2 casi su 20 (f.r. = 0.10; I.C. 0.012 - 0.317) vengono effettuati confronti multipli (test di Mann–Whitney U e test t di Student) senza correzione del livello di significatività.

Tabella 1. Caratteristiche dei gruppi inclusi nello studio

Numero totale di gruppi	frequenza assoluta	Numerosità per gruppo	Numerosità totale per articolo
1	1	7	7
4	2	3 - 25	16 - 175*
6	1	6 - 20	120
7	1	3 - 10	49
8	1	4 - 5	40
9	1	3 - 50	156
10	3	4 - 24	40 - 149
12	1	6	72
14	1	1 - 13	77
15	1	4 - 6	78
20	1	6 - 10	176
38	1	3 - 8	342
50	1	8 - 17	475-552
54	1	3 - 11	346 - 364
57	1	3 - 9	321-479
59	1	6	342
88	1	2 - 8	368-484

* > 175, per quanto riportato.

Discussione

I risultati ottenuti, sia pure preliminari, confermano quanto era atteso sulla scarsa o praticamente nulla attenzione nel riportare accuratamente la numerosità campionaria e i metodi statistici utilizzati nel contesto della ricerca farmacologica effettuata sull'animale. È evidente che la presenza di un "Editor" biostatistico costituirebbe un primo provvedimento, assolutamente irrinunciabile per cercare di portare i ricercatori ad un doveroso livello di sensibilità nei confronti dell'impiego e della descrizione dei metodi statistici adottati nell'analisi dei dati rilevati nelle loro ricerche. Infine, stante la totale interazione tra il metodo scientifico e quello statistico, si ritiene anche che il metodo scientifico sia non adeguatamente perseguito nella pianificazione, conduzione e descrizione delle ricerche; forse, può essere riconosciuto, ma senz'altro confinato alla mera applicazione delle metodiche di ricerca nel contesto della sperimentazione in laboratorio e su animali da laboratorio. Gli Autori si augurano, tra l'altro, che l'utilizzo degli animali, in particolare di mammiferi, nell'ambito della sperimentazione preclinica si riduca fino quasi a scomparire grazie all'utilizzo

di tecniche alternative tra cui anche i cosiddetti studi clinici di fase 0 a dosi minimali nell'uomo. Nell'attesa che questo auspicabile miglioramento si avveri, la valutazione della dimensione campionaria negli studi preclinici sugli animali appare di fondamentale importanza etica e, quindi, scientifica e, pertanto, deve essere effettuata e appropriatamente riportata.

Bibliografia

[1] Russell W MS, Burch RL. The Principles of Humane Experimental Technique, special edition. Potters Bar: Universities Federation for Animal Welfare. 1959.
[2] Balls M. The Principles of Humane Experimental Technique: Timeless Insights and Unheeded Warnings Altex 27, Special Issue 2010.
[3] Guide to the Care and Use of Experimental Animals. Volume 1 1993. Edited by: Ernest D Olfert, DVM; Brenda M Cross, DVM; A Ann McWilliam. Canadian Council on Animal Care (CCAC).
[4] Guide for the Care and Use of Laboratory Animal. National Research Council (US) Institute for Laboratory Animal Research. Washington (DC): National Academies Press (US); 1996. Eighth Edition.
[5] Allgoewer A, Mayer B. Sample size estimation for pilot animal experiments by using a Markov Chain Monte Carlo approach. Alternatives to Laboratory Animals 2017;45(2):83-90.
[6] Kilkenny C, Parsons N, Kadyszewski E, et al. Survey of the quality of experimental design, statistical analysis and reporting of research using animals. PLoS One 2009;4:e7824.
[7] Festing MFW. Statistics and animals in biomedical research. Significance 2010;7(4):176-7.
[8] Charan J. and Kantharia N. D. How to calculate sample size in animal studies? Journal of Pharmacology and Pharmacotherapeutics 2013 ; Vol 4: Issue 4, 303-6.
[9] Kilkenny C, Browne WJ, Cuthill IC, et al. Improving Bioscience Research Reporting: The ARRIVE Guidelines for Reporting Animal Research. PLoS Biol 2010;8(6): e1000412.

RETROSPECTIVE STUDY ON ITALIAN ADPKD DISEASE MANAGEMENT COST

Lanati Elena Paola[1], Iorio Arianna[1], Marcelli Alessandro[1], Ruggeri Matteo[2], Romana Rolli Francesca[2], Caramuscio Raffaele[3], Ballantini Fabrizio[3], Scolari Francesco[4]

1. *MA Provider S.r.l., Milan, Italy*
2. *School of Health Economics and Management - Università Cattolica del Sacro Cuore, Rome, Italy*
3. *Otsuka Pharmaceutical Italy, Milan, Italy*
4. *U.O di Nefrologia e Dialisi, Ospedale di Montichiari - AO Spedali Civili di Brescia*

Introduction

The ADPKD (Autosomal Dominant Polycystic Kidney Disease) is the most common genetic kidney disease. It is associated to development of end-stage kidney disease requiring dialysis or transplantation and it is caused by PKD1 and PKD2 genes mutations. Recent studies have demonstrated that patients with ADPKD have a worsening in quality of life and a burden of disease similar to cancer patients [***1, 2***].
The aim of this study was to estimate the burden of ADPKD in Italy, analyzing only direct costs and evaluating the healthcare costs according to progression stage of chronic kidney disease (CKD). The primary outcome was the average annual cost per patient with ADPKD in Italy. The secondary outcome was represented by the average annual cost per patient with ADPKD for CKD1, CKD2, CKD3, CKD4, CKD5 (not in dialysis), dialysis and post-transplant stage.

Methods

This retrospective, observational study was carried out by gathering data through a CRF (Case Report Form) in six hospitals in Italy. We estimated costs associated with polycystic kidney disease (ADPKD). On the basis of identified cost drivers, the analysis has been performed using Activity Based Costing method. Resource consumptions were collected for each patient based on the outpatients and/or hospital admission notes during the period 2012-2015. The cost of the direct resources has been calculated from the perspective of the Italian National Health Service (NHS). Data have been collected on outpatient visits, laboratory biochemical and genetic tests, diagnostic and therapeutic procedures, drug treatments, hospitalization in ordinary or day hospital (DH) regimen attributable to the chronic kidney disease and comorbidities-related. The national tariffs defined by the list of charges for medical and outpatient services, as updated by 18 October 2012 Ministerial Decree, are used to estimate the cost of outpatient specialist visits and diagnostic tests. Costs of hospitalizations and DH are calculated according to the Diagnosis Related Group (DRG) system, currently in use, applied as a proxy of the costs for kidney transplant and secondary diagnoses. Drug therapy has been valorised through retail or ex-factory price of each drug, according to reimbursement class.

Results

A sample of 191 patients, with a mean age of 52.5 years, was considered in this evaluation. The analysis estimated a mean total cost associated with ADPKD management equal to € 7,921. Approximately 40% of management costs of the whole sample is related to patients in dialysis, followed by CKD V and post-transplant (about 20% each) patients, CKD IV (13.5%) and finally CKD I, CKD II, CKD III (about 1% each). Costs increased with the disease progression, except for post-transplant stage. This stage presented a lower cost compared to dialysis, as post-transplant patients generally do not perform dialysis and six patients did not perform transplant during the period 2012-15, and therefore the transplantation costs have not been

allocated (cod. DRG 302: € 33,162). The outpatient specialist care (which includes dialysis) showed the highest impact on total costs, followed by drug therapies and hospitalizations.

Conclusions

The study has underlined the relevant burden of Autosomal Dominant Polycystic Kidney Disease, especially in the end-stage, and implicitly the importance of slowing the disease progression, both from patient and NHS perspectives.

References

[**1**] Saini T, Murtagh FE, Dupont PJ, McKinnon PM, Hatfield P, Saunders Y. Comparative pilot study of symptoms and quality of life in cancer patients and patients with end stage renal disease. Palliat Med 2006;20(6):631-6.

[**2**] Torres VE, Harris PC, Pirson Y. Autosomal dominant polycystic kidney disease. Lancet 2007;369(9569):1287-301.

OTTIMIZZAZIONE DELLA QUOTA DI DROP-OUT NELLE SPERIMENTAZIONI CLINICHE: UN APPROCCIO INNOVATIVO IN UN TRIAL ONCOLOGICO

Maglietta Giuseppe[12], Malagoli Tagliazucchi Guidantonio[1], Musolino Antonino[3], Maglietta Gianluca[4], Caminiti Caterina[1]

1. *UO Ricerca e Innovazione, Azienda Ospedaliero-Universitaria di Parma*
2. *Dipartimento di Statistica, Informatica, Applicazioni- Università degli Studi di Firenze*
3. *UO Oncologia Medica, Azienda Ospedaliero - Universitaria di Parma*
4. *Dipartimento di Farmacia - Università degli Studi di Salerno-Fisciano*

Introduzione

Nell'ambito delle sperimentazioni, la determinazione statistica della numerosità campionaria rappresenta una fase cruciale nella pianificazione di uno studio [***1***]. Infatti, tale numerosità, è indispensabile per la valutazione dell'end-point primario a prestabiliti livelli di significatività (α) e di potenza statistica (1-β).
La numerosità campionaria ha inoltre importanti implicazioni etiche ed è determinante nella valutazione della feasibility economico/organizzativa di un progetto di ricerca [***2***]. La sua rilevanza etica è determinata dall' indicare quanti pazienti potranno beneficiare, ed essere al tempo stesso esposti a rischi, non tutti noti, derivanti dalla terapia sperimentale. In merito alle ricadute economiche, un sottodimensionamento potrebbe inficiare i risultati dello studio rendendolo statisticamente non valido, mentre, un sovradimensionamento comporterebbe uno spreco delle risorse economiche [***3***].
Nel determinare la numerosità di soggetti da arruolare in una sperimentazione, oltre al tener conto delle importanti implicazioni su menzionate, occorre considerare il fenomeno del drop-out [***4***], ovvero l'uscita prematura dei pazienti dallo studio o la loro non conformità nella valutazione dell'end-point primario [***5***]. La metodologia classica suggerisce di specificare nei protocolli scientifici, una stima della quota di drop-out che si presume osservare nel corso della sperimentazione. Tale quota, dovrà essere infine integrata alla numerosità statisticamente determinata [***6***].
In realtà, in alcune tipologie di studi la quota di dati mancanti potrebbe essere ingente, specialmente in quelli sperimentali con trattamenti ad alti livelli di drop-out legati alla tossicità, oppure, in quelli dove sono necessarie misurazioni laboratoristiche o di bio-marcatori [***7***]. I trial oncologici sono un esempio di studi con le sopracitate caratteristiche e per i quali, rispetto alle altre aree terapeutiche principali, è registrata la percentuale più alta di fallimento [***8***]. Inoltre in tale ambito, gli alti costi delle terapie e la necessità di un'attività di farmacovigilanza rendono il drop-out estremamente oneroso [***9***].
D'altra parte, negli studi di superiorità di un trattamento, il ricercatore è ragionevolmente motivato a proporre al maggior numero di pazienti la nuova terapia ed è quindi intenzionato nel definire quote di drop-out più elevate di quelle che ci si aspetterebbe di osservare.
Diviene così importante individuare una soglia di drop-out che bilanci la perdita d'informazioni necessarie alla valutazione dell'end-point primario dello studio ma che allo stesso tempo non aumenti drasticamente i costi della sperimentazione.
Nonostante la numerosità aggiunta legata alla quota di drop-out argini in qualche modo tale criticità, questa non è sempre sufficiente al raggiungimento della numerosità statisticamente determinata, comportando una riduzione della potenza statistica oppure la richiesta di proroghe dello studio.
L'obiettivo è confrontare un nuovo approccio per la gestione di drop-out, basato sui fondamenti del calcolo delle probabilità, con l'approccio proposto dalla metodologia classica.
In particolare si intende:

1) dimostrare i limiti della semplice integrazione campionaria mediante applicazione di una presunta percentuale di drop-out
2) valutare il livello di ottimizzazione del processo decisionale del nuovo approccio e le performance economiche ad esso associato.

Metodi

Nell'ambito della sperimentazione oncologica di fase IIb a due bracci non comparativi "Immun-Her" la dimensione campionaria $\hat{n}$ è stata determinata mediante Disegno Ottimale di Simon a due stadi.
Definita la quota di drop-out δ che si assume osservare durante la sperimentazione, è stata calcolata la numerosità integrata $\widehat{n^*}=(\hat{n} * (1+\delta))$così come suggerito dalla metodologia classica.
Sotto assunzione di distribuzione Binomiale, attraverso la funzione di massa di probabilità $f(x) = f(x;n,p) = \binom{n}{x} p^x(1-p)^{n-x}$, è stata calcolata la probabilità di raggiungere esattamente la numerosità stimata $\hat{n}$ dovendo arruolare $\widehat{n^*}$pazienti.
Attraverso la funzione di probabilità cumulata $F(X)$è stata determinata la probabilità di riuscire a valutare almeno $\hat{n}$ pazienti dovendone arruolare $\widehat{n^*}$.
Graficamente sarà illustrato l'aumento probabilistico di successo (inteso come raggiungimento della numerosità statisticamente determinata $\hat{n}$) indotto dalla osservazione di un paziente in più oltre quelli $\widehat{n^*}$ richiesti.
Il calcolo del campione è stato ottenuto mediante software PASS v.15, mentre le analisi e le simulazioni sono state condotte mediante il software statistico R-v.2.15.0.

Risultati

Attraverso il Disegno Ottimale di Simon a due stadi, per entrambi i gruppi di trattamento, risulta necessaria una numerosità campionaria di 15 soggetti. Essendo il drop-out supposto esser pari al 50% [10]; $\widehat{\mathbf{n}^*}$ = 60 pazienti dovranno essere arruolati in totale. Tuttavia, la probabilità di osservare esattamente $\hat{\boldsymbol{n}}$ = 30 soggetti valutabili avendone arruolati 60 è pari al 10%, mentre di quella di osservarne almeno 30 è il 55%. Incrementando $\widehat{\mathbf{n}^*}$ di un ulteriore 5% ($\widehat{\mathbf{n}^{**}}$=63), la probabilità di osservare almeno 30 soggetti valutabili per l'end-point primario aumenta di 14 punti percentuali (dal 55% al 69%).
Oltre a tale incremento, nel grafico di figura 1 viene mostrato lo scostamento della curva binomiale (in blu) da quella ad incremento costante di probabilità di successo (in rosso). La divergenza diviene accentuata dal quarto paziente in più oltre il 60-esimo.
Nella presente sperimentazione, i costi per paziente legati all'operazione chirurgica (mastectomia e ricostruzione), alla terapia chemio-immunoterapica neoadiuvante e quella adiuvante ammontano a 85003,00 USD, di questi $\frac{1}{3}$ sono della sola terapia sperimentale neo adiuvante che eccedono rispetto alla standard di 11335,00 USD.
Se a questo valore economico si sommano i costi sperimentali delle valutazione bioptiche (1000,00 USD) e sierologiche (1500,00 USD) di laboratorio, l'aggiunta di questi tre pazienti si misura anche come ulteriore finanziamento nel pubblico da parte del privato pari a 41505,00 USD (valore economico al netto dei costi di farmacovigilanza).

Conclusioni

Attraverso i risultati ottenuti, è possibile constatare come l'integrazione di soli tre pazienti assicuri alla azienda farmaceutica di ridurre sensibilmente (dal 45% al 31%) la possibilità di incorrere nel mancato raggiungimento della numerosità campionaria utile alla valutazione oggettiva della performance del trattamento. L'ottenimento di questo risultato ha un forte impatto in termini economici, infatti potrebbe

consentire di risparmiare costi e tempi relativi ad una eventuale proroga. Inoltre, per gli enti pubblici si traduce in un guadagno pari a 41505,00 USD inteso come investimento da fondi privati, legittimati da un processo metodologico.
Un ulteriore vantaggio offerto dal nostro approccio consiste nella sua duttilità ai campi di applicazione diversi da quello oncologico.

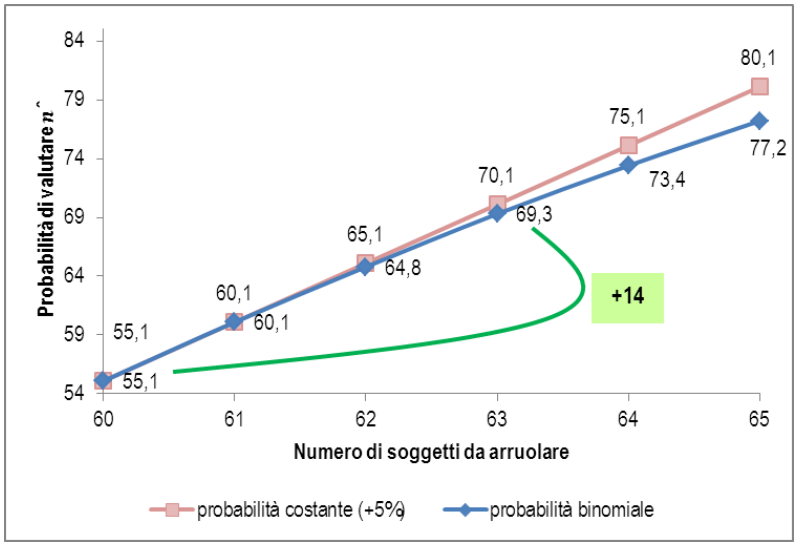

Figura 1. Probabilità di valutare n in funzione del numero di soggetti da arruolare

Bibliografia

[1] Charles P, Giraudeau B, Dechartres A, et al. Reporting of sample size calculation in randomized controlled trials: review. BMJ, 2009; 338:b1732.
[2] Lenth RV. Some practical guidelines for effective sample size determination. The American Statistician, 2001; Vol. 55, No. 3.
[3] Ioannidis JPA, Greenland S. Hlatky M, et al, Increasing value and reducing waste in research design, conduct, and analysis. Lancet, 2014; 383:166–75.
[4] Shih WJ. Problems in dealing with missing data and informative censoring in clinical trials. Current Controlled Trials in Cardiovascular Medicine, 2002;3:4.
[5] Bacchieri A, Della Cioppa G. Fondamenti di ricerca clinica. Springer, 2004.
[6] Little RJ, D'Agostino R, Cohen ML, et al. The prevention and treatment of missing data in clinical trials. The New England Journal of Medicine, 2012;367:14
[7] Ibrahim JG, Chu H. Chen M., Missing data in clinical studies: Issues and methods. Journal of clinical oncology, 2012;30:3297-3303
[8] Harrison RK. Phase II and phase III failures, Nature Reviews Drug Discovery, 2016; Vol 15: 817:818
[9] Sertkaya A, Wong H-H, Jessup A, Belech T. Key cost drivers c
United States. Clinical Trials,2016;13(2):117–26.
[10] Biomarker Study of Immune-mediated Mechanism of Action of Neoadjuvant Trastuzumab in Patients With HER2-positive Breast Cancer (IMMUN-HER). Clinical Trial.gov 2017;NCT03144947

INDICE PER AUTORE