

EBPH – 2026, Volume 21, Supplementary Issue 1

<https://riviste.unimi.it/index.php/ebph/index>

Volume 21

Supplementary Issue 1

SISMEC Congress 2026

24-25 September 2026

Parma, Italy

Book of Abstracts

© 2026 Milano University Press

SISMEC Congress 2026

24-25 September 2026
Parma, Italy



Book of Abstracts

Vol. 21 Suppl. 1



This Book of Abstracts is a publication by
Epidemiology Biostatistics and Public Health (EBPH),
formerly Italian Journal of Public Health,

published in Diamond Open Access with CC BY-SA license

by Milano University Press.

DOI: 10.54103/ /2282-0930/sismec14

© 2026, The Authors

ISSN: 2282-0930

Contacts: ebph@unimi.it

<https://riviste.unimi.it/index.php/ebph>

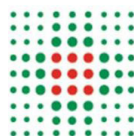
Guest editors of the volume were Simona Villani, Università degli Studi di
Pavia, Giovanni Veronesi, Università degli Studi dell'Insubria, Giuseppe
Maglietta, Azienda Ospedaliero-Universitaria di Parma, under the supervi-
sion of the Editor-in-Chief, who retained the final decision on the publication.



Congress sponsored by



**UNIVERSITÀ
DI PARMA**



**SERVIZIO SANITARIO REGIONALE
EMILIA-ROMAGNA**
Azienda Unità Sanitaria Locale di Parma
Azienda Ospedaliero - Universitaria di Parma

Table of Contents

EBPH & SISMEC COMMITTEES	4
SISMEC CONGRESS 2026	5
USEFUL INFORMATION	6
CONGRESS PROGRAM	8
ABSTRACTS	10
GENERAL EPIDEMIOLOGY	10
BIostatistical METHODS	102

EBPH Committee

EDITORS IN CHIEF

Carlo La Vecchia, Università degli Studi di Milano, Milan, Italy

Leonardo Villani, Università Cattolica del Sacro Cuore, Rome, Italy

MANAGING EDITOR

Rossella Bonzi, Università degli Studi di Milano, Milan, Italy

FOUNDING EDITORS

Giovanni Corrao, Università di Milano-Bicocca, Milan, Italy

Walter Ricciardi, National Institute of Health, Rome, Italy

EDITORIAL BOARD

Stefania Boccia, Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy

Paolo Boffetta, Stony Brook University, Stony Brook, NY, USA

Flavia Carle, Università Politecnica delle Marche, Ancona, Italy

John P.A. Ioannidis, Stanford University, Stanford, CA, USA

Martin Mc Kee, London School of Hygiene and Tropical Medicine, London, UK

Cristina Montomoli, Università degli Studi di Pavia, Pavia, Italy

Luigi Palmieri, Istituto Superiore di Sanità, Rome, Italy

Walter Ricciardi, National Institute of Health, Rome, Italy

SISMEC

BOARD OF DIRECTORS

Simona Villani

Giovanni Veronesi

Paola Borrelli

Simona Signoriello

Alessandro Marcon

Giuseppe Maglietta

Valentina Panetta

Lorenza Scotti

ORGANIZING COMMITTEE

Giuseppe Maglietta - President

Matteo Puntoni

Caterina Caminiti

Stefano Bendoni

Francesca Diodati

Denisa Balan

UOC Ricerca Clinica ed Epidemiologica

Azienda Ospedaliero Universitaria di Parma

SISMEC Congress 2026

PATIENT-REPORTED OUTCOME MEASURES (PROMs) IN THE AGE OF ARTIFICIAL INTELLIGENCE

The Role and Challenges of the Biostatistician in Elevating the Patient's Voice

24-25 SEPTEMBER 2026 – PARMA

NH PARMA

Viale Paolo Borsellino, 31 - 43122 Parma

Artificial Intelligence is revolutionizing how we evaluate clinical outcomes and quality of care.

In this scenario, Patient-Reported Outcome Measures (PROMs) represent a valuable and still partially unexplored resource. The congress brings together Italian experts to explore the potential of AI applied to patient-centric data, focusing on predictive methodologies, integration into care pathways, and ethical and clinical perspectives.

The intersection of PROMs and AI offers significant opportunities, but it also poses methodological, ethical, and practical challenges that require a critical and rigorous contribution from biostatisticians. The program pays special attention to data quality, model validation, reproducibility, transferability into clinical practice, and national and international research projects.



Useful Information

CONGRESS VENUE

NH Parma, Viale Paolo Borsellino 31 - 43122 Parma

CONTACTS

congresso2026@sismec.info

HOW TO GET TO THE CONGRESS VENUE

The hotel is located within the train station complex, just a short walk from the historic center.

Parma is easily reachable via:

A1 Milan–Bologna motorway (Parma Centro exit)

A15 Parma–La Spezia motorway

From the ring road, follow the signs for Stazione FS (train station).

The hotel has a paid parking available.

ORGANIZATIONAL OFFICE

MI&T srl - Provider n. 1509

Valeria Cavaliere

Centro Direzionale Bolomnia, Via Guelfa 5 - 40138 Bologna

Cell. 331 8088131 - Tel. 051 220427

valeria.cavaliere@mitcongressi.it

CONGRESS WEBSITE

www.sismec.info/congresso2026/

Public transport

The train station is the main hub for urban transport.

Several bus lines connect the station with all areas of the city; the stop is just a few meters from the hotel.

On foot

Parma's historic center (Cathedral, Baptistery, Pilot-ta) is about a 10-minute walk away.

ONLINE PROGRAM AND UPDATES



NH Parma - External view



Congress Program

THURSDAY 24 SEPTEMBER

WORKSHOPS, MEETINGS, AND PLENARY SESSIONS

09:30-12:30	WORKSHOP SISMEC - Collaborare per imparare: il Team-Based Learning nella didattica della statistica medica Free event reserved for the first 30 registered participants <i>Lecturer: Prof. Marina di Carro - University of Genoa</i>
13:00-13:30	Assembly of the Board of Full Professors of Medical Statistics
13:30-14:30	Assembly of the Board of Professors and Researchers in Medical Statistics
13:45-14:00	Congress Participant Registration
13:30-15:00	Coffee Station
14:30-15:30	SISMEC Members' Assembly
15:30-16:00	Opening of the Conference Opening Remarks – Welcome Message – Greetings from the Authorities <i>Simona Villani, Giuseppe Maglietta</i>

PLENARY SESSION

DALLA VOCE DEL PAZIENTE ALLA PREDIZIONE: PROMs E AI TRA EVIDENZE E LIMITI

Curators: Danila Azzolina and Antonino Musolino

16:00-16:25	Dai dati del paziente alla decisione clinica PROMs e AI nella pratica reale <i>Massimo Di Maio - Presidente AIOM</i>
16:25-16:50	I Discrete Choice Experiments dalla teoria alla pratica clinica Misurare le preferenze di pazienti e clinici per ottimizzare l'alleanza terapeutica <i>Patrizio Pasqualetti - Università La Sapienza</i>
16:50-17:15	L'AI può prevedere la voce del paziente? <i>Evidenze, limiti e prospettive</i> <i>Fabio Efficace - GIMEMA</i>
17:15-17:40	Dai PROMs alle decisioni Integrare preferenze e modelli decisionali <i>Paola Berchialla - Università Torino</i>
17:40-18:00	General discussion
20:00	Social dinner

FRIDAY 25 SEPTEMBER

PROJECTS, APPLICATIONS, AND SCIENTIFIC CONTRIBUTIONS

09:00-09:30

Assembly of Full Professors

PLENARY SESSION

DAI MODELLI AI SISTEMI REALI: ESPERIENZE DA PROGETTI EUROPEI E INTERNAZIONALI SU PROMs E AI

Curators: Matteo Puntoni and Elena Giovanna Bignami

09:00-09:25	Dai dati alle decisioni Piattaforme integrate e AI per una medicina patient-centered <i>Dario Gregori - Università di Padova</i>
09:25-09:50	Intelligenza artificiale e salute della donna Integrare PROMs per una medicina più equa e personalizzata <i>Giorgia Garganese and Stefania Boccia - Policlinico Gemelli</i>
09:50-10:15	Dalla raccolta dei PROMs ai sistemi intelligenti Architetture per l'integrazione e l'analisi <i>Giuseppe Fico - Universidad Politécnica de Madrid</i>
10:15-10:40	PROMs e intelligenza artificiale Global PROMs Initiative per una ricerca centrata sulla persona <i>Giampaolo Brichetto - AISM</i>
10:40-11:00	General discussion
11:00-12:45	Parallel Sessions Oral & Short Presentations
12:45-13:45	Lunch Break
13:45-15:30	Parallel Sessions Oral & Short Presentations
15:30-16:00	Best Scientific Contribution Award Closing Remarks & Conference Conclusion

GENERAL EPIDEMIOLOGY

Municipal-Scale Mapping of Health-Related Ecosystem Services in Italy

Sofia Tagliaferro^(1,2), Mauro Masiero⁽³⁾, Pierpaolo Marchetti⁽¹⁾, Luisa Fernanda Eusse Villa⁽³⁾, Francesca Locatelli⁽¹⁾, Nicola Caranci⁽⁴⁾, Alessandro Marcon⁽¹⁾

 Tagliaferro Sofia 0000-0002-9602-970X, Marcon Alessandro 0000-0002-2778-658X

(1) Department of Diagnostics and Public Health, University of Verona, Verona, Italy

(2) Institute of Clinical Physiology, National Research Council, Pisa, Italy

(3) Department of Land, Environment, Agriculture and Forestry, University of Padua, Padua, Italy

(4) Health and Social Systems Research and Innovation Sector, Directorate General for Personal Care, Health and Welfare, Bologna, Italy

CORRESPONDING AUTHOR: Alessandro Marcon, Department of Diagnostics and Public Health, University of Verona, Verona, Italy.

E-Mail: alessandro.marcon@univr.it

INTRODUCTION

Ecosystem services (ES) are the benefits people derive from ecosystems and are increasingly recognised as important determinants of human health and well-being. ES flow represents the services actually delivered and used, while demand reflects societal needs. Spatial information on ES is essential to support land management and prioritise interventions where services are most needed.

OBJECTIVES

We mapped municipal-scale flow and demand for health-related ES in Italy, analysed geographic patterns, and developed a dataset to support targeted interventions.

METHODS

We selected eight health-related regulating ES, i.e., regulation of local climate, air quality, waste, erosion, water flow, pest and disease control, natural hazard protection, and water purification. We used the matrices from Burkhard et al. (2014) [1] to assign expert-based estimated levels of ES flow and demand to 44 Land Cover classes across Italy using the 2018 CORINE dataset. We assessed ES distribution at the municipal scale using two approaches: 1) percentage of municipal areas with high ES flow (or demand) based on 2021 census tract boundaries; 2) a population-weighted measure

using the 2023 Global Human Settlement (GHS) dataset, estimating for each municipality the average percentage of areas within 300 m of GHS grid centroids with high ES flow (or demand).

RESULTS

Nationwide, the percentage of municipal areas with high ES flow was generally low, with median values ranging from 19% (air quality regulation) to 47% (pest and disease control), decreasing to 3–27%, respectively, when weighted for the population. The lowest flows occurred in lowland, coastal, and urban areas. In contrast, ES demand was high across most municipalities, with a median of 67–68% of municipal areas showing high demand (31% for water purification), rising to 94% (78% for water purification) when weighted for the population. High-demand municipalities were predominantly in urban areas and lowlands.

CONCLUSIONS

This study provides the first national-scale mapping of health-related ES flow and demand in Italy. Results highlight low ES availability and high demand in most municipalities, particularly in urban and lowland areas, underscoring the need for planning strategies to improve ecosystem services and population well-being.

Extrapancreatic Cancer Aggregation in Familial Pancreatic Cancer: A Registry-Based Analysis Integrating IRFARPC and AIRTUM Data

Erica Secchettin^(1,2), Silvia Carrara⁽³⁾, Maria Terrin⁽³⁾, Arianna Dal Buono⁽³⁾, Massimo Falconi⁽⁴⁾, Livia Archibugi⁽⁵⁾, Raffaele De Luca⁽⁶⁾, Valentina Arcangeli⁽⁷⁾, Salvatore Paiella^(1,2)

 Erica Secchettin 0000-0001-7924-5577

(1) Department of Surgery, Dentistry, Paediatrics and Gynaecology, University of Verona, Verona, Italy

(2) University of Verona, Verona, Italy

(3) Department of Endoscopy, Humanitas Research Hospital, IRCCS, Rozzano 20089, Milano, Italy

(4) Vita-Salute San Raffaele University, Milan, Italy and Pancreas Translational and Clinical Research Center, IRCCS San Raffaele Scientific Institute, Milan, Italy

(5) Pancreatico-Biliary Endoscopy and Endosonography Division, Pancreas Translational and Clinical Research Center, IRCCS San Raffaele Scientific Institute, Milan, Italy

(6) Department of Surgical Oncology Istituto di Ricovero e Cura a Carattere Scientifico (IRCCS) Istituto Tumori "Giovanni Paolo II," Bari, Italy

(7) Romagna Cancer Registry IRCCS Istituto Romagnolo per lo Studio dei Tumori "Dino Amadori" (IRST), Meldola, Forlì-Cesena, Italy

CORRESPONDING AUTHOR: Erica Secchettin, Department of Surgery, Dentistry, Paediatrics and Gynaecology, University of Verona, Verona, Italy. E-Mail: secchettin.e@gmail.com

INTRODUCTION

Familial pancreatic cancer (PC) accounts for approximately 5–10% of all PC cases and is frequently associated with aggregation of extrapancreatic malignancies (EMs) within affected families. Nevertheless, the epidemiologic profile of EMs in familial PC kindreds, including clustering patterns, parental lineage distribution, and age at onset, remains incompletely characterized in large registry-based familial cohorts.

OBJECTIVES

To investigate the spectrum, clustering patterns, parental lineage distribution, and early-onset occurrence of extrapancreatic malignancies in families enrolled in the Italian Registry of Families at Risk for Pancreatic Cancer (IRFARPC), and to contextualize age-at-onset patterns against population-based cancer registry estimates (AIRTUM).

METHODS

We conducted a registry-based analysis of baseline pedigree data from the IRFARPC (2019–2025), including 1,228 index subjects with familial PC and 3,765 unique relatives

(1,846 first-degree and 1,618 second-degree relatives). EMs were evaluated at the family level, and clustering was assessed according to the number of distinct EM sites within each pedigree. Lineage-specific analyses were restricted to second-degree relatives with documented maternal or paternal attribution. Maternal-to-paternal (M/P) ratios were calculated for each EM site to evaluate lineage asymmetry. Age-at-diagnosis distributions and proportions of early-onset cancers (<50 years) were descriptively compared with age- and sex-specific estimates from Italian population-based cancer registries (AIRTUM 2013–2017).

RESULTS

Overall, 47.9% of index subjects reported at least one EM within their family history, while 20.5% showed clustering of ≥ 2 distinct extrapancreatic cancer sites. Breast ($n=306$; 19.3% of families), colorectal ($n=178$; 11.8%), lung ($n=124$; 9.0%), and prostate cancer ($n=105$; 7.7%) were the most frequently reported EMs. Breast cancer represented the most common EM in both maternal (5.0%) and paternal (4.7%) lineages and showed a near-symmetric distribution (M/P ratio: 1.07), suggesting a possible autosomal susceptibility pattern. In contrast, lung cancer showed a higher representation along paternal lineages (M/P ratio: 0.61), whereas uterine–ovari-

an cancers (M/P ratio: 2.12) and hematologic malignancies (M/P ratio: 2.38) showed relative maternal enrichment. PC remained predominantly late onset, with 3.6% of cases diagnosed before age 50, consistent with AIRTUM estimates (3.9%). Conversely, early-onset occurrence was descriptively higher than population-based estimates for uterine–ovarian cancers (+3.8 percentage points vs. AIRTUM) and hematologic malignancies (+2.8 percentage points).

CONCLUSIONS

Familial PC is associated with a heterogeneous spectrum of EMs characterized by differential clustering across cancer sites and parental lineages. The balanced maternal–paternal distribution observed for breast cancer supports an autosomal susceptibility pattern, whereas the maternal predominance of uterine–ovarian and hematologic malignancies may be compatible with inherited predisposition pathways, including BRCA1/2-related mechanisms. The paternal predominance observed for lung cancer may also be influenced by shared environmental or behavioral exposures.

Association between Green Exercise and Adherence to the Mediterranean Diet in Adolescents: Preliminary Results from a Multicenter Cross-Sectional Study

Alice Mannocci⁽¹⁾, Alice Fognani⁽¹⁾, Andrea Cattaneo⁽¹⁾, Angelo Pascariello⁽¹⁾, Valentina Merolle⁽²⁾, Daniele Masala⁽²⁾, Barbara Picconi⁽¹⁾

 Alice Mannocci 0000-0002-8527-6940

(1) Department for the Promotion of Human Sciences and Quality of Life, Telematic University "San Raffaele", Rome, Italy

(2) Italian Committee for Sport Against Drugs (CISCoD), Italy

CORRESPONDING AUTHOR: Alice Mannocci, Department for the Promotion of Human Sciences and Quality of Life, Telematic University "San Raffaele", Rome, Italy. E-Mail: alice.mannocci@uniroma5.it

INTRODUCTION

Adolescence represents a critical period for the adoption of healthy behaviors that may influence health throughout the life course. Among these, regular physical activity and adherence to the Mediterranean Diet (MD) are recognized as important determinants of health. In recent years, growing attention has been paid to green exercise (GE), defined as physical activity performed in natural environments such as urban parks, beaches, mountains, rural areas, or sports facilities characterized by a high presence of green spaces, due to its potential benefits for physical and psychological well-being.

OBJECTIVES

To assess the association between physical activity, participation in GE, and adherence to the MD among adolescents.

METHODS

A multicenter cross-sectional study was conducted in 10 secondary schools located across five Italian regions (Piedmont, Lombardy, Tuscany, Lazio, and Apulia). Sociodemographic, anthropometric, and lifestyle-related information was collected through a self-administered online questionnaire. Physical activity was defined as engaging in moderate-to-vigorous physical activity for at least 50 minutes, at

least three times per week. GE was defined as participation in physical activity in natural environments at least once per week. Adherence to the MD was assessed using the KIDMED questionnaire and classified as low, medium, or high adherence. Associations between variables were evaluated using the χ^2 test, with statistical significance set at $p < 0.05$. The study was reported according to the STROBE Statement recommendations for observational studies. Statistical analyses were performed using SPSS software. Ethical acknowledgment was obtained from the CET Lazio Area 5 Ethics Committee (No. 122/SR/24).

RESULTS

A total of 1,714 adolescents were included, with a mean age of 16.1 years ($SD = 1.7$) and a mean BMI of 21.2 kg/m^2 ($SD = 3.2$). Preliminary analyses showed a statistically significant association between GE participation (at least once per week) and MD adherence levels ($p < 0.001$), with a higher prevalence of high MD adherence among GE participants compared with non-participants (30.7% vs. 18.9%).

A significant association was also observed between engaging in moderate-to-vigorous physical activity at least three times per week and MD adherence ($p < 0.001$). Specifically, the prevalence of high MD adherence was greater among physically active adolescents than among their inactive peers (24.8% vs. 11.3%).

No significant associations were found between GE participation and sex ($p = 0.175$) or BMI categories ($p = 0.053$).

Similarly, no significant associations were observed between MD adherence levels and sex ($p=0.644$) or BMI categories ($p=0.578$).

CONCLUSIONS

Preliminary findings suggest that regular physical activity and participation in activities performed in natural environments are associated with greater adherence to the Mediterranean Diet among adolescents. These findings support the importance of integrated health promotion strategies aimed at simultaneously encouraging physical activity, contact with natural environments, and healthy dietary behaviors during adolescence

Association of Vitamin C Supplementation and Genetic Susceptibility with Multiple Sclerosis Risk: A Prospective Population-Based Cohort Study

Teresa Fazio⁽¹⁾, Andrea Nova⁽¹⁾, Giovanni Di Caprio⁽¹⁾, Alice Cinquepalmi⁽¹⁾, Simone Perna⁽²⁾,
Mariangela Rondanelli⁽³⁾, Luisa Bernardinelli⁽¹⁾

 Teresa Fazio 0000-0002-9577-8450

(1) Department of Brain and Behavioral Sciences, University of Pavia, 27100 Pavia, Italy

(2) Division of Human Nutrition, Department of Food, Environmental and Nutritional Sciences (DeFENS), University of Milan, Milan, Italy

(3) Department of Public Health, Experimental and Forensic Medicine, University of Pavia, Pavia, Italy

CORRESPONDING AUTHOR: Teresa Fazio, Department of Brain and Behavioral Sciences, University of Pavia, 27100 Pavia, Italy. E-mail: teresa.fazio@unipv.it

INTRODUCTION

Multiple sclerosis (MS) is a chronic immune-mediated neurological disorder of the central nervous system and a leading cause of neurological disability. Vitamin C, due to its antioxidant and neuroprotective properties, has been hypothesized to reduce MS risk, but epidemiological evidence remains limited and inconsistent. Furthermore, it is unclear whether its association varies according to an individual's genetic susceptibility to MS.

OBJECTIVES

This study aimed to evaluate the association between vitamin C supplementation and MS risk, and to assess whether this relationship is modified by the MS polygenic risk score (MS-PRS) for MS.

METHODS

To this aim we conducted a prospective cohort study using UK Biobank data, including 486,908 adults aged 37-70 years free of neurological disease at baseline. Vitamin C supplementation was self-reported at recruitment. Incident MS cases were identified during a median follow-up of 13.5 years. Propensity score-based Inverse Probability of Treatment Weighting (IPTW) was used to balance a wide range of demographic, lifestyle, dietary, health-related, and supplement-use confounders. Weighted Cox proportional hazards

models were used to estimate hazard ratios (HRs). Furthermore, we examined whether the association differed according to genetic susceptibility to MS, measured using MS-PRS.

RESULTS

During follow-up, 452 participants were diagnosed with MS. Vitamin C supplementation was reported by 8.8% of participants and was associated with a 49% lower risk of MS (HR=0.51, [95% CI: 0.32; 0.82], $p=0.004$). The estimate was robust across multiple sensitivity analyses, including a negative control outcome analysis. The association was stronger among individuals with higher MS-PRS (HR=0.31, [95% CI: 0.17; 0.55], $p<0.001$), whereas no significant association was observed among those with lower MS-PRS. No interaction was observed by sex or age, although estimates suggested a potentially stronger association in younger individuals.

CONCLUSIONS

In this large prospective cohort, vitamin C supplementation was associated with a lower risk of incident MS, particularly among individuals with higher genetic susceptibility. Although residual confounding cannot be excluded, the consistency of the association in this large prospective cohort and its biological plausibility support further investigation of vitamin C as a candidate modifiable factor for MS prevention.

Rapid Sepsis Detection Using Routine Laboratory Data: A Logistic Regression Approach

Annalisa Belli⁽¹⁾, Chiara Della Franca⁽²⁾, Daniela Ligi⁽²⁾, Francesca Salvatori⁽²⁾, Marisol Palomino Huaman⁽²⁾, Marc Vasse⁽³⁾, Rocchi Marco B. L.⁽¹⁾, Davide Sisti⁽¹⁾, Ferdinando Mannello⁽²⁾

 Annalisa Belli 0009-0008-8561-6088

(1) Department of Biomolecular Sciences, Section of Biostatistics, University of Urbino Carlo Bo, Urbino, Italy

(2) Department of Biomolecular Sciences, Section of Biochemistry and Biotechnology, Unit of Clinical Biochemistry, University of Urbino Carlo Bo, Urbino, Italy

(3) Clinical Biology Department, Hôpital Foch, Suresnes, France

CORRESPONDING AUTHOR: Annalisa Belli, Department of Biomolecular Sciences, Section of Biostatistics, University of Urbino Carlo Bo, Urbino, Italy. E-Mail: a.belli7@campus.uniurb.it

INTRODUCTION

Sepsis is a life-threatening condition associated with high morbidity and mortality, in which early diagnosis is crucial for improving clinical outcomes. Routine haematological parameters generated by automated analyzers have emerged as promising biomarkers for sepsis detection due to their rapid availability and low cost.

OBJECTIVES

The principal aim was to investigate the predictive value of routine hematochemical parameters for sepsis diagnosis and to construct a validated multivariable logistic regression model. Particular attention was given to variable selection, model validation, and cut-off optimization through the weighted Youden index to maximize sensitivity while maintaining high specificity.

METHODS

A total of 597 subjects admitted to the Emergency Department were included and the sample was randomly divided into training ($n = 427$; 71.5%) and validation datasets ($n = 170$; 28.5%). Dataset homogeneity was assessed through independent samples t-tests and chi-square tests with effect size estimation (Cohen's d and Cramer's V). Multicollinearity was evaluated using variance inflation factors ($VIF > 10$ threshold) and univariate logistic regression analyses were initially performed to identify candidate predictors associated with sepsis, with variable preselection based on statistical significance and

sensitivity > 0.70 . A forward stepwise multivariable logistic regression approach was subsequently applied (entry $p = 0.05$; removal $p = 0.10$). Model fit was evaluated using Nagelkerke's R^2 and Akaike Information Criterion (AIC). Discriminative ability was assessed using ROC curves, area under the curve (AUC), sensitivity, specificity, and weighted Youden index (Jw). Since minimizing false negatives represented the main clinical priority, multiple probability thresholds (0.25–0.75) were tested to optimize classification performance.

RESULTS

Most hematochemical parameters differed significantly between sepsis and non-sepsis groups. Nine variables demonstrated strong discriminatory performance in univariate analyses (sensitivity > 0.70). The final multivariable model included ten predictors and demonstrated excellent goodness-of-fit (Nagelkerke's $R^2 = 0.947$; AIC = 84.689). At the conventional threshold (0.50), sensitivity and specificity reached 97.12% and 97.26% in the training dataset, and 92.96% and 96.67% in validation. Threshold optimization identified 0.31 as the optimal cut-off (Jw = 0.9796), increasing sensitivity to 99.04% while maintaining specificity at 96.35%, substantially reducing false negatives.

CONCLUSIONS

The proposed statistical framework enabled the development of a highly discriminative and robust predictive model for sepsis detection. The prioritization of sensitivity-driven threshold optimization improved clinical applicability by minimizing missed diagnoses while preserving high specificity.

Organization Determinants of Patient-Reported Experience of Integrated Care in Multiple Sclerosis: A Multilevel Analysis of Italian MS Clinical Centres

Michela Ponzio⁽¹⁾, Pasquale Paletta⁽¹⁾, Marco Salivetto⁽¹⁾, Tommaso Manacorda⁽²⁾

 Michela Ponzio 0000-0001-5245-0474

(1) Public Health Research Area, Italian Multiple Sclerosis Foundation (FISM), Genoa, Italy

(2) Observatory on the Rights of People with Multiple Sclerosis and Related Disorders, Italian Multiple Sclerosis Society (AISM), Genoa, Italy

CORRESPONDING AUTHOR: Michela Ponzio, Public Health Research Area, Italian Multiple Sclerosis Foundation (FISM), Genoa, Italy.
E-Mail: michela.ponzio@unipv.it

INTRODUCTION

People with multiple sclerosis (MS) require coordinated care across multiple professionals, with potential risks of fragmentation and discontinuity. Within the activities of the Observatory on MS Policy at AISM, the Italian MS Foundation conducted two national surveys among people with MS (PwMS) and MS Clinical Centres (MSCCs). Patient experience of integrated care was measured using the Integrated Care Experience Scale for Multiple Sclerosis (ICES-MS), a validated MS-specific patient-reported experience measure. Linking patient- and MSCC-level data enables assessment of how MS centre characteristics relate to patients' experience of integrated care.

OBJECTIVES

To assess the association between organizational characteristics of MSCCs and patient-reported experience of integrated care, measured using the ICES-MS, while accounting for individual patient characteristics and within-centre correlation.

METHODS

Data were derived from two national cross-sectional surveys conducted by the Italian Multiple Sclerosis Foundation in 2026: one among Italian MS Clinical Centres and one among people with MS. The primary outcome was the total ICES-MS score, obtained by summing 8 items rated on a 0–3 Likert scale and rescaled to 0–100, with higher scores indicating integrated and person-centred care. Multilevel

linear regression models with a centre-specific random intercept were fitted to account for patients nested within centres. Models were adjusted for centre-level and individual patient characteristics. The intraclass correlation coefficient (ICC) was estimated to quantify between-centre variability.

RESULTS

Overall, 966 people with MS from 135 MSCCs were included. In the null model, the ICC was 0.046, indicating that 4.6% of ICES-MS variability was attributable to the centre level. The fully adjusted model was statistically significant (Wald $\chi^2(20)=127.75$, $p<0.001$). Compared with the reference area (South and Islands), ICES-MS scores were higher in PwMS treated by MSCC in Central Italy ($\beta=8.77$; 95% CI: 4.52 to 13.02), North-East Italy ($\beta=9.66$; 95% CI: 4.84 to 14.49), and North-West Italy ($\beta=9.80$; 95% CI: 5.40 to 14.21). Those treated in medium-sized centres showed higher scores than small centres ($\beta=8.60$; 95% CI: 3.28 to 13.93), whereas the association for large centres was not statistically significant ($\beta=4.72$; 95% CI: -0.91 to 10.34). PwMS treated in centres providing care based on a local integrated care pathway had higher scores than other centres ($\beta=4.16$; 95% CI: 0.10 to 8.22). Critical waiting times reported by clinicians were associated with lower scores ($\beta=-5.11$; 95% CI: -9.75 to -0.48). Increasing EDSS levels were associated with lower ICES-MS scores: $\beta = -5.46$ for the moderate group (95% CI: -9.11 to -1.82, $p = 0.003$) and $\beta = -8.51$ for the high group (95% CI: -13.87 to -3.15, $p = 0.002$). In the adjusted model, the ICC decreased to approximately 0.012.

CONCLUSIONS

Patient-reported experience of integrated and person-centred care showed limited but detectable variability across MS Clinical Centres. Higher ICES-MS scores were associated with medium-sized centres, a local integrated care pathway, and follow-up at centres located in Northern or Central Italy, whereas critical waiting times and higher disability levels were associated with lower scores. These findings suggest that both organizational/contextual characteristics of MS Clinical Centres and individual patient factors may influence patients' experience of integrated care.

Acceptance, Perception, and Technostress of Generative AI among Radiologists at the University Hospital “Paolo Giaccone” of Palermo

Domenica Matranga⁽¹⁾, Giovanni Di Stefano⁽²⁾, Manuela Lodico⁽¹⁾, Dario Monzani⁽²⁾, Valeria Seidita⁽³⁾, Roberto Cannella⁽⁴⁾, Laura Maniscalco⁽¹⁾, Silvana Miceli⁽²⁾, Sergio Salerno⁽⁴⁾

 Domenica Matranga 0000-0002-2466-937X

(1) Dept of Health Promotion, Mother and Child Care, Internal Medicine and Medical Specialties, University of Palermo, Italy

(2) Dept of Psychology, Educational Science and Human Movement, University of Palermo, Italy

(3) Dept of Engineering, University of Palermo

(4) Dept of Biomedicine, Neuroscience and Advanced Diagnostics, University of Palermo

CORRESPONDING AUTHOR: Domenica Matranga, Dept of Health Promotion, Mother and Child Care, Internal Medicine and Medical Specialties, University of Palermo, Italy. E-Mail: domenica.matranga@unipa.it

INTRODUCTION

The widespread adoption of generative artificial intelligence (AI) tools across multiple areas of medical practice, together with their predictive, diagnostic, and therapeutic applications, has raised growing interest in understanding healthcare professionals' acceptance, use, and perceptions of this technology. At the same time, evidence suggests that AI implementation, particularly in the absence of adequate competencies, may contribute to technology-related stress (technostress).

OBJECTIVES

This study aimed to investigate the mechanism through which AI perceptions influence technostress through the mediating role of AI acceptance among radiologists, a professional group particularly exposed to AI applications in clinical practice. In addition, the study examined whether self-efficacy, defined as individuals' beliefs in their capabilities to execute behaviors necessary to achieve specific performance outcomes (Bandura, 1982), moderates the relationship between AI perceptions and acceptance.

METHODS

A sample of 71 physicians working at the Department of Radiology and Diagnostic Imaging of the University Hospital “Paolo Giaccone” of Palermo completed an online ques-

tionnaire administered through Google Forms. The survey included the AI Perception Scale (Shinners et al., 2022), the Acceptance Scale (Venkatesh et al., 2012), the Self-Efficacy Scale (Venkatesh et al., 2003), and the Technostress Scale (Tarafdar et al., 2007), which assesses overall technostress, techno-overload, techno-complexity/insecurity, and techno-uncertainty.

For each construct, exploratory factor analysis (EFA) was conducted to identify the underlying factor structure, followed by confirmatory factor analysis (CFA) to evaluate the fit of the measurement model against competing alternatives. Model fit was assessed using the Chi-square/df ratio (≤ 3), Root Mean Square Error of Approximation (RMSEA ≤ 0.08), Comparative Fit Index (CFI ≥ 0.90), Tucker–Lewis Index (TLI ≥ 0.90), and Standardized Root Mean Square Residual (SRMR ≤ 0.08). Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were used to compare competing models, with lower values indicating better fit. Conditional indirect effects of Perceptions of AI on technostress outcomes through AI Acceptance, were estimated at low (-1 SD), mean, and high ($+1$ SD) levels of Self-Efficacy, using the moderated mediation model implemented through the PROCESS macro Model 7 (Hayes, 2022) with 5000 bootstrap samples and bias-corrected 95% confidence intervals.

RESULTS

AI perception was directly and positively associated with overall technostress ($b = 0.57$, 95% CI = 0.21–0.93), tech-

no-overload ($b = 0.58$, 95% CI = 0.16–1.00), and techno-complexity/insecurity ($b = 0.83$, 95% CI = 0.42–1.23), but not with techno-uncertainty ($b = -0.02$, 95% CI = -0.46–0.43). AI acceptance was directly and negatively associated with the same outcomes (e.g., overall technostress: $b = -0.55$, 95% CI = -0.91 to -0.18). Furthermore, significant indirect effects of AI perception on technostress through AI acceptance emerged at both low ($b = -0.36$, 95% CI = -0.65 to -0.11) and moderate ($b = -0.28$, 95% CI = -0.53 to -0.09) levels of self-efficacy.

CONCLUSIONS

The findings suggest that AI perceptions influence technostress both directly and indirectly through AI acceptance. Therefore, the implementation of AI as a support tool for radiologists in routine clinical practice should aim to mitigate technostress by promoting AI acceptance, strengthening professional competencies, and addressing workload demands and organizational clarity.

Temporal Trends in Antimicrobial Resistance of *Acinetobacter Baumannii*: A Mixed-Effects Analysis of Clinical Isolates Collected at Palermo University Hospital (2013–2025)

Manuela Lodico⁽¹⁾, Salvatore Antonino Distefano⁽²⁾, Laura Maniscalco⁽¹⁾, Marco Enea⁽¹⁾, Giovanni Giammanco^(1,2), Domenica Matranga⁽¹⁾, Domenico Graceffa⁽²⁾

 Manuela Lodico 0009-0000-4226-0714

(1) Dept of Health Promotion, Mother and Child Care, Internal Medicine and Medical Specialties, University of Palermo, Italy

(2) Microbiology and Virology Unit, Azienda Ospedaliera Universitaria Policlinico "Paolo Giaccone", Palermo, Italy

CORRESPONDING AUTHOR: Manuela Lodico, Dept of Health Promotion, Mother and Child Care, Internal Medicine and Medical Specialties, University of Palermo, Italy. E-Mail: manuelalodico0@gmail.com

INTRODUCTION

Acinetobacter baumannii is recognized by the World Health Organization (WHO) as one of the ESKAPE priority pathogens because of its environmental persistence and ability to acquire antimicrobial resistance. Carbapenem-resistant *A. baumannii* (CRAB) has been included in the WHO Global Priority List of Antibiotic-Resistant Bacteria since 2017 and remains a critical threat. In Europe, CRAB is responsible for a substantial burden of healthcare-associated infections and mortality, particularly among critically ill patients.

OBJECTIVES

To evaluate temporal trends in antimicrobial resistance among clinical isolates of *Acinetobacter baumannii* collected at the Microbiology Laboratory of the University Hospital "Paolo Giaccone" of Palermo over a 13-year period (2013–2025).

METHODS

A retrospective analysis was conducted on all clinical isolates of *A. baumannii* collected between 2013 and 2025 ($N = 3,010$). One isolate per patient was included. Data collected comprised year of isolation, patient demographics (age and sex), hospital ward (Intensive Care Unit/Emergency, Medical Area, Surgical Area), and biological sample type.

Antimicrobial susceptibility testing was performed against polymyxins, fluoroquinolones, aminoglycosides, and carbapenems. Isolates were classified as multidrug-resistant (MDR) when resistant to at least three antimicrobial classes.

Comparisons between MDR and non-MDR isolates were performed using chi-square or Fisher's exact tests for categorical variables and the Kruskal–Wallis test for age. To identify independent predictors of MDR isolation and evaluate temporal variability, a mixed-effects logistic regression model was applied, with isolates nested within years. Age, sex, clinical area, and sample type were included as fixed effects, while year of isolation was modelled as a random effect. Results were reported as adjusted odds ratios (ORs) with 95% confidence intervals (95% CIs).

RESULTS

Among the 3,010 isolates analyzed, 2,727 (90.6%) were classified as MDR. Patients were predominantly male (60.3%), with a mean age of 71.0 ± 15.7 years; 71.4% were aged 65 years or older. Most isolates originated from Intensive Care/Emergency Units (43.8%), followed by Medical (33.9%) and Surgical Wards (21.0%).

Resistance rates remained consistently high throughout the study period, reaching 95.8% for carbapenems, 95.1% for fluoroquinolones, and 91.0% for aminoglycosides. Polymyxin resistance remained low (1.6%). Univariable analyses did not identify significant associations between patient or clinical characteristics and MDR status. In the mixed-effects model,

age emerged as the only independent predictor of MDR isolation (OR 1.01; 95% CI 1.00–1.02; $p = 0.019$). Significant between-year variability was observed (random intercept SD = 0.404; 95% CI 0.046–0.526). Year-specific analyses showed a significantly increased probability of MDR isolation in 2020, whereas significantly lower-than-expected odds were observed in 2022 and 2023.

CONCLUSIONS

The observed temporal variability suggests that health-care-system factors, beyond individual patient characteristics, influenced MDR *A. baumannii* circulation. The increase observed in 2020 coincided with the COVID-19 pandemic, characterized by increased intensive care admissions, broad-spectrum antibiotic use, prolonged mechanical ventilation, and critical care overcrowding. Conversely, the reduction observed in 2022–2023 may reflect strengthened infection prevention and control measures and reduced antibiotic selective pressure associated with declining COVID-19-related critical care admissions.

Mapping Patient and Public Involvement in Pediatric Palliative Care: When, Where and How

Sara Pessano^(1,2), Chiara Rebuffi⁽¹⁾, Vera Morsellino⁽³⁾, Maria Luisa Cristina^(4,5), Marina Sartini^(4,5), Alessio Carbone^(4,6)

 Sara Pessano 0009-0009-6862-979X, Alessio Carbone 0009-0008-9907-3299

(1) Scientific Direction, Clinical Epidemiology Unit, IRCCS Istituto Giannina Gaslini, Genova, Italy

(2) Department of Public Health, Experimental and Forensic Medicine, Unit of Biostatistics and Clinical Epidemiology, Medical Statistics and Biometry Postgraduate School, Pavia, Italy

(3) Pediatric Pain and Palliative Care Service, IRCCS Istituto Giannina Gaslini, Genoa, Italy.

(4) Department of Health Sciences, University of Genoa, Genoa, Italy

(5) Operating Unit Hospital Hygiene, Galliera Hospital, Genoa, Italy

(6) Department of Medicine and Division of Medical Oncology, Galliera Hospital, Genoa, Italy

CORRESPONDING AUTHOR: Alessio Carbone, Department of Medicine and Division of Medical Oncology, Galliera Hospital, Genoa, Italy. E-Mail: alessio.carbone@galliera.it

INTRODUCTION

Paediatric palliative care (PPC) is a medical subspecialty that aims to improve the quality of life for children with life-limiting or life-threatening illnesses. PPC focuses on the needs of patients and families and is rooted in communication and collaboration between healthcare teams, children, and their relatives to align medical interventions with family goals and values. Given the specific features of this field, patient and public involvement (PPI) is expected to significantly improve the relevance of systematic reviews on this topic by identifying outcomes that are important to patients and their families, and by making results more accessible to stakeholders.

OBJECTIVES

Quantify and describe PPI in PPC reviews.

METHODS

A librarian performed a systematic search, with the objective of mapping PPI in PPC reviews, on PubMed, EMBASE and Cochrane databases, retrieving systematic reviews from inception to March 2026. We used a free-text search and thesaurus descriptors search (MeSH -PubMed and Emtree-Embase) to select papers following a PICO model. Then we managed the retrieved results using a specific tool. Following this preliminary data, we are conducting an umbrella review to further analyze the topic.

RESULTS

Our systematic literature search retrieved 731 references, reduced to 581 after deduplication. Three independent reviewers screened the references and, after consensus, 123 satisfied inclusion criteria; three full texts couldn't be retrieved. Nine over 120 papers benefited from PPI at various stages of the review, mostly during planning and dissemination. These reviews were all from European countries and were published in the last decade. Some of the reviews without PPI explored technical and management aspects. A big majority of them were focused on patients' and families' feelings and needs.

CONCLUSION

PPI involvement in PPC reviews is still limited in numbers and kind of contribution, with differences between countries. A bigger work on the topic is going on with the involvement of personnel of a PPC team and representatives of PPC parents' association to better quantify and describe PPI in PPC reviews, while evaluating its effects on enhancing the relevance, quality, and dissemination of research. Geographical and cultural differences in the topic will also be explored.

Epidemiology of Workplace Accidents in Liguria (Italy), 2023–2025: Incidence, Severity, and Inequalities between Italian and Foreign Workers

Sara Pessano^(1,2), Maria Luisa Cristina^(3,4), Alessio Carbone^(4,5), Chiara Dupont⁽⁴⁾, Gianluca Ottria⁽⁴⁾, Elisa Schinca⁽⁴⁾, Marina Sartini^(3,4)

 Sara Pessano 0009-0009-6862-979X

(1) Scientific Direction, Clinical Epidemiology Unit, IRCCS Giannina Gaslini, Genoa, Italy

(2) Department of Public Health, Experimental and Forensic Medicine, Unit of Biostatistics and Clinical Epidemiology, Medical Statistics and Biometry Postgraduate School, Pavia, Italy

(3) Operating Unit Hospital Hygiene, Galliera Hospital, Genoa, Italy

(4) Department of Health Sciences, University of Genoa, Genoa, Italy

(5) Department of Medicine and Division of Medical Oncology, Galliera Hospital, Genoa, Italy

CORRESPONDING AUTHOR: Sara Pessano, Scientific Direction, Clinical Epidemiology Unit, IRCCS Giannina Gaslini, Genoa, Italy.

E-Mail: sarapessano@gaslini.org

INTRODUCTION

Workplace accidents are defined as unintended, unforeseen events that occur during employment and result in physical or psychological injury, illness, or death. They are a significant source of morbidity and mortality worldwide, with major consequences for public health, the economy, and workers' quality of life.

In Italy, nearly 600,000 events were documented in 2024, resulting in over 1,000 deaths. Analysing the settings and dynamics of these events, as well as the sociodemographic characteristics of the victims, would provide a deeper understanding of the phenomenon and identify potential preventive measures.

OBJECTIVES

This study aims to provide a detailed epidemiological analysis of workplace accidents in Liguria between 2023 and 2025. It will draw up a risk profile of the victims, paying particular attention to equity and considering the potential impact of differential vulnerability on foreign workers.

METHODS

A retrospective evaluation was conducted on data from the DATIS/SINIACA registry, considering all Emergency

Care Unit (ECU) admissions in Liguria due to workplace accidents between 1 January 2023 and 31 December 2025. The victims and events were described using absolute and percentage frequencies for categorical variables and the mean and standard deviation (SD) for quantitative data. Seasonal and geographical trends were also studied. Risk ratios were calculated for specific occupational risk categories, using direct standardisation to compare different populations. The data were analysed using Stata 18 SE® software.

RESULTS

A total of 53,382 workplace accidents were recorded during the three-year study period (2023–2025). The geographical distribution showed a higher incidence of events in the province of Genoa. Accidents occur more frequently on weekdays. A seasonal pattern was observed, with a higher number of incidents during spring and early summer, consistent with previous findings, and a peak in July. Most injured workers were male (66%), and the age group most frequently affected was 45–54 years (25.1%). Foreign nationals accounted for 17.2% of all cases.

Among Italian workers, incidence rates were 11.04, 11.26, and 10.51 per 1,000 people at risk in 2023, 2024, and 2025, respectively. Age- and sex-standardized incidence rates among foreign workers, using the Italian working population as the reference population, were 17.43, 17.69, and 16.36 per 1,000 workers at risk in 2023, 2024, and

2025, respectively. Foreign workers experienced a significantly higher risk of occupational injury compared with Italian workers for all the years of surveillance (2023: IRR = 1.58; 2024: IRR 1.57; 2025: IRR 1.56).

CONCLUSIONS

This study reports on the epidemiological characteristics of workplace accidents in Liguria over a three-year period. The results reveal different risk profiles for Italian and foreign citizens in relation to these events. Unfortunately, details on workplace settings and the dynamics of the accidents were unavailable. Nevertheless, such data warrant further studies to identify the reasons for these inequalities and to determine preventative measures that could be adopted to address the problem.

Determinants and Inequalities of Interpersonal Violence: A 10-Year Analysis of Emergency Department Visits before and during Covid-19

Marina Sartini^(1,2), Alessio Carbone^(2,3), Sara Pessano^(4,5), Chiara Dupont⁽²⁾, Gianluca Ottria⁽²⁾, Elisa Schinca⁽²⁾, Arianna Orgiana⁽²⁾, Paolo Romairone⁽⁶⁾, Maria Luisa Cristina^(1,2)

 Marina Sartini 0000-0002-7127-2893

(1) Operating Unit Hospital Hygiene, Galliera Hospital, Genoa, Italy

(2) Department of Health Sciences, University of Genoa, Genoa, Italy

(3) Department of Medicine and Division of Medical Oncology, Galliera Hospital, Genoa, Italy

(4) Scientific Direction, Clinical Epidemiology Unit, IRCCS Giannina Gaslini, Genoa, Italy

(5) Department of Public Health, Experimental and Forensic Medicine, Unit of Biostatistics and Clinical Epidemiology, Medical Statistics and Biometry Postgraduate School, Pavia, Italy

(6) Liguria Digitale – Servizi per la Salute – Regione Liguria, Italy

CORRESPONDING AUTHOR: Marina Sartini, Department of Health Sciences, University of Genoa, Genoa, Italy. E-Mail: sartini@unige.it

INTRODUCTION

Violence is a complex phenomenon with significant physical and psychological consequences and multifactorial determinants, requiring a multidisciplinary approach. Epidemiological data is often fragmented and underestimated. According to the World Health Organization, there are around 475,000 homicides and hundreds of millions of non-fatal violent incidents occur each year, resulting in substantial long-term consequences.

OBJECTIVE

This study aimed to quantify violence events in Liguria and to describe their relationship with sociodemographic characteristics of people involved. Routinely collected healthcare data can inform surveillance and prevention strategies. Temporal trends were analyzed with a focus on the effect of the pandemic on the phenomenon.

METHODS

A retrospective study was conducted, collecting data on all emergency department (ED) visits for violence and self-harm recorded in a northern Italian region between January 2016 and September 2025 from the regional healthcare data warehouse. Descriptive analyses were performed to charac-

terize temporal and sociodemographic patterns. An Interrupted Time Series based on a Poisson model was used to assess changes associated with the lockdown measures implemented during the pandemic.

RESULTS

During the study period, 5,541,712 ED visits were recorded, of which 54,321 (0.98%) were related to violence. Rates varied across provinces. Foreign nationals accounted for approximately one-third of visits, with an increasing trend over time, suggesting possible inequalities in risk exposure. Visits were more frequent among males, during evening and nighttime hours, on weekends, and in the summer months, consistent with contexts of increased social interaction.

The ITSA model demonstrated no significant pre-intervention trends both for Italian (IRR 0.999; $p = 0.605$) and foreign residents (IRR 0.997; $p = 0.191$), indicating stable incidence prior to the lockdown. In a similar way, there was no evidence of an immediate level change at the onset of COVID pandemic restrictions, while a marked and statistically significant reduction in incidence was observed during the lockdown period in both groups, with an approximate 59% and 57% decrease in admissions respectively for Italians and foreign residents. The post-COVID trend highlighted differences among the two population groups, as Italians' trajectory of violence-related ED admissions did not vary (IRR 1.002; $p = 0.488$), while a significant change in the post-intervention trend was ob-

served for foreign residents (IRR 1.007; $p = 0.027$), indicating a modest but statistically significant increase in incidence over time following the lockdown in this group.

CONCLUSIONS

This study provides a reliable, population-based picture of violence-related ED visits in a Northern Italian region over the last decade. It identifies population groups at greater risk -particularly young adult males and foreign nationals- as well as temporal and geographical patterns, suggesting clear targets for preventive and supportive interventions. The differential impact of the pandemic on interpersonal violence, and the diverging post-lockdown trends between Italian and foreign residents, highlight the role of social contexts and structural inequalities in shaping violence-related outcomes. These findings underscore the epidemiological relevance of violence as a public health issue and call for evidence-based policies, targeted prevention strategies, and the strengthening of surveillance systems within healthcare and social services.

Controlled Mechanical Ventilation and Indoor Air Quality in Classrooms: Evidence from a Quasi-Experimental Study

Francesco Barci⁽¹⁾, Pierpaolo Marchetti⁽¹⁾, Clara Peretti⁽²⁾, Gianmaria Fulici⁽³⁾, Luca Verdi⁽³⁾, Alessandro Marcon⁽¹⁾

(1) Department of Diagnostics and Public Health, University of Verona, Verona, Italy

(2) Consultant, Air Analysis and Radioprotection Laboratory, Autonomous Province of Bolzano, Bolzano, Italy

(3) Air Analysis and Radioprotection Laboratory, Autonomous Province of Bolzano, Bolzano, Italy

CORRESPONDING AUTHOR: Francesco Barci, Department of Diagnostics and Public Health, University of Verona, Verona, Italy. E-Mail: francesco.barci@univr.it

INTRODUCTION

Indoor air quality control in school environments is a relevant public health issue. Effective ventilation contributes to reducing health risks associated with pollutants and infectious aerosols. Controlled mechanical ventilation (CMV) systems, designed for air exchange in confined spaces, may be a strategy to improve microclimatic parameters.

OBJECTIVES

To evaluate the effect of CMV systems specifically designed for air exchange on indoor carbon dioxide (CO₂) concentrations (as an indirect indicator of ventilation and system effectiveness) and indoor air temperature (a parameter related to comfort and energy efficiency) in classrooms of a primary school in Bolzano.

METHODS

In the necessARIA project, coordinated by the Autonomous Province of Bolzano and funded by the Ministry of Health (PREV-A-2022-12377013 project), 18 classrooms were monitored using Aranet PRO sensors. Measurements of CO₂ (ppm) and temperature (°C) were recorded every minute during October and November 2024. Among these, 11 classrooms had no CMV systems, 5 were equipped with CMV designed for radon mitigation, and 2 were equipped with CMV designed to guarantee ventilation rates in accordance with current building regulations. The activation of the CMV systems in the latter two classrooms on November 1st pro-

vided a quasi-experimental setting, allowing for controlled pre-post comparison between classrooms with air-exchange CMV activation and classrooms without such activation, serving as control groups. Analyses were limited to actual occupancy hours (8:00 AM – 1:00 PM on school days). To account for classroom-specific temporal patterns and to deal with autocorrelation, the analyses were conducted separately for each classroom, using generalized least-squares models with autoregressive moving-average error structures: ARMA 2,1 for log CO₂ and ARMA 1,1 for indoor air temperature. Predictors for both CO₂ and indoor air temperature included month, day of the week, and a spline-modeled time function. Indoor air temperature and external air temperature were also considered in the analyses of CO₂ and air temperature, respectively. For each outcome, a meta-analysis stratified by CMV type was conducted on the coefficient for month (November vs October) derived from the models.

RESULTS

Descriptive analysis showed that, between October and November, the median CO₂ concentration rose from 1097 ppm to 1306 ppm (+19.1%) in classrooms without CMV, and from 1042 ppm to 1279 ppm (+22.7%) in classrooms with radon CMV. In contrast, in classrooms with CMV designed for air exchange, the median CO₂ concentration decreased from 1072 ppm to 979 ppm (-8.7%) after activation on 1 November. Indoor air temperature decreased slightly across all groups. The meta-analysis of coefficients confirmed these results: the activation of CMV designed to meet ventilation standards was associated with a reduction in indoor CO₂ levels in the post-intervention phase compared to the pre-intervention

phase (−7% in geometric mean), while during the same period, an increase was observed in classrooms without CMV (+17%) and in those with radon CMV (+31%). No relevant differences between groups were observed for indoor air temperature.

CONCLUSIONS

These findings suggest that systems specifically designed for air exchange contribute to reducing indoor CO₂ levels in educational settings without compromising thermal comfort.

Impact of Propensity Score-Based Methods on Effect Estimates in Two Large Observational Endovascular Stroke Cohorts

Stefano Bondoni⁽¹⁾, Alessandro Pezzini⁽²⁾, Valentina Saia⁽³⁾, Giovanni Pracucci⁽⁴⁾, Salvatore Mangiafico⁽⁵⁾, Danilo Toni⁽⁶⁾, Ilaria Casetta⁽⁷⁾, Fabrizio Sallustio⁽⁸⁾, Patrizia Nencini⁽⁹⁾, Enrico Fainardi⁽¹⁰⁾, Andrea Saletti⁽¹¹⁾, Andrea Zini⁽¹²⁾, Manuel Cappellari⁽¹³⁾, Rossana Tassi⁽¹⁴⁾, Mauro Bergui⁽¹⁵⁾, Stefano Vallone⁽¹⁶⁾, Caterina Caminiti⁽¹⁾, Giuseppe Maglietta⁽¹⁾, Matteo Puntoni⁽¹⁾; on behalf of the Italian Registry of Endovascular Treatment in Acute Stroke (IRETAS) Group

 Stefano Bondoni 0009-0005-3048-8531

(1) Clinical Research and Epidemiology Unit, Azienda Ospedaliero-Universitaria di Parma, Parma, Italy

(2) Department of Medicine and Surgery, University of Parma, Parma, Italy

(3) Neurology and Stroke Unit, S. Corona Hospital, Pietra Ligure (SV), Italy

(4) Department of NEUROFARBA, Neuroscience Section, University of Florence, Florence, Italy

(5) Interventional Neuroradiology Consultant, IRCCS Neuromed, Pozzilli (IS), Italy; Adjunct Professor of Interventional Neuroradiology, Tor Vergata University, Sapienza University and S. Andrea Hospital, Rome, Italy

(6) Department of Human Neurosciences, Sapienza University of Rome, Rome, Italy

(7) Neurology Unit, University Hospital Arcispedale S. Anna, Ferrara, Italy

(8) Unità di Trattamento Neurovascolare, Ospedale dei Castelli, ASL Roma 6, Rome, Italy

(9) SOD Stroke Unit, Azienda Ospedaliero-Universitaria Careggi, Florence, Italy

(10) Department of Biomedical, Experimental and Clinical Sciences, Neuroradiology Unit, University of Florence and Careggi University Hospital, Florence, Italy

(11) UO Neuroradiologia, Azienda Ospedaliero-Universitaria di Ferrara, Ferrara, Italy

(12) IRCCS Istituto delle Scienze Neurologiche di Bologna, Department of Neurology and Stroke Center, Maggiore Hospital, Bologna, Italy

(13) Stroke Unit, Azienda Ospedaliera Universitaria Integrata di Verona, Verona, Italy

(14) UOC Stroke Unit, Azienda Ospedaliero-Universitaria Senese, Siena, Italy

(15) Department of Neuroscience, University of Turin, Turin, Italy

(16) Neuroradiology Unit, Ospedale Civile di Baggiovara, Azienda Ospedaliero-Universitaria di Modena, Modena, Italy

CORRESPONDING AUTHOR: Stefano Bondoni, Clinical Research and Epidemiology Unit, Azienda Ospedaliero-Universitaria di Parma, Parma, Italy. E-Mail: stefano.bondoni98@hotmail.com

INTRODUCTION

Observational studies in acute ischemic stroke are vulnerable to confounding by indication and treatment-selection bias, potentially distorting treatment-effect estimates. Propensity score methods can improve comparability between exposure groups and support more reliable interpretation of registry-based evidence. The Italian Registry of Endovascular Treatment in Acute Stroke (IRETAS) is an ongoing multicenter registry of patients undergoing endovascular thrombectomy (EVT).

OBJECTIVE

To assess the impact of propensity score-based methods on exposure–outcome effect estimates in two large observational EVT cohorts.

METHODS

Data were obtained from the IRETAS registry. Two independent cohorts were analyzed: a tandem occlusion cohort (n=2,148), comparing dissection-related versus atherosclerotic etiology, and an antithrombotic therapy cohort (n=10,662),

comparing patients receiving any antithrombotic therapy before stroke onset with untreated patients. Associations were first estimated in unmatched cohorts using regression models appropriate for binary and ordinal outcomes.

Propensity score approaches were tailored to the exposure-group structure of each cohort. In the tandem cohort, dissection-related cases represented approximately 11% of patients; therefore, different propensity score matching strategies were evaluated to identify the approach best suited to the case-mix, balancing covariate comparability, sample-size retention, and precision. In the antithrombotic cohort, the smaller exposure group represented approximately 37% of patients; therefore, full matching was used to retain the entire study population while improving covariate balance and preserving precision. In this cohort, a hierarchical testing strategy was applied, first evaluating any pre-stroke antithrombotic therapy versus no therapy and subsequently assessing whether associations differed between antiplatelet therapy and oral anticoagulants. Balance was assessed using standardized mean differences and Love plots. After propensity score adjustment, associations were re-estimated using weighted generalized estimating equation models, incorporating propensity score-derived weights and accounting for the matched or weighted design.

RESULTS

Substantial baseline imbalances were observed in both cohorts before adjustment. In the tandem cohort, propensity score matching generated a balanced matched population of 713 patients: 224 with dissection-related and 489 with atherosclerotic etiology. Before matching, dissection-related etiology was associated with higher odds of functional independence at 90 days, defined as mRS 0–2, compared with atherosclerotic etiology (OR 2.31, 95% CI 1.75–3.06). After matching, the association was attenuated and no longer statistically significant (OR 1.33, 95% CI 0.94–1.83; $p=0.115$).

In the antithrombotic cohort, full matching retained all patients while substantially improving covariate balance. Before matching, any pre-stroke antithrombotic therapy was associated with a less favorable 90-day mRS distribution compared with no therapy (OR 0.85, 95% CI 0.80–0.92). After matching, this association disappeared (OR 1.03, 95% CI 0.96–1.11). In the hierarchical analysis by antithrombotic class, APLT changed from a negative association before matching (OR 0.90, 95% CI 0.83–0.98; $p=0.011$) to a positive association after matching (OR 1.10, 95% CI 1.01–1.19; $p=0.038$), whereas the association for OAC was attenuated after matching.

CONCLUSIONS

In two EVT registry applications, propensity score-based methods materially changed the interpretation of crude exposure–outcome associations. Apparent clinically relevant differences were markedly attenuated or reversed after balancing baseline characteristics, highlighting the role of confounding and treatment-selection bias in observational stroke research.

Exploring Gender-Based Intersectionality in Patient-Reported Experience of Integrated Care in Multiple Sclerosis

Pasquale Paletta⁽¹⁾, Michela Ponzio⁽¹⁾, Marco Salivetto⁽¹⁾, Tommaso Manacorda⁽²⁾

 Pasquale Paletta 0000-0001-7559-0170

(1) Public Health Research Area, Italian Multiple Sclerosis Foundation (FISM), Genoa, Italy

(2) Italian Multiple Sclerosis Association (AISM), Genoa, Italy

CORRESPONDING AUTHOR: Pasquale Paletta, Public Health Research Area, Italian Multiple Sclerosis Foundation (FISM), Genoa, Italy.
E-Mail: pasquale.paletta@fismets.it

INTRODUCTION

People with multiple sclerosis (pwMS) require coordinated care across multiple professionals and services. Patient-reported experience measures (PREMs) can help assess whether care is perceived as integrated and person-centred. Gender-related differences in healthcare needs, disease burden, and interactions with healthcare systems may contribute to inequalities in perceived care experience. However, these differences may vary according to individual characteristics.

OBJECTIVES

To investigate whether gender is associated with PREMs of integrated care, measured using the Integrated Care Experience Scale for Multiple Sclerosis (ICES-MS), and to explore potential intersectional effects between gender and individual-level characteristics.

METHODS

Data were obtained from a national survey on pwMS conducted by the Italian MS Foundation in 2026. The primary outcome was the ICES-MS total score, rescaled to 0-100, with higher scores indicating better experience of integrated and person-centred care. Multilevel linear regression models with a random intercept for MS Clinical Centre were fitted to account for clustering of patients within centres. The main individual-level model included gender, age, disability level (mild, moderate, severe), area of residence (North, Center, South/Islands), occupational status (employed, unemployed), and disease duration. To explore intersectionality,

separate interaction models were fitted between gender and age, disability level, area of residence, occupational status, and disease duration. Adjusted predicted means were estimated for relevant subgroups.

RESULTS

Overall, 1,520 pwMS from 222 MS Clinical Centres were included. In the main individual-level model, being a man was associated with a slightly higher ICES-MS scores compared to with being a woman (β 2.69; 95% CI: 0.20 to 5.17; $p=0.034$). Compared with severe disability, mild disability was strongly associated with higher ICES-MS scores (β 9.76; 95% CI: 5.90 to 13.61; $p<0.001$), while moderate disability showed a borderline association (β 3.97; 95% CI: -0.03 to 7.97; $p=0.052$). Area of residence was also significantly associated with ICES-MS scores, with higher scores observed in all areas compared to South/Islands. Longer disease duration was associated with lower scores (β -0.15 per year; 95% CI: -0.28 to -0.03; $p=0.017$), whereas age and occupational status were not significantly associated with the outcome.

No statistically significant interactions were found between gender and age, disability level, area of residence, occupational status, or disease duration. Adjusted predicted means showed consistent descriptive patterns: ICES-MS scores tended to be higher among men, among people with mild disability, in the North and the Centre, and in patients with shorter disease duration. However, since the interaction terms were not statistically significant, these findings suggest additive differences rather than clear evidence of intersectional effects between gender and individual characteristics.

CONCLUSIONS

Gender, disability level, area of residence, and disease duration were associated with PREMs of integrated care among pwMS. However, no robust evidence of gender-based intersectional effects was observed across the individual characteristics examined. Multilevel modelling provided an appropriate framework to account for the hierarchical structure of the data and to estimate adjusted subgroup differences. These findings suggest that gender differences in ICES-MS scores exist, suggesting that interventions aimed at improving integrated care experience should primarily address clinical complexity and disease progression, while monitoring potential gender-related differences.

The First Human Biomonitoring Study in the Marche Region: Assessment of Environmental Exposure and Health Effects

Alessandro Fontanarosa⁽¹⁾, Alessandro Catalini⁽¹⁾, Antonello Lupi⁽¹⁾, Alessandra Chiappini⁽²⁾, Andrea Filonzi⁽³⁾, Cristina Mancini⁽³⁾, Vittorio Marconi⁽³⁾, Silvia Bartolacci⁽⁴⁾, Marco Baldini⁽⁴⁾

 Alessandro Fontanarosa 0009-0009-6349-892X

1) Regional Health Agency, Marche Region, Ancona, Italy

(2) Department of Industrial Engineering and Mathematical Sciences, Polytechnic University of Marche, 60131 Ancona, Italy

(3) UOC ISP Environment and Health AST Ancona, Ancona, Italy

(4) Environmental Epidemiology Service, ARPA Marche, Ancona, Italy

CORRESPONDING AUTHOR: Alessandro Fontanarosa, Regional Health Agency, Marche Region, Ancona, Italy. E-Mail: alessandro.fontanarosa97@gmail.com

BACKGROUND

As part of the INSINERGIA national project, funded under the National Complementary Plan – “Health, Environment, Biodiversity and Climate” Program, a human biomonitoring study has been launched in several Sites of National Interest (SINs) for environmental remediation. Human biomonitoring represents a key tool for public health and prevention tool, allowing the measurement of environmental contaminants or their metabolites in biological matrices and providing an integrated and directed assessment of individual exposure from multiple sources and pathways. Among the SINs involved the site in Falconara Marittima (Ancona, Italy) is of particular relevance due to its long-standing industrial activities and the potential risk of environmental contamination associated with the sources of pressure present in the area.

OBJECTIVES

The study aims to assess exposure to environmental contaminants among residents of the municipality of Falconara Marittima, an area included within the SIN.

METHODS

This is a national multicenter cross-sectional study with a low level of intervention. The project involves 13 SINs across 9 Italian regions and plans to recruit approximately 2,600

participants, about 200 subjects for each operational unit. Eligible subjects are adults aged 20–50 years who have continuously resided in the study area for at least five years. Individuals with occupational exposure to the contaminants under investigation and those unable to provide informed participation will be excluded.

For the Falconara Marittima SIN, the population was classified according to environmental exposure levels using a threshold derived from a PM₁₀ dispersion model based on pollutant iso-concentration estimates. A balanced stratified random sampling was implemented, with strata defined by sex (male/female), age group (20–34 and 35–50 years), and environmental exposure level (high/low), using data from the regional healthcare beneficiary's database. To account for expected non-response, additional subjects were randomly selected within each stratum.

Data collection will include questionnaires on sociodemographic characteristics, lifestyle factors, and environmental risk perception. Clinical and anamnestic information, anthropometric and vital parameters, and blood and urine samples will also be collected.

Laboratory analyses will include biomarkers of exposure, such as heavy metals, aromatic and aliphatic compounds, polycyclic aromatic hydrocarbons (PAHs), per- and polyfluoroalkyl substances (PFAS), and polychlorinated biphenyls (PCBs). Biomarkers of effect, including inflammatory cytokines and biochemical-clinical indicators (hematological, metabolic, and endocrine profiles), as well as genetic susceptibility biomarkers (single nucleotide polymorphisms), will also be evaluated.

Descriptive and appropriate group comparison tests will be used to evaluate differences in exposure levels and biological indicators. Associations between exposure and effect biomarkers will be assessed using generalized linear models (GLMs) adjusted for potential confounders, including age, sex, smoking status and other covariates.

RESULTS

The study is currently in the implementation phase. The organizational framework, site logistics, and biological sample management procedures have been established. Data collection is scheduled for June–July 2026.

CONCLUSIONS

The study represents the first integrated human biomonitoring model in the field of environmental-health of this complexity and scope to be developed in the Marche Region. The study is expected to provide valuable evidence for exposure assessment and support the planning of public health interventions in environmentally contaminated areas.

Ovarian Cancer Survival across the Introduction of Targeted Treatment: A Nationwide Analysis by Histotype and Stage

Silvia Mignozzi^(1,2), Victoria Cooley^(2,3), Cassia B. Trewin-Nybråten⁽⁴⁾, Renée Turzanski Fortner^(2,5)

 Silvia Mignozzi 0009-0003-5703-0534

(1) Department of Clinical Sciences and Community Health, Department of Excellence 2023-2027, University of Milan, Milan, Italy

(2) German Cancer Research Center (DKFZ), Division of Cancer Epidemiology, Heidelberg, Germany

(3) Heidelberg University, Medical Faculty Heidelberg, Heidelberg, Germany

(4) Department of Registration, Cancer Registry of Norway, Norwegian Institute of Public Health, Oslo, Norway

(5) Department of Research, Cancer Registry of Norway, Norwegian Institute of Public Health, Oslo, Norway

CORRESPONDING AUTHOR: Silvia Mignozzi, Department of Clinical Sciences and Community Health, Department of Excellence 2023-2027, University of Milan, Milan, Italy. E-mail: silvia.mignozzi@unimi.it

INTRODUCTION

Bevacizumab was approved for targeted therapy of ovarian cancer in 2011 by the European Medicines Agency, with the subsequent introduction of PARP (poly[ADP]-ribose polymerase) inhibitors in 2014. Improvements in outcomes of cytoreductive surgery occurred in parallel in many regions.

OBJECTIVES

Population-based studies evaluating survival trends in the context of modern treatment are sparse. Our primary objective was to assess variations in net survival trends across grouped calendar periods during the period in which targeted therapies were introduced.

METHODS

Nationwide data were derived from the Center for Cancer Registry Data at the Robert Koch Institute. Expected mortality rates were derived from German life tables provided from the Federal Statistical Office of Germany and stratified by age, calendar year and federal state. Cases were diagnosed between January 1, 2010 and December 31, 2023. Survival time was computed as the period between a patient's cancer diagnosis and their death, or December 31, 2023. We defined the following four calendar periods for diagnosis: 2010–2012: largely

pre-targeted therapy period; 2013–2016: period following the initial introduction of bevacizumab; 2017–2019: period following the initial introduction of PARP inhibitors; 2020–2023: targeted therapy period. Excess hazard and net survival across the selected periods were estimated using flexible parametric models within a relative survival framework. The models were stratified by histology and included period of diagnosis, stage group (I–II, III–IV), natural splines of age at diagnosis, and area of residence. Net survival and excess hazard were predicted from models for females with median age at diagnosis, residing in the most common area (Central Area) and for each diagnosis period, histological subtype, and stage group.

RESULTS

A total of 62,447 ovarian cancer cases were included. A total of 74% of cases were of high-grade serous histology and 66% of cases were diagnosed at stage III or IV. Excess hazards differed substantially by histotype, and also within histotype by time since diagnosis. Among individuals with the predominant high-grade serous tumours, as well as endometroid tumours, the excess mortality hazard was notably lower in the most recent diagnosis period (2020–2023), as compared to the earliest period (2010–2012). For high-grade serous and endometroid tumours, 3- and 5-years net survival increased over the entire period. Survival patterns over time differed among the other histotypes, with more modest differences across the evaluated time periods.

CONCLUSIONS

Ovarian cancer survival in Germany has improved between 2010 and 2023, specifically for the high-grade serous and endometrioid histotypes. This coincides with the introduction of targeted therapies, as well as with improvements in debulking surgery. There remains a need for further studies aimed at enabling earlier diagnosis and improving prognosis in ovarian cancer, as well as for research to develop and refine novel treatments plans tailored to histotypes.

Association of Prebiotic Fiber Intake with Colorectal Cancer Risk in an Italian Case-Control Study

Arianna Natale⁽¹⁾, Federica Fiori⁽²⁾, Maria Parpinel⁽²⁾, Federica Turati⁽¹⁾, Carlo La Vecchia⁽¹⁾, Francesca Bravi⁽¹⁾, Marta Rossi⁽¹⁾

 Arianna Natale 0000-0001-5939-1999

(1) Department of Clinical Sciences and Community Health, Dipartimento di Eccellenza 2023–2027, University of Milan, Milan, 20133, Italy

(2) Department of Medicine, University of Udine, Udine, 33100, Italy

CORRESPONDING AUTHOR: Arianna Natale, arianna.natale@unimi.it

INTRODUCTION

Colorectal cancer (CRC) is inversely associated with dietary fiber intake. Prebiotic fibers are a subgroup of dietary fibers defined as non-digestible compounds selectively utilized by the gut microbiota, potentially influencing intestinal homeostasis. Some prebiotics such as inulin-type fructans (ITF) and galacto-oligosaccharides (GOS) have been investigated in relation to CRC risk, but epidemiological evidence is inconsistent. A large case-control study reported inverse associations for GOS, while findings for ITF are generally null. Overall, data on specific prebiotic fibers and CRC risk remain limited.

OBJECTIVES

To investigate the association between intake of specific prebiotic fibers, including ITF, fructo-oligosaccharides (FOS) and GOS and CRC risk in an Italian case-control study.

METHODS

We analyzed data from a multicentric case-control study conducted in Milan (Italy) between 2017 and 2019, including 100 histologically confirmed CRC cases and 200 controls, of whom 100 had intestinal adenomas (IA) and 100 were IA- and cancer-free. Cases and controls were individually matched by study center, sex, and age (± 5 years). Diet was assessed using a validated food frequency questionnaire. Intake of prebiotic fibers, namely ITF, total and specific FOS (kestose, nystose and fructofuranosyl-nystose), total and specific GOS (stachyose and raffinose) was estimated using ad hoc composition databases. Energy intake was derived from the Italian food composition database for epidemiological studies.

Using a logistic regression model conditioned on the matching variables and including terms for education (<7 , 7–11, ≥ 12 years) and energy intake tertiles, we estimated the odds ratios (OR) and the corresponding 95% confidence intervals (CI) of CRC for sex-specific tertiles of prebiotic intake (computed on controls). We also estimated the OR for an increment of one SD of prebiotics. P values for trend across tertiles were computed through the Wald β test. Analyses were additionally stratified by cancer site and sex.

RESULTS

Higher intake of selected prebiotic fibers was associated with lower CRC risk. Nystose showed an inverse association for the highest versus lowest tertile (OR 0.44; 95% CI 0.22–0.89), with a significant trend across tertiles. Raffinose was inversely associated (continuous OR 0.63; 95% CI 0.43–0.91), with a stronger inverse association for rectal cancer (OR 0.30; 95% CI 0.15–0.63). No consistent associations were observed for the other prebiotics.

In stratified analyses, for nystose an inverse association was observed in males (OR 0.55; 95% CI 0.43–0.89) but not in females (OR 0.95; 95% CI 0.61–1.47; p-heterogeneity = 0.053). Conversely, for stachyose, an inverse association was found in females (OR 0.53; 95% CI 0.31–0.90) but not in males (OR 1.24; 95% CI 0.92–1.66; p-heterogeneity = 0.010).

CONCLUSIONS

Nystose and raffinose were inversely associated with overall CRC risk, and raffinose showed an inverse association with rectal cancer. Sex-specific inverse associations were observed for nystose and stachyose. These findings support a potential role of selected prebiotic fibers in CRC etiology.

Temporal Patterns of Cancer Incidence and Mortality in Italy: Cohort Effects and Young-Adult Trends

Nicolò Lentini⁽¹⁾, Claudia Santucci⁽¹⁾, Carlo La Vecchia⁽²⁾, Stefania Boccia^(1,3), Roberta Pastorino^(1,3)

 Nicolò Lentini 0009-0008-3111-1535

(1) Section of Hygiene, Department of Life Sciences and Public Health, Università Cattolica del Sacro Cuore, Rome, Italy

(2) Department of Clinical Sciences and Community Health, University of Milan, Milan, Italy

(3) Department of Woman and Child Health and Public Health, Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy

CORRESPONDING AUTHOR: Nicolò Lentini, nicolo.lentini@unicatt.it

BACKGROUND

Over the past decades, cancer mortality has declined in many high-income countries owing to reductions in exposure to tobacco and other risk factors, screening programs, and advances in diagnosis and treatment. However, ongoing changes in environmental, behavioral, and metabolic risk factors continue to reshape the epidemiology of cancer.

OBJECTIVES

To provide a comprehensive assessment of long-term cancer incidence and mortality trends in Italy and investigate cohort patterns through age-period-cohort (APC) analyses.

METHODS

We used certified mortality data for the period 1980-2023 from the World Health Organization and the Italian National Institute of Statistics databases, incidence data for 1995-2018 from the European Cancer Information System. We computed age-standardized mortality rates (ASMRs) and age-standardized incidence rates (ASIRs) using the Segi-Doll world standard population. We described long-term trends using locally estimated scatterplot smoothing (LOESS) regression. We quantified overall percentage changes between 2023-2010 for mortality and between 2018-2008 for incidence, with 95% confidence intervals (CI) derived from a parametric bootstrap assuming Poisson-distributed counts. We performed APC analyses using generalized additive models with a quasi-Poisson specification. Age, period, and

cohort effects were modeled using cubic regression splines, allowing flexible estimation of nonlinear temporal patterns while addressing the identifiability problem inherent to APC models.

RESULTS

Between 2010 and 2023, overall mortality declined by 25.5% (95%CI: -26.3%; -24.6%) in males (to 103.5/100,000) and 13.9% (95%CI: -14.9%; -12.7%) in females (to 71.4/100,000), with major reductions for stomach (~35%), lung (~30%) (except for a 15.4% increase in females aged 50+), and colorectal cancers (~33%). Mortality increased for testicular (8.3%), uterine (2.8%), and pancreatic cancer in both sexes (2.0% males, 6.9% females). Among young adults (25-49 years), mortality from all cancers declined by 22.1% (95%CI: -26.5%; -17.6%) in males and 19.2% (95%CI: -23.2%; -14.9%) in females but increased in testicular (36.6%), and uterine cancer (16.6%). In this age group pancreatic cancer decreased by 21.1%, whereas it remained stable in males (2.0%). APC analyses showed declining cohort-specific risks for most cancers among cohorts born after approximately 1925. Pancreatic cancer exhibited largely stable cohort risks, whereas female lung cancer risk increased up to the 1960 birth cohort before declining in more recent generations.

Between 2008 and 2018, overall cancer incidence decreased by 10.9% (95%CI: -11.8%; -10.0%) in males (to 305.7/100,000) but increased by 2.9% (95%CI: 2.0%; 4.1%) in females (to 273.7/100,000), with rising trends for skin (~70%), thyroid (~25%), testicular (24.2%), and lung cancer in females aged 50+ (27.2%). Rates were stable for

pancreatic cancer (0.3% in males, 6.7% in females) and increased modestly for breast cancer (3.9%). Among young adults, incidence increased in both males by (3.6%, 95%CI: -0.4%; 7.7%) and females (10.3%, 95%CI: 7.0%; 13.3%).

CONCLUSIONS

Cancer mortality in Italy has declined substantially over recent decades, reflecting the combined effects of smoking cessation in men, screening, and therapeutic advances. Nevertheless, less favorable trends persist for selected cancers, particularly pancreatic cancer and lung cancer among older women. The increasing incidence of cancers associated with metabolic, environmental, and lifestyle-related factors highlights emerging public health challenges.

Questionnaire-Based Tools for Colorectal Cancer Screening and Risk Assessment: A Scoping Review

Bruna De Oliveira Ascef⁽¹⁾, Elisa Camussi⁽²⁾, Francesco Venturelli⁽³⁾, Arena Stefania⁽¹⁾, Francesca Mori⁽²⁾, Valeria Virzi⁽²⁾, Alessio David⁽²⁾, Massimo Vicentini⁽³⁾, Gianluigi Ferrante⁽²⁾, Carlo Senore⁽²⁾, Vittorio Simeon⁽¹⁾

 Bruna De Oliveira Ascef 0000-0002-2744-3974

(1) *Unità di Statistica Medica, Università degli Studi della Campania Luigi Vanvitelli, Napoli, Italy*

(2) *AOU Città della Salute e della Scienza di Torino – CPO Piemonte, Torino, Italy*

(3) *Servizio di Epidemiologia, Azienda USL-IRCCS di Reggio Emilia, Reggio Emilia, Italy*

CORRESPONDING AUTHOR: Bruna De Oliveira Ascef, *Unità di Statistica Medica, Università degli Studi della Campania Luigi Vanvitelli, Napoli, Italy. E-Mail: bruna.deoliveiraascef@unicampania.it*

INTRODUCTION

Colorectal cancer (CRC) is one of the most common malignancies worldwide. The implementation of screening programs has been shown to significantly reduce mortality. Individual risk stratification using structured tools may improve the efficiency of screening programs by directing resources to individuals at higher risk while reducing unnecessary procedures for those at lower risk. However, the available literature reveals substantial heterogeneity among existing tools, with limited clinical validation and poor integration into screening programs.

OBJECTIVES

To systematically map questionnaire-based tools developed for CRC risk assessment and stratification in average-risk populations, evaluating their methodological characteristics, applicability, and validation status.

METHODS

A scoping review was conducted following the Arksey and O'Malley framework and the Joanna Briggs Institute methodology, and reported according to the PRISMA-ScR guidelines. Searches were performed in Medline/PubMed, Embase, Scopus, and Web of Science up to July 24, 2025. Study selection was conducted using the Rayyan platform by three reviewers working in parallel. Studies describing, developing, or validating questionnaire-based tools (paper-based, digital, app-based, or risk algorithms) for CRC risk assessment in average-risk adults were included. Data extraction was

performed using a standardized form. ChatGPT was used to support preliminary data extraction, which was subsequently verified by human reviewers.

RESULTS

A total of 4,554 records were identified. After duplicate removal ($n=2,467$), 2,087 records underwent screening, and 106 full-text articles were assessed for eligibility. Nineteen studies met the inclusion criteria. Studies were published between 2006 and 2025 and originated from nine countries, primarily China, the Netherlands, Norway, and the United States. Ten studies (52.6%) developed study-specific tools, whereas nine (47.4%) used, adapted, or modified previously published instruments, most commonly the Asia-Pacific Colorectal Screening (APCS) score and its variants. Questionnaires contained between 4 and 29 items (median: 6), covering demographic, clinical, behavioral, and lifestyle-related risk factors. Measures of discrimination were reported in 13 studies, with AUC/C-index values ranging from 0.60 to 0.89 (median: 0.67). Overall, 40 validation analyses were identified across multiple outcomes, including incident CRC, advanced neoplasia, and precancerous lesions. Predictive performance varied substantially across target populations, assessed outcomes, and questionnaire compositions. External validation was reported for only a minority of tools.

CONCLUSIONS

This scoping review provides a comprehensive mapping of questionnaire-based tools for colorectal cancer risk assessment in average-risk populations. Although several instruments demonstrated acceptable discriminatory performance,

substantial heterogeneity was observed in questionnaire content, target outcomes, modeling approaches, and validation strategies. The limited availability of external validation and the limited integration of these tools into organized screening programs represent important barriers to implementation. Future research should prioritize standardization, external validation, and evaluation of these tools in real-world screening settings to support risk-based colorectal cancer screening strategies.

Patient-Reported Outcome Measures after Out-of-Hospital Cardiac Arrest: A Prospective Methodological Analysis of Retention and Follow-Up Completeness within the Revive Pathway

Roberto Primi^(1,2,3), Laura Calabrese^(2,3), Marco Mion^(2,3,5,6), Sara Bendotti^(2,3), Leila Ulmanova^(2,3), Alessia Currao^(2,3,4), Simona Villani⁽¹⁾, Simone Savastano^(2,3), Enrico Baldi^(2,3)

 Roberto Primi 0000-0002-4271-6248

(1) Department of Public Health, Experimental and Forensic Medicine. University of Pavia, Pavia, Italy

(2) Division of Cardiology, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy

(3) Cardiac Arrest and Resuscitation Research Team (RESTART), Fondazione IRCCS Policlinico San Matteo, Pavia, Italy

(4) Department of Internal Medicine and Medical Therapy, University of Pavia, Pavia, Italy

(5) Anglia Ruskin School of Medicine & MTRC, Chelmsford, United Kingdom

(6) Anglia Ruskin University, Chelmsford, United Kingdom

CORRESPONDING AUTHOR: Roberto Primi, Department of Public Health, Experimental and Forensic Medicine. University of Pavia, Pavia, Italy. E-Mail: roberto.primi01@universitadipavia.it

INTRODUCTION

Cardiac arrest is the loss of mechanical cardiac function and the absence of systemic circulation. When it occurs out-of-hospital it is defined out-of-hospital cardiac arrest (OHCA). Survival after OHCA does not fully capture recovery, as survivors often experience persistent cognitive, psychological, and functional consequences. Patient-reported outcome measures (PROMs) are essential to assess these domains, but their longitudinal collection in routine survivorship care depends on patient engagement and retention over time. REVIVE is a regional hub-and-spoke survivorship pathway implemented within the Cardiac Arrest Registry of the Lombardy Region (LombardiaCARE) network to coordinate post-discharge follow-up after OHCA and standardize the assessment of patient-reported outcomes.

OBJECTIVES

To describe the structure of PROM collection within REVIVE and to evaluate patient retention, evaluability, and follow-up completeness over the 2025 calendar year.

METHODS

We conducted a descriptive analysis of the REVIVE study between January and December 2025. During this period, 1,663 OHCA events were admitted to participating hospitals, 127 patients were eligible for the REVIVE pathway, 99 were contacted by the hub psychologist after signing informed consent, and 67 completed at least one assessment. REVIVE is structured as a hub-and-spoke pathway, with a specialist hub coordinating post-discharge follow-up and outcome assessment across participating hospitals. PROMs and related outcomes were collected at predefined timepoints and included discharge, 1 month, 3 months, and 6 months after the OHCA event. The PROMs assessed were anxiety and depression (HADS), post-traumatic stress symptoms (IES-R), cognitive performance (MoCA), and health-related quality of life (EQ-5D-5L), together with a clinical semi-structured interview. The analysis focused on the number of patients contributing at each timepoint and on the availability of at least one evaluable PROM per patient.

RESULTS

Among the 99 consenting patients, 67 (67.7%) were evaluable and formed the analytic denominator. At the visit level, 46 patients had at least one PROM recorded at discharge, 49 at 1 month, 52 at 3 months, and 49 at 6 months.

Domain-level completeness was high: HADS was available in 66 of 67 evaluable patients (98.5%), while MoCA and EQ-5D VAS were available in all evaluable patients (67/67, 100%). These results suggest that the main methodological challenge was not baseline collection, but maintaining patient engagement across repeated assessments, which in PROM-based survivorship care is inherently patient-driven rather than passive.

CONCLUSIONS

PROMs add value beyond survival by capturing outcomes that matter to OHCA survivors, but their validity depends on sustained patient engagement throughout follow-up. In REVIVE, the main challenge was not baseline collection, but maintaining retention in a patient-driven and psychologically sensitive assessment process, where readiness to participate may vary over time. These findings support clear definition of the evaluable cohort, transparent reporting of timepoint-specific completeness, and retention-oriented strategies to improve the quality and interpretability of longitudinal PROM data.

Determinants of Televisits in Diabetology in the Tuscany Region

Rebecca Drabczyk⁽¹⁾, Sabrina Giometto⁽¹⁾, Matilde Razzanelli⁽²⁾, Neda Parsa⁽²⁾, Paolo Francesconi⁽²⁾, Ersilia Lucenteforte⁽¹⁾

(1) University of Florence, Florence, Italy

(2) Regional Health Agency, Florence, Italy

CORRESPONDING AUTHOR: Rebecca Drabczyk, University of Florence, Florence, Italy. E-Mail: rebecca.drabczyk@edu.unifi.it

INTRODUCTION

Telemedicine represents a fundamental innovative tool for optimizing the management of chronic diseases such as diabetes. However, to effectively support the planning and governance of healthcare services, it is crucial to identify the profiles of patients accessing this care modality in the real world. These analyses are intended to support the planning, organization, and governance of healthcare services.

OBJECTIVES

This study aimed to identify factors associated with the use of televisits rather than outpatient follow-up visits among patients with diabetes in Tuscany between 2021 and 2025.

METHODS

We conducted a retrospective cohort study based on administrative data from the Tuscany Region harmonized by the Regional Health Agency (ARS). The population included patients with a first specialist diabetology visit between January 1, 2021 and December 31, 2025, defined as the index date. To define the characteristics, a one-year observation period prior to the index date was considered, defined as the look-back period. A logistic regression model was used, considering the type of follow-up visit as the outcome (televisit vs outpatient visit as reference). Three models were estimated: crude, adjusted for age and gender, and further adjusted for the number of comorbidities and the number of different fourth level Anatomical Therapeutic Chemical (ATC) codes. Since the same individual could contribute multiple observations, standard errors were estimated using a cluster-robust variance estimators at the individual level.

RESULTS

The cohort included 539,645 diabetology follow-up visits, 16% of which were conducted as televisits. In the fully adjusted model, female gender was directly associated with follow-up televisits compared to male gender (odds ratio (OR) 1.12, 95% confidence interval (CI) 1.08–1.16). Compared with the 66–85 age group, all other age groups showed a direct association with televisits, particularly patients aged 0–18 (OR: 5.29, CI: 4.76–5.87). Regarding pharmacological complexity, patients with 11–20 different fourth level ATC codes and those with more than 20 codes presented a direct association with televisits, compared to patients with 0–5 ATCs (OR: 1.13, CI: 1.07–1.19, and OR: 1.60, CI: 1.47–1.73, respectively). Among therapies, insulin, antipsychotics, antidepressants, and anti-dementia drugs were directly associated with televisits, while ACE inhibitors, SGLT2 inhibitors, beta-blockers, and MRA diuretics were inversely associated. Indicators of higher care complexity (admission to long-term care facilities, pre-admission to long-term care facilities, hospitalizations in the previous year, and 21–100 home care visits) also showed a direct association. Among comorbidities, dementia, Parkinson's disease, and stroke were directly associated with televisits, while peripheral arterial occlusive disease, acute myocardial infarction, and hypertension were inversely associated. For other conditions (including COPD, heart and kidney failure, atrial fibrillation), the association disappeared after adjustment.

CONCLUSIONS

Overall, televisits were associated with socio-demographic characteristics, particularly gender and age, as well as specific clinical-care profiles. In particular, direct associations emerged between televisits and female gender, younger age groups, higher pharmacological complexity, certain indicators of care complexity, and some neurological or chronic comorbidities.

Healthcare Utilization Databases for Real-World Outcome Analysis of Ruxolitinib Treatment in Myelofibrosis

Barbara Mora⁽¹⁾, Matteo Franchi^(2,3), Ludovica Margotto⁽²⁾, Olivia Leoni⁽⁴⁾, Daniela D'ippoliti⁽⁵⁾, Emanuela Carloni⁽⁵⁾, Ilaria Cozzi⁽⁵⁾, Enrica Santelli⁽⁵⁾, Fabrizio Gemmi⁽⁶⁾, Claudia Szasz⁽⁶⁾, Margherita Maffioli⁽⁷⁾, Carmelo Gurnari⁽⁸⁾, Enrico Attardi⁽⁸⁾, Daniele Cattaneo⁽¹⁾, Marta Bortolotti⁽¹⁾, Nicola Stefano Fracchiolla⁽¹⁾, Alessandra Iurlo⁽¹⁾, Giovanni Corrao^(2,3), Matteo Giovanni Della Porta^(9,10), Alessandro Maria Vannucchi⁽¹¹⁾, Maria Teresa Voso⁽⁸⁾, Paola Guglielmelli⁽¹¹⁾, Francesco Passamonti^(1,12)

 Ludovica Margotto 0009-0005-4684-9530

(1) Hematology, Fondazione IRCCS Ca' Granda, Ospedale Maggiore Policlinico, Milan, Italy

(2) National Center for Healthcare Research and Pharmacoepidemiology, University of Milano-Bicocca, Milan, Italy

(3) Unit of Biostatistics, Epidemiology and Public Health, University of Milano Bicocca, Milan, Italy

(4) Regional Epidemiological Observatory, Lombardy Region, Milan, Italy

(5) Department of Epidemiology, Lazio Regional Health Service, Rome, Italy

(6) Quality and Equity Unit, Regional Health Agency of Tuscany, Florence, Italy

(7) Hematology, Ospedale di Circolo, ASST Sette Laghi, Varese, Italy

(8) Department of Biomedicine and Prevention, University of Rome Tor Vergata, Rome, Italy

(9) Comprehensive Cancer Center, IRCCS Humanitas Research Hospital, Rozzano, Italy

(10) Department of Biomedical Sciences, Humanitas University, Pieve Emanuele, Italy

(11) Centro di Ricerca e Innovazione Malattie Mieloproliferative (CRIMM), AOU Careggi, University of Florence, Florence, Italy

(12) Department of Oncology and Hemato-Oncology, University of Milan, Milan, Italy

CORRESPONDING AUTHOR: Ludovica Margotto, National Center for Healthcare Research and Pharmacoepidemiology, University of Milano-Bicocca, Milan, Italy. E-Mail: ludovica.margotto@unimib.it

INTRODUCTION

Myelofibrosis (MF) is a BCR::ABL1-negative myeloproliferative neoplasm characterised by a heterogeneous clinical phenotype encompassing abnormalities in blood counts, bone marrow fibrosis, splenomegaly, constitutional symptoms and vascular complications. MF may progress to accelerated phase (AP) and blast phase (BP). MF includes primary MF, comprising prefibrotic and overt-PMF, and secondary MF, evolving from polycythemia vera or essential thrombocythemia. Survival estimates vary from 7 to 14 years according to MF subtype. Prognostic scores define patients' suitability to stem cell transplant (SCT). Apart from SCT, disease burden control relies on JAK inhibitors (JAKis), among which Ruxolitinib (RUX) was the first approved in Italy. Because of its high cost, RUX dispensing data are routinely collected by regional health authorities through healthcare utilization databases (HCUDs), enabling the generation of real-world evidence.

OBJECTIVES

To assess overall survival (OS), time to RUX discontinuation and healthcare resource use in a population of intermediate- and high-risk MF patients who initiated RUX outside a clinical trial, using regional HCUD.

METHODS

We conducted a retrospective population-based cohort study using HCUD from three Italian regions (Lombardy, Lazio and Tuscany) accounting for 20 million inhabitants. The study included 652 intermediate- and high-risk MF patients who initiated RUX between October 2014 and December 2017, followed through regional data availability (up to December 2023 in Lombardy and Lazio, July 2021 in Tuscany). We estimated OS and time to RUX discontinuation using the Kaplan-Maier estimator and assessed red blood cell (RBC)

transfusions, SCT, hospitalization-related events, and average individual healthcare costs. We evaluated which factors among sex, age classes, year of RUX initiation, RUX starting dose, SCT (as time-dependent), and Multisource Comorbidity Score (MCS) classes were associated with mortality using a multivariable Cox proportional hazard model. Analyses were performed separately within each region, and region-specific estimates were combined using a two-stage meta-analysis.

RESULTS

Over a median follow-up of 36.8 months, median OS was 48.0 months (95%CI: 43.2-51.6), and median time to RUX discontinuation was 31.2 months (95%CI: 26.4-36.0). In the first 6 months, 408 (69%) patients required no RBC transfusions, 172 (29%) received 1-5 units, and 14 (2%) received ≥ 6 units. 10.9% of patients underwent SCT. Incidence rates per 100 person-years were 10.30 for infections, 5.47 for solid tumors, 5.22 for AP/BP, 3.47 for bleeding events, and 1.56 for thrombosis. Mean annual cost per patient was 30,675€. Independent predictors of mortality were male sex (vs female: HR: 1.59, 95%CI: 1.12-2.23), age 70-79 years (vs <70: HR: 2.00, 95%CI: 1.60-2.49), age ≥ 80 years (vs <70: HR: 3.31, 95%CI: 2.11-5.19), poor MCS (≥ 15 vs 0-4: HR: 1.97, 95%CI: 1.45-2.68), and lower starting RUX dose (<20 vs ≥ 20 mg BID: HR: 1.42, 95%CI: 1.10-1.83).

CONCLUSIONS

HCUD provided information on RUX treatment in an unselected, real-world population of intermediate- and high-risk MF patients. While disease-specific details (e.g., blood counts, mutation profile, karyotype) are unavailable in HCUD, these data are pivotal for understanding real-world drug use and supporting pharmacoeconomic evaluations.

Tissue-Specific Epigenetic Drift in Parkinson's Disease: A Multi-Region DNA Methylation Meta-Analysis

Davide Fernando Dragani^(1,2), Luciano Calzari⁽¹⁾, Sabrina Germani⁽²⁾, Davide Sacco⁽¹⁾, Erika Carbone⁽²⁾, Davide Gentilini^(1,2)

 Davide Fernando Dragani 0009-0007-2882-4481

(1) Bioinformatics and Statistical Genomics Unit, IRCCS Istituto Auxologico Italiano, Cusano Milanino, Milan, Italy

(2) Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy

CORRESPONDING AUTHOR: Dragani Davide Fernando, Bioinformatics and Statistical Genomics Unit, IRCCS Istituto Auxologico Italiano, Cusano Milanino, Milan, Italy. E-mail: davidefernando.dragani01@universitadipavia.it

INTRODUCTION

Parkinson's disease (PD) is a neurodegenerative disorder whose genetics explains only part of the disease risk, suggesting the involvement of non-genetic factors. Epigenetic-drift, defined as the accumulation of stochastic epigenetic mutations (SEMs) in DNA, has been implicated in other neurodegenerative diseases but remains poorly characterized in PD.

OBJECTIVES

To investigate the role of global and gene-specific epigenetic-drift in PD and assess its relevance across different tissues and brain regions affected by the disease.

METHODS

A meta-analytic study design was adopted by integrating public DNA methylation datasets. Eligibility criteria for dataset inclusion required the use of Illumina 450K or EPIC arrays, phenotypic metadata, clinical or neuropathological diagnosis of PD, at least 15 subjects per study group and cross-sectional study design. Applying these criteria yielded a final cohort comprising 3,425 peripheral blood and 378 brain tissue samples from case-control designs, alongside a multi-region dataset containing intra-subject repeated measures across caudate nucleus, prefrontal cortex and substantia nigra from PD individuals (N=97). SEMs were identified as extreme outlier using Tukey's rule, defining reference ranges from control beta values at each CpG site or through a leave-one-out resampling procedure for the multi-region dataset.

Differences in global SEM burden between cases and controls were evaluated using regression models and were meta-analyzed via a random-effects meta-analysis. Differences in methylation deviation distributions between cases and controls were evaluated using regression models with subject as a random effect. The multi-region dataset was analyzed using mixed-effects models, accounting intra-subject correlation, followed by post-hoc comparisons to identify region-specific increases in epigenetic drift. Gene-specific associations were explored using kernel-based methods (SKAT), treating SEMs as rare variants within an aggregation framework. All models were adjusted for standard EWAS covariates and validated via sensitivity analyses.

RESULTS

Meta-analysis revealed a significant increase in global SEM burden in the frontal lobe of PD patients compared with HCs ($\beta = 0.6$; $p = 0.03$), whereas no significant differences were observed in peripheral blood or the parietal cortex. In the frontal lobe, PD patients exhibited an enrichment of SEMs characterized by smaller methylation deviations, while HCs showed a relative enrichment of high-delta SEMs ($p < 0.01$). Gene-based analysis in blood identified genes differentially enriched for SEMs (Bonferroni-corrected p -value < 0.01), involved in axon guidance pathway. In the multi-region dataset, SEM burden increased progressively from the caudate nucleus to the prefrontal cortex, reaching its highest level in the substantia nigra, where it was significantly greater than in the other regions.

CONCLUSIONS

Our findings indicate that PD is associated with a tissue-specific increase in epigenetic-drift, particularly in brain regions most affected by neurodegeneration. The enrichment of low-delta SEMs suggests increased local epigenetic heterogeneity, consistent with the accumulation of SEMs that are poorly shared among cells. These findings support a potential involvement of epigenetic-drift in PD pathophysiology, although they do not allow us to determine whether this phenomenon contributes to disease development or arises as a consequence of neurodegenerative processes. In particular, increased cell epigenetic heterogeneity may either enhance the vulnerability of brain tissue to neurodegeneration or represent a downstream consequence of disease-associated pathological processes

Digital PROMs, Psychometric Quality and Unsupervised Profiling in a Digital Cohort of University Students: Results from the DiCoBENE Study

Andrea Maugeri⁽¹⁾, Martina Barchitta⁽¹⁾, Marco Enea⁽²⁾, Domenica Matranga⁽²⁾, Antonella Agodi⁽¹⁾

 Andrea Maugeri 0000-0003-2655-8574

(1) Dipartimento di Scienze Mediche, Chirurgiche e Tecnologie Avanzate "GF Ingrassia", Università degli Studi di Catania, Italy

(2) Dipartimento di Promozione della Salute, Materno-infantile, Medicina interna e specialistica di eccellenza, Università di Palermo, Italy

CORRESPONDING AUTHOR: Andrea Maugeri, Dipartimento di Scienze Mediche, Chirurgiche e Tecnologie Avanzate "GF Ingrassia", Università degli Studi di Catania, Italy. E-Mail: andrea.maugeri@unict.it

INTRODUCTION

Digital cohorts represent a relevant infrastructure for the remote and scalable collection of patient-reported outcome measures (PROMs). In university populations, these tools allow for a multidimensional assessment of psychophysical well-being, including sleep quality, perceived stress, anxiety and depressive symptoms, and quality of life. Beyond the evaluation of psychometric properties, the characterization of the latent structure of outcomes through unsupervised approaches is also of particular relevance.

OBJECTIVES

To describe the baseline results of the DiCoBENE study (Digital Cohort per la Valutazione del BENEssere degli Studenti Universitari), a digital cohort of university students, pursuing three objectives: to assess data completeness and internal reliability of digital PROMs; to quantify the frequency of adverse outcomes and the associations among the measured domains; and to explore the multidimensional structure of outcomes using latent profile analysis (LPA), principal component analysis (PCA), and network analysis.

METHODS

DiCoBENE is an observational, prospective, repeated-measures digital cohort launched in Catania. At baseline, the Pittsburgh Sleep Quality Index (PSQI), Perceived Stress

Scale-10 (PSS-10), Generalized Anxiety Disorder-7 (GAD-7), Patient Health Questionnaire-9 (PHQ-9), and WHO-QOL-BREF were administered digitally. Analyses included descriptive statistics, Cronbach's alpha, McDonald's omega, Spearman correlations, and multivariate analyses using LPA, PCA on standardized scores, and network analysis based on partial correlations.

RESULTS

The baseline sample included 442 participants, with completion rates of 84.2% for GAD-7, PSS-10, and WHO-QOL-BREF domains, 83.9% for PHQ-9, and 83.7% for PSQI. The frequency of adverse outcomes was high: poor sleep quality (PSQI > 5) was observed in 58.8%, moderate-to-severe anxiety (GAD-7 ≥ 10) in 46.0%, moderate-to-severe depressive symptoms (PHQ-9 ≥ 10) in 29.5%, moderate stress in 54.0%, and high stress in 19.8%. Internal reliability was good to excellent for GAD-7 ($\alpha = 0.877$; $\omega = 0.881$), PHQ-9 ($\alpha = 0.818$; $\omega = 0.820$), and PSS-10 ($\alpha = 0.929$; $\omega = 0.931$), and moderate for PSQI ($\alpha = 0.654$; $\omega = 0.681$). Correlations among domains were consistent with theoretical expectations. In particular, PSS-10 correlated with PHQ-9 ($\rho = 0.759$) and GAD-7 ($\rho = 0.736$), while PHQ-9 showed an inverse correlation with the WHOQOL-BREF psychological domain ($\rho = -0.578$). Multivariate analyses, conducted on 365 complete observations, identified an optimal four-profile latent solution through LPA, interpretable as a gradient of psychophysical burden. PCA showed that the first principal component explained 55.6% of the total variance, while the

first two components explained 67.4% overall. Network analysis revealed a coherent structure, with greater centrality for WHOQOL-BREF Physical, PHQ-9, WHOQOL-BREF Psychological, and PSS-10.

CONCLUSIONS

These findings document the feasibility and methodological robustness of a multidimensional assessment of psychophysical well-being using digital PROMs in a university cohort. The high frequency of adverse outcomes, the adequate reliability of most instruments, and the convergence of findings obtained through LPA, PCA, and network analysis indicate that sample heterogeneity can be described primarily along a continuum of psychophysical burden rather than through distinct phenotypes. These findings support the integrated use of psychometric and unsupervised approaches in the analysis of digital PROMs in complex observational studies.

Violence against Healthcare Workers: Key Characteristics of the Phenomenon in the Marche Region in 2025

Natascia Puliani⁽¹⁾, Edlira Skrami^(1,2), Sonia Tonucci⁽³⁾, Anna Carrabs Valleverdina⁽³⁾, Domenico Lombardi Fortino⁽³⁾, Rosaria Gesuita^(2,4), Flavia Carle⁽²⁾

(1) Postgraduate School of Medical Statistics and Biometry, Department of Biomedical Sciences and Public Health, Politecnica delle Marche University, Ancona, Italy

(2) Centre for Epidemiology, Biostatistics and Medical Informatics, Politecnica delle Marche University, Ancona, Italy

(3) Regional Health Agency Marche Region, Ancona, Italy

(4) IRCCS INRCA, Ancona, Italy

CORRESPONDING AUTHOR: Natascia Puliani, Postgraduate School of Medical Statistics and Biometry, Department of Biomedical Sciences and Public Health, Politecnica delle Marche University, Ancona, Italy. E-Mail: n.puliani@pm.univpm.it

INTRODUCTION

Workplace violence (WPV) in healthcare settings is a growing social and structural problem. However, the lack of adequate methods for active monitoring and surveillance, based on systematic incident reporting, makes it difficult to quantify the extent of the problem and assess the effectiveness of measures aimed at preventing it.

OBJECTIVES

The aim of this study was to estimate the risk of violent incidents against healthcare workers in health care services in the Marche Region and evaluate its association with professional and socio-demographic variables.

METHODS

In this cross-sectional study, we used questionnaires from personnel of Regional Health System involved in the program monitoring violent incidents in the healthcare setting as required by Ministry of Health and collected by Regional Centre for Health Risk Management and Patient Safety of the Marche region in 2025. The survey form for incidents of violence was developed by the National Observatory on the Safety of Healthcare and Social-Health Professions (ONSEPS). Workers' demographic and professional characteristics as well as the types of incidents were summarized using absolute and percentage frequencies. The overall and by professional role (Physicians, Nurses, and non-clinical

support staff) risk of violent incidents (per 1000 workers), with 95% CI, were estimated as the ratio between the number of workers reporting at least one WPV incidents and the total number of workers in service during 2025. The total number of employees in the Regional Health System in 2025 was derived from the Annual Staff Report of the Marche Region by the New Health Information System. The RRs, with 95% CI, were calculated for comparing the violent incidents across the professional role. All data were processed in compliance with the European (GDPR, EU 2016/679) and national privacy laws (D.lgs. 196/2003 and subsequent amendments).

RESULTS

In 2025, 397 healthcare workers reported violence, of whom 69.5% were women, 31.0% aged 50-59 years, and 71.8% nurses. Violence was more frequently verbal (71.0%), more frequently involved patient-on staff violence (66.9%) and occurred more frequently in the Emergency Department (25.1%).

The WPV overall risk was 17.9 ‰ (95% CI 16.2 - 19.7), 11.9 ‰ (95% CI 8.2 - 15.5) for doctors and 30.3 ‰ (95% CI 26.8 - 33.8) for nurses. Nurses had 2.6 times higher risk of experiencing violence than doctors (RR = 2.6, 95% CI 2.5 - 2.7) and more than four times higher than non-clinical support staff (RR = 4.1, 95% CI 4.0 - 4.3), while doctors were 60% more likely to experience violence than non-clinical support staff (RR = 1.6, 95% CI 1.6 - 1.7).

CONCLUSION

Our findings are consistent with the latest ONSEPS report, based on data collected in 2022–23. According to that data, the Marche Region has the lowest risk rates compared to those recorded in the northern and central regions but has not yet implemented specific systems for identifying or recording cases. It is necessary to continue implementing prevention and management strategies aimed at countering the belief among workers that such accidents constitute an inevitable “occupational hazard.”

Predicting Methodological Quality of Non-Profit Clinical Trial Protocols: A Multicentre Extension

Giulia Gambini^(1,3), Daniela Bencardino⁽²⁾, Lorenza Sanfilippo⁽¹⁾, Eleonora Fresi⁽³⁾, Annalisa De Silvestri⁽³⁾, Catherine Klersy⁽³⁾, Virginia Valeria Ferretti⁽³⁾, Riccardo Braggion⁽⁴⁾, Elvia Malo⁽⁴⁾, Federico Giovanni Gobbi^(5,6)

 Giulia Gambini 0009-0003-0795-7972

(1) Scientific Directorate, Biostatistics Service, IRCCS Sacro Cuore Don Calabria Hospital, 37024 Negrar di Valpolicella, Verona, Italy

(2) Scientific Directorate, Data Management Service, IRCCS Sacro Cuore Don Calabria Hospital, 37024 Negrar di Valpolicella, Verona, Italy

(3) Biostatistics Unit, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy

(4) Clinical Trials Research Unit, IRCCS Sacro Cuore Don Calabria Hospital, 37024 Negrar di Valpolicella, Verona, Italy

(5) Department of Infectious, Tropical Diseases and Microbiology, IRCCS Sacro Cuore Don Calabria Hospital, 37024 Negrar di Valpolicella, Verona, Italy

(6) Department of Clinical and Experimental Sciences, University of Brescia, 25123 Brescia, Italy

CORRESPONDING AUTHOR: Giulia Gambini, Scientific Directorate, Biostatistics Service, IRCCS Sacro Cuore Don Calabria Hospital, 37024 Negrar di Valpolicella, Verona, Italy. E-Mail: giulia.gambini@sacrocuore.it

INTRODUCTION

The SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) Statement has long been the international reference for the content of clinical trial protocols, ensuring transparency, methodological rigour and ethical soundness [1]. The recent release of SPIRIT 2025 has updated and expanded its recommendations [2]. Nonetheless, numerous studies continue to document suboptimal adherence, particularly regarding study design and statistical analysis. Assessing real-world adherence helps identify recurrent shortcomings and areas requiring targeted support or training.

OBJECTIVES

In a previous single-centre study we assessed whether it is possible to predict which protocols show high adherence to the methodological SPIRIT items, identifying study-level characteristics associated with quality. The main limitation was the single-centre nature of the data. The present work extends the analysis to a multicentre design, combining protocols from two Italian research centres: Fondazione IRCCS Policlinico San Matteo (Pavia) and IRCCS Sacro Cuore Don Calabria (Negrar di Valpolicella, Verona), aiming to confirm the predictors of adherence and improve their generalisability.

METHODS

Design and methodological features of no-profit interventional protocols submitted between 2021 and 2025 to the reference Ethics Committees of the two centres were collected and recorded in a centralised REDCap (Research Electronic Data Capture) database. Adherence was assessed on the SPIRIT 2013 items concerning methods, design, data collection and management, and analysis (items 9-21b); the 2013 version was used because all protocols predate SPIRIT 2025. Each item was scored as fulfilled or not, and an individual score was computed as the number of satisfied items, dichotomised at the median into good and poor adherence. For the predictive objective, machine learning algorithms (Random Forest, eXtreme Gradient Boosting, Boosted Logistic Regression) were applied to a pool of candidate predictors; performance was evaluated through accuracy, area under the ROC curve (AUC) and F1 score. Variable importance was then summarised and the most influential predictors were entered into a multivariable logistic regression model, reporting odds ratios and 95% confidence intervals. Covariates included sponsorship, methodological features, submission timing and thematic area. Analyses were performed using Stata 19 and R 4.4.3.

RESULTS

Data collection and harmonisation across the two centres are currently ongoing. The IRCCS Sacro Cuore Don Calabria

contributed 107 no-profit interventional protocols submitted to its reference Ethics Committee between 2021 and 2025. Combined with the 132 protocols of the previous single-centre study, the pooled cohort reaches 239 protocols. This larger and more heterogeneous sample is expected to produce more stable predictive models and to enable a formal assessment of the reproducibility of the predictors identified in the single-centre analysis. We expect to confirm the availability of statistical and methodological support as a factor associated with good adherence, while the relative contribution of the remaining candidate predictors (sponsorship, methodological features, submission timing and thematic area) will be clarified in the pooled cohort. The machine learning algorithms are expected to discriminate good from poor adherence and to converge on a consistent subset of influential predictors, which the multivariable logistic regression should retain as independent. The final estimates will be presented.

CONCLUSIONS

The multicentre extension addresses the main limitation of the previous work, strengthening the external validity of the findings. By integrating conventional statistical approaches with machine learning models, the study aims to clarify which protocol features predict high adherence to SPIRIT and to confirm the role of multidisciplinary teams, including biostatistics, in improving methodological quality. The findings are expected to provide evidence to inform training strategies aimed at improving the quality of clinical trial protocols.

Annual Variations in Asthma Attacks in the Italian Adult Population

Tibor András Nyári⁽¹⁾, Mamoon Arshad⁽²⁾, Alessandro Marcon⁽²⁾, Tamás Lantos⁽¹⁾, Francesca Locatelli⁽²⁾, Pierpaolo Marchetti⁽²⁾, Lorenzo Di Spazio⁽²⁾, Salvatore Battaglia⁽³⁾, Maria Beatrice Bilò⁽⁴⁾, Roberto Bono⁽⁵⁾, Angelo Guido Corsico⁽⁶⁾, Nicola Murgia⁽⁷⁾, Pietro Pirina⁽⁸⁾, Giuseppe Verlato⁽²⁾

 Giuseppe Verlato 0000-0001-5262-8818

(1) Department of Medical Physics and Informatics, Albert Szent-Györgyi Medical School, University of Szeged – Szeged, Hungary

(2) Unit of Epidemiology and Medical Statistics, Department of Diagnostics & Public Health, University of Verona – Verona, Italy

(3) Division of Respiratory Medicine, PROMISE Department, Paolo Giaccone University Hospital, University of Palermo – Palermo, Italy

(4) Department of Clinical and Molecular Sciences, Università Politecnica Delle Marche – Ancona, Italy

(5) Department of Public Health and Pediatrics, University of Turin – Turin, Italy

(6) Department of Internal Medicine and Therapeutics, University of Pavia – Pavia, Italy

(7) Department of Environmental and Preventive Sciences, University of Ferrara – Ferrara, Italy

(8) Clinical and Interventional Pulmonology, University Hospital Sassari (AOU) – Sassari, Italy

CORRESPONDING AUTHOR: Giuseppe Verlato, Unit of Epidemiology and Medical Statistics, Department of Diagnostics & Public Health, University of Verona – Verona, Italy. E-Mail: giuseppe.verlato@univr.it

INTRODUCTION

Seasonal variation plays a crucial role in the occurrence and exacerbation of allergic diseases, yet real-world, large-scale epidemiological evidence remains limited. Asthma attacks can be triggered by pollen and often worsen during the pollen season. However, seasonal patterns may vary across different climatic areas.

OBJECTIVES

The primary aim of the present study was to assess the seasonality of asthma attacks in Italian adults.

METHODS

We analyzed GEIRD (Gene Environment Interaction in Respiratory Diseases) data, collected from 2008 to 2014 in seven Italian centers, three (Verona, Pavia, Turin) located in the Po Valley (sub-continental area) characterized by sub-continental climate, and four (Sassari, Ancona, Terni, Palermo) in the Mediterranean area.

Seasonality was assessed using a Poisson regression framework incorporating Fourier terms (extended Stolwijk

model). Both single-peak (unimodal) and double-peak (bimodal) cyclic trends were investigated. The model selection was based on both pseudo R-squared values (R^2 ; explanatory power) and the Goodness-of-Fit (GoF) test. The model providing the best overall fit was retained.

RESULTS

The study included 204 subjects (137 and 67 from sub-continental and Mediterranean areas, respectively) who reported having suffered from asthma attacks for a median of 4 months in the previous year. Both unimodal and bimodal Poisson regression models were fitted to this cyclic pattern of asthma attacks. All unimodal models had acceptable GoF. However, the bimodal models had higher pseudo R-squared values with acceptable GoF. Overall, the seasonal pattern of asthma attacks was characterized by a main peak in April and a secondary peak in September ($p < 0.001$; $R^2 = 0.503$; $\text{GoF} = 0.95$).

In the sub-continental area, a primary peak was detected in May and a secondary one in September ($p < 0.001$; $R^2 = 0.42$; $\text{GoF} = 0.89$), with relative increases of 34% and 19%, respectively. In contrast, the seasonality of asthma attacks in the Mediterranean area was characterized by the first peak in May, whereas the second peak occurred in October

($p=0.004$; $R^2=0.36$; $GoF=0.99$) with relative increases of 37% and 22%, respectively.

CONCLUSION

Asthma attacks in Italy mainly occur during spring. The bi-modal cyclic pattern explained 50% of the monthly variability in asthma attacks. Nevertheless, a weaker secondary peak of asthma attacks was observed in autumn. The sub-continental (Po Valley) and Mediterranean areas had slightly different autumn peaks, which may reflect differences in ragweed and Urticaceae exposure, and warrant further investigation. In addition, the modeling approach identified significant seasonal variation.

Socio-Material Vulnerability and Youth Mental Disorders in Lombardy: A Bayesian Mapping Study

Asia Filosa⁽¹⁾, Matteo Franchi^(2,3), Maria Antonella Costantino⁽⁴⁾, Renato Borgatti⁽⁵⁾, Martina Maria Mensi⁽⁵⁾, Elisa Fazzi⁽⁶⁾, Jessica Galli⁽⁶⁾, Vanna Poli⁽⁷⁾, Olivia Leoni⁽⁷⁾, Ivan Limosani⁽⁷⁾, Laura Alessia Reale⁽⁴⁾, Giovanni Corrao^(3,7,8), Cristina Montomali⁽¹⁾

 Asia Filosa 0009-0005-7381-0869

(1) Biostatistics and Clinical Epidemiology Unit, Department of Public Health, Experimental and Forensic Medicine, University of Pavia, Pavia, Italy

(2) Biostatistics, Epidemiology and Public Health Unit, Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy

(3) National Centre for Healthcare Research and Pharmacoepidemiology, University of Milano-Bicocca, Milan, Italy

(4) Child and Adolescent Neuropsychiatric Unit (UONPIA), Foundation IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy

(5) Child Neuropsychiatry Unit, Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy

Welfare Department, Lombardy Region, Milan, Italy – Child Neuropsychiatry Unit, IRCCS Mondino Foundation, Pavia, Italy

(6) Department of Clinical and Experimental Sciences, University of Brescia, Brescia, Italy – Child Neurology and Psychiatry Unit, ASST Spedali Civili of Brescia, Brescia, Italy

(7) Welfare Department, Lombardy Region, Milan, Italy

(8) Professor Emeritus of Medical Statistics, University of Milano-Bicocca, Milan, Italy

CORRESPONDING AUTHOR: Asia Filosa, Biostatistics and Clinical Epidemiology Unit, Department of Public Health, Experimental and Forensic Medicine, University of Pavia, Pavia, Italy. E-mail: asia.filosa@unipv.it

INTRODUCTION

Socio-economic inequalities are well-documented drivers of a mental health gradient, with children and adolescents from low socio-economic backgrounds being disproportionately vulnerable to mental disorders. Evidence suggests that childhood deprivation not only increases the prevalence of mental health conditions but also restricts access to routine healthcare, contributing to poorer health and economic outcomes in adulthood. While this association is well-established, small-area spatial analyses examining the geographic distribution of mental disorders in children and adolescents in relation to composite deprivation indices remain scarce, particularly in the Italian context.

OBJECTIVES

To assess whether the geographical distribution of children and adolescents with at least one hospitalisation or Emergency Department (ED) visit for mental disorders in Lombardy is associated with the Social and Material Vulnerability Index (SMVI).

METHODS

A population-based spatio-temporal ecological study was conducted using regional administrative healthcare data from the Lombardy (2016–2022). The study population included all residents aged 0–18 years as of December 31st of each year. A mental disorder case was defined as any individual with at least one ED visit or hospital discharge recording an ICD-9-CM code in the range 290–319 in any diagnostic field during the reference year. The SMVI, developed by the Italian National Institute of Statistics (Istat), was categorized into quintiles ranging from Q_1 (lowest vulnerability) to Q_5 (highest vulnerability). Observed and expected case counts were modelled using Bayesian hierarchical scaled Besag–York–Mollié (BYM2) models to estimate age- and sex-adjusted area-specific relative risks, interpretable as smoothed SMRs. Three model specifications were compared: a purely spatial model, a spatio-temporal model with a Random Walk of order 1 (RW1) temporal component, and a spatio-temporal model with a Knorr–Held Type I interaction. Area-specific relative risks (RR) and posterior probabilities (PP) were estimated. Analyses were performed in R using INLA framework.

RESULTS

Findings were consistent across all three Bayesian BYM2 model specifications. Compared with the lowest SMVI quintile, only areas in the highest vulnerability quintile showed a credible association with increased risk of mental disorders diagnoses among youths, with posterior relative risks of 1.096 (95% CrI: 1.017-1.182), 1.097 (1.018-1.182), and 1.101 (1.019-1.190) in the spatial, spatio-temporal, and spatio-temporal interaction models, respectively. No statistically credible associations were found for intermediate quintiles. After adjusting for SMVI, residual spatial risk showed a marked west-east gradient across the region. Areas with high posterior probability of excess risk ($PP > 0.80$) were concentrated in the Milan metropolitan area, the north-western pre-alpine belt, the Pavia area, and a few localized hotspots near Lake Garda, whereas central-eastern Lombardy and the alpine area of Sondrio showed consistently lower-than-expected risk ($PP < 0.20$).

CONCLUSIONS

The study highlights that the SMVI captures a meaningful share of the phenomenon's variability, showing an increased risk of mental disorders in areas with the highest vulnerability. However, the persistence of strong spatial clusters in the residual risk maps suggests that the SMVI does not fully explain the heterogeneous territorial distribution of cases in Lombardy. This indicates the presence of additional local, environmental, or healthcare supply and access determinants that warrant further investigation.

Long-Term Global Trends and Sex Disparities in Homicide Mortality: A Comparative Analysis of 39 Countries over 40 Years

Linia Patel⁽¹⁾, Margherita Pizzato⁽²⁾, Silvia Mignozzi⁽¹⁾, Eva Negri⁽²⁾, Carlo La Vecchia⁽¹⁾, Gianfranco Alicandro^(3,4)

 Gianfranco Alicandro 0000-0002-0430-2714

(1) Department of Clinical Sciences and Community Health, Università degli Studi di Milano, Milano, Italy

(2) Department of Medical and Surgical Sciences, Alma Mater Studiorum Università di Bologna, Bologna, Italy.

(3) Department of Pathophysiology and Transplantation, Università degli Studi di Milano, Milano, Italy

(4) Mother and Child Department, Cystic Fibrosis Center, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milano, Italy

CORRESPONDING AUTHOR: Gianfranco Alicandro, Department of Pathophysiology and Transplantation, Università degli Studi di Milano, Milano, Italy, Via Commenda 9, 20122 Milano, Italy. Tel.: +39 0255032456. e-mail address: gianfranco.alicandro@unimi.it

INTRODUCTION

The homicide rate is widely regarded as a key indicator of societal violence. Although homicides account for less than 1% of all deaths globally and have generally declined over time, they remain a critical public health and social concern. Homicide continues to attract sustained attention from both the media and the public, contributing to a perception that violence is increasing, despite empirical evidence.

OBJECTIVES

This study aims to analyse long-term trends in homicide mortality across multiple countries, with particular attention to variations by sex and geographic region. Specifically, it aims to evaluate whether the perceived increase in insecurity and violence, especially violence against women, is supported by empirical evidence derived from homicide data.

METHODS

We extracted data on homicide deaths from the WHO Mortality Database (1979 to the latest available year) and mid-year population estimates from the United Nations Population Division. Countries with at least medium-quality data and a minimum of five years of observations were included. Age-standardized mortality rates (ASRs) were calculated by sex, and male-to-female ratios were used to assess sex differ-

ences. Annual percent changes (APCs) were estimated using joinpoint regression models.

RESULTS

A total of 39 countries were included in the analysis. Among males, ASRs ranged from 0.22 per 100,000 person-years in Japan to 46.6 in Colombia; among females, rates ranged from 0.17 in Japan to 4.86 in Guatemala. Most countries showed declining trends over time, although recent increases were observed in France among males (APC 2017–2022: +12.8%, 95% CI: 5.9–20.1), Sweden among males (APC 2011–2023: +5.6%, 95% CI: 3.1–8.2), and in the United States among both males (APC 2014–2022: +5.2%, 95% CI: 3.1–7.3) and females (APC 2014–2022: +4.4%, 95% CI: 2.8–5.9).

In Europe, the median male-to-female ASR ratio was 1.9, whereas higher ratios were observed in the United States (4.0), South Africa (7.2), and Latin America (ranging from 3.2 to 12.0).

Italy was among the countries with the lowest homicide mortality rates in both males (0.46 per 100,000 person-years in 2021) and females (0.26 per 100,000 person-years in 2021), with a male-to-female ASR ratio of 1.8.

CONCLUSIONS

Homicide mortality has declined in both sexes across most countries over the past four decades. Nevertheless, substan-

tial geographical disparities persist, with particularly high burdens observed in Latin America, South Africa, and the United States. Recent increases reported in several high-income countries underscore the fragility of these gains and highlight the need for sustained surveillance and context-specific prevention strategies. The persistence of marked sex differences, where males consistently account for the majority of victims, suggests that the underlying drivers of male vulnerability to lethal violence remain largely unchanged. Although female homicide rates have remained relatively stable overall, they exhibit a distinct epidemiological profile, often associated with intimate partner violence. This pattern calls for continued, targeted policy responses addressing gender-specific risk factors.

Effectiveness on Health and Well-Being of an eHealth-Based Rehabilitation Program in Older People with Acute Heart Failure and/or Acute Coronary Syndrome: The Flip Side of the Delphic Study

Elena Tagliabue⁽¹⁾, Andrea Passantino⁽²⁾, Roberto FE Pedretti⁽¹⁾, Gabriele M Maglio⁽¹⁾, Eros Patuzzo⁽¹⁾, Denise Bruno⁽¹⁾, Francesco Prattichizzo⁽¹⁾, Marta Rigoni⁽³⁾, Fabiola Olivieri^(4,5), Gaia Cattadori⁽¹⁾

 Elena Tagliabue 0000-0002-8153-2599

(1) IRCCS MultiMedica, Milan, Italy

(2) Cardiac Rehabilitation Unit of Bari Institute, Istituti Clinici Scientifici Maugeri IRCCS, Bari, Italy

(3) Department of Biomedical, Surgical and Dental Sciences, University of Milan, Milan, Italy

(4) Clinic of Laboratory and Precision Medicine, IRCCS INRCA, Ancona, Italy

(5) Department of Clinical and Molecular Sciences, Università Politecnica Delle Marche, Ancona, Italy

CORRESPONDING AUTHOR: Elena Tagliabue, IRCCS MultiMedica, Milan, Italy. E-Mail: elena.tagliabue@multimedica.it

INTRODUCTION

Frailty is a clinical syndrome highly prevalent among elderly individuals with cardiovascular disease and is characterized by a progressive decline in the physiological reserves of multiple organ systems. Preliminary evidence suggests that frailty may not represent an irreversible condition, raising the possibility that targeted interventions could improve resilience and potentially alter the course of the syndrome. Cardiac rehabilitation (CR) programs are established effective tools to improve cardiovascular prognosis and quality of life. However, few studies have explored the efficacy of eHealth, home-based CR interventions on frailty. The DELPHIC (eLDerly hEart faiLure Prevention reHAbilitation biomarkers prognostiC) study was designed to evaluate the effectiveness on frailty status of an eHealth-based, home-delivered CR program in older adults with cardiovascular disease.

OBJECTIVES

Primary endpoint was to evaluate frailty, in terms of prevalence and mean Essential Frailty Toolset (EFT) score, after 3-6 months of follow-up. Secondary outcomes regarded other aspects of life measured by scales/scores as depression (Geriatric Depression Scale - GDS), nutrition (Mini Nutritional

Assessment - MNA), quality of life (EuroQoL score), daily activity (Barthel index), therapeutic adherence (Morisky Medication Adherence Scale - MMAS) and comorbidities (Cumulative Illness Rating Scale - CIRS).

METHODS

We designed a multicenter, randomized, parallel-group trial enrolling people ≥ 65 years recently hospitalized for acute heart failure and/or acute coronary syndrome. Participants were randomized 1:1 to an eHealth home-based tele-rehabilitation program or the usual care. Intervention program included 3 rehabilitation sessions per week, 1 weekly telephone consultation with a research nurse who remotely monitored patients' vital signs (i.e. weight, blood pressure and ECG), and 14 psychoeducational and nutritional sessions (7 with a psychologist and 7 with a nutritionist).

RESULTS

206 total patients were equally distributed to intervention or usual care. At the end of the follow-up, the prevalence of frailty was significantly lower in the intervention group than in the control group (2.6% vs 11.9%, $p=0.04$), accompanied by a significant reduction in the mean EFT score ($p<0.01$). Fur-

thermore, Barthel index significantly increased in the intervention group than in control group ($p=0.01$). Psychoeducational intervention led to a significantly lower GDS ($p<0.01$) and higher EuroQol ($p<0.01$) in treated subjects, indicating an improved self-perceived quality of life. In addition, the nutritional/educational intervention had a positive and measurable effect on participants, resulting in a significantly increased MNA score compared with the control group ($p<0.01$).

CONCLUSIONS

Telemedicine represents a form of continuity of care, allowing patients to feel monitored and supported from their homes, thereby reducing the sense of abandonment that may follow hospital discharge. The opportunity to receive health-care services at home may contribute to making patients feel more comfortable during the intervention. Furthermore, the involvement of caregivers, who are often excluded from hospital-based care settings, may be crucial in remote rehabilitation pathways, providing practical assistance and fostering greater engagement in the care process.

Migration History and Self-Reported Health among Migrants in Italy: Evidence from the MIGHTY Project

Ayman Lakhla⁽¹⁾, Edlira Skrami⁽¹⁾, Elisa Barbiano Di Belgiojoso⁽²⁾, Flavia Carle⁽¹⁾, Rosaria Gesuita⁽¹⁾, on behalf of the MIGHTY Project working group*

(1) Center of Epidemiology, Biostatistics and Medical Information Technology, Polytechnic University of Marche, Ancona, Italy

(2) Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milano, Italy

*MIGHTY Project working group: Elisa Barbiano di Belgiojoso, Rosaria Gesuita, Stefania M.L. Rimoldi, Dario Pescini, Paolo Berta, Nadia Solaro, Paola Chiodini, Alessandro Avellone, Antonella Zambon, Flavia Carle, Edlira Skrami, Andrea Faragalli, Marica Iommi, Silvana Antonietta Romio

CORRESPONDING AUTHOR: Ayman Lakhla, Center of Epidemiology, Biostatistics and Medical Information Technology, Polytechnic University of Marche, Ancona, Italy. E-Mail: ayman.lakhla01@gmail.com

INTRODUCTION

The scientific literature has extensively documented the so-called “immigrant health paradox,” whereby newly arrived immigrants often exhibit better health status than native-born populations with comparable demographic and socioeconomic characteristics. However, this advantage tends to diminish over time leading to a process of “unhealthy assimilation” in which health status deteriorates with longer residence in the host country.

OBJECTIVES

Building on this framework, this study investigates whether a similar dynamic exists within the migrant population in Italy. Specifically, after adjusting for socio-demographic characteristics, we examine whether newly arrived immigrants maintain a health advantage compared with more established migrants and second-generation individuals (those born in Italy of foreign origin). Additionally, this study aims to identify the main determinants associated with self-reported health.

METHODS

This study analysed cross-sectional data from the MIGHTY project survey, conducted in the Lombardy and Marche regions, between 2024 and 2025. The sample was stratified into four groups based on migration status and length of residence: recent immigrants (less or equal than 5 years), established immigrants (6–10 years), long-term immigrants (more

than 10 years), and second-generation individuals born in Italy to foreign parents. A binary logistic regression model was used to assess the association between self-reported health and various socio-demographic, socio-economic, and health-related covariates. For this analysis, self-reported health was dichotomized from its original categories into good and not good, defining good health as the primary outcome. Health related covariates included the number of chronic diseases and the presence of depressive symptoms assessed through the PHQ8 questionnaire. All analyses were performed using R-studio and results were reported as Odds Ratio (OR) and 95% Confidence intervals (95%CI).

RESULTS

3625 (out 3660 recruited) participants with complete data on self-reported health status were included in this analysis, 75% reported good health status, 80% younger than 50 years. The regression model showed a non-linear relationship between migration status and self-reported health. Using long-term migrants as the reference category, second-generation individuals (OR = 1.93, 95%CI: 1.20–3.20) and established immigrants (OR = 1.51, 95%CI: 1.05–2.02) showed a significantly higher probability of reporting good health. Comparison between recent immigrants and long-term residents was not statistically significant (OR = 1.34, 95%CI: 0.88–2.03). Regarding other main determinants, the probability of reporting good health was significantly reduced by increasing age(years) (OR = 0.96, 95%CI: 0.95–0.97), depressive symptoms (OR = 0.86, 95%CI: 0.83–0.89), financial struggles compared to the ability to cover expenses

(OR = 0.54, 95%CI: 0.41–0.71), and the presence of chronic diseases (OR = 0.19, 95%CI: 0.14–0.24).

CONCLUSIONS

The findings provide partial support to the unhealthy assimilation hypothesis among immigrants in Italy. Rather than indicating a progressive deterioration in health with increasing length of residence, they suggest that the relationship between migration history and self-reported health is non-linear. Migration status, increasing age, depressive symptoms, financial difficulties, and chronic conditions were independently associated with a lower likelihood of reporting good health. These findings highlight that both migration-related and socio-economic and health-related factors should be considered when addressing health inequalities among migrant populations.

Knowledge Gaps and Digital Readiness among Patients with COPD: Preliminary Results from the European JARED Project

Arianna Baldassari⁽¹⁾, Ayman Lakhla⁽²⁾, Andrea Faragalli⁽³⁾, Martina Bonifazi⁽⁴⁾, Brigid Unim⁽⁵⁾, Flavia Carle⁽²⁾, Edlira Skrami^(1,2), Jelena Ostojic⁽⁶⁾, Andrea Vukić Dugac⁽⁶⁾, Rosaria Gesuita^(2,7)

(1) Postgraduate School of Medical Statistics and Biometry, Department of Biomedical Sciences and Public Health, Politecnica delle Marche University, Ancona, Italy

(2) Centre for Epidemiology, Biostatistics and Medical Informatics, Politecnica delle Marche University, Ancona, Italy

(3) Department of Life Sciences, Health and Healthcare Professions, Link Campus University, Rome, Italy

(4) Interstitial Lung Diseases, Pleural Diseases and Bronchiectasis Unit, Azienda Ospedaliero-Universitaria "Ospedali Riuniti", Ancona, Italy

(5) Department of Cardiovascular, Endocrine-metabolic Diseases and Aging, Istituto Superiore di Sanità, Italy

(6) Clinic for respiratory diseases, University Hospital Centre Zagreb, Croatia

(7) IRCCS INRCA, Ancona, Italy

CORRESPONDING AUTHOR: Arianna Baldassari, Postgraduate School of Medical Statistics and Biometry, Department of Biomedical Sciences and Public Health, Politecnica delle Marche University, Ancona, Italy. E-Mail: a.baldassari@pm.univpm.it

INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is a progressive chronic respiratory disease and major cause of morbidity. Critical objectives of COPD management are to preserve patients' functional status and quality of life.

Knowledge and awareness on the disease are critical for self-management but evidence on these aspects remains limited.

OBJECTIVES

This study, conducted within the European project 'Joint Action on Chronic Respiratory Diseases' (JARED), aimed to assess COPD-related knowledge and self-management awareness and willingness to use digital tools in patients with COPD.

METHODS

This cross-sectional study included patients attending routine outpatient visits at the Pulmonology Department of the AOU Ospedali Riuniti delle Marche (Italy) and the Department of Pulmonary Diseases Jordanovac, University Hospital Centre Zagreb (Croatia). Patients completed a questionnaire consisting of patients' knowledge on the disease in general (5 items), risk factors (4 items), pharmacological treatment (5

items), non-pharmacological treatments (4 items), and acute exacerbations (4 items). Two additional sections assessed willingness to use digital tools (4 items) and self-management awareness (5 items). Knowledge was dichotomized at optimal level if patients gave at least three or four correct answers in sections with four items or five items respectively, according to a predefined cut-off. Likewise, willingness and awareness were set with at least three or four positive answers respectively. The associations between knowledge, willingness, awareness, and demographic and anamnestic characteristics were evaluated by Chi-square or Fisher exact test when appropriate.

RESULTS

Overall, 231 COPD patients were recruited, 53% were male, with a mean age of 68 years, 78% retired, 73% lived in urban areas, 37% had high education, 38% current smoker, 52% former smoker, mean smoking duration 36 years and 19 cigarettes per day, 66% with a disease duration ≥ 4 years.

The percentage of patients with an optimal knowledge ranged between 55% in COPD risk factors and 87% in acute exacerbations. Current or former smokers had more frequently an optimal general knowledge than a sub-optimal one (96% vs 86%, $p=0.009$); patients with disease duration ≥ 4 years had more frequently optimal knowledge in general, on risk factors, on non-pharmacological treatments than sub-optimal

(respectively, 71% vs 52%, $p=0.009$; 73% vs 56%, $p=0.011$; 69% vs 51%, $p=0.036$); optimal knowledge in non-pharmacological treatments was more frequent than sub-optimal when patients had high education (41% vs 19%, $p=0.010$).

Sixty percent of participants were willing to use digital tools. They were younger than those not willing (66 vs 71 years, $p<0.001$), more frequently employed (22% vs 7.6%, $p=0.006$), and more likely to live in urban areas (79% vs 66%, $p=0.031$).

Forty percent of participants were aware of own disease and self-management, and they had longer smoking duration than those who were not (38 vs 34 years on average, $p=0.035$).

CONCLUSIONS

Although overall COPD-related knowledge was satisfactory, substantial gaps persist in key aspects of COPD, particularly regarding disease risk factors and self-management awareness. Patients less inclined to use digital health tools may require tailored educational strategies. These findings support the need for tailored educational interventions to improve disease management and promote patient engagement in COPD care.

The Collection of PROMs as Part of Real-World Federated Analytics: Lessons Learned from the REWINDER Project in Diabetes in Emilia-Romagna

Fabrizio Carinci⁽¹⁾, Rossella Messina⁽²⁾, Martina Mencarelli⁽²⁾, Seyedmohammad Khoshroo⁽³⁾, Massimo M. Michelini⁽³⁾, Alessandra Dei Cas^(4,5), Elisa Iezzi^(4,5), Raffaella Aldigeri^(4,5), Paolo Di Bartolo⁽⁶⁾

 Fabrizio Carinci 0000-0002-7524-227X

(1) Unicamillus International Medical University, Rome, Italy

(2) AUSL Romagna, Ravenna, Italy

(3) AUSL Reggio Emilia, Italy

(4) University of Parma, Parma, Italy

(5) University Hospital of Parma, Parma, Italy

(6) University of Bologna, Bologna, Italy

CORRESPONDING AUTHOR: Fabrizio Carinci, Unicamillus International Medical University, Rome, Italy. E-Mail: fabrizio.carinci@unicamillus.org

OBJECTIVES

Patient-reported outcome measures (PROMs) have become an essential component of value-based and person-centred diabetes care. Nevertheless, their routine implementation remains uncommon because longitudinal PROMs collection, linkage with clinical information, and privacy protection pose substantial organisational and technical challenges. Federated analytics has been proposed as a strategy to overcome these barriers by enabling distributed analyses without transferring identifiable patient-level data. The REWINDER study ("Real-World data and patient-reported outcomes IN Diabetes") aimed at implementing a regional federated infrastructure integrating PROMs into routine diabetes care and exploring scalability beyond the pilot setting.

METHODS

A federated infrastructure was implemented across three areas of Emilia-Romagna to combine clinical, administrative and PROMs data while preserving local data ownership. The design included the collection of target sample size of $n=150$ questionnaires in each area, using validated instruments of the ICHOM Diabetes Standard Set (PAID-5, PHQ-9 and WHO-5). Data were digitized locally using different means: Romagna and Reggio-Emilia used an existing application (Health Meeting and Smarty, respectively),

while Parma performed data entry in an Excel sheet. PROMs were linked locally to other data sources included in the REWINDER database through unique patient identifiers and analysed using the NeuBIRO federated statistical platform. We evaluated data completeness, linkage quality, temporal consistency and operational feasibility prior to assessing patient outcomes.

RESULTS

Baseline PROMs were available for $n=521$ individuals, mostly collected by Romagna (51.1%), followed by Reggio-Emilia (33.8%) and Parma (15.2%). However, the proportion of patients contributing complete longitudinal information decreased considerably after applying protocol requirements and successful linkage across data sources. The percentage of patients with a valid follow-up (between 4 and 10 months from baseline) was similar between Romagna (86.5%) and Parma (88.6%), but considerably lower in Reggio-Emilia (30.1%), reflecting organisational differences in data collection processes. Additional discrepancies emerged from asynchronous patient recruitment, incomplete availability of active patient lists and incomplete linkage of telephone follow-up interviews. Further results regarding the evaluation of linkage with demographical, clinical and population-based data available through the REWINDER cohort are currently in progress and will be available at the Conference.

CONCLUSIONS

The study demonstrates that federated collection and analysis of PROMs can be successfully embedded within routine diabetes care while maintaining data privacy. However, technical feasibility alone is insufficient to support sustainable large-scale implementation. Although federated analysis remained technically feasible despite these constraints, the experience demonstrated that organizational processes, rather than analytical methods, may represent the principal determinant of data quality and effective interoperability. Manual data collection procedures, heterogeneous local workflows and poorly synchronized follow-up substantially reduce the completeness and comparability of longitudinal patient-reported outcomes. A companion abstract presents the detailed results related to the statistical analysis of PROMs for the subset of people with Type 2 diabetes with valid follow-up.

Future federated infrastructures should prioritise native digital PROMs collection, automated linkage with clinical information systems, standardised operational protocols and integrated scheduling mechanisms. These organisational prerequisites are likely to have a greater impact on the quality and scalability of federated real-world evidence than further methodological refinements of distributed analytical models.

Psychological Burden and Clinical Risk Profiling in Type 2 Diabetes: Evidence from PROMs in the REWINDER Project

Fabrizio Carinci⁽¹⁾, Rossella Messina⁽²⁾, Martina Mencarelli⁽²⁾, Seyedmohammad Khoshroo⁽³⁾, Massimo M. Michelini⁽³⁾, Alessandra Dei Cas^(4,5), Elisa Iezzi^(4,5), Raffaella Aldigeri^(4,5), Paolo Di Bartolo⁽⁶⁾

 Fabrizio Carinci 0000-0002-7524-227X

(1) Unicamillus International Medical University, Rome, Italy

(2) AUSL Romagna, Ravenna, Italy

(3) AUSL Reggio Emilia, Italy

(4) University of Parma, Parma, Italy

(5) University Hospital of Parma, Parma, Italy

(6) University of Bologna, Bologna, Italy

CORRESPONDING AUTHOR: Fabrizio Carinci, Unicamillus International Medical University, Rome, Italy. E-Mail: fabrizio.carinci@unicamillus.org

OBJECTIVES

Patient-Reported Outcome Measures (PROMs) are rarely embedded in routine diabetes care, although they provide information not captured by clinical and administrative data alone. The REWINDER project aimed to integrate PROMs with real-world diabetes data from Emilia-Romagna to describe psychological burden in people with type 2 diabetes and explore its relationship with metabolic and clinical indicators, particularly glycaemic control and patient risk profiling.

METHODS

REWINDER is a multicentre observational study with a mixed retrospective/prospective design involving AUSL Romagna, AUSL Reggio Emilia and AOU Parma. A sample including N=521 people with Type 2 diabetes was enrolled by diabetes clinics to monitor the psychological evolution through the administration of three PROMs: PAID-5 for diabetes distress, PHQ-9 for depressive symptoms and WHO-5 for psychological well-being. Questionnaires collected at baseline were expected to be evaluated after 6 months, linked with demographic, clinical and service-related data, including HbA1c, disease duration, blood pressure, body mass index and lifestyle information. Descriptive analyses were conducted overall and according to HbA1c target status. A multivariable logistic regression model estimated the

probability of reporting a high total PROMs score, indicating psychological compromise, according to sex, age, disease duration and HbA1c target status.

RESULTS

A total of N=353 subjects included in the baseline cohort reported a valid follow-up between 4-10 months. Most participants were men (60.6%), prevalently aged between 50-59 years. Disease duration was 1-5 years in 34.3% and 6-15 years in 44.1%. The sample reported a clinical profile with relevant cardiometabolic and lifestyle difficulties: 54.1% had uncontrolled HbA1c (>7%), 69.1% had systolic blood pressure ≥ 130 mmHg and 66.0% had diastolic blood pressure ≥ 80 mmHg. Among those with lifestyle data, 25.2% reported no physical activity, 30.0% were current smokers and 24.7% former smokers. Approximately 80% were overweight or obese. PROMs identified substantial psychological burden: diabetes distress was present in 30.3%, depressive symptoms in 36.5% and poor psychological well-being in 22.4%. Diabetes distress and depressive symptoms were more frequent among participants outside glycaemic target. Being in HbA1c target was associated with a lower probability of reporting a high total PROMs score (OR=0.59; $p=0.0496$; 95% CI 0.35-0.99), whereas female sex was strongly associated with higher psychological risk (OR=2.67; $p=0.0002$; 95% CI 1.61-4.47). Age showed a small inverse association, while disease duration was not significant. Probability curves con-

firmed higher predicted risk among women and among those outside HbA1c target.

CONCLUSIONS

The REWINDER study provided a practical demonstration of how PROMs can help clinicians integrating patient experience and psychological risk assessment into diabetologic consultations. These findings support the inclusion of diabetes distress assessment in routine care to activate timely and targeted psychological, educational and integrated interventions, as recommended by recent international guidelines. The gender difference suggests areas of concern among women at higher risk. Further studies are needed to clarify the relationship between glycaemic values and psychological burden. A companion abstract discusses critical issues regarding the scalability of the approach within the complex federated infrastructure of REWINDER and the hurdles experienced in data collection and regular follow-up of PROMs.

Impact of Commercially-Available Hybrid Closed Loop Systems on Sleep Outcomes in Youths and Adults with Type 1 Diabetes: A Systematic Review and Meta-Analysis

Evelina Maines⁽¹⁾, Stefania Fanti⁽²⁾, Camilla Buffo⁽³⁾, Francesco Favaretto⁽³⁾, Roberto Franceschi⁽¹⁾, Sergio Di Molfetta⁽⁴⁾, Matteo Franchi⁽³⁾, Lorenzo Menichelli⁽³⁾

 Lorenzo Menichelli 0009-0005-3472-356X

(1) Pediatric Department, Santa Chiara Hospital of Trento, Largo Medaglie d'Oro 9, 38122, Trento, Italy

(2) Department of Pediatrics, Santa Chiara Hospital of Trento, Azienda Councillor, Lombardy Region, Milano, Lombardia, Italy

(3) Unit of Biostatistics, Epidemiology and Public Health, Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy

(4) Department of Precision and Regenerative Medicine and Ionian Area, Section of Internal Medicine, Endocrinology, Andrology and Metabolic Diseases, University of Bari Aldo Moro, Bari, Italy

CORRESPONDING AUTHOR: Lorenzo Menichelli, Unit of Biostatistics, Epidemiology and Public Health, Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy. E-Mail: lorenzo.menichelli@unimib.it

INTRODUCTION

Closed-loop insulin delivery systems, including hybrid closed-loop (HCL) technologies, are now considered the standard of care for individuals with type 1 diabetes (T1D) and are increasingly adopted by both adults and younger populations. Beyond improving glycemic outcomes, their potential impact on patient-reported outcomes such as sleep has gained increasing attention. However, evidence regarding their effects on sleep quality and duration in individuals with T1D and their caregivers remains inconsistent.

OBJECTIVES

The objective of this study was to assess whether the use of commercially available hybrid closed-loop systems improves subjective sleep quality and objective sleep parameters in youths and adults with type 1 diabetes and in their caregivers compared with standard diabetes management.

METHODS

This systematic review and meta-analysis included randomized and non-randomized studies in children, adolescents,

and adults with T1D and their caregivers, evaluating commercially available hybrid closed-loop (HCL) systems. The primary outcome was the PSQI; secondary outcomes included actigraphy-derived parameters (TST, TIB, SE, SOL, NWAK, WASO). Analyses were performed separately by study design and population using a random-effects model (REML; Hartung-Knapp adjustment with IQWiG6 ad hoc correction); risk of bias was assessed with RoB 2 and ROBINS-I, and certainty of evidence with GRADE.

RESULTS

For the PSQI, RCTs showed no statistically significant effect of HCL in subjects (6 studies, $n=573$; MD -0.30 [95% CI -0.92 ; 0.32]) or caregivers (3 studies, $n=216$; MD -1.39 [95% CI -3.66 ; 0.88]), despite a trend toward improvement in both groups. Non-randomized pre-post studies, encompassing a larger pooled sample, suggested a statistically significant improvement in sleep quality in both subjects (11 comparisons, $n=672$; MD -0.40 [95% CI -0.63 ; -0.17]) and caregivers (7 comparisons, $n=372$; MD -1.05 [95% CI -1.54 ; -0.56]).

CONCLUSIONS

Overall, the PSQI findings are directionally favorable to HCL, but the strength of inference differs by study design. The lack of a statistically significant effect in RCTs suggests that any benefit on perceived sleep quality is likely modest and still uncertain. By contrast, non-randomized pre-post studies show a clearer signal, though these results should be interpreted cautiously given their vulnerability to confounding, selection bias, and regression to the mean. The caregiver effect appears more pronounced than the subject effect, possibly reflecting a reduction in nocturnal diabetes-related burden, but this remains hypothesis-generating rather than definitive.

Temporal Dynamics, Demographic Determinants and Clinical Complexity of Heart Failure Hospitalizations in Italy across the Covid-19 Pandemic Era

Marco Enea⁽¹⁾, Giovanni Tinervia⁽¹⁾, Sofia Bellomo⁽¹⁾, Michela Conti⁽²⁾, Dario Genovese⁽¹⁾, Gabriele Biagio Marrella⁽¹⁾, Domenica Matranga⁽¹⁾, Aurelio Seidita⁽¹⁾, Giuseppe Vella⁽¹⁾, Emanuele Amodio⁽¹⁾

 Marco Enea 0000-0002-9281-5746

(1) Department of Health Promotion, Mother and Child Care, Internal Medicine and Medical Specialties "G. D'Alessandro" (PROMISE), University of Palermo, Palermo, Italy

(2) Departmental Simple Operating Unit of Adult and Oncology-Oriented Endocrinology, Ospedali Riuniti Villa Sofia-Cervello Hospital, Palermo, Italy

CORRESPONDING AUTHOR: Marco Enea, Department of Health Promotion, Mother and Child Care, Internal Medicine and Medical Specialties "G. D'Alessandro" (PROMISE), University of Palermo, Palermo, Italy. E-Mail: marco.enea@unipa.it

INTRODUCTION

Heart failure (HF) is one of the leading causes of hospitalization, disability and healthcare resource use in high-income countries, particularly in ageing populations such as Italy. The COVID-19 pandemic profoundly disrupted the organization of healthcare services, raising concerns about the continuity of chronic disease management and the true long-term epidemiological impact on HF admissions.

OBJECTIVES

To analyse the long-term temporal trends of HF hospitalizations in Italy over a fifteen-year period (2008–2022), assessing the contribution of demographic factors, seasonality and clinical complexity. As a secondary objective, we tested whether the pandemic produced a sustained, structural shift in the epidemiological curve of admissions.

METHODS

Retrospective observational study based on national Hospital Discharge Records (SDO) from both public and private facilities. Admissions with HF as the primary diagnosis were identified through ICD-9-CM codes. Cases were stratified by age, sex, season, DRG relative weight and clinical sever-

ity according to the Elixhauser Comorbidity Index. Temporal trends were estimated using a Poisson Inverse Gaussian Generalized Linear Mixed Model, with a random intercept for region and the time component modelled through a segmented regression with a structural break in 2020, distinguishing the pre-pandemic trend, the level change and the slope change in the late-pandemic phase. The Annual Percent Change (APC) and the Average Annual Percent Change (AAPC) were also computed. The analysis is a refinement of a recently published manuscript (Amodio et al., 2026). All analyses were conducted with R (version 4.5.0).

RESULTS

Over the study period, 3,162,075 HF admissions were recorded, corresponding to a crude rate of 35.11 per 10,000 person-years. A slight male preponderance emerged, and age was confirmed as the strongest determinant of risk, with rates increasing exponentially up to 228.31 per 10,000 person-years in subjects aged ≥ 75 years. A marked seasonal pattern was observed, peaking in winter. The largest share of admissions (59.3%) fell within the intermediate comorbidity tier (Elixhauser 13–20). Temporal analysis revealed a long-term structural decline, with a sharp drop in 2020 (–24%; RR 0.80) followed by a partial rebound in 2021–2022 (RR 1.06). Highly complex patients (Elixhauser >20) showed no signs of recovery, with an accelerated decline in the late-pandemic phase.

demic phase, whereas resource-intensive admissions (DRG weight >1.05) displayed a steadily increasing trend. The model incorporated region-specific random slopes, demonstrating that regional temporal trajectories were highly consistent with national aggregated estimates across all comorbidity profiles.

CONCLUSIONS

The burden of HF hospitalizations in Italy remains substantial and evolving, driven by population ageing, sex differences, seasonality and rising clinical complexity. Population-level data show no sustained, structural increase in admissions during the late-pandemic phase: the observed fluctuations appear attributable to organizational and behavioural factors, such as service disruption and care avoidance, rather than to a true epidemiological shift. These findings highlight the need for integrated and resilient care models capable of ensuring continuity of HF management even during health emergencies.

Early Identification of Low Bone Density in Primary Ovarian Insufficiency: A Machine Learning Approach

Lorenzo Messeri⁽¹⁾, Agostino Ruotolo⁽²⁾, Francesca Pampaloni⁽²⁾, Chiara Marzi⁽¹⁾

(1) Department of Statistics, Computer Science and Applications "G. Parenti," University of Florence, Florence (FI), Italy

(2) Pediatric and Adolescent Gynecology Unit, Careggi University Hospital, Florence (FI), Italy

CORRESPONDING AUTHOR: Lorenzo Messeri, Department of Statistics, Computer Science and Applications "G. Parenti," University of Florence, Florence (FI), Italy. E-Mail: lorenzo.messeri09@gmail.com

INTRODUCTION

Primary Ovarian Insufficiency (POI) is a condition characterized by the premature loss of ovarian function before the age of 40, leading to estrogen deficiency and several systemic complications, including reduced bone mineral density (BMD). Although dual-energy X-ray absorptiometry is considered the gold standard for assessing BMD, its use may be limited by costs and waiting times. In this context, machine learning models may help identify patients at higher risk of low BMD at an earlier stage.

OBJECTIVES

The aim of this study was to develop a machine learning model for the early prediction of reduced bone mineral density in women with POI, using medical history, clinical, laboratory, genetic, and instrumental data collected in clinical practice. A secondary objective was to identify the variables most strongly associated with the outcome through model interpretability techniques.

METHODS

The analysis was conducted on a clinical dataset of 250 patients with POI from the Pediatric and Adolescent Gynecology Unit, Careggi University Hospital, including demographic, clinical, hormonal, genetic, and autoimmune characteristics. After a priori variable selection, based on literature and redundancy criteria, and multiple imputation of missing data, the dataset was split into a development (80%) and a test set (20%). Subsequently, the development set was split into a training (80%) and a validation (20%) sets; several su-

pervised classification algorithms including linear penalized models, Support Vector Machines, Random Forest, Gradient Boosting Machine, AdaBoost, XGBoost, and Gaussian Process, were trained on the training set. For each algorithm, hyperparameter tuning was performed using 5-fold cross-validation within the training set, and the optimal hyperparameter configuration was selected based on the area under the receiver operating characteristic curve (AUC). The resulting models, fitted with their optimal hyperparameters, were then evaluated on the validation set for model selection. Finally, the best-performing model was assessed on the test set to obtain an unbiased estimate of its predictive performance. Interpretability was assessed using Shapley values.

RESULTS

The analyzed population is relatively young, with a mean age of 27 years, and it's characterized by marked hormonal variability, reflecting the underlying heterogeneity of the clinical condition. In this context, ensemble and boosting-based machine learning models demonstrated the best overall discriminative performance compared to other approaches. In particular, Random Forest achieved the highest predictive accuracy on the independent test set, with an AUC of 0.803, followed closely by Gradient Boosting Machine (AUC = 0.793), XGBoost (AUC = 0.769), and AdaBoost (AUC = 0.746), confirming the ability of tree-based ensemble methods in this setting.

Calibration analysis further indicated a good agreement between predicted probabilities and observed outcomes. Finally, SHAP analysis highlighted the most relevant predictors of the outcome, including key clinical and laboratory variables such as FSH levels, body mass index, and age at menarche, in line with established clinical evidence.

CONCLUSIONS

Machine learning models based on clinical and laboratory variables represent a promising tool for the early identification of POI patients at risk of reduced bone mineral density, improving risk stratification and clinical decision-making support.

HLA-DQ Allele Combinations and Antibody Profiles in Children with Type 1 Diabetes: A Retrospective Study

Donato Martella^(1,2), Giulia Colombo⁽³⁾, Simona Ferraro^(4,7), Anna Ippoliti⁽⁴⁾, Benedetta Panero⁽⁴⁾, Antonella Zambon^(1,5), Gianvincenzo Zuccotti^(4,6), Alberto Dolci^(3,6), Cristina Cereda^(6,7), Chiara Mameli^(4,6)

 Donato Martella 0009-0004-4398-0747

(1) Laboratory of Quantitative Methods for Life, Health and Society, Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy

(2) Unit of Epidemiological Psychiatry and Digital Mental Health, IRCCS St. John of God Clinical Research Center, Brescia, Italy

(3) Clinical Pathology Unit, ASST Fatebenefratelli-Sacco, Milan, Italy

(4) Pediatric Department, Buzzi Children's Hospital, Milan, Italy

(5) Biostatistics Unit, IRCCS Istituto Auxologico Italiano, Milan, Italy

(6) Department of Biomedical and Clinical Sciences, Università di Milano, Milan, Italy

(7) Center of Newborn Screening, Functional Genomics and Rare diseases, Pediatric Department, Buzzi Children's Hospital, Milan Italy

CORRESPONDING AUTHOR: Donato Martella, Laboratory of Quantitative Methods for Life, Health and Society, Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy. E-Mail: donatomartella97@gmail.com

INTRODUCTION

Type 1 diabetes (T1D) is an autoimmune disease where T cells gradually destroy pancreatic beta cells. The strongest genetic factors influencing T1D susceptibility are HLA-DQ haplotypes, which affect the variety of autoantigens presented to T cells. Islet autoantibodies, such as insulin (IAA), glutamic acid decarboxylase (GAD), islet antigen-2 (IA-2A), and zinc transporter 8 (ZnT8A), are established markers of autoimmune activity, with varying prevalence and clinical implications among patients.

OBJECTIVES

To assess the relationship between HLA-DQ single alleles and pairs of allele combinations with the presence of four islet autoantibodies at the diagnosis of T1D in a pediatric cohort, we employed a data-driven variable selection approach followed by robust regression modeling.

METHODS

A retrospective study was conducted involving 180 children with Type 1 Diabetes (T1D) recruited at Buzzi Hospital between October 2014 and March 2025. All patients under-

went HLA-DQ genotyping, and individual allele carrier status was coded as binary variables. From the cohort, all pairwise allele combinations with sufficient frequency were identified and tested alongside single alleles as potential predictors. Variable selection for each autoantibody outcome was performed independently using optimized Random Forest models, with hyperparameter tuning to minimize out-of-bag error. The top predictors identified were subsequently entered into Poisson regression models with robust sandwich variance estimation, yielding prevalence Risk Ratios (RR) and valid confidence intervals. All p-values were adjusted for multiple comparisons using the Benjamini-Hochberg false discovery rate (FDR) procedure, applied globally across all tests.

RESULTS

Preliminary results indicate that specific alleles, namely DQ8 and DQ7, are independently associated with positivity for IA-2A after applying false discovery rate (FDR) correction. The relative risks (RR) for DQ8 and DQ7 are 1.48 (95% confidence interval [CI] 1.13–1.96, $p = 0.023$) and 1.49 (95% CI 1.12–1.97, $p = 0.023$), respectively. However, no single allele was found to be significant for GADA, ZnT8A, or IAA positivity after adjustment. Among allele pairs with sufficient frequency, the combination of DQ5 and DQ6 showed a significant association with IAA positivity, yielding an RR of 3.22

(95% CI 1.37–7.55, $p = 0.011$). Additionally, the allele pairs DQ8×DQ7 and DQ2×DQ7 were associated with IA-2A positivity, both showing an RR of 1.46 (95% CI 1.08–1.99, $p = 0.021$). No allele pairs reached significance for GADA or ZnT8A positivity after FDR correction.

CONCLUSIONS

This study shows that HLA-DQ alleles influence islet autoantibody profiles in children with T1D, both individually and in combination. Distinct signatures emerge across different autoantibody specificities. DQ8 and DQ7 are identified as independent predictors of IA-2A positivity, while combinations of alleles reveal additional associations that single alleles do not capture. Notably, the combination of DQ5 and DQ6 is linked to insulin autoantibodies (IAA), and the combination of DQ8 and DQ7 is associated with IA-2A. These findings support a combinatorial model of how HLA-DQ influences autoimmune targeting and provide a foundation for further validation in larger prospective studies.

Identifying Existing Methods for Assessing Indirectness in Evidence Synthesis: A Scoping Review

Stefano Giuseppe Lazzarini^(1,2), Matteo Johann Del Furia^(3,4), Chiara Arienti^(5,6), Jens Schünemann Holger^(5,6,7), Antonella Zamboni^(1,8)

 Stefano Giuseppe Lazzarini 0000-0002-9576-2459

(1) Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy

(2) IRCCS Fondazione Don Carlo Gnocchi, Milan, Italy

(3) Department of Biomedical, Surgical and Dental Sciences, University "La Statale", Milan, Italy

(4) Department of Mental and Physical Health and Preventive Medicine, University of Campania Luigi Vanvitelli, Naples, Italy

(5) Clinical Epidemiology and Research Center (CERC), Department of Biomedical Sciences, Humanitas University, Pieve Emanuele, Milan, Italy

(6) IRCCS Humanitas Research Hospital, Rozzano, Milan, Italy

(7) Charité, Universitätsmedizin-Berlin, Allergy and Immunology, Berlin, Germany

(8) Biostatistics Unit, IRCCS Istituto Auxologico Italiano, Milan, Italy

CORRESPONDING AUTHOR: Stefano Giuseppe Lazzarini, Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy. E-Mail: s.lazzarini11@campus.unimib.it

INTRODUCTION

Given the characteristics of the literature in the field of rehabilitation, which is marked by substantial clinical heterogeneity, poor reporting, and complexity in terms of interventions and outcomes, conducting meta-analyses can be complex or, in some cases, impossible. In particular, the concept of clinical heterogeneity is addressed within the indirectness domain of the GRADE (Grading of Recommendations Assessment, Development, and Evaluation) approach, which is assessed "by evaluating how closely the population, intervention, comparison, and outcomes in the available evidence align with those specified in the question of interest".

OBJECTIVES

As a first step toward developing a structured tool to assist systematic review authors in determining whether downgrading for indirectness is warranted, assessing the severity of indirectness, and translating these judgments into appropriate levels of certainty downgrading in rehabilitation evidence synthesis according to GRADE's threshold approach, this study aims to identify existing structured methods (i.e., tools, checklists and scales) for assessing indirectness in any biomedical field.

METHODS

This is a scoping review. MEDLINE, Embase, and CINAHL were searched using specifically designed search strategies up to May 20th, 2026. In addition, the GRADE Book, the Cochrane Handbook for Systematic Reviews of Interventions, the Joanna Briggs Institute Manual for Evidence Synthesis – 2024 Edition, and the Agency for Healthcare Research and Quality (AHRQ) website were consulted. Eligible studies were articles or book chapters describing the development, implementation, empirical validation, or use of any tool, checklist, or scale, aimed at assessing, measuring, or generally investigating indirectness (or related concepts) in evidence synthesis. Study selection and data extraction were performed by two independent authors.

RESULTS

From 2593 individual records identified through database searching, two tools were identified, namely the Clinical Diversity In Meta-analyses (CDIM) tool and the Clinical Diversity in Pediatric Meta-analyses (CDIPM) tool. The latter is a modified version of the former (specifically designed for paediatric populations), available as a conference proceeding only. Both tools include four domains (i.e., setting, population, intervention, and outcome) and 11 items, each to be scored

as low, moderate, or high clinical heterogeneity/diversity (0-2 points). Additionally, indirectness tables for both a single trial and the entire body of evidence, and a summary table for applicability were identified from the GRADE Book and the AHRQ Methods Guide for Comparative Effectiveness Reviews on applicability, respectively. The former included four subdomains (i.e., population, intervention, comparison and outcome) and four possible options regarding whether the evidence is sufficiently direct (i.e., Yes, Probably Yes, Probably No, and No), informing an overall judgement of indirectness across the subdomains. The latter requires a narrative assessment of the applicability of the evidence across five domains (Population, Intervention, Comparators, Outcomes, and Setting).

CONCLUSIONS

The findings highlight the paucity of structured methods available for assessing indirectness. At the same time, the identified methods largely share the same domains, likely suggesting that these may represent the key components that a tool for assessing indirectness should address. This work represents the first step that will inform the development of a tool to assess indirectness in the field of rehabilitation evidence synthesis.

Investigation of Acceptance and Behavioral Intention to Use AI Chatbots among Medical Students at the Universities of Palermo and Catania: A Preliminary Study of the DIADEMA Project

Chiara Di Mitri⁽¹⁾, Laura Maniscalco⁽¹⁾, Marco Enea⁽¹⁾, Giovanni Di Stefano⁽²⁾, Silvana Miceli⁽²⁾, Antonella Agodi⁽³⁾, Andrea Giuseppe Maugeri⁽³⁾, Domenica Matranga⁽¹⁾

 Chiara Di Mitri 0000-0002-1269-3313

(1) Department of Health Promotion, Mother and Child Care, Internal Medicine and Medical Specialties"- PROMISE - University of Palermo, 90127 Palermo, Italy

(2) Department of Psychology, Educational Science, Exercise Science, and Training. University of Palermo, 90128 Palermo, Italy

(3) Department of Medical, Surgical and Advanced Technologies "Gian Filippo Ingrassia", University of Catania, 95123 Catania, Italy

CORRESPONDING AUTHOR: Chiara Di Mitri, Department of Health Promotion, Mother and Child Care, Internal Medicine and Medical Specialties"- PROMISE - University of Palermo, 90127 Palermo, Italy. E-Mail: chiara.dimitri@you.unipa.it

INTRODUCTION

Artificial Intelligence (AI), particularly generative AI, is increasingly transforming educational and information-seeking practices. Medical students are among the most frequent users of AI chatbots for learning and accessing medical information. However, due to their direct and dialogic mode of interaction, concerns have been raised regarding chatbots use and their potential impact on cognitive autonomy, critical thinking, and dependency-related attitudes.

OBJECTIVES

This study presents the preliminary findings of the broader DIADEMA Project, which investigates AI chatbot dependency among first-year medical students at the Universities of Palermo and Catania. Specifically, this initial phase focuses on assessing the acceptance and behavioural intention to use AI chatbots in this population.

MATERIALS AND METHODS

Following informed consent, participants completed a questionnaire assessing sociodemographic, academic characteristics and AI chatbot use for educational activities. Furthermore, an adapted Italian UTAUT2 model assessing AI

chatbot acceptance and behavioural intention, and the Dependency on AI (DAI) Scale, were used. Quantitative variables were summarized as means and standard deviations (SDs), whereas categorical variables were reported as frequencies and percentages. The reliability of the UTAUT2 constructs was assessed using Cronbach's alpha and McDonald's omega coefficients. Confirmatory factor analysis (CFA) was performed to evaluate the factorial validity of the UTAUT2 measurement model, and the model fit was assessed using the RMSEA and SRMR indices. Structural equation modelling (SEM), based on the UTAUT2 theoretical framework, was used to examine the relationships between latent constructs and behavioural intention to use AI chatbots, and results were reported as standardized path coefficients (β). Statistical significance was set at $p < 0.05$.

RESULTS

A total of 248 first-year medical students were included in the analyses, with a mean age of 19.9 ± 2.7 years, 71.0% were female. Average grade was 22.8 ± 3.6 with mean CFU of 18.2 ± 2.9 . Most students reported extensive experience with AI chatbots, 73.0% used them for more than one year. ChatGPT (75.4%) and Gemini (66.5%) were the most commonly used tools and 61.3% reporting chatbot use at least 4 days per week. The scope of use was information retrieval (68.1%), academic tasks (46.4%), and exam preparation

(45.6%). All UTAUT2 latent constructs showed satisfactory reliability, Cronbach's α between 0.74 (Facilitating Conditions) and 0.92 (Performance Expectancy), McDonald's ω between 0.78 (Facilitating Conditions) and 0.93 (Price Value). CFA supported the validity of the measurement model (RMSEA = 0.048; SRMR = 0.049). Effort expectancy (5.75 ± 0.92), facilitating conditions (5.70 ± 0.77), and performance expectancy (5.44 ± 1.02) showed the highest mean scores. SEM identified performance expectancy ($\beta = 0.784$, $p < 0.001$) and habit ($\beta = 0.267$, $p < 0.001$) as significant predictors of behavioural intention, whereas no significant effects were observed for effort expectancy, social influence, facilitating conditions, hedonic motivation, or price value. The model explained 82.2% of the variance in behavioural intention ($R^2 = 0.822$).

CONCLUSIONS

These preliminary findings suggest that AI chatbots are widely adopted and positively perceived among first-year medical students. Performance expectancy and habit emerged as the main determinants of behavioural intention to use AI chatbots.

Patient-Reported Outcomes and Care Experience in Chronic Patients Enrolled in a Family and Community Nurse Follow-Up Programme: Pre-Post Analysis of the “SENECA” Study (ASL Alessandria)

Selene Narcisi^(1,3), Luca Caltagirone⁽¹⁾, Erica Picasso⁽¹⁾, Gloria Spataro⁽¹⁾, Riccardo Mazzucco⁽¹⁾, Laura Lazzarino⁽⁴⁾, Calogero Volante⁽⁴⁾, Maria Elisena Focati⁽⁴⁾, Marinella Bertolotti⁽²⁾, Antonio Maconi⁽¹⁾, Guglielmo Pacileo⁽¹⁾

 Selene Narcisi 0009-0008-8083-7439

(1) Research and Innovation Department (DAIRI), Local Health Authority of Alessandria (ASL AL), Clinical Governance Quality Research operative unit, Alessandria, Italy

(2) Research Training Innovation Infrastructure, Research and Innovation Department (DAIRI), Azienda Ospedaliero-Universitaria SS. Antonio e Biagio e Cesare Arrigo, Alessandria, Italy

(3) Department of Public Health, Experimental and Forensic Medicine, University of Pavia, Pavia, Italy

(4) Health Professions Directorate (Di.P.Sa.), Alessandria Local Health Authority (ASL AL), Alessandria, Italy

CORRESPONDING AUTHOR: Selene Narcisi, Research and Innovation Department (DAIRI), Local Health Authority of Alessandria (ASL AL), Clinical Governance Quality Research operative unit, Alessandria, Italy. E-Mail: snarcisi@aslal.it

INTRODUCTION

Chronic obstructive pulmonary disease (COPD) and heart failure (HF) are leading causes of hospitalisation and reduced quality of life in elderly populations. In Italy, the progressive implementation of the National Chronic Disease Plan and the emerging role of the Family and Community Nurse (IFeC), call for evidence on nurse-led care models' impact on patient-centred outcomes. The “Evaluation of the effectiveness of a new management model for patients with Chronic obstructive pulmonary disease and/or heart failure” (SENECA) study evaluate a new IFeC-based care model at ASL Alessandria, incorporating systematic collection of Patient-Reported Outcome Measures (PROMs) and Patient-Reported Experience Measures (PREMs) at enrolment and at six-month follow-up.

OBJECTIVES

The study aims to assess pre-post changes in health-related quality of life, care experience and therapeutic adherence after six months of IFeC-led follow-up; secondary, to explore the presence of socio-demographic and clinical predictors of change in patient-reported outcomes.

METHODS

SENECA is a prospective observational study enrolling patients discharged with diagnosis of COPD, HF or both from hospitals of Casale Monferrato, Acqui Terme, Novi Ligure, Tortona, and from Community Health Center of Moncalvo.

Quality of life, patient-reported outcomes and medication adherence were assessed at baseline (T0) and at six-month follow-up (T1) using the SF-12 Standard Version 1, IC-PREM-I, and the 8-item Morisky Medication Adherence Scale (MMAS-8). Normality of score changes was assessed using Shapiro-Wilk test to guide the choice between parametric and non-parametric methods; changes between T0 and T1 were evaluated using the Wilcoxon signed-rank test for continuous outcomes, while changes in dichotomised medication adherence categories were assessed using McNemar's test. Independent predictors of change in the SF-12 physical (Δ PCS) and mental (Δ MCS) component scores were estimated via robust multiple linear regression, with baseline scores included as covariates. Analyses were performed in Stata v17.

RESULTS

A total of 237 patients (median age 81 years; 50.2% female) were enrolled: 47.3% with diagnoses of COPD, 26.6%

with heart failure, 26.1% with both.

The PREM score increased significantly from T0 to T1 across all sites (mean 12.62 vs 13.36; Wilcoxon $z = -6.02, p < 0.001$), reflecting greater satisfaction with the care experience. Therapeutic adherence increased significantly (McNemar $p = 0.014$; OR = 3.0, 95% CI: 1.14–9.23).

Both SF-12 components improved significantly at follow-up (PCS: $p < 0.001$; MCS: $p = 0.001$); 43% of patients reached the minimum clinically important difference (Minimal Clinically Important Difference-MCID ≥ 3 points) for the physical component. Diagnosis-stratified analyses showed significant SF-12 improvements in both the COPD and HF subgroups; in the comorbidity group, improvement reached significance only for the mental component. Regression models identified female sex ($\beta = -3.97$ for Δ PCS; $p < 0.004$), comorbidity diagnosis ($\beta = -4.91$ for Δ PCS; $p < 0.002$) and baseline adherence (medium: $\beta = +7.50, p < 0.001$; high: $\beta = +4.72, p = 0.018$, for Δ MCS) as independent predictors of PROMs change.

CONCLUSIONS

The IFeC-based care model evaluated in the SENECA study is associated with clinically and statistically significant improvements in PROMs and PREMs in an elderly chronic patient population, although the reduced score variability observed for PREM at follow-up may indicate a ceiling effect warranting further investigation. Baseline therapeutic adherence emerged as a key predictor of mental health gain, supporting the need for adherence-stratified interventions. The Seneca study, initiated concurrently with the launch of this care model, represents a first step toward future larger-scale studies aimed at identifying the factors most strongly associated with overall patient satisfaction, encompassing physical health, psychological well-being, and quality of life.

Diagnostic Accuracy and Alignment with Human Reader Responses across Five Large Language Models on Complex Clinical Case Challenges

Rino Bellocco^(1,2), Luca Soraci⁽³⁾, Lorenzo Lo Cicero⁽⁴⁾, Salvatore Silipigni⁽⁵⁾, Giuseppe Lanfranchi⁽⁶⁾, Savino Sciascia⁽⁷⁾, Wisit Cheungpasitporn⁽⁸⁾, Andrea Corsonello^(3,9), Domenico Santoro⁽⁴⁾, Guido Gembillo⁽⁴⁾

 Luca Soraci 0000-0002-0171-3358

(1) Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy

(2) Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

(3) Unit of Geriatric Medicine and Clinical Epidemiology, Italian National Research Center on Aging (IRCCS INRCA), Cosenza, Italy

(4) Unit of Nephrology, Department of Clinical and Experimental Medicine, University of Messina, Messina, Italy

(5) Radiology Unit, Biomorf Department, University of Messina, Messina, Italy

(6) MIFT Department, University of Messina, Messina, Italy

(7) University Center of Excellence on Nephrologic, Rheumatologic and Rare Diseases, ASL Città di Torino, San Giovanni Bosco Hub Hospital, Turin, Italy

(8) Division of Nephrology and Hypertension, Mayo Clinic, Rochester, Minnesota, USA

(9) Department of Pharmacy and Health and Nutritional Sciences, University of Calabria, 87036 Rende, Italy

CORRESPONDING AUTHOR: Luca Soraci, Unit of Geriatric Medicine and Clinical Epidemiology, Italian National Research Center on Aging (IRCCS INRCA), Cosenza, Italy. E-Mail: l.soraci@inrca.it

INTRODUCTION

Large language models (LLMs) match or exceed physician accuracy on benchmark diagnostic tasks. Whether higher accuracy reflects human-like diagnostic behavior, which bears on how safely clinicians can supervise these systems, is rarely assessed.

OBJECTIVES

To determine whether diagnostic accuracy and alignment with human reader responses move together across widely accessible LLMs under default, point-of-care conditions.

METHODS

This cross-sectional evaluation submitted the 29 most recent New England Journal of Medicine (NEJM) case challenges to 5 publicly accessible LLMs (GPT-5, Claude, Google Gemini, DeepSeek, and Grok) through each model's web interface (November 5-8, 2025) at default settings, one query per case, no system prompt. Each case offered a clinical vignette and 6 candidate diagnoses; published hu-

man reader response distributions (4386-17 714 per case; median, 7732) were the comparator. Diagnostic accuracy against the published gold-standard diagnosis, compared with the human reader majority by the exact McNemar test; agreement by Cohen κ ; alignment by a human-error similarity index (the share of model errors on cases the reader majority also missed) and by the Jensen-Shannon divergence (JSD) between each model's response and the reader distribution; and ordinary least squares regression of per-case JSD on normalized reader-disagreement entropy.

RESULTS

Reader majority accuracy was 48.3% (95% CI, 31.4-65.6). All 5 reached or exceeded this value, from 51.7% for DeepSeek (95% CI, 34.4-68.6) to 82.8% for Grok (95% CI, 65.5-92.4); Grok ($P = .02$), GPT-5 ($P = .004$), and Gemini ($P = .04$) significantly exceeded the reader majority, whereas Claude and DeepSeek did not. Grok, the most accurate model, showed the weakest alignment (similarity index, 0.40; mean JSD, 0.50 bits), with 3 of 5 errors on cases readers answered correctly. GPT-5 showed the strongest alignment (index, 1.000), failing only on cases that also challenged readers. For 4 of 5 models, divergence rose with reader-disa-

greement entropy; DeepSeek was the exception.

CONCLUSIONS

In this evaluation, diagnostic accuracy and alignment with human reader responses diverged across LLMs. Because a model can outperform clinicians in aggregate yet fail where oversight is least likely to detect it, evaluation of clinical AI should weigh error structure alongside accuracy.

From Health Impact Assessment to Decision Support: An Optimization Framework for Prioritizing Urban Greening Interventions under Spatial Constraints

Giorgia Burbui⁽¹⁾, Costanza Borghi⁽²⁾, Stefano Mancuso⁽²⁾, Gherardo Chirici⁽²⁾, Michela Baccini⁽³⁾

 Giorgia Burbui 0009-0005-3080-555X

(1)Department of Architecture, University of Florence, Florence, Italy

(2)Department of Agricultural, Food, Environmental and Forestry Sciences and Technologies, University of Florence, Florence, Italy

(3)Department of Statistics, Computer Science, Applications G. Parenti, University of Florence, Florence, Italy

CORRESPONDING AUTHOR: Giorgia Burbui, Department of Architecture, University of Florence, Florence, Italy. E-Mail: giorgia.burbui@unifi.it

INTRODUCTION

Urban greening has become a key strategy for promoting population health and supporting sustainable urban planning. Although Health Impact Assessments (HIAs) can quantify the potential health benefits of alternative greening scenarios, they rarely provide practical guidance on how limited green infrastructure should be spatially allocated to maximize public health gains under real-world planning constraints.

OBJECTIVES

This study aimed to develop a spatial optimization framework to prioritize urban greening interventions at the census tracts level in the municipality of Florence (Italy). The proposed algorithm allocates additional tree cover while accounting for land availability constraints and maximizing the reduction in attributable deaths (AD) associated with greenness exposure.

METHODS

Deaths in 2025 among residents aged ≥ 35 years were redistributed from the municipal to the census tracts level using a multinomial allocation weighted by the local age and sex population structure. Exposure in 2023 was measured as mean Tree Cover Density (TCD). Since the exposure-response function, derived from the literature, was estimated

for the Normalized Difference Vegetation Index (NDVI), a generalized additive model (GAM) was used to calibrate TCD into NDVI. Uncertainty in the exposure-response function and stochastic mortality allocation was propagated through Monte Carlo simulations and incorporated into the optimization framework. An iterative greedy optimization algorithm was then developed to allocate additional tree cover by maximizing the marginal reduction in AD per square meter of new tree cover. At each iteration, tree cover was incrementally assigned to the census tract with the highest expected marginal health benefit, subject to realistic spatial constraints, including the available plantable area and a maximum target of 10% TCD, consistent with the 2024 Nature Restoration Law.

RESULTS

Among the 2,168 census tracts in Florence, 395 had already achieved the 10% TCD target in 2023, leaving 1,773 sections eligible for additional greening interventions. The proposed optimization framework allocated new tree cover to 1,740 census tracts while using approximately 8.5% of the total potentially plantable area available across the municipality. Overall, the optimized allocation increased TCD and NDVI across eligible census tracts and was associated with an estimated reduction of approximately 48 deaths among residents aged ≥ 35 years. The prioritization strategy consistently selected census sections with the highest expected marginal health benefit while satisfying all spatial constraints.

CONCLUSIONS

The proposed framework extends conventional HIA by integrating statistical modelling with spatial optimization to support evidence-based urban greening policies. Although still largely theoretical, this approach represents a step toward a practical decision-support tool for identifying where limited greening resources can generate the greatest public health benefit under realistic implementation constraints.

Sugar-Sweetened Soft Drinks Consumption and Gastric Cancer: Results from the StoP Project

Claudia Santucci⁽¹⁾, Arianna Natale⁽²⁾, Carlo La Vecchia⁽²⁾, Marta Rossi⁽²⁾, for the Stomach Cancer Pooling Project (StoP) Consortium

 Claudia Santucci 0000-0002-9845-5307

(1) Section of Hygiene, Department of Life Sciences and Public Health, Università Cattolica del Sacro Cuore, Rome, Italy

(2) Department of Clinical Sciences and Community Health, Dipartimento di Eccellenza 2023–2027, Università degli Studi di Milano, 20133 Milano, Italy

CORRESPONDING AUTHOR: Claudia Santucci, Section of Hygiene, Department of Life Sciences and Public Health, Università Cattolica del Sacro Cuore, Rome, Italy. E-Mail: claudia.santucci@unicatt.it

BACKGROUND

Soft drinks, and specifically sugar-sweetened non-alcoholic beverages, have been associated with metabolic disease, all-cause and cancer mortality in selected studies. However, findings for gastric cancer are inconsistent.

OBJECTIVES

To evaluate the association between sugar sweetened soft drinks and the risk of gastric cancer using data from the StoP Project, a global consortium of gastric cancer including 34 case-control and nested cohort studies from the Americas, Asia and Europe.

METHODS

We conducted a pooled analysis within the StoP Project, using harmonized data from 17 case-control and nested cohort studies with available information on soft drinks. In the present work we did not consider low-calorie sweetened beverages. Overall, 5,929 gastric cancer cases and 16,092 controls were included. Soft drink intake was categorized according to frequency of consumption. Odds ratios (OR) for gastric cancer, and the corresponding 95% confidence intervals (CI), were computed using logistic regression models adjusted for relevant confounders, including total energy intake. Pooled ORs were computed using random-effects models. For dose-risk assessment, we fitted restricted cubic spline models in a two-stage dose-risk meta-analysis.

RESULTS

Compared with non-consumers, the pooled OR for gastric cancer was 1.18 (95% CI 0.98–1.42) for 2 servings/week to less than 1 serving/day, 1.42 (95% CI 1.06–1.91) for 1 serving/day, and 1.64 (95% CI 1.17–2.30) for more than 1 serving/day. Dose-risk analysis showed a progressive increase in gastric cancer risk with higher soft drink intake: OR 1.09 (95% CI 0.92–1.30) at 0.5, 1.25 (95% CI 0.99–1.59) at 1, 1.50 (95% CI 1.13–2.00) at 1.5, and 1.81 (95% CI 1.22–2.69) at 2.0 servings/day. No significant heterogeneity was found in stratified analyses by age, sex, alcohol consumption and total energy intake.

CONCLUSIONS

Sugar sweetened soft drink consumption was associated with an increased risk of gastric cancer, with a dose-risk relationship.

Systolic Blood Pressure and Acute High Altitude Exposure: Which Relationship?

Zeng Ziqian⁽¹⁾, Davide Soranna⁽²⁾, Antonella Zambon^(2,3), Fabrizio Tucci⁽⁴⁾, Gianfranco Parati^(4,5), Bilo Grzegorz^(4,5)

 Davide Soranna 0000-0001-8067-2779

(1) Department of Epidemiology and Biostatistics, Chengdu Medical College, Chengdu, China

(2) Biostatistics Unit, IRCCS Istituto Auxologico Italiano, Milan, Italy

(3) Laboratory of Quantitative methods for Life, Health and Society, Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy

(4) Department of Medicine and Surgery, University of Milano-Bicocca, Milan, Italy

(5) Department of Cardiology, IRCCS Istituto Auxologico Italiano, Milan, Italy

CORRESPONDING AUTHOR: Davide Soranna, Biostatistics Unit, IRCCS Istituto Auxologico Italiano, Milan, Italy. E-Mail: d.soranna@auxologico.it

INTRODUCTION

Exposure to high altitude (HA) is associated with hypobaric hypoxia, a physiological condition that may induce significant cardiovascular adaptations, including changes in systolic blood pressure (SBP). Although several studies have investigated the effect of acute HA exposure on SBP, their findings remain heterogeneous, and the magnitude of SBP changes appears to vary across different settings and populations. In particular, the shape of the relationship between altitude and SBP have not yet been clearly established.

AIM

The primary aim of this meta-analysis was to investigate the relationship between altitude levels and the SBP changes.

METHODS

We searched PubMed, Embase, and the Cochrane Library to identify studies published up to February 28, 2025, using the search terms "(altitude OR hypoxia) AND blood pressure AND humans". We included studies enrolling adults (>18 years) permanently living below 1000 m above sea level (lowlanders), who were assessed both at sea level and during acute exposure to high altitude (>2500 m above sea level). Eligible studies reported conventional SBP as a continuous outcome before and during high-altitude exposure (or the corresponding mean difference) and provided

a measure of variability for the SBP values. To investigate the relationship between SBP changes (related to sea level) and altitude, we applied the one-stage dose-response meta-analysis proposed by Orsini and Crippa, restricted to studies including at least two different high-altitude exposure levels. We additionally explored the altitude-SBP relationship using a more flexible multilevel mixed-effects meta-regression model, including all available SBP measurements. Within both approaches, we evaluated linear and restricted cubic splines and compared their goodness of fit using the Akaike Information Criterion (AIC).

RESULTS

We included 10 studies that reported SBP measurements at two or more high-altitude levels, whereas 31 reported a single SBP measurement during high-altitude exposure. In the dose-response meta-analysis, the restricted cubic spline model with knots at the 5th, 50th, and 95th percentiles appeared to fit the available data slightly better than the alternative models (AIC = 461.62), suggesting a non-linear relationship between altitude and SBP, with larger SBP increases observed at higher altitude levels. Consistent findings were observed in the multilevel mixed-effects meta-regression, where the same spline specification achieved the lowest AIC (512.39) and showed a similar pattern of progressively larger SBP increases with increasing altitude.

CONCLUSIONS

Our preliminary analyses suggest a possible non-linear relationship between altitude and SBP, with a tendency toward greater SBP increases at higher altitudes. Although both modelling approaches showed broadly consistent results, the existence and location of a potential threshold point remain uncertain. The clinical relevance of these BP increases, particularly in individuals at high cardiovascular risk, remains to be established.

Knowledge, Clinical Experience and Perceived Competence in Gender Dysphoria Care: Preliminary Results of a Survey of Physicians in the Marche Region

Carmela Centola⁽¹⁾, Sonia Tonucci⁽²⁾, Grazia Anna Rita Martino⁽²⁾, Andrea Evangelisti⁽²⁾, Luca Santarelli⁽²⁾, Edlira Skrami^(1,3), Rosaria Gesuita^(3,4), Flavia Carle⁽³⁾

(1) Postgraduate School of Medical Statistics and Biometry, Department of Biomedical Sciences and Public Health, Politecnica delle Marche University, Ancona, Italy

(2) Territory and Socio-Health Integration Sector, Regional Health Agency Marche Region, Ancona, Italy

(3) Centre for Epidemiology, Biostatistics and Medical Informatics, Politecnica delle Marche University, Ancona, Italy

(4) IRCCS INRCA, Ancona, Italy

CORRESPONDING AUTHOR: Carmela Centola, Postgraduate School of Medical Statistics and Biometry, Department of Biomedical Sciences and Public Health, Politecnica delle Marche University, Ancona, Italy. E-Mail: carmelacentola@gmail.com

INTRODUCTION

The American Psychiatric Association defines Gender Dysphoria (GD) as a psychiatric diagnosis that refers to the psychological distress that results from an incongruence between one's sex assigned at birth and one's gender identity. Physicians play a crucial role in supporting people experiencing this psychological distress, by providing healthcare, information, and referrals. To ensure this support, both practitioners and specialists must possess specific knowledge and skills.

OBJECTIVES

To assess physicians' knowledge in GD, their clinical experience, and their perceived competence in managing patients with DG.

METHODS

The data were collected by a web-based survey performed between May and July 2025. All physicians working in Marche Region in 2025 were invited to participate. The survey consisted of an anonymous questionnaire containing open and closed-ended questions, covering four areas of interest: information on the participants' professional role, clinical experience with patients with GD, specific training on DG, institutional support for DG management. Variables

were summarized using frequencies and percentages. Associations were evaluated using Chi-square or Fisher's Exact test, when appropriate. Statistical analyses were conducted using R.

RESULTS

A total of 553 physicians took part in the survey (16.4% of the target population), 55.2% were general practitioners or pediatricians (42.9% and 12.3%, respectively), 45.8% had been working for more than 20 years, 54.4% had experience with people exhibiting GD, and 71% of these reported having encountered no more than 5 such individuals during the course of their professional practice. Among physicians with experience with GD cases, about 60% reported the presence of comorbidities, in particular symptoms of anxiety, depression and other disorders, in their patients.

Considering only the complete questionnaires (n=442), 87% did not feel prepared to manage complications resulting from hormone therapies, 81% believed they had not received specific training on GD during their medical education and professional practice, 74% did not feel supported by their colleagues and their institution in managing patients with DG, and 85% were unaware of a specific facility for patients wishing to begin treatment. Physicians with more than 10 years of professional practice were more likely to have clinical experience with patients with GD ($p = 0.009$). In turn, physicians with previous GD-related clinical experience reported

significantly higher perceived competence in managing these patients ($p=0.001$).

CONCLUSIONS

This survey revealed substantial gaps in physicians' knowledge, training, and perceived preparedness regarding the management of patients with gender dysphoria in the Marche Region. Most respondents reported clinical exposure, insufficient training, lack of institutional support, and poor awareness of referral services. General practitioners and paediatricians may play a pivotal role because of their involvement in first contact and long-term care. Although the findings should be interpreted considering the low response rate and the self-reported nature of the data, they suggest the need for structured educational initiatives and improved healthcare pathways to support the management of patients with GD.

Implementation of Vaccination Strategies for Fragile Patients in Hospital: The Experience of a Tertiary Care Center with Hematology Patients

Antonella Ippolito⁽¹⁾, Gabriele Giorgianni⁽²⁾, Alessandra Romano⁽³⁾, Vittorio Grieco⁽⁴⁾, Alessia Scapellato⁽⁵⁾, Antonina Torrisi⁽⁶⁾, Alessia Di Prima⁽⁶⁾, Rosalia Ragusa⁽⁷⁾

 Antonella Ippolito 0000-0002-9616-2957

(1) Scuola di Specializzazione Statistica Sanitaria e Biometria – Dipartimento di Sanità Pubblica, Medicina Sperimentale e Forense-Università di Pavia - Dipartimento di Prevenzione Medica - UOSD - Registro Tumori, Pavia, Italy

(2) UOS Attività Vaccinali metropolitane e provinciali ASP 3 Catania, Catania, Italy

(3) UOC Ematologia - AOU Policlinico "G. Rodolico – San Marco", Catania, Italy

(4) Dipartimento di Prevenzione Asp 5 Messina, Italy

(5) Scuola di Specializzazione in Igiene e Medicina Preventiva, Università di Catania, Catania, Italy

(6) Registro Tumori Integrato Catania-Messina-Enna - AOU Policlinico "G. Rodolico – San Marco", Catania, Italy

(7) Coordinatore Commissione HTA – AOU Policlinico "G. Rodolico-San Marco" Catania, Italy

CORRESPONDING AUTHOR: Antonella Ippolito, Scuola di Specializzazione Statistica Sanitaria e Biometria – Dipartimento di Sanità Pubblica, Medicina Sperimentale e Forense-Università di Pavia, Pavia, Italy. E-Mail: antonellaippolito81@gmail.com

INTRODUCTION

Multiple myeloma is a hematological malignancy associated with a marked immunosuppression state and a consequent increased risk of vaccine-preventable infections. In this vulnerable population, the main guidelines recommend specific vaccination strategies, including influenza, pneumococcal, herpes zoster, and SARS-CoV-2 vaccinations, but in clinical practice, the implementation of these recommendations is often incomplete and uneven for several reasons.

OBJECTIVES

To describe vaccination coverage among patients with multiple myeloma treated at a referral center and to evaluate the level of implementation of recommended vaccination strategies in clinical practice. The impact of SARS-CoV-2 vaccination on overall coverage estimates and its relationship with survival was also analyzed.

METHODS

A retrospective observational study was conducted on a cohort of 379 patients diagnosed with multiple myeloma followed at the AOU CT. For each patient, information re-

garding SARS-CoV-2, influenza, pneumococcal, herpes zoster, and meningococcal vaccinations was reconstructed from the clinical and vaccination database. Binary variables were created for each vaccine and an overall vaccination coverage variable, calculated both including and excluding SARS-CoV-2 vaccination. The demographic characteristics of the cohort and individual vaccine coverage were described. Associations between vaccination status, age groups (<50, 50–64, 65–74, ≥75 years), and gender were assessed using the chi-square test. An exploratory survival analysis using the Kaplan–Meier method was also performed, excluding SARS-CoV-2 vaccination.

RESULTS

The cohort included 379 patients with multiple myeloma, with a slight male predominance (54.5%) and a median age at diagnosis of 66 years (IQR 57–74). Overall vaccination coverage was 75.5%; however, excluding SARS-CoV-2 vaccination, it dropped to 47.0%, highlighting the crucial role of COVID-19 vaccination campaigns. SARS-CoV-2 vaccination was the most frequent (73%), followed by influenza vaccination (42%), while pneumococcal (23%), herpes zoster (15%), and meningococcal (4%) coverage were significantly lower. Considering overall coverage, the association with age groups was significant ($\chi^2 = 8.57$; $p = 0.036$), but lost signifi-

cance when excluding the anti-SARS-CoV-2 component ($\chi^2 = 4.16$; $p = 0.245$). Anti-pneumococcal and anti-Herpes Zoster coverage showed significant differences between age groups ($p = 0.018$ and $p = 0.027$), with lower levels in older subjects; no differences emerged by gender ($p = 0.680$). Exploratory survival analysis, conducted excluding the anti-SARS-CoV-2 vaccine, showed better survival in vaccinated patients compared to unvaccinated ones (log-rank $\chi^2 = 67.03$; $p < 0.001$).

CONCLUSIONS

The results show that, in patients with multiple myeloma, overall vaccination coverage may mask insufficient implementation of recommended vaccinations other than the SARS-CoV-2 vaccine. The persistence of low coverage for some key vaccines and the observed improved survival rate among vaccinated subjects support the need to strengthen dedicated vaccination programs that are systematically integrated into hematology practice in this population.

A Comparison of the 2009 CKD-EPI and EKFC Equations for Estimating GFR in a Large Italian Outpatient Cohort, and the Potential Consequences for Patients' Management

Dino Gibertoni⁽¹⁾, Luca Campanini^(2,3), Mariacristina Gregorini⁽²⁾, Marcora Mandreoli⁽⁴⁾, Antonio Santoro⁽⁴⁾

 Dino Gibertoni 0000-0002-1722-3724

(1) SS Epidemiologia e Statistica, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy

(2) S.C. Nefrologia e Dialisi, Azienda USL - IRCCS di Reggio Emilia, Reggio Emilia, Italy

(3) Scuola di Specializzazione in Statistica Sanitaria e Biometria, Università di Pavia, Pavia, Italy

(4) Comitato Scientifico del Registro PIRP

CORRESPONDING AUTHOR: Dino Gibertoni, SS Epidemiologia e Statistica, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy. E-Mail: dino.gibertoni@aosp.bo.it

INTRODUCTION

An accurate estimation of the glomerular filtration rate (eGFR) is essential for chronic kidney disease (CKD) management and clinical decision-making. The CKD-EPI 2009 has long been the standard but it is currently being replaced by new equations designed to correct some of its shortcomings. Specifically, in 2021 the European Kidney Function Consortium (EKFC) developed a full age spectrum equation aimed at improving accuracy by introducing age-specific parameters. Unlike the CKD-EPI equation, which applies a single fixed age coefficient, the EKFC equation theory supports a non-linear decline of kidney function, where GFR remains quite stable until 40 years of age before progressively decreasing with ageing.

OBJECTIVES

To understand how eGFR estimates obtained by CKD-EPI and EKFC differ according to age and sex, and to assess the impact of CKD categories reclassification and subsequent clinical decision making induced by shifting from the CKD-EPI 2009 to the EKFC equation.

METHODS

Data were retrospectively collected from the Emilia-Romagna regional PIRP registry of individuals with chronic kid-

ney disease, which was established in 2004 and includes data from patients attending the 12 hospital nephrology units of the region. For this analysis, we selected patients who entered PIRP from 2009 onwards and medical records with creatinine range 0.45-5.54 mg/dl to ensure compatibility with creatinine calibration methods used by EKFC. After excluding records with known data quality issues, we obtained a cohort of 127,876 medical records from 30,549 patients. Calculated eGFR was grouped into categories (G1 to G5) according to the KDIGO CKD guidelines. Statistical analyses were all descriptive and included gender-specific scatterplots to compare the delta between formulae (CKD-EPI - EKFC) and patients' age.

RESULTS

The EKFC equation provided lower eGFR estimates than CKD-EPI in 80.7% of records. This was particularly evident in younger and older age groups, while higher estimates from EKFC concentrated around 40 years of age. Grouping according to the baseline eGFR placed 87.9% of patients in the same category. EKFC estimation using baseline data classified 254 patients (0.8%) in a less advanced CKD group and 3418 patients (11.2%) in a more advanced CKD group. Dialysis initiation was disproportionately lower (2.5% vs. 8.9% overall) and higher (36.2%) in patients assigned to a more or less advanced CKD group by EKFC, respectively. Deaths were disproportionately lower when EKFC assigned patients

to a less advanced CKD group (10.2% vs. 30.6% overall).

CONCLUSIONS

Consistent with previous literature, our results showed that EKFC equation tend to attribute lower eGFR values than CKD-EPI equation. However, this was not uniformly distributed across all age brackets, due to the age dependency of GFR hypothesized in the EKFC equation. For 10% of patients, switching to the EKFC estimate would have resulted in a change of CKD grouping, which could have impacted disease management. Assuming that the hypothesis of age-GFR relationship made in EKFC is more accurate than that underlying CKD-EPI, this figure indicates that in absolute terms a great number of patients would benefit from the switch. Nevertheless, the observed discrepancies in clinical outcomes warrant careful further investigation.

Trends in Pancreatic Cancer Mortality in Europe and the USA, 1990-2024, and Predictions to 2030

Riccardo Zanzi⁽¹⁾, Gianfranco Alicandro^(2,3), Carlo La Vecchia⁽¹⁾, Silvia Mignozzi⁽¹⁾, Claudia Santucci⁽⁴⁾

 Riccardo Zanzi 0009-0008-7952-7046

(1) Department of Clinical Sciences and Community Health, Department of Excellence 2023-2027, University of Milan, Milan, Italy

(2) Department of Pathophysiology and Transplantation, University of Milan, Milan, Italy

(3) Mother and Child Department, Cystic Fibrosis Centre, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy

(4) Department of Life Sciences and Public Health, Hygiene Section. Università Cattolica del Sacro Cuore, Rome, Italy

CORRESPONDING AUTHOR: Riccardo Zanzi, Department of Clinical Sciences and Community Health, Department of Excellence 2023-2027, University of Milan, Milan, Italy. E-Mail: riccardo.zanzi@unimi.it

INTRODUCTION

Pancreatic cancer is the fourth leading cause of cancer-related mortality in Europe, accounting for approximately 140,000 deaths in 2024. While mortality for many major cancers has declined over recent decades, pancreatic cancer did not decline and is characterized by a poor prognosis and high fatality rates. This reflects the lack of population screening, the high proportion of diagnoses at advanced stages due to nonspecific symptoms, and the limited efficacy of current treatment. Monitoring mortality trends and geographical inequalities across Europe is therefore essential to assess the evolving disease burden.

OBJECTIVES

The aim of this study was to describe pancreatic cancer mortality trends in Europe and the USA from 1990 to 2024 and to investigate differences across countries and sexes.

METHODS

Pancreatic cancer death counts for the period 1990-2024 were obtained from the World Health Organization, while population estimates were retrieved from the United Nations Population Division. Analyses were restricted to European countries with at least 1 million inhabitants, no ongoing armed conflicts, medium-to-high mortality data quality (usability $\geq 60\%$), and at least 20 years of available data. The USA was also considered. Age-standardised mortality rates (ASMRs) per 100,000 were computed for each country and the EU-27 using the world standard population. Joinpoint

regression models were used to identify significant changes in mortality trends, estimating annual percentage changes (APCs) for each time segment and average annual percentage changes over the entire study period. For the EU-27 and the USA, mortality rates were projected up to 2030.

RESULTS

Twenty-nine European countries and the USA were included. In the triennium 2020-2022, ASMRs for pancreatic cancer in the EU-27 were 8.06 per 100,000 males and 5.15 per 100,000 females, while in the USA they were 7.35 in males and 5.59 in females. The highest mortality rates were observed in Hungary, Latvia, and the Czech Republic, where ASMRs exceeded 10 per 100,000 males and 7 per 100,000 females. Recent trends showed decreasing mortality among males in the EU-27 (APC 2018-2023: -0.6, 95% CI: -1.0, -0.2) and no significant changes among females (APC 2019-2023: -0.3, 95% CI: -1.0, 0.4). In the USA, mortality rates did not significantly change among males (APC 2001-2023: -0.0, 95% CI: -0.1, 0.1), whereas an upward trend was observed among women (APC 2002-2023: +0.2, 95% CI: 0.1, 0.3). By 2030, ASMRs in the EU-27 is predicted to reach 7.64 (95% Prediction interval, PI: 7.38, 7.91) per 100,000 males and 5.56 (95% PI: 5.23, 5.91) per 100,000 females in the EU-27. Corresponding figure for the USA were 7.36 (95% PI: 7.26, 7.46) per 100,000 males and 5.69 (95% PI: 5.59, 5.78) females.

CONCLUSIONS

Overall, pancreatic cancer mortality in the EU-27 and the USA has stopped rising among men since the mid-2020s and

is predicted to moderately decline through 2030. However, several European countries showed stable or increasing mortality, while trends remained less favourable among women. The results confirm the limited progress in reducing the burden of this highly lethal malignancy and reveal important geographical inequalities across Europe, reflecting differences in tobacco smoking, obesity, and diabetes.

BIOSTATISCAL METHODS

Agentic AI for Automated Sensitivity Analysis of Missing Not at Random Data in Patient-Reported Outcome Measures: A Simulation Study

Nicolò Cosimo Albanese⁽¹⁾

(1) Amazon Web Services

CORRESPONDING AUTHOR: Nicolò Cosimo Albanese, Amazon Web Services. E-mail: nicoloac@amazon.it

INTRODUCTION

Patient-Reported Outcome Measures (PROMs) are central to evaluating surgical effectiveness. In the NHS England PROMs programme, only 27% of hip replacement episodes yielded post-operative responses in 2023/24. When missingness depends on the unobserved outcome (Missing Not At Random, MNAR), standard methods yield biased estimates. Sensitivity analysis via pattern-mixture models and tipping-point approaches is recommended (ICH E9(R1)), yet requires expert judgment for parameter specification and interpretation, making it underutilized. Agentic AI systems, capable of autonomous reasoning and tool use, may automate this workflow while preserving statistical rigor.

OBJECTIVES

To develop and evaluate an agentic AI system that autonomously conducts MNAR sensitivity analysis on PROMs data, comparing its accuracy and efficiency against exhaustive grid-based approaches.

METHODS

We used the NHS England PROMs 2023/24 dataset (Open Government Licence v3), selecting $N=11,629$ primary hip replacements with complete Oxford Hip Score (OHS) responses (mean health gain 22.6, $SD=10.2$). A marginal-benefit synthetic scenario (gain ≈ 5.5) was generated to test tipping-point identification. Non-response was simulated via a selection model $P(R=1 | X, Y) = \text{expit}(\alpha_0 + \alpha_1 X + \delta Y^*)$, with δ controlling MNAR severity and target missingness of 15%, 25%, and 35%.

The agent was built using the open-source Strands Agents SDK with Amazon Bedrock (Claude Sonnet 4). It was equipped with five statistical tools for baseline analyses,

pattern-mixture evaluation, grid-based tipping-point search, bisection refinement, and clinical context retrieval. The agent autonomously decides which δ values to evaluate, when to refine, and when to conclude.

The comparator was exhaustive search over a fixed 10-point δ -grid with $m=20$ imputations and Rubin's rules. Performance was assessed via concordance, tipping-point accuracy, number of δ -evaluations, and computation time across 30 Monte Carlo replications (Phase 1 and 2) and 90 agent runs per scenario (Phase 3), with $n=2,000$ bootstrap samples.

RESULTS

On real NHS data at 25% simulated missingness, bias increased monotonically with MNAR severity (MI-MAR bias -1.24 at $\delta=-1$) and was negligible under MAR (bias < 0.04 at $\delta=0$). In the high-effect scenario (gain=22.6), both the exhaustive grid (360 replications) and the agent (90 runs) confirmed robustness in 100% of cases. In the marginal-benefit scenario (gain=5.5), the exhaustive grid found tipping points only at 35% missingness (tipping point -4.03 at true $\delta=-3.0$, found in 97% of reps; tipping point -4.63 at true $\delta=-2.0$, found in 53%).

The agent identified tipping points in 100% of runs. At 35% missing with true $\delta=-2.0$, the agent found mean $\hat{\delta}_{\text{tip}}=-4.52$ ($SD=0.47$), concordant with the grid, using 5.6 tool calls in 71 seconds. The agent achieved higher detection than the fixed grid (100% vs 53%) because adaptive exploration extends beyond predefined boundaries. The agent consistently provided clinically contextualized interpretation.

CONCLUSIONS

An agentic AI system can autonomously conduct MNAR sensitivity analysis on real PROMs data, identifying tipping points with accuracy concordant to exhaustive methods and

higher detection sensitivity due to adaptive exploration. The approach reduces the expertise barrier for routine sensitivity analysis in PROMs programmes where high non-response makes MNAR a concrete concern. Source code is publicly available.

Methods for Single and Multilayer Network Analysis for Biomedical Research: The MixMashNet R Package

Maria De Martino^(1,2), Federico Triolo⁽²⁾, Adrien Perigord⁽²⁾, Alice Margherita Ornago⁽²⁾, Miriam Isola⁽¹⁾, Davide Liborio Vetrano⁽²⁾, Caterina Gregorio⁽²⁾

(1) Department of Medicine, University of Udine, Udine, Italy

(2) Aging Research Center, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet and Stockholm University, Stockholm, Sweden

CORRESPONDING AUTHOR: Maria De Martino, Department of Medicine, University of Udine, Udine, Italy. E-Mail: maria.demartino@uniud.it

INTRODUCTION

Network medicine provides a framework through which many biomedical phenomena are currently understood, reflecting the multidimensional nature of available data (e.g., genes, diseases, functional parameters). This is particularly relevant in aging research, where aging is driven by interconnected biological processes spanning multiple physiological domains. However, most existing network analysis tools are limited to single layer or homogeneous data structures, and uncertainty quantification is rarely addressed.

OBJECTIVE

Develop a statistical framework for the estimation and interpretation of heterogeneous single and multilayer networks and to demonstrate its application in aging-related biomedical data.

METHODS

We developed an R package, MixMashNet, for the estimation and analysis of heterogeneous single and multilayer networks using Mixed Graphical Models (MGMs). The MGM framework enables the joint modelling of continuous, categorical, and binary variables, allowing the definition of distinct yet interconnected layers. Importantly, the estimation procedure incorporates regularization penalties specifically designed to address the challenges of high-dimensional data. MixMashNet supports both intra and interlayer connections, offering control over which layers can interact and

which remain independent, while allowing adjustment for selected covariates.

The package provides tools for community detection using different algorithms, estimation of centrality and bridge-centrality indices, and computation of community scores. In addition, non-parametric bootstrap procedures are implemented to evaluate uncertainty of community assignments, edge weights and node centralities. MixMashNet also enables the identification of excluded nodes that may act as bridging nodes between biological processes. Visualization tools based on ggplot2 and Shiny are included for interactive network exploration.

RESULTS

The proposed framework enables the estimation and characterization of complex dependency structures in biomedical data while incorporating uncertainty quantification through bootstrap procedures. MixMashNet supports the identification of stable network organization, community structure, and node-level properties, facilitating the interpretation of multidimensional biological systems. In particular, the framework enables the identification of bridge nodes connecting different biological domains, including nodes excluded from stable community assignment that may still contribute to network connectivity. MixMashNet was applied to several biomedical datasets, including aging-related biomarker data from the Swedish National Study on Aging and Care in Kungsholmen (SNAC-K) cohort, illustrating its potential for investigating interactions across different domains such as chronic diseases and blood-based biomarkers.

CONCLUSION

MixMashNet provides a unified framework for the estimation, visualization, and interpretation of heterogeneous single and multilayer networks with integrated uncertainty quantification. By supporting the investigation of complex biological mechanisms and their connections, MixMashNet may facilitate the generation of clinically interpretable hypotheses in network medicine, aging research and precision medicine.

Impact of Pre-Existing Depression on In-Hospital and 1-Year Outcomes of Acute Myocardial Infarction: A Population-Based Study

Elisa Teruzzi⁽¹⁾, Lorenzo Cangiano⁽¹⁾, Nicola Cosentino⁽¹⁾, Matteo Franchi⁽²⁾, Giancarlo Marenzi⁽¹⁾, Alice Bonomi⁽¹⁾

 Elisa Teruzzi 0009-0008-3125-1060

(1) Centro Cardiologico Monzino, IRCCS, Milan, Italy

(2) Laboratory of Healthcare Research and Pharmacoepidemiology, Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy; National Centre for Healthcare Research and pharmacoepidemiology, University of Milano-Bicocca, Milan, Italy

CORRESPONDING AUTHOR: Elisa Teruzzi, Centro Cardiologico Monzino, IRCCS, Milan, Italy. E-mail: elisa.teruzzi@cardiologicomonzino.it

INTRODUCTION

Depression has been identified as a risk factor for Acute Myocardial Infarction (AMI); however, uncertainty persists regarding the prognostic implications of pre-existing depression and its treatment in patients with AMI.

OBJECTIVES

This study aimed to investigate the prognostic role of depression in patients with AMI, evaluating in-hospital mortality, one-year mortality and one-year rehospitalization for recurrent AMI or acute heart failure.

METHODS

A population-based cohort study was conducted using administrative healthcare databases from Lombardy, Italy. The study included all adult patients (≥ 18 years) hospitalized between 2009 and 2022 with a diagnosis of acute myocardial infarction. Pre-existing depression was identified from patients' clinical history recorded in the two years preceding the index myocardial infarction. Associations were estimated using multivariable models, including log-binomial regression for in-hospital mortality, Cox proportional hazards models for one-year mortality, and Fine and Gray competing risks models for one-year rehospitalization. All models were adjusted for baseline comorbidities and medication use.

RESULTS

The study included 213,057 adult patients with acute myocardial infarction, 29,254 (13.73%) of whom had pre-existing depression. Compared with patients without depression, those with depression were older, more likely to be female, and had a higher prevalence of baseline comorbidities and medication use. Pre-existing depression was independently associated with increased risk of in-hospital mortality (adjusted RR: 1.23; 95% CI: 1.19–1.27) and one-year mortality (adjusted HR: 1.26; 95% CI: 1.23–1.30). After adjustment for baseline characteristics, pre-existing depression was not associated with an increased risk of one-year rehospitalization (adjusted HR: 0.94; 95% CI: 0.91–0.97). Within the subgroup of patients with depression, no significant differences in the study outcomes were observed between treated and untreated individuals.

CONCLUSIONS

These findings highlight the prognostic relevance of pre-existing depression in patients with acute myocardial infarction. An independent association with adverse outcomes was observed, with no significant differences between treated and untreated patients.

Short-Term Exposure to Air Pollutants and Acute Cardio-Cerebrovascular Events in Lombardy Region: A Case-Time Series Study (2016-2023)

Ginevra Torrigiani^(1,2), João Vasco Santos^(3,4), Luisa Ojeda Fernandez⁽²⁾, Marta Baviera⁽²⁾, Ida Fortino⁽⁵⁾, Mauro Tettamanti⁽⁶⁾, Antonio Gasparri⁽⁷⁾, Antonella Zamboni⁽¹⁾

 Ginevra Torrigiani 0009-0009-9987-7657

(1) Laboratory of Quantitative methods for Life, Health and Society (QmLHS-Lab), Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy

(2) Laboratory of Cardiovascular Prevention, Department of Health Policy, Istituto di Ricerche Farmacologiche Mario Negri IRCCS, Milan, Italy

(3) MEDCIDS@RISE-Health, Faculty of Medicine, University of Porto, Porto, Portugal

(4) Public Health Unit, ULS Santo António, Porto, Portugal

(5) Directorate General for Health, Lombardy Region, Milan, Italy

(6) Laboratory of Geriatric Epidemiology, Istituto di Ricerche Farmacologiche Mario Negri IRCCS, Milan, Italy

(7) Environment & Health Modelling (EHM) Lab, Department of Public Health Environments and Society, London School of Hygiene & Tropical Medicine, London, UK

CORRESPONDING AUTHOR: Ginevra Torrigiani, Laboratory of Quantitative methods for Life, Health and Society (QmLHS-Lab), Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy. E-Mail: ginevra.torrigiani@unimib.it

INTRODUCTION

Air pollution is a leading environmental risk factor for morbidity and mortality worldwide, with fine particulate matter ($PM_{2.5}$) among the most harmful pollutants for cardio-cerebrovascular health. The World Health Organization (WHO) estimates that air pollution contributes to millions of premature deaths annually, largely through cardiovascular diseases.

While Europe generally maintains lower average air pollution levels than other global regions, a substantial proportion of its population remains exposed to concentrations exceeding WHO guidelines. This pattern is evident in the Lombardy region due to Po Valley meteorology and intense anthropogenic emissions.

A large body of evidence has shown that short-term exposure to $PM_{2.5}$ is associated with an increased risk of acute cardio-cerebrovascular events in the days following exposure. Proposed mechanisms include systemic inflammation, oxidative stress, endothelial dysfunction, autonomic imbalance, and pro-thrombotic pathways, which may trigger acute cardio-cerebrovascular events in susceptible individuals. However, traditional epidemiological designs have limited ability to capture exposure-lag-response dynamics, which require more flexible approaches.

OBJECTIVES

To evaluate the association between short-term exposure to $PM_{2.5}$ and the risk of incident acute cardio-cerebrovascular events among residents of Lombardy.

METHODS

This population-based study used administrative health-care databases from Lombardy (~10 million inhabitants) and included residents aged ≥ 45 years with a first acute hospitalization for cardio-cerebrovascular event during 2016–2023. Incident cases were identified using a 5-year washout period.

Short-term effects were assessed using a case-time series design combined with Distributed Lag Non-linear Models to characterize exposure-lag-response associations. A lag window of 5 days was considered in the main analysis, with sensitivity analyses at 3 or 7 days.

Single- and multi-pollutant models were implemented to estimate the independent effect of $PM_{2.5}$, adjusting for NO_2 and O_3 . Effect estimates were expressed per 10 $\mu g/m^3$ increase in $PM_{2.5}$. Models were adjusted for day of the week and meteorological variables (mean temperature,

relative humidity, wind speed) using natural splines, plus a binary indicator for rainfall.

RESULTS

During the study period, a total of 403,266 incident acute cardio-cerebrovascular events were recorded within a regional source population of approximately 5 million residents aged ≥ 45 years. In the main analysis, $[(PM)]_{2.5}$ was significantly associated with an increased risk of acute cardio-cerebrovascular events.

In the single-pollutant model, a $10 \mu\text{g}/\text{m}^3$ increase in $[(PM)]_{2.5}$ was associated with a 3.2% increase in risk (95% CI: 2.0–4.4%). The lag structure showed a peak effect at lag 0 (1.7%; 95% CI: 0.8–2.7%), declining to baseline by lag 3.

In multi-pollutant models, adjustment for O_3 slightly strengthened the association (4.0%; 95% CI: 2.8–5.2%), whereas adjustment for $[(NO)]_2$ attenuated the effect (1.9%; 95% CI: 0.8–3.1%), likely due to correlation between pollutants sharing emission sources. All results were stable across lag windows.

CONCLUSIONS

Short-term exposure to $[(PM)]_{2.5}$ was associated with an increased risk of acute cardio-cerebrovascular events, with effects concentrated mainly in the first days after exposure. These findings reinforce the evidence that air pollution acts as a trigger of cardio-cerebrovascular hospitalizations and highlight the need for effective strategies to reduce population exposure to fine particulate matter.

Length of Stay and Comorbidities in Patients Hospitalized for Herpes Zoster: A Real-World Italian Study (2011–2023)

Arianna Viarchi⁽¹⁾, Antonella Mattei⁽¹⁾

 Arianna Viarchi 0009-0007-9571-3765

(1) Department of Life, Health and Environmental Sciences, University of L'Aquila, L'Aquila, Italy

CORRESPONDING AUTHOR: Arianna Viarchi, Department of Life, Health and Environmental Sciences, University of L'Aquila, L'Aquila, Italy. E-Mail: arianna.viarchi@graduate.univaq.it

INTRODUCTION

Herpes zoster (HZ) represents a significant public health burden and a frequent cause of hospital admission, especially among elderly and clinically vulnerable individuals. While the disease is generally self-limiting, concomitant comorbidities may contribute to worse clinical outcomes and increased healthcare resource use. Hospital length of stay (LOS) is an important measure of disease burden and healthcare utilization, yet the determinants of prolonged hospitalization among patients with HZ have not been thoroughly investigated at the national level.

OBJECTIVES

To identify comorbidity-related risk factors for prolonged hospitalization among patients admitted with herpes zoster in Italy between 2011 and 2023.

METHODS

A retrospective population-based study was conducted using the Italian Hospital Discharge Records database. All hospitalizations associated with HZ between January 1, 2011 and December 31, 2023 were included using ICD-9-CM diagnostic codes. Clinical variables included chronic respiratory diseases, kidney diseases, diabetes mellitus, malignant neoplasms, and autoimmune diseases. Length of stay was analyzed as the main outcome and categorized as short versus prolonged hospitalization according to the overall mean LOS. Poisson regression models adjusted for age and sex were used to estimate the association between comorbidities and prolonged hospitalization, expressed as incidence rate ratios (IRRs) and 95% confidence intervals (95% CIs).

The number of hospitalizations with short and long LOS was included as an offset in the models. Data analysis was performed using STATA/BE 18.0.

RESULTS

Between 2011 and 2023, 47,933 hospitalizations with HZ were recorded. Overall, 34.56% of patients experienced a prolonged hospital stay (16,564 hospitalizations). Comorbidities were identified in 27.71% of hospitalizations (13,280 cases). The analysis revealed that HZ patients hospitalized with malignant neoplasms had a 1.46-times higher risk of prolonged length of stay than those without malignant neoplasms (IRR 1.46, 95% CI 1.40–1.53). The risk was also increased among patients with kidney diseases (IRR 1.24, 95% CI 1.17–1.31) and COPD (IRR 1.11, 95% CI 1.09–1.27). Age was the strongest predictor of prolonged hospitalization, with patients aged ≥ 80 years with malignant neoplasms showing a 2.48-fold higher risk compared with the reference group (95% CI 2.23–2.77).

CONCLUSIONS

This nationwide real-world study highlights the substantial impact of selected comorbidities on hospitalization burden among patients with HZ. Malignant neoplasms, kidney diseases, and COPD were independently associated with prolonged hospitalization, while advanced age emerged as the strongest predictor of extended length of stay. These findings underscore the importance of identifying high-risk patients and support targeted preventive strategies, including HZ vaccination, particularly among older adults and clinically vulnerable populations.

Canine Olfactory Detection of Parkinson's Disease from Sebum Samples: Longitudinal Modelling of Training Performance in the PrOCRI Study

Denisa Gabriela Balan⁽¹⁾, Giulia Pedretti⁽²⁾, Eleonora Biffi⁽²⁾, Paola Maria Valsecchi⁽²⁾, Anna Negrotti⁽³⁾, Giuseppe Maglietta⁽¹⁾

 Denisa Gabriela Balan 0009-0001-7799-2066

(1) Clinical Research and Epidemiology Unit, Azienda Ospedaliero-Universitaria di Parma, Parma, Italy

(2) Department of Chemistry, Life Sciences and Environmental Sustainability, University of Parma, Parma, Italy

(3) Neurology Unit, Azienda Ospedaliero-Universitaria di Parma, Parma, Italy

CORRESPONDING AUTHOR: Denisa Gabriela Balan, Clinical Research and Epidemiology Unit, Azienda Ospedaliero-Universitaria di Parma, Parma, Italy. E-mail: denisa.balan@ao.pr.it

INTRODUCTION

Parkinson's disease (PD) diagnosis remains mainly clinical and may be challenging in early disease stages. Previous observations suggest that patients with PD may have a distinctive skin odour, possibly related to volatile organic compounds present in sebum. A preliminary step is therefore to assess whether trained dogs can learn to recognize PD-associated sebum samples under controlled training conditions.

OBJECTIVES

To pilot canine olfactory learning during the preliminary PrOCRI training phase, assessing recognition of sebum samples from patients with PD versus healthy controls and changes in performance across increasing training complexity.

METHODS

We analysed 4,364 training trials performed by three dogs, including positive-target and blank trials. The primary outcome was a correct response, defined as correct identification of a positive sample or correct rejection of a blank trial. Overall accuracy was decomposed into sensitivity and specificity.

The training pathway was progressive: the first two kits represented imprinting/familiarization, subsequent kits a softer learning phase, and from the 13th kit onward a higher-complexity phase aimed at approximating more realistic recognition conditions through greater sample mixing and task difficulty.

Multivariable logistic regression models included dog identity, training phase, target identity, and dog-by-phase interaction. Generalized additive models with binomial link were used to describe non-linear learning trajectories over training days. The model was trained using data up to the 21st kit and then used to forecast performance on the last two kits, corresponding to kits 22–23. These were considered an internal holdout phase. Uncertainty was assessed using simulation-based prediction intervals and bootstrap resampling.

RESULTS

The transition to the higher-complexity training phase, starting from the 13th kit, was associated with reduced odds of a correct response for Dylan, the reference dog (OR=0.63; 95% CI 0.50–0.80; $p<0.001$), reflecting increased task difficulty. Performance changes differed across dogs. The dog-by-phase interaction was not significant for Kappa, whereas Olivia showed a significant positive interaction (OR=1.86; 95% CI 1.35–2.57; $p<0.001$), suggesting preserved or improved performance despite increased training complexity. Olivia's model-based accuracy increased from approximately 0.65 in the early phase to approximately 0.78 in the higher-complexity phase. GAMs confirmed dog-specific learning trajectories, with higher odds of correct detection for Olivia (OR=1.73; $p=0.006$). In the internal holdout phase, observed performance exceeded predictions for Dylan in both trial types, Olivia in positive trials, and Kappa in both trial types (Δ range: 0.101–0.262; $p\leq 0.040$), suggesting that increased training complexity did not produce a measurable performance decline.

CONCLUSIONS

This pilot analysis suggests that canine olfactory recognition of PD-associated sebum samples is feasible within a structured and progressively complex training setting. However, these findings should be interpreted as evidence of learning and preliminary recognition ability, not as clinical diagnostic validation. Further studies using independent, previously unseen samples from new patients and controls are required before any diagnostic application can be considered.

Improving Machine Learning and Deep Learning Models for 30-Day ICU Readmission Prediction Using Ensemble Bayesian Model Averaging

Emanuele Koumantakis⁽¹⁾, Konstantina Remoundou⁽²⁾, Ioanna Roussaki⁽²⁾, Alessia Visconti⁽¹⁾, Paola Berchialla⁽¹⁾

 Emanuele Koumantakis 0000-0001-6994-8375

(1) Center for Biostatistics, Epidemiology, and Public Health, Department of Clinical and Biological Sciences, University of Turin, Italy

(2) Institute of Communication and Computer Systems, National Technical University of Athens, Athens, Greece

CORRESPONDING AUTHOR: Emanuele Koumantakis, Center for Biostatistics, Epidemiology, and Public Health, Department of Clinical and Biological Sciences, University of Turin, Italy. E-Mail: emanuele.koumantakis@unito.it

INTRODUCTION

Intensive Care Unit (ICU) readmissions increase morbidity, mortality, and healthcare costs. Existing deep learning (DL) models for predicting 30-day ICU readmission show limited clinical utility due to poor calibration and overconfident predictions that can mislead clinicians. While combining multiple predictive models to improve overall performance is a well-established and often effective strategy, the most widely used ensemble approaches overlook model uncertainty, which can still lead to overconfident decision-making. In contrast, Ensemble Bayesian Model Averaging (EBMA) provides a robust framework for incorporating model uncertainty. While EBMA rewards discriminative power, it also penalizes models that provide overconfident yet inaccurate predictions, resulting in more reliable and clinically meaningful risk estimates.

OBJECTIVES

To develop an EBMA-based framework to integrate predictions of 30-day ICU readmission from multiple DL, machine learning (ML), and traditional scoring models.

METHODS

We analyzed first ICU admissions for adult patients (≥ 18 years) from the MIMIC-III database ($n=32,784$ stays). We randomly split the data into training (80%), calibration (13%), and test (7%) sets. Drawing from our own and other systematic reviews, we adapted 17 DL models — including RNN-based architectures, CNN-LSTM combinations, and a transformer-based model — and 6 ML models (two logistic regression

models, Naïve Bayes, random forest, support vector machine, and XGBoost), training all models on the same training set. Their predictions, together with the traditional SWIFT and APACHE-II scores, served as input for our EBMA-based framework. Model weights were then estimated on the calibration set. We assessed the calibration and discrimination of the obtained ensemble on the test set using the Brier Score and AUROC, respectively. We shared the code supporting this study at <https://github.com/ekoumantakis/ebma-icu-readmission>.

RESULTS

Our EBMA-based framework achieved a Brier Score of 0.051 and an AUROC of 0.728, outperforming all individual models in calibration while matching or exceeding their discrimination performance. These results remained robust after excluding lower-weighted models and, notably, even after removing the highest-weighted and best-calibrated model. Compared with individual models, similarly well-calibrated models (i.e., with similar Brier Scores) achieved AUROC values of at most 0.683, indicating inferior discrimination. In contrast, the most discriminative model (AUROC = 0.716) exhibited poor calibration (Brier Score 0.148). While our EBMA-based framework achieved only a marginal improvement in AUROC over this model (+1.7%), it substantially improved calibration, yielding more reliable and clinically valuable predictions.

CONCLUSIONS

Well-calibrated predictions are more valuable for clinical decision-making than high discrimination alone. Our EBMA-based framework reliably integrates multiple models,

delivering superior calibration while maintaining a discrimination comparable to the best-performing individual model. Our framework shifts away from the “winner-takes-all” paradigm of standalone models toward a more robust, ensemble-based strategy. Future work should focus on external validation and on boosting model diversity to further improve discriminative performance.

Targeting Clinically Relevant Causal Estimands in Trials with Latent Exposure-Induced Intermediate Confounding through Structural Equation Modeling and Propensity Score Weighting

Pasquale Dolce⁽¹⁾, Domenico Iervolino⁽²⁾, Mohd Rashid⁽²⁾, Antonio Iaconelli^(3,4), Danila Azzolina⁽¹⁾

 Pasquale Dolce 0000-0002-7588-6067

(1) Biostatistics and Clinical Trial Methodology Unit, CRC DEMeTra, Department of Translational Medical Science, University of Naples Federico II, Naples, Italy

(2) Unit of Biostatistics, Epidemiology and Public Health, Department of Cardiac, Thoracic, Vascular Sciences and Public Health, University of Padova, Padova, Italy

(3) School of Cardiovascular and Metabolic Health, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, UK.

(4) Department of Cardiovascular Medicine, Fondazione Policlinico Universitario A. Gemelli IRCCS, L. go A. Gemelli, 1, 00168 Rome, Italy

CORRESPONDING AUTHOR: Pasquale Dolce, Biostatistics and Clinical Trial Methodology Unit, CRC DEMeTra, Department of Translational Medical Science, University of Naples Federico II

INTRODUCTION

Evaluating therapeutic efficacy in heart failure with reduced ejection fraction (HFrEF) is often challenging because improvements in the primary outcome, such as left ventricular ejection fraction (LVEF), may be influenced by complex mechanistic pathways. A methodological issue in HFrEF trials arises when congestion acts as an exposure-induced intermediate confounder of the relationship between rescue therapy, a post-randomization intercurrent event, and the primary outcome. Separating the true treatment effect from the impact of clinical rescue is therefore essential to understand the therapy's underlying mechanistic pathways. This objective is closely aligned with the ICH E9(R1) addendum, which emphasizes the need to translate clinical questions into clearly defined causal estimands in order to appropriately address intercurrent events. Moreover, since congestion is not directly observed as a single variable, it may be conceptualized as a latent variable measured via correlated indicators subject to measurement error.

OBJECTIVES

The primary objective of this study is to evaluate the performance of Structural Equation Modeling (SEM) in estimating the Controlled Direct Effect (CDE) in the presence of latent in-

termediate confounding and measurement error. We identify the Marginal Structural Model (MSM) as a standard causal inference benchmark utilizing Inverse Probability Weighting (IPW) to compare its performance against the SEM in terms of bias, efficiency, and statistical power.

METHODS

We employed a simulation study to assess estimator performance across a range of sample sizes ($n = 100$ to 1000) and target CDE scenarios. The data-generating mechanism was designed to either satisfy or violate the Positivity Assumption. The CDE was defined as the effect of heart failure therapy on LVEF not operating through rescue therapy (M): $CDEM = E[Y_{1,m}] - E[Y_{0,m}]$, where M is fixed to the same value under both treatment levels. The comparative framework contrasted SEM, which explicitly incorporates latent variables and measurement error for indicators of congestion, against MSM. Performance metrics included Mean Bias, RMSE, Statistical Power, and Type I Error Rate.

RESULTS

Analysis of estimator bias revealed that SEM demonstrated superior performance, maintaining significantly lower mean bias than MSM weighting approaches, as the strength

of the intermediate confounding effect increased. In terms of statistical power, SEM showed increasing performance as the sample size increased, reaching the target 80% power threshold and properly identifying the CDE across all simulated effect sizes. SEM successfully avoided practical positivity violations observed in MSM. While MSM struggled with the high dimensionality of the intermediate confounders, SEM employed latent-variable dimension reduction to maintain estimation stability.

CONCLUSIONS

SEM provides a robust and flexible unified framework for estimating clinically meaningful causal effects in complex HFrEF trials. Our findings suggest that, when combined with a well-specified causal graph and explicit identification assumptions, SEM is a powerful tool for pre-specified causal estimands rather than merely an exploratory path-analysis tool. SEM is highly effective for addressing intercurrent events, such as rescue therapy, in compliance with ICH E9(R1) guidelines. Treatment–mediator interaction will be addressed in future studies.

AI-Enhanced PROMs and Drug-Drug Interaction Screening in Advanced Breast Cancer: Protocol of a Prospective Observational Pilot Study

Ornella Tinelli⁽¹⁾, Claudia Von Arx⁽²⁾, Roberta Caputo⁽²⁾, Stefania Cocco⁽²⁾, Sara Parola⁽²⁾, Claudia Calderaio⁽²⁾, Martina Pagliuca⁽¹⁾, Luisa Cirillo⁽²⁾, Martina Autiero⁽²⁾, Teresa Basileo⁽³⁾, Michelino De Laurentiis⁽²⁾, Michela Piezzo⁽²⁾

 Ornella Tinelli 0009-0007-9003-4915

(1) Clinical and Translational Oncology, Scuola Superiore Meridionale, Naples, Italy

(2) Department of Breast and Thoracic Oncology, Istituto Nazionale Tumori – IRCCS - “Fondazione G. Pascale”, Naples, Italy

(3) Department of Biology, University of Naples Federico II, Naples, Italy

CORRESPONDING AUTHOR: Ornella Tinelli, Clinical and Translational Oncology, Scuola Superiore Meridionale, Naples, Italy. E-Mail: o.tinelli@ssmeridionale.it

INTRODUCTION

In advanced breast cancer, treatment pathways are becoming increasingly prolonged and complex across different biological subtypes. New systemic and targeted therapies require long-term monitoring and careful personalization according to patient-specific factors that demand increasingly individualized therapeutic pathways, such as comorbidities, lifestyle factors, quality of life (QoL), and polypharmacy risks like drug-drug interactions (DDIs). Current real-world pathways often suffer from delays in personalization, suboptimal adherence, and unmet needs in QoL. Remote monitoring platforms like Cureety TechCare may help clinicians to transform reactive care into proactive management. The integration of patient-reported data with DDI screening may further support personalized monitoring in patients receiving long-term systemic treatments.

OBJECTIVES

This pilot study proposes a digital workflow framework to optimize diagnostic-therapeutic paths: Cureety for QoL/lifestyle monitoring and DDI integration for comorbidity-adjusted alerts. The primary objective is to assess the feasibility of the procedure by measuring patient compliance with ePRO questionnaires and completeness of DDI screening during the first month of monitoring. Secondary objectives include describing early patient-reported needs, monitoring QoL and lifestyle-related information over time, and exploring whether

the integration of comorbidities and concomitant medications may support more individualized patient monitoring.

METHODS

This is a prospective, non-interventional observational pilot study. Eligible patients will be adults with advanced breast cancer receiving systemic therapy in routine clinical practice. Patients will be monitored through Cureety TechCare for ePROs, QoL, symptoms, and lifestyle-related information. Cureety TechCare is an AI-driven digital remote monitoring platform specifically designed for oncology patients, enabling remote PRO collection through CTCAE-based questionnaires and four-level clinical classification. In parallel, a DDI module will be used to document concomitant medications and identify potential pharmacological interaction risks in the context of each patient's clinical history, comorbidities, and reported symptoms. The primary endpoint will be Composite Platform Compliance, defined as the proportion of patients achieving both at least 70% of expected ePRO questionnaire completion and at least 85% completeness of DDI screening for documented concomitant medications after the first month.

RESULTS

As this is a protocol, results are not yet available. The study will estimate the feasibility of the integrated workflow and describe early monitoring outputs during the first 30 days of treatment, including first-month tolerance as an exploratory

integrated digital health status indicator. Patients will be classified as “Good Health Status” (GHS) or “Poor Health Status” (PHS) based on the proportion of days with green/yellow Cureety status and low-risk DDI profile versus orange/red Cureety status and/or higher-risk DDI profile.

DISCUSSION

This protocol addresses a key methodological challenge in integrating patient-reported data with pharmacological risk assessment in real-world oncology care. Its methodological relevance lies in the definition of composite feasibility endpoints, the management of longitudinal patient-reported data, and the interpretation of digital alerts generated by the combination of PROMs and pharmacological risk information.

CONCLUSIONS

This pilot protocol may generate preliminary evidence on the feasibility of AI-supported integration of PROMs and DDI screening in advanced breast cancer care. The study may inform the design of future larger studies evaluating the implementation of patient-centered digital monitoring models in routine oncology practice.

Bayesian Shrinkage versus Frequentist Variable Selection in Small-Sample Translational Research: A Monte Carlo Simulation Study Grounded in Regenerative Medicine Biomarker Discovery

Michele Francesco Di Tolla⁽¹⁾, Domenico Iervolino⁽²⁾, Vittorio Scisciola⁽¹⁾, Martina D'Angelo⁽¹⁾, Pietro Fusco⁽¹⁾, Pasquale Dolce⁽¹⁾, Vittoria D'Esposito⁽³⁾, Pietro Formisano⁽¹⁾, Danila Azzolina⁽¹⁾

 Michele Francesco Di Tolla 0000-0002-1376-3782

(1) Biostatistics and Clinical Trial Methodology Unit, Clinical Research Center DEMeTra, Department of Translational Medical Science, University of Naples Federico II, Naples, Italy

(2) PhD Program in Translation Specialistic Medicine "G.B. Morgagni" Curriculum "Biostatistics and Clinic Epidemiology" University of Padua, Padua, Italy

(3) Institute of Endotypes in Oncology, Metabolism and Immunology "G. Salvatore" – National Research Council (IEOMI-CNR), Naples, Italy

CORRESPONDING AUTHOR: Michele Francesco Di Tolla, Biostatistics and Clinical Trial Methodology Unit, Clinical Research Center DEMeTra, Department of Translational Medical Science, University of Naples Federico II, Naples, Italy. michelefrancesco.ditolla@unina.it

INTRODUCTION

Observational studies in regenerative medicine represent a promising frontier for the discovery of predictive biomarkers of therapeutic response, yet are structurally constrained by small sample sizes. This low-sample-size regime exposes variable selection methods to a high risk of overfitting and false discovery, with direct consequences for the reproducibility of findings. Although frequentist penalised approaches such as LASSO are widely adopted in this setting, their properties under small-sample conditions have not been systematically characterized relative to Bayesian shrinkage alternatives.

OBJECTIVES

To evaluate and compare the performance of frequentist and Bayesian variable selection methods across a range of sample sizes, with the aim of identifying which approach achieves the most robust trade-off between statistical power (TPR) and false discovery control (FDR) in small-sample translational research settings.

METHODS

A Monte Carlo simulation study was conducted (2,000 replicates per sample size level) grounded in empirical pa-

rameters derived from a published clinical case study involving patients with Grade I Knee Osteoarthritis treated with Autologous Protein Concentrate (APC), classified as responders or non-responders based on NPRS variation. Five true signal variables (NPRS, IL-9, MIP-1 β , RANTES, VEGF) and four noise variables were simulated across nine sample size levels ($n = 20-100$ per group). Seven methods were compared: standard logistic regression (GLM), LASSO with lambda.1se criterion, and five Bayesian approaches (Bayesian Ridge, Bayesian Lasso, Horseshoe, Horseshoe+, Spike-and-Slab). Performance metrics included TPR, FDR, Matthews Correlation Coefficient (MCC), Jaccard Index, and Exact Recovery rate.

RESULTS

Bayesian Ridge demonstrated the most favourable and stable TPR/FDR trade-off across the full sample size range, maintaining FDR < 0.025 at all levels, systematically lower than LASSO (FDR 0.04–0.07) and GLM (FDR 0.03–0.06), while achieving TPR comparable to frequentist methods at $n \geq 60$. MCC converged between Bayesian Ridge and LASSO at $n \geq 70$; however, LASSO exhibited an anomalous FDR peak at $n = 50-60$, indicative of unstable signal-noise discrimination at the critical power threshold. Per-variable analysis confirmed a detectability hierarchy: NPRS, RANTES, and VEGF reached selection frequencies > 0.90 at $n = 60$ under Bayesian Ridge, whereas IL-9 and MIP-1 β required larger samples.

CONCLUSIONS

In small-sample translational research, Bayesian Ridge regression offers a methodologically superior alternative to frequentist penalised approaches, combining adequate statistical power with robust false discovery control. Findings indicate a minimum sample size of $n = 40\text{--}50$ per group for reliable variable selection regardless of the method adopted and support the pre-specification of Bayesian shrinkage models in biomarker discovery studies in regenerative medicine.

A Bayesian Adaptive Platform Trial for Evaluating Mechanical Ventilation Strategies Using Patient-Centered Endpoints in Heterogeneous ICU Populations SMART-Vent

Domenico Iervolino^(1,2), Pasquale Dolce⁽²⁾, Vittorio Scisciola⁽²⁾, Fusco Pietro⁽²⁾, Michele Francesco Di Tolla⁽²⁾, Mohd Rashid Khan^(1,2), Savino Spadaro⁽³⁾, Danila Azzolina⁽²⁾

 Domenico Iervolino 0009-0002-1682-3792

(1) PhD Program in Translation Specialistic Medicine "G.B. Morgagni", Curriculum "Biostatistics and Clinic Epidemiology", University of Padua, Padua, Italy

(2) Biostatistics and Clinical Trial Methodology Unit, Clinical Research Center DEMeTra, Department of Translational Medical Science, University of Naples Federico II, Naples, Italy

(3) Department Morphology, Surgery and Experimental Medicine Faculty of Medicine, University of Ferrara, Ferrara, Italy

CORRESPONDING AUTHOR: Domenico Iervolino, domenico.iervolino608@gmail.com

INTRODUCTION

Acute respiratory distress syndrome (ARDS) and acute respiratory failure exhibit profound heterogeneity in clinical presentation, lung mechanics, and individual response to intensive care treatments. Traditional fixed-design randomized trials are poorly tailored to evaluate the efficacy of dynamic, multistage therapeutic strategies, failing to account for both the sequential nature of clinical decision-making and the underlying biological heterogeneity of patients.

OBJECTIVES

To introduce SMART-Vent, an innovative design that embeds a Sequential Multiple Assignment Randomized Trial (SMART) within a Bayesian adaptive platform. The objective is to optimize phenotype-driven mechanical ventilation sequences in the intensive care unit, addressing the heterogeneity of treatment effects in an efficient and ethically sound manner.

METHODS

A Monte Carlo simulation engine was developed to evaluate the operating characteristics of the SMART-Vent design across complex clinical scenarios with total sample sizes

ranging from N=200 to N=1500. Simulated patients were stratified into two primary clinical phenotypes (ARDS and non-ARDS) and assigned to a two-stage ventilatory pathway (first-line therapies, and rescue strategies for non-responders). The statistical framework relied on Bayesian hierarchical models estimated via MCMC algorithms. The design incorporated Response-Adaptive Randomization (RAR) to dynamically update allocation weights and a decision rule driven by a 95% Highest Density Interval (HDI). Performance metrics were benchmarked directly against a standard, multiplicity-adjusted frequentist design.

RESULTS

Simulations demonstrated that SMART-Vent significantly outperforms frequentist benchmarks in identifying optimal dynamic treatment regimes. Utilizing the 95% HDI decision rule successfully circumvented the severe sample-size inflation imposed by classical α -error corrections, achieving target statistical power with smaller sample cohorts. Despite sequential interim analyses and adaptive allocation, the Bayesian False Discovery Rate (FDR) remained inherently self-calibrated and strictly bound around the nominal 5% threshold across all strata. From an ethical perspective, the RAR algorithm optimized patient care, progressively allocating up to 80% of responsive subjects to the truly superior ventilatory strategy. Sensitivity analyses confirmed that incorpo-

rating informative prior distributions accelerated early signal detection in initially underpowered subgroups (e.g., ARDS non-responders) without compromising the data-driven asymptotic integrity of the model.

CONCLUSIONS

By seamlessly merging SMART methodology with a Bayesian adaptive platform, the SMART-Vent design maximizes statistical efficiency, maintains rigorous control over false-positive rates, and dynamically minimizes patient exposure to sub-optimal ventilatory pathways. This framework provides a methodologically robust and ethically grounded solution for conducting future precision medicine trials and Complex Innovative Designs (CID) in critical care.

Targeting What Matters in Pediatric Trials: Bayesian Basket and Umbrella Trials for Heterogeneous Airway Disease

Mohd Rashid Khan⁽¹⁾, Angelo Capaso⁽²⁾, Mariateresa Russo⁽²⁾, Domenico Iervolino⁽¹⁾, Pasquale Dolce⁽²⁾, Danila Azzolina^(1,2)

 Mohd Rashid Khan 0000-0001-8855-3272

(1) Unit of Biostatistics, Epidemiology and Public Health, Department of Cardiac, Thoracic and Vascular Sciences, University of Padova, 35131, Padova, Italy

(2) Biostatistics and Clinical Trial Biometry, Clinical Research Center DEMeTra, Department of Translational Medicine, University of Naples Federico II, 80138, Naples, Italy

CORRESPONDING AUTHOR: Mohd Rashid Khan, Unit of Biostatistics, Epidemiology and Public Health, Department of Cardiac, Thoracic and Vascular Sciences, University of Padova, 35131, Padova, Italy. Email: mohdrashid.khan@ubep.unipd.it

INTRODUCTION

Children with chronic inflammatory airway disease often present with similar signs and symptoms despite markedly different underlying biology. Nasal cytokine profiling routinely reveals five endotypes: eosinophilic, IL-5-high, IL-13-high, mixed, and pauci-inflammatory, each associated with differential response to targeted therapies with biologics. In trials with small sample sizes and challenging long-term spirometric follow-up, early change in small-airway resistance ($\Delta R5-R20$ at 14 days) is a common surrogate outcome. Traditional parallel-group RCTs struggle to capture this heterogeneity. To address this, we developed a Bayesian basket-and-umbrella trial framework tailored to the pediatric context and evaluated its performance through an extensive simulation study.

OBJECTIVES

To develop and evaluate Bayesian basket and umbrella clinical trial designs for pediatric chronic inflammatory airway disease, assessing their ability to improve the efficiency, precision, and reliability of treatment-effect estimation across cytokine-defined endotypes under realistic pediatric trial conditions.

METHODS

Two designs were constructed around the same motivat-

ing example. In the basket configuration, one biologic agent was tested across the five cytokine-defined endotypes. In the umbrella configuration, three mechanism-specific agents, anti-IL-5, anti-IL-13, and dual-pathway modulation, were compared within a shared clinical population, with treatment assignment informed by cytokine signatures at enrollment.

Hierarchical Bayesian models were used to estimate subgroup-specific treatment effects, allowing structured information borrowing across biologically related endotypes. Hyperparameters were elicited from pediatric pulmonology experts to reflect realistic variability, and weakly informative priors ensured interpretability. Posterior probabilities guided interim rules for early stopping (superiority/futility), adaptive randomization, and dynamic modification of subgroup enrollment.

A simulation study of 10,000 trials per scenario incorporated constraints typical of pediatric research: unequal subgroup prevalence, rare strata (<5%), biomarker misclassification, early dropout due to symptom flares, and heterogeneous effect sizes. Operating characteristics included posterior power, false-positive/false-negative rates, precision of subgroup effect estimation, probability of correct endotype identification, expected sample size, and time to decision.

RESULTS

Across most clinically plausible scenarios, Bayesian basket and umbrella designs outperformed classical RCTs. Precision in treatment-effect estimation improved by 20–50% in rare endotypes, and adaptive randomization reduced ex-

pected sample size by 15–30% while maintaining strict control of false discoveries. Posterior decision rules enabled earlier identification of beneficial therapies without compromising robustness. When substantial biomarker misclassification was introduced, hierarchical models automatically reduced inappropriate borrowing across endotypes, preventing misleading shrinkage and preserving inferential reliability. Clinically, the Bayesian designs provided sharper, endotype-specific therapeutic signals than achievable with parallel-group trials of comparable size.

CONCLUSIONS

Starting with a realistic pediatric airway disease example, this simulation study shows that Bayesian basket and umbrella trials offer a flexible, efficient, and biologically aligned approach for evaluating targeted therapies in children. By integrating mechanistic knowledge, hierarchical modeling, and adaptive decision rules, these designs address key limitations of conventional RCTs, small samples, heterogeneity, and operational constraints, while delivering clinically interpretable subgroup insights. Their use may substantially accelerate treatment development in pediatric settings where traditional designs are underpowered or insufficiently granular.

Geo Data in Healthcare: A Methodological Framework for Analyzing Spatial Inequalities in Out-of-Hospital Cardiac Arrest

Roberto Primi^(1,2,3), Bertazzon Stefania⁽⁴⁾, Simona Villani⁽¹⁾, Sara Bendotti^(2,3), Leila Ulmanova^(2,3), Alessia Currao^(2,3,5), Enrico Baldi^(2,3), Simone Savastano^(2,3)

 Roberto Primi 0000-0002-4271-6248

(1) Department of Public Health, Experimental and Forensic Medicine. University of Pavia, Pavia, Italy

(2) Division of Cardiology, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy

(3) Cardiac Arrest and Resuscitation Research Team (RESTART), Fondazione IRCCS Policlinico San Matteo, Pavia, Italy

(4) Department of Geography, University of Calgary, Calgary, Alberta, Canada

(5) Department of Internal Medicine and Medical Therapy, University of Pavia, Pavia, Italy

CORRESPONDING AUTHOR: Roberto Primi, Department of Public Health, Experimental and Forensic Medicine. University of Pavia, Pavia, Italy. E-mail: roberto.primi01@universitadipavia.it

INTRODUCTION

Cardiac arrest is the loss of mechanical cardiac function and systemic circulation; when it occurs outside hospital settings, it is defined as out-of-hospital cardiac arrest (OHCA). OHCA is influenced by both individual and contextual factors, including population structure, socioeconomic disadvantage, and access to emergency care. Geographic approaches offer a framework to investigate spatial inequalities and support place-based decision making.

OBJECTIVES

This study assesses spatial inequalities in OHCA incidence in relation to demographic (ageing index, population density) and socioeconomic (low income) determinants in Lombardy between 2021 and 2024.

METHODS

A municipal-level retrospective ecological study was conducted in Lombardy (2021–2024). The outcome was an OHCA event. Data were obtained from the Lombardia Cardiac Arrest Registry (CARE), covering 924 municipalities across seven provinces. The eligible population included residents of municipalities in the CARE registry (cumulative at-risk population: 17,612,296 inhabitants). Due to incomplete registry cov-

erage in 296 municipalities, analyses focused on 628 municipalities forming the South-East (SE) macro-area (Pavia, Lodi, Cremona, Mantua, Brescia). Determinants included: proportion of taxpayers with annual income $\leq \text{€}10,000$ (proxy for socioeconomic vulnerability), ageing index (proxy for demographic imbalance), and population density (proxy for rurality/urbanization). Contextual covariates were: bystander cardiopulmonary (CPR) rate (proxy for community response) and EMS response time (minutes from emergency call to first vehicle arrival, proxy for geographic isolation). The analytical strategy combined global and local spatial autocorrelation (Global Moran's I ; Local Indicators of Spatial Association, LISA), hotspot analysis (Getis-Ord G_i^*), and geographically weighted regression (GWR) to identify clustering patterns and spatially varying associations between OHCA and determinants, adjusting for covariates.

RESULTS

The cumulative OHCA incidence was 126 per 100,000 inhabitants. OHCA incidence was not randomly distributed, but showed significant spatial heterogeneity. Global Moran's I confirmed clustering ($I = 0.27$; $p < 0.001$). Municipalities characterized by high- and low-incidence aggregation patterns were identified (Getis-Ord G_i^* z -scores $> |1.96|$): a hot-spot cluster located in the south-western hilly area of the region and a cold-spot cluster in the south-eastern flat area. GWR showed that the strength and direction of associations

varied across the region, highlighting spatial non-stationarity in the relationship between OHCA incidence and the area-level determinants, independently of community response, response time, and geographic isolation. Specifically, increasing the ageing index and the proportion of low-income individuals significantly increased OHCA incidence, whereas population density was inversely correlated with OHCA incidence.

CONCLUSIONS

OHCA incidence hot spots were concentrated in mountainous and hilly areas, especially in remote areas distant from hospitals. Indicators of social inequality showed spatially heterogeneous associations with OHCA, confirming that territorial disparities extend beyond demographic composition alone. Considering both structural determinants and contextual covariates improved interpretation of spatial patterns and highlighted the contribution of place-based factors to OHCA burden. The integration of geographic data with health information represents a robust methodological approach for studying spatial inequalities in healthcare. Cardiac arrest findings may be translated into public health impact and guide equity-oriented planning by identifying priority areas and informing geographically targeted interventions.

Optimizing Pediatric Outcomes: Bayesian Modeling of Complex Ventilation-Free Outcomes in Pediatric Respiratory Trials

Mariateresa Russo⁽¹⁾, Pasquale Dolce⁽²⁾, Danila Azzolina⁽³⁾

 Mariateresa Russo 0009-0005-8761-5191

(1) University of Naples Federico II, Naples, Italy

(2) Biostatistics and Clinical Trial Methodology Unit, Clinical Research Center DEMeTra, Department of Translational Medical Science, University of Naples Federico II, Naples, Italy

(3) Biostatistics and Clinical Trial Methodology Unit, Clinical Research Center DEMeTra, Department of Translational Medical Science, University of Naples Federico II, Naples, Italy

CORRESPONDING AUTHOR: Mariateresa Russo, University of Naples Federico II, Naples, Italy. E-Mail: mariateresa.russo9@studenti.unina.it

INTRODUCTION

Clinical trials in pediatric intensive care increasingly rely on complex endpoints, such as duration of mechanical ventilation or ventilation-free days, which are bounded, highly skewed, and characterized by a large proportion of structural zeros. Conventional statistical approaches frequently assume normality or simplify these outcomes through dichotomization, resulting in substantial information loss and potentially masking clinically meaningful treatment effects. This work illustrates a Bayesian framework for modeling complex respiratory outcomes using data from a randomized noninferiority trial comparing high-flow nasal cannula (HFNC) and non-invasive ventilation (NIV) in 252 infants hospitalized with bronchiolitis.

OBJECTIVES

To evaluate Bayesian models tailored to semi-continuous, zero-inflated, and bounded respiratory outcomes and compare their predictive performance with conventional approaches.

METHODS

Rather than reducing the outcome to binary endpoints such as intubation, we modeled the complete distribution of mechanical ventilation duration using Bayesian models specifically designed for the underlying data structure, including Hurdle, Zero-One Inflated Beta (ZOIB), Bayesian linear, and cumulative logistic regression models. These approaches ac-

commodate excess zeros, skewness, bounded outcomes, and ordinal representations while naturally incorporating uncertainty through posterior probability distributions. Predictive performance was evaluated using Leave-One-Out cross-validation.

RESULTS

Advanced Bayesian models provided substantially better characterization of the observed outcome distribution than conventional Gaussian models. The ZOIB model demonstrated the best predictive performance, accurately representing both the probability of requiring invasive ventilation and the distribution of ventilation duration among treated patients. Hurdle and cumulative logistic models similarly captured clinically relevant aspects of the outcome while preserving interpretability. Across all Bayesian analyses, HFNC remained noninferior to NIV regarding intubation risk, with a lower expected duration of mechanical ventilation and high posterior probability of noninferiority.

CONCLUSIONS

The choice of statistical model should be driven by the biological and clinical characteristics of the endpoint rather than analytical convenience. Bayesian models tailored to semi-continuous, zero-inflated, and bounded outcomes provide richer clinical interpretation, avoid the loss of information associated with dichotomization, and support probability-based decision making. This framework is broadly

applicable to pediatric critical care and other clinical trials involving complex composite or duration-based outcomes, offering a principled approach for improving treatment effect estimation and clinical inference.

From Small Samples to Better Predictions: Data Augmentation in Muscle Disease Classification

Fabiana Pia Vitiello⁽¹⁾, Chiara Scarfi⁽¹⁾, Santina Caliri⁽¹⁾, Danilo Caudo⁽¹⁾, Davide Sattin⁽²⁾, Antonio Musarò⁽³⁾, Nunzio Iraci⁽⁴⁾, Christian Lunetta⁽²⁾, Maria Cristina De Cola⁽¹⁾

 Fabiana Pia Vitiello 0009-0005-5594-1427

(1) IRCCS Centro Neurolesi "Bonino Pulejo", Messina, Italy

(2) Istituti Clinici Scientifici Maugeri Spa Società Benefit, Milano, Italy

(3) Dipartimento SAIMLAL, Sapienza Università di Roma, Roma, Italy

(4) Dipartimento BIOMETEC, Università degli studi di Catania, Catania, Italy

CORRESPONDING AUTHOR: Fabiana Pia Vitiello, IRCCS Centro Neurolesi "Bonino Pulejo", Messina. E-Mail: fabiana.vitiello@irccsme.it

INTRODUCTION

Recent advances in genomics, neuroimaging, and other technological methods for biomedical data acquisition have led to a significant increase in the complexity and volume of information available in medical research. In this context, machine learning and deep learning approaches are increasingly used; however, their application in clinical and biomedical settings is often limited by the presence of many variables and relatively small sample sizes. This condition may lead to unstable estimates, reduced model performance, and an increased risk of overfitting. For this reason, data augmentation techniques may represent a useful strategy to artificially increase dataset size and evaluate whether synthetic data expansion can improve predictive model performance.

OBJECTIVES

The aim of this study was to compare three data augmentation techniques (RandomOverSampler with shrinkage, Jittering, and SMOTE) to improve performance of a Machine Learning model in muscle disease classification.

METHODS

We used bioelectrical impedance analysis (BIA) parameters and muscle thickness data measured by ultrasound collected in the M-Brain project (PNRR-MCNT1-2023-12378447). The original dataset included 50 subjects (20 diagnosed with sarcopenia according to EWGSOP2 criteria, and 30 without) and 17 variables (nine for BIA and eight for ultrasound). At the beginning, we used the real dataset to identify the most

effective classifier among five supervised classification algorithms (Logistic Regression with L2 regularization, K-Nearest Neighbors, Support Vector Machine with a radial basis function kernel, Linear Discriminant Analysis, and Random Forest). For each augmentation technique, we performed a classification using the KNN (the best classifier) previously trained on real data, on the augmented dataset. We retained only the synthetic samples that were correctly classified or had a probability of being a true positive equal to 1.0, and rejected the rest. Model validation was performed using Leave-One-Out Cross-Validation. Performance was assessed using confusion matrix, ROC-AUC, balanced accuracy, recall, precision, and F1-score. This procedure was also performed on subsets that included either ultrasound or BIA variables.

RESULTS

The KNN model without data augmentation achieved a ROC-AUC of 0.640, balanced accuracy of 0.650, recall of 0.600, precision of 0.571, and F1-score of 0.585. The application of filtered data augmentation improved model performance. RandomOverSampler with shrinkage achieved a ROC-AUC of 0.817, balanced accuracy of 0.758, recall of 0.741, precision of 0.690, and F1-score of 0.714. Jittering achieved a ROC-AUC of 0.826, balanced accuracy of 0.770, recall of 0.815, precision of 0.667, and F1-score of 0.733. SMOTE achieved a ROC-AUC of 0.830, balanced accuracy of 0.764, recall of 0.778, precision of 0.677, and F1-score of 0.724. Additional analyses performed on subsets that included either ultrasound or BIA variables confirmed an overall improvement in performance metrics compared with the original data.

CONCLUSIONS

All the data augmentation techniques considered improved classification performance compared with the analysis performed without augmentation. However, Jittering showed slightly better overall performance. These findings suggest that data augmentation is a suitable strategy for improving classification performance with this type of data, though it must be handled carefully.

Can Data Augmentation Change the Best-Performing Classifier? Evidence from a Machine Learning Study on Sarcopenia

Maria Cristina De Cola⁽¹⁾, Fabiana Pia Vitiello⁽¹⁾, Chiara Scarfi⁽¹⁾, Elisabetta Morini⁽¹⁾, Mattia Papa⁽¹⁾, Danilo Caudo⁽¹⁾, Davide Sattin⁽²⁾, Antonio Musarò⁽³⁾, Nunzio Iraci⁽⁴⁾, Christian Lunetta⁽²⁾

 Maria De Cola 0000-0002-7509-3833

(1) IRCCS Centro Neurolesi "Bonino Pulejo", Messina, Italy

(2) Istituti Clinici Scientifici Maugeri Spa Società Benefit, Milano, Italy

(3) Dipartimento SAIMLAL, Sapienza Università di Roma, Roma, Italy

(4) Dipartimento BIOMETEC, Università degli studi di Catania, Catania, Italy

CORRESPONDING AUTHOR: Maria Cristina De Cola, IRCCS Centro Neurolesi "Bonino Pulejo", Messina, Italy. E-Mail: mariacristina.decola@ircssme.it

INTRODUCTION

One of the M-Brain project's aims is to identify biomarkers that can distinguish between healthy and pathological ageing. The choice of classifier can have a significant impact on predictive performance due to the challenges posed by limited sample size, class imbalance and heterogeneous features. Using data augmentation techniques can improve the performance of different machine learning algorithms.

OBJECTIVES

Compare different supervised classification models to assess whether the best-performing classifier on real data remained the most effective after data augmentation.

METHODS

We used bioelectrical impedance analysis (BIA) parameters and muscle thickness data measured by ultrasound collected. The original dataset included 50 subjects (20 diagnosed with sarcopenia according to EWGSOP2 criteria, and 30 without) and 17 variables (nine for BIA and eight for ultrasound). We trained five supervised classification algorithms—Logistic Regression with L2 regularization (LR-L2), K-Nearest Neighbors (KNN), Support Vector Machine (SVM) with a radial basis function kernel, Linear Discriminant Analysis, and Random Forest—on the original dataset. The dataset was subsequently augmented using jittering, a technique that gener-

ates synthetic samples through small random perturbations of the training data. To preserve the underlying data structure, synthetic instances were generated using the learned weights of each classifier. Any new instance was added to the dataset according to a given threshold. The five algorithms were then re-evaluated to assess the impact of data augmentation on classification performance. Model validation was performed using Leave-One-Out Cross-Validation, and performance by balanced accuracy, recall, precision, and F1-score measures.

RESULTS

In the real-data analysis, KNN showed the best overall performance, achieving a balanced accuracy of 0.650, recall of 0.600, precision of 0.571, and F1-score of 0.585. All classifiers performance improved after data augmentation, although gains were not consistent across all models and settings. The KNN + SVM pipeline performed best, i.e. the feature weights derived from KNN guided the selection of the most suitable synthetic instances generated through jittering, when we used a conservative threshold (1.0–0.7). The five classifiers were then trained and evaluated on the augmented dataset, and SVM achieved the best performance, reaching a balanced accuracy of 0.795, recall of 0.741, precision of 0.769, and F1-score of 0.755. Surprisingly, the LR-L2 + KNN pipeline achieved the highest classification metrics at the lower threshold (0.6–0.5), reaching a balanced accuracy of 0.807, recall of 0.815, precision of 0.733, and F1-score of 0.772 at a threshold of 0.5.

CONCLUSIONS

The results suggest that the classifier that performs best on the original dataset may not remain the best after data augmentation. KNN achieved the highest performance on the original data, whereas SVM outperformed all other classifiers after synthetic data generation. However, when the threshold for generating synthetic instances was relaxed, KNN once again became the top-performing classifier. These findings underscore the importance of reevaluating classification models after data augmentation rather than assuming that the model selected on the original dataset will continue to perform best. They also highlight the importance of balancing class distributions and verifying that the statistical properties of the original variables are preserved following data augmentation.

SARS-CoV-2 Transmission in Italy and Vaccination Status: A Novel Approximate Bayesian Computation Approach for Inferring Compartmental Models from Multiple Data Sources

Chiara Sguanci⁽¹⁾, Cecilia Viscardi⁽²⁾, Michela Baccini⁽¹⁾

(1) Department of Statistics, Computer Science, Applications "G. Parenti", University of Florence, Italy

(2) Department of Economics and Statistics, University of Salerno, Italy

CORRESPONDING AUTHOR: Chiara Sguanci, Department of Statistics, Computer Science, Applications "G. Parenti", University of Florence, Italy. E-Mail: chiara.sguanci@unifi.it

INTRODUCTION

Infectious disease spread is commonly modelled using stochastic compartmental models. However, often model parameters cannot be fully identified from a single data source. Therefore, reliable inference requires the integration of multiple epidemiological datasets, each providing partial and complementary information. The joint analysis of heterogeneous and interdependent data sources, combined with the complexity of realistic compartmental models, typically leads to an analytically intractable likelihood function, making standard Bayesian inference infeasible. We therefore resort to likelihood-free methods, specifically Approximate Bayesian Computation (ABC).

OBJECTIVES

We study SARS-CoV-2 transmission dynamics in Italy during May–October 2022, when the vaccination campaign was in its final phase and vaccination coverage in the population was stable. The aim is to estimate the basic reproduction number stratified by vaccination status, exploiting two complementary data sources: daily COVID-19 deaths from the Protezione Civile (PC) and COVID-19 mortality records stratified by vaccination status from the Istituto Superiore di Sanità (ISS). To this end, we develop a (SIRD)² compartmental model and an ABC algorithm that jointly assimilates both data sources.

METHODS

The (SIRD)² compartmental model partitions the Italian population into vaccinated and unvaccinated sub-popu-

lations with distinct transition rates. Specifically, infection transmission is governed by different reproduction numbers, depending on the vaccination status of the person who is infected. The estimation methodology relies on an ABC algorithm based on a Double Nested Simulator (DNS). The DNS employs a macro-simulator to generate the whole latent data-generating process, i.e. the trajectories of all model compartments, including the daily time series of COVID-19 deaths, and a second simulator that, using the output of the macro-simulator, produces the number of COVID-19 deaths occurring in different time windows, stratified by vaccination status. While the macro-simulator mimics daily data made available by PC, the second simulator mimics contingency tables reported in periodic reports issued by ISS. The ABC algorithm returns samples from an approximate posterior distribution of model parameters conditioned on both data sources simultaneously.

RESULTS

Pooling the PC and ISS sources within the DNS-ABC framework substantially improves parameter identifiability relative to single-source analyses, reducing posterior uncertainty on the vaccination-status-specific reproduction numbers. The stratified (SIRD)² model successfully disentangles the transmission dynamics of vaccinated and unvaccinated individuals over the study period. To ensure parameter identifiability, we fix mortality and recovery rates and estimate $R_{(0,V)}$ and $R_{(0,N)}$, the basic reproduction numbers among susceptible vaccinated and unvaccinated individuals, respectively. The posterior mean of $R_{(0,V)}$ is 2.567 (95% credible interval: 2.358–2.778), whereas of $R_{(0,N)}$ is 1.213 (95%

credible interval: 0.876–1.551). Predictive performance further confirms the advantage of multi-source integration.

CONCLUSIONS

The proposed DNS-ABC framework, together with the (SIRD)² model, enables the estimation of group-specific reproduction numbers by stratifying transmission according to vaccination status. The multi-source ABC approach is well suited to settings lacking a single comprehensive data source, where heterogeneous partial data can be integrated to enable inference. Since the basic reproduction numbers capture both biological transmissibility and behavioural factors, the results suggest that the latter may have contributed to the higher reproduction number estimated among vaccinated than among unvaccinated individuals.

Likelihood-Free Bayesian Reinforcement Learning for Response-Adaptive Randomization in Clinical Trials

Stefano Masini⁽¹⁾, Cecilia Viscardi⁽²⁾, Michela Baccini⁽³⁾

(1) Department of Computer Science, University of Pisa, Italy

(2) Department of Economics and Statistics, University of Salerno, Italy

(3) Department of Statistics, Computer Science, Applications "G. Parenti", University of Florence, Italy

CORRESPONDING AUTHOR: Stefano Masini, Department of Computer Science, University of Pisa, Italy. E-Mail: stefano.masini@unifi.it

INTRODUCTION

Response-adaptive randomization is a class of adaptive clinical trial designs in which treatment allocation probabilities are updated during the study according to accumulating patient outcomes. By assigning future patients with higher probability to the treatment arm that appear more beneficial, these designs aim to improve trial efficiency while addressing ethical concerns related to patient benefit. However, adaptive allocation requires sequential decision-making under uncertainty, especially when the treatment effect and the response-generating mechanism are only partially known. Bayesian Reinforcement Learning (BRL) provides a natural framework for this problem, but standard approaches usually rely on the availability of an explicit likelihood, which may be unavailable or computationally intractable in realistic clinical settings.

OBJECTIVES

We propose a novel likelihood-free full BRL approach for response-adaptive randomization in clinical trials. The objective is to update, online, the posterior distribution of both the unknown environment parameters and the resulting optimal allocation policy, thereby quantifying uncertainty not only on treatment effects but also on the adaptive randomization rule itself.

METHODS

The proposed method combines Approximate Bayesian Computation with Iterated Batch Importance Sampling (IBIS), from which the name Likelihood-Free IBIS (LF-IBIS) is

derived. At each stage of the trial, the parameters that govern the distribution of the unknown response are propagated and reweighted according to their plausibility, which is assessed using simulated pseudo-trial histories rather than an explicit likelihood. For each posterior particle, the corresponding optimal allocation policy is computed, yielding an empirical posterior distribution over response-adaptive policies. We evaluated the method in a two-arm clinical trial simulation, where patient responses are sequentially observed and allocation probabilities are updated as new outcomes become available. A conjugate Bayesian setting was used as a validation scenario, allowing comparison with the exact posterior distribution.

RESULTS

The likelihood-free procedure closely reproduced the sequential posterior learning of the treatment response probabilities obtained in the validation scenario. The resulting posterior distribution of the optimal response-adaptive policy evolved coherently as patient outcomes accumulated, progressively favouring the treatment arm associated with higher expected benefit while retaining policy uncertainty when evidence was still limited. Additional simulations in settings without closed-form posterior updates showed that the method can maintain online Bayesian learning and adaptive policy updating even when standard likelihood-based BRL is not directly applicable.

CONCLUSIONS

LF-IBIS provides a flexible Bayesian framework for response-adaptive randomization in clinical trials, that does not require explicit likelihood specification. The flexibility of the

method enables the straightforward incorporation of adverse events and individual patient characteristics into the implicit model. By producing posterior distributions over both model parameters and allocation policies, the method supports transparent uncertainty quantification. This is particularly relevant for modern clinical research settings in which complex data-generating mechanisms, patient-centered outcomes and sequential decision-making require statistically principled, interpretable and ethically aware adaptive methods.

A Machine Learning Framework for Binary and Multiclass Classification of Cognitive Impairment Using Structural MRI Biomarkers

Alessia Venuto⁽¹⁾, Ambra Faraci⁽¹⁾, Alessandro Grimaldi⁽¹⁾, Giovanni Restuccia⁽¹⁾, Alessandro Sturniolo⁽¹⁾, Annalisa Militti⁽¹⁾, Giuseppa Maresca⁽¹⁾, Angelo Quartarone⁽¹⁾, Angela Marra⁽¹⁾, Lilla Bonanno⁽¹⁾

 Alessia Venuto 0009-0001-5016-1991

(1) IRCCS Centro Neurolesi "Bonino Pulejo"

CORRESPONDING AUTHOR: Alessia Venuto, IRCCS Centro Neurolesi "Bonino Pulejo", Messina, Italy. E-Mail: alessia.venuto@irccsme.it

INTRODUCTION

Alzheimer's disease (AD) is the leading cause of dementia worldwide and is characterized by progressive decline in cognitive functions. Mild Cognitive Impairment (MCI) represents an intermediate stage between normal aging and dementia, with measurable cognitive decline but largely preserved functional independence. However, overlapping clinical features and heterogeneous disease patterns may limit the accuracy of conventional classification approaches. Machine learning (ML) methods can support clinical decision-making by identifying complex neuroanatomical patterns in neuroimaging data and improving the characterization of different stages of cognitive impairment.

OBJECTIVES

This study aimed to compare different supervised ML algorithms for binary and multiclass classification of cognitive impairment using structural MRI-derived volumetric biomarkers. Specifically, we evaluated the performance of models trained using neuroanatomical features alone and models integrating MRI biomarkers with demographic variables. In addition, SHAP-based explainability analysis was applied to identify the features contributing most to model predictions and to improve the interpretability of the classification framework.

METHODS

The study included 204 participants, comprising 54 patients with AD, 79 individuals with MCI, and 71 HC.

T1-weighted structural MRI scans were processed using automated pipelines to extract cortical and subcortical volumetric measures. Five supervised ML algorithms were evaluated: Logistic Regression, Random Forest, XGBoost, LightGBM, and Support Vector Machine (SVM). All models were implemented within a standardized pipeline including feature scaling, feature selection using SelectKBest with ANOVA F-test, hyperparameter tuning with GridSearchCV, and stratified 5-fold cross-validation. Two experimental settings were compared: models trained only on MRI-derived volumetric features and models trained on a combination of MRI and demographic features. Performance was evaluated using accuracy, F1-score, ROC-AUC, and confusion matrices. Finally, SHAP analysis was applied to the best-performing models to assess the contribution of individual features to classification decisions.

RESULTS

In the binary classification task distinguishing MCI/AD from HC, all models achieved higher performance than in the multiclass setting. Using only MRI-derived volumetric features, Logistic Regression showed the best performance, reaching an accuracy of 0.70 and a ROC-AUC of 0.76. The inclusion of demographic variables further improved classification performance, with SVM achieving the highest results, corresponding to an accuracy of 0.75 and a ROC-AUC of 0.80. In the multiclass classification task involving AD, MCI, and HC, Random Forest achieved the best performance when using only volumetric biomarkers, with an accuracy of 0.53 and a ROC-AUC of 0.69. When MRI and demographic features were combined, SVM achieved the best performance, reaching an accuracy of 0.56 and a ROC-AUC of 0.72. SHAP

analysis identified the amygdala, hippocampus, corpus callosum subdivisions, and age as the most influential features driving model predictions.

CONCLUSIONS

The findings support the potential role of structural MRI biomarkers combined with interpretable ML approaches for the classification of cognitive impairment. The integration of demographic variables improved model performance, particularly in binary discrimination tasks, whereas multiclass classification remained more challenging, likely due to the clinical and neuroanatomical overlap between MCI and AD patterns. This framework may represent a promising approach for future clinical decision-support tools aimed at improving diagnostic stratification, although further validation in larger independent cohorts is required.

Patient-Reported Outcome Measures to Inform HPV Vaccination Strategies in University Students

Martina Barchitta^(1,2), Andrea Maugeri^(1,2), Erminia Di Liberto⁽¹⁾, Giuliana Favara⁽¹⁾, Roberta Magnano San Lio⁽¹⁾, Marco Enea⁽²⁾, Laura Maniscalco⁽²⁾, Domenica Matranga⁽²⁾, Antonella Agodi⁽¹⁾

 Martina Barchitta 0000-0002-0905-5003

(1) Department of Medical, Surgical Sciences and Advanced Technologies "GF Ingrassia", University of Catania, Catania, Italy

(2) Department of Health Promotion, Mother and Child Care, Internal Medicine and Medical Specialties, School of Medical Statistics and Biometry, University of Palermo, Palermo, Italy

CORRESPONDING AUTHOR: Martina Barchitta, Department of Medical, Surgical Sciences and Advanced Technologies "GF Ingrassia", University of Catania, Catania, Italy. E-Mail: martina.barchitta@unict.it

INTRODUCTION

Patient-reported outcome measures (PROMs) are essential tools to inform improvements in healthcare quality, to determine what matters to individuals in care and to guide clinical and prevention decision-making. Timely vaccination of university students is a cost-effective strategy to reduce infection rates and the burden of Human Papillomavirus (HPV)-associated consequences including various types of cancer. In fact, a recent meta-analysis found that university students exhibit limited awareness of HPV and vaccination, with about half willing to be vaccinated.

OBJECTIVES

The present study aims to assess university students' knowledge, attitudes, and perceptions regarding HPV infection and HPV vaccination, as well as the factors influencing vaccine uptake and vaccination intention. The study seeks to identify the sociodemographic, behavioral and psychosocial determinants associated with both actual vaccination adherence and future vaccination intentions. By providing a comprehensive understanding of the drivers and barriers to HPV vaccine acceptance among university students, the findings are expected to support the development of targeted educational and public health interventions aimed at improving HPV vaccination coverage in this population.

METHODS

A cross-sectional observational study was conducted on a sample of students at the University of Catania enrolled during the three open days conducted at the University of Catania to

promote HPV vaccination in students, during the years 2025 and 2026. A structured questionnaire was administered to collect sociodemographic and behavioral information as well as to assess participants' knowledge and attitudes. The data were analyzed using descriptive statistics, logistic regression, and clustering to identify behavioral determinants and profiles.

RESULTS

A total of 270 students (mean age 22.1 years $\pm$ 3.16), 69,5 female, were enrolled. A marked gender difference in HPV vaccination coverage was shown with approximately 66% of female students having received the vaccine compared with only about 33% of male students. A total of 75% of vaccinated individuals came from families with at least a high school diploma or university degree. Knowledge of the HPV-cervical cancer link exceeded 90%, while knowledge relating to other HPV associated pathologies were at around 40-45%. Logistic regression analysis identified trust in the vaccine and health authorities as the main predictors of vaccination intention, while low risk perception and confidence in natural immunity represented barriers. This model showed good predictive power (accuracy 73%; AUC 0.78). Cluster analysis identified three distinct student profiles: a group characterized by positive attitudes toward vaccination and high levels of knowledge, a group that was well-informed but hesitant, and a third group consisting of skeptical and poorly informed individuals.

CONCLUSIONS

PROMs are useful tools to explore students' self-reported

awareness of HPV-related diseases, perceived susceptibility and severity, beliefs about vaccine effectiveness and safety, and willingness to receive or recommend HPV vaccination. Results of the present study emphasize the importance of developing targeted, evidence-based interventions, including personalized educational initiatives and multidisciplinary collaborations, to reduce disparities and enhance vaccine uptake among university students.

A Multivariate Age-Course Framework for Senescence-Related Gene Expression Profiling in a Mouse Model of Osteogenesis Imperfecta

Camilla Torriani^{(1)*}, Wendy Franco Perez^{(2)*}, Roberta Besio⁽²⁾, Simona Villani⁽¹⁾, Antonella Forlino⁽²⁾

 Camilla Torriani 0009-0006-5879-6795

(1) Department of Public Health, Experimental and Forensic Medicine, Unit of Biostatistics and Clinical Epidemiology, University of Pavia, Pavia, Italy

(2) Department of Molecular Medicine, Biochemistry Unit, University of Pavia, Pavia, Italy

* These authors contributed equally to this work

CORRESPONDING AUTHOR: Camilla Torriani, Department of Public Health, Experimental and Forensic Medicine, Unit of Biostatistics and Clinical Epidemiology, University of Pavia, Pavia, Italy. E-Mail: camilla.torriani01@universitadipavia.it

INTRODUCTION

Osteogenesis imperfecta is primarily characterised by skeletal fragility, but collagen alterations may also affect extracellular tissues, making multitissue ageing-related molecular profiling biologically relevant. Ageing is accompanied by molecular changes that do not occur in isolation, but emerge as coordinated shifts across biological pathways, tissues and time. In experimental studies, however, the common practice of analysing molecular data marker by marker may overlook coordinated expression programmes and tissue-specific developmental trajectories. This issue is particularly relevant in small age-course studies, where complex experimental designs require methods able to preserve the structure of the data while providing interpretable summaries of multigene variation.

OBJECTIVES

This study aimed to evaluate whether senescence-related gene expression showed coordinated age-course trajectories in BrlIV mice carrying the heterozygous Col1a1G349C/+ mutation, a model of Osteogenesis Imperfecta, compared with wild-type Col1a1+/+ mice. Specifically, we evaluated the joint expression profile of P16, P21, P53 and LMNB1 across age, genotype and tissue, with the goal of identifying tissue-specific molecular trajectories rather than isolated gene-level differences.

METHODS

Expression levels of P16, P21, P53 and LMNB1 were analysed in lung, bone and heart samples collected at 6, 12 and 18 months from wild-type and heterozygous mice. Gene-expression values were $\log_{10}(x+1)$ -transformed and z-standardised to obtain comparable marker profiles. The four genes were then analysed as a multivariate molecular signature. Differences in coordinated expression patterns were assessed using permutational multivariate analysis of variance based on distance matrices, with permutations constrained within experimental batch identifiers to account for potential batch-related variability. A global model including age, genotype, tissue and their interactions was complemented by tissue-specific models. Homogeneity of multivariate dispersion and sensitivity analyses using an alternative distance metric were used to assess robustness. Ordination-based visualisation, heatmaps and age-course plots were used to support interpretation. Gene-specific mixed-effects models were fitted as complementary analyses to identify the markers contributing to the multivariate signal.

RESULTS

The global multivariate model showed that senescence-marker expression differed across experimental conditions, with age, genotype and tissue acting jointly rather than independently. The model explained approximately 38% of the total multivariate variation, indicating a substan-

tial reorganisation of the combined gene-expression profile. Tissue-specific analyses revealed that these trajectories were not uniform across organs. Lung and heart showed clearer genotype-dependent age-course patterns, whereas bone displayed more heterogeneous expression dynamics. Higher-order interaction terms supported the presence of tissue-specific age-related differences between wild-type and heterozygous animals. These findings were consistent in sensitivity analyses using an alternative distance metric and were not explained by major global differences in multivariate dispersion. Visualisation of the molecular profiles showed relatively stable age-course trajectories in wild-type animals and more dynamic, tissue-dependent changes in heterozygous animals. Gene-specific models provided complementary evidence that individual markers contributed differently to the coordinated multivariate patterns.

CONCLUSIONS

This exploratory study shows that small-scale molecular datasets can be analysed as coordinated age-course profiles rather than as independent marker-level outcomes. Integrating constrained permutation-based multivariate testing, visualisation of molecular trajectories and gene-specific follow-up models provides an interpretable framework for studying molecular profiles across tissues and biological conditions. This approach may help bridge classical biostatistical modelling and modern molecular data analysis in experimental studies with limited sample size but complex temporal structure.

Estimating the PM_{2.5}–Mortality Concentration–Response Function in the Rome Longitudinal Study (ROLS): A Methodological Comparison of Linear, Spline, and SCHIF Approaches

Annafrancesca Smimmo⁽¹⁾, Federica Nobile⁽²⁾, Paola Schiattarella⁽¹⁾, Teresa Speranza⁽¹⁾, Piergiacomo Di Gennaro⁽¹⁾, Mario Fordellone⁽¹⁾, Carla Pollastro⁽¹⁾, Vittorio Simeon⁽¹⁾, Simona Signoriello⁽¹⁾, Massimo Stafoggia⁽²⁾, Paolo Chiodini⁽¹⁾

 Annafrancesca Smimmo 0009-0002-3877-8398

(1) Medical Statistics Unit, University of Campania “Luigi Vanvitelli”, Caserta, Italy

(2) Environmental, Occupational Epidemiology and Tumor Registry Unit, Department of Epidemiology of the Regional Health Service - ASL Roma 1 Lazio Region, Roma, Italy

CORRESPONDING AUTHOR: Annafrancesca Smimmo, Medical Statistics Unit, University of Campania “Luigi Vanvitelli”, Caserta, Italy.
E-Mail: annafrancesca.smimmo@unicampania.it

INTRODUCTION

In health impact assessments (HIAs) of air pollution, a crucial issue is the correct specification of the concentration-response function (CRF) between exposure and outcome. Traditional approaches often assume a linear relationship, with the risk of introducing bias when the association is actually non-linear. In this context, Shape Constrained Health Impact Functions (SCHIF) represent a methodological advancement developed specifically for HIA. They allow flexible curves that are consistent with biological plausibility constraints to be modelled, while at the same time they estimate a single effect coefficient, since non-linearity is incorporated into the transformation of exposure in the risk function. SCHIF returns three estimates of the CRF from the optimal model, an ensemble of the best three models, and an ensemble of all estimated models. Simulation studies show that SCHIF better describe the association between PM_{2.5} and the outcome compared with linear models, but this has not yet been verified using real data.

OBJECTIVES

To evaluate, in a real cohort, the methodological robustness of the estimated CRF between chronic exposure to fine particulate matter (PM_{2.5}) and non-accidental mortality

obtained with the optimal SCHIF model; to compare the estimates with those obtained from Cox models with exposure specified as a linear term or by natural splines (NS).

METHODS

The administrative cohort of residents (30+ years) in Rome on 09/10/2011 (2011 Census) was analysed, with follow-up to 31/12/2019. The following were estimated: (i) a Cox model with a linear term for PM_{2.5} (L), (ii) a Cox model with NS with 3 knots (S), and (iii) the SCHIF model, adjusted for individual- and area-level variables. The Bayesian Information Criterion (BIC) was used for model selection.

RESULTS

All analytical approaches highlight the presence of a significant effect of PM_{2.5} exposure on the outcome. In L, the estimated exposure effect is $\beta=0.012$ (SE=0.029). However, the comparison with model S is statistically significant, indicating that the association is not adequately described by a linear relationship. In model S, the first spline term is significant ($\beta=0.023$; SE=0.009), whereas the additional non-linear terms are not. The optimal SCHIF model identifies a non-linear relationship, with deviation from linearity at the lowest PM_{2.5} levels, with an estimated coefficient of $\beta(z)=0.140$

$(SE=0.032) \times \omega(z | \min(z), 0.1) \times \log(1+z)$. In the comparison between models, the minimum BIC is observed for the SCHIF model, supporting a better fit to the data.

CONCLUSIONS

The SCHIF model shows the best fit and is confirmed as a promising approach for estimating more realistic and robust CRFs, useful in health impact assessments.

ACKNOWLEDGMENTS

Project carried out with the technical and financial support of the Ministry of Health - PNC PREV-A-2022-12376.

Estimating the Temperature-Mortality Relationship Using a Novel Local Estimand: Application to the Metropolitan Area of Florence

Giulio Biscardi⁽¹⁾, Alessandra Mattei⁽¹⁾, Chiara Marzi⁽¹⁾, Daniela Nuvolone⁽²⁾, Michela Baccini⁽¹⁾

 Giulio Biscardi 0009-0000-2702-4238

(1) Department of Statistics, Computer Science and Applications "Giuseppe Parenti" (DiSIA), University of Florence, Florence, Italy

(2) Tuscan Regional Health Agency, Florence, Italy

CORRESPONDING AUTHOR: Giulio Biscardi, Department of Statistics, Computer Science and Applications "Giuseppe Parenti" (DiSIA), Florence, Italy. E-Mail: giulio.biscardi@unifi.it

INTRODUCTION

The Average Dose-Response Function (aDRF) is a standard global estimand for evaluating the causal effect of a continuous exposure. The estimation of the aDRF relies, among others, on the assumption that all units share common support. However, when studying the short-term effects of temperature, some units (days) cannot realistically be exposed to certain exposure doses. The lack of common support can lead to problematic extrapolations.

OBJECTIVE

We propose a method to estimate local aDRFs within a Potential Outcomes (PO) framework, and use it to identify the short-term relationship between average temperature (lag 0-3) and natural deaths in the Florence metropolitan area, Italy (2010-2019).

METHODS

Under the Stable Unit Treatment Value, we define for each day the number of deaths that would have occurred for different temperature levels, i.e. the POs. Only the PO corresponding to the actually observed exposure is available, whereas the remaining POs are missing and must be imputed. We impute them using the following procedure. We do it according to the following procedure. Under local weak unconfoundedness, we first estimate the Generalized Propensity Score (GPS), defined as the conditional density of temperature given confounders. Then, for each day, we estimate the GPS at the observed temperature, R_i , and over a pre-specified grid of temperature val-

ues. Finally, we fit a flexible model on daily mortality given temperature and the R_i and use it to impute the POs for every day at each temperature level in the grid. Most existing methods rely on a global overlap assumption and target a global aDRF, estimated by averaging the imputed potential outcomes across all days for each temperature level. In contrast, we rely on a local overlap assumption and estimate local aDRFs for subgroups of days for which local overlap holds. We define temperature intervals and for each interval we identify the subgroup of days based on the overlap between the R_i estimated for the days with temperature within the interval and the GPS predicted in the middle of the interval for the days with temperature outside the interval. The aDRF is then estimated for each interval, by averaging the imputed POs over the selected subgroup at each temperature level within the interval. We perform a simulation study to evaluate the performance of our method.

RESULTS

Our procedure has good performance in terms of Absolute Bias and Root Mean Square Error. For the study area, we estimate 5 local aDRFs for the temperature intervals [0-6], [6-12], [12-18], [18-24], [24-30]. The local aDRFs are consistent with the literature, highlighting the U-shaped risk related to extreme temperatures.

CONCLUSIONS

Our local approach addresses the common support problem, avoiding inappropriate extrapolation. It yields dose-response curves that have causal interpretation and can be readily used to conduct impact evaluations under counterfactual scenarios.

Missing Data Imputation Strategies for Real-World Clinical Registry Data in Patients with Substance Use Disorder: A Comparative Analysis Preceding the Estimation of Cardiovascular Outcomes

Letizia Lorusso⁽¹⁾, Nicola Bartolomeo⁽²⁾, Francesco Ciccirillo⁽³⁾, Antonio Boschini⁽⁴⁾, Paolo Ottogalli⁽⁴⁾, Paolo Trerotoli⁽²⁾

 Letizia Lorusso 0000-0002-2058-9284

(1) Department of Precision and Regenerative Medicine - (DiMePre-J), University of Bari Aldo Moro, Bari, Italy

(2) Interdisciplinary Department of Medicine, University of Bari Aldo Moro, Bari, Italy

(3) UO Cardiology Unit, "V. Fazzi" Hospital, Lecce, Italy

(4) San Patrignano Community Medical Center, Coriano (RN), Italy (A.B. Medical Director; P.O. Health Director)

CORRESPONDING AUTHOR: Letizia Lorusso, Department of Precision and Regenerative Medicine - (DiMePre-J), University of Bari Aldo Moro. E-Mail: l.lorusso80@phd.uniba.it

INTRODUCTION

Real-world clinical registry data from addiction treatment centers are a valuable source for investigating the association between substance use disorder (SUD) and cardiovascular disease (CVD) but are frequently affected by missing data arising from irregular attendance and heterogeneous data collection. Inadequate handling of missing data may introduce substantial bias in the estimation of epidemiological associations, particularly when missingness varies markedly across variables.

OBJECTIVES

To compare the performance of non-parametric missing data imputation methods in a mixed-type real-world clinical registry and evaluate their impact on cardiovascular risk estimates across variables with increasing levels of missingness.

METHODS

A retrospective cohort of 4,895 unique patients admitted to the San Patrignano addiction treatment center (Bologna, Italy) between 2015 and 2024 was analyzed using 40 variables (8 continuous, 32 categorical). The outcome variable was withheld from the imputation model throughout.

Three imputation methods were assessed: missForest, kNN, and guided hot-deck, all tuned by grid search. MissForest was evaluated using built-in error estimates: NRMSE for quantitative and PFC for categorical variables, whereas kNN and guided hot-deck were assessed through method-specific proxy indicators. A binary cardiovascular outcome was defined as the presence of cardiovascular disease-related

conditions. Odds ratios (ORs) and 95% confidence intervals (95%CI) were estimated using complete-case analysis and after imputation for variables characterized by different levels of missingness: cocaine use (1%), obesity (12%), total cholesterol (39%) and overdose history (73%).

To quantify the sensitivity of epidemiological estimates to the imputation strategy, two summary indicators were calculated:

$$\Delta OR = \frac{(OR_{max} - OR_{min})}{OR_{CC}}$$

and

$$\Delta IC = \frac{(IC_{maxwidth} - IC_{minwidth})}{IC_{widthCC}}$$

where OR_{CC} and $IC_{widthCC}$ denote the odds ratio and confidence interval width obtained from the complete-case analysis. All statistical analyses were done in R (version 4.5.0).

RESULTS

kNN showed a more favorable proxy-based profile than guided hot-deck, with lower continuous and categorical proxy values (0.032 and 0.072 vs. 0.091 and 0.226, respectively), while missForest yielded NRMSE=0.604, indicating a relatively high value, and PFC=0.110. The impact of imputation on epidemiological estimates increased with missingness

proportion. For low-missingness variables (cocaine 1%, obesity 12%), OR estimates were stable across all methods (OR range: 4.0–4.1; $\Delta\text{OR}=0.26$; $\Delta\text{IC}=0.44$). For total cholesterol (39% missing), modest differences emerged across imputation strategies ($\Delta\text{OR}=0.62$; $\Delta\text{IC}=0.55$), although the epidemiological interpretation remained unchanged. The largest variability was observed for overdose history (73% missing), where OR estimates ranged from 1.8 to 3.3 ($\Delta\text{OR}=0.62$; $\Delta\text{IC}=0.55$). Nevertheless, overdose was consistently identified as a positive risk factor across all methods.

CONCLUSIONS

The practical impact of missing data imputation depended strongly on the amount of information recovered from incomplete variables. For low-to-moderate missingness, complete-case and imputed analyses yielded consistent estimates across algorithms. Differences in point estimates and confidence interval widths emerged only above 70% missingness, without changing the direction or substantive interpretation. Because reconstruction-performance metrics may be structurally heterogeneous across methods, they are best restricted to within-method tuning. In this setting, the operational value of imputation was more appropriately assessed through its impact on downstream epidemiological inference, with ΔOR and ΔIC providing an interpretable framework for comparing imputation strategies in real-world registry studies.

Simultaneous Prediction of Multiple Functional Outcomes in Post-Stroke Rehabilitation Using Machine Learning

Valeriano Videtta^{(1)*}, Chiara Marzi^{(1)*}, Sofia Partilora⁽¹⁾, Leonardo Corsi⁽²⁾, Julieta Giacani⁽²⁾, Jorge Navarro⁽³⁾, Lucia Falco⁽⁴⁾, Caterina Tramonti⁽⁵⁾, Ersilia Lucenteforte^{(1)*}, Francesca Cecchi^{(2,6)†}, Andrea Mannini^{(2)†}

 Ersilia Lucenteforte 0000-0001-5608-5902

(1) Dipartimento di Statistica, Informatica, Applicazioni "G. Parenti", Università degli studi di Firenze, Firenze, Italy

(2) IRCCS Fondazione Don Carlo Gnocchi ETS, Firenze, Firenze, Italy

(3) IRCCS Fondazione Don Carlo Gnocchi ETS, Milano, Italy

(4) Fondazione Don Carlo Gnocchi ETS, Parma, Italy

(5) Fondazione Don Carlo Gnocchi ETS, La Spezia, Italy

(6) Dipartimento di Medicina Sperimentale e Clinica, Università degli studi di Firenze, Firenze, Italy

* These authors have contributed equally

† These authors have contributed equally

CORRESPONDING AUTHOR: Lucenteforte Ersilia, Dipartimento di Statistica, Informatica, Applicazioni "G. Parenti", Università degli studi di Firenze, Firenze, Italy. E-Mail: ersilia.lucenteforte@unifi.it

INTRODUCTION

Stroke is a leading cause of long-term disability worldwide. Despite the central role of rehabilitation to functional recovery, accurately predicting outcomes remains difficult due to the marked variability in post-stroke recovery trajectories. Machine learning may help modelling complex, non-linear relationships between clinical variables and outcomes. However, most existing approaches focus on single outcomes, neglecting the multidimensional nature of post-stroke recovery. Multi-output models may represent a more clinically coherent and statistically efficient alternative.

OBJECTIVES

This study aims to compare single-output and multi-output machine learning approaches for predicting rehabilitation outcomes at discharge in post-stroke patients undergoing intensive rehabilitation. Specifically, we trained and compared four algorithms in both their single-output and multi-output formulations to predict three functional outcomes and evaluated whether joint modelling improves prediction accuracy over independent single-output machine learning models.

METHODS

Data were obtained from the STRATEGY registry, a multicentre Italian prospective cohort of 743 post-stroke patients admitted for intensive inpatient rehabilitation at IRCCS Fondazione Don Carlo Gnocchi centres from 2022 onward. The registry systematically collects standardised demographic, clinical, neurological, cognitive, and functional data from adults admitted within 30 days of ischemic or haemorrhagic stroke, assessed at admission and discharge according to the PMIC2020 framework and additional measures of clinical complexity and comorbidity. Outcomes of interest were Modified Barthel Index (mBI), Trunk Control Test (TCT), and modified Rankin Scale (mRS) scores at discharge. After data cleaning, 702 patients were included. Model development and evaluation were conducted using a nested validation scheme. In the outer loop, the dataset was split into a development set ($n = 561$, 80%) and a held-out test set ($n = 141$, 20%). Within the development set, an inner 5-fold cross-validation (CV) procedure was used for hyperparameter tuning and model selection. Missing data were imputed using k-Nearest Neighbours, with the imputation model fitted exclusively on training folds to prevent information leakage. Each training fold and the final development set were augmented via synthetic data

generation using the R package *synthpop*.

Four machine learning algorithms were evaluated in both single-output and multi-output formulations: Random Forest, Elastic Net, CatBoost, and XGBoost. For each algorithm, hyperparameters were optimised through grid search within the inner 5-fold CV. Model selection was based on a composite metric defined as the average Normalised Root Mean Square Error (NRMSE) across the three outcomes. The best-performing single-output and multi-output models were retrained on the full development set and evaluated on the held-out test set.

RESULTS

The best-performing multi-output model was the multi-output Random Forest, whereas Elastic Net achieved the best performance among single-output models, with comparable performance during cross-validation. On the held-out test set, the multi-output approach achieved better performance on all three outcomes, with the largest improvements observed for mBI and TCT.

CONCLUSIONS

Multi-output modelling, by leveraging the shared structure among rehabilitation outcomes, improved prediction accuracy across all three targets compared to independent single-output models. These findings suggest that multi-output machine learning approaches may provide a more effective framework for prognostic modelling in post-stroke rehabilitation.

Sample Size Determination in Studies Involving Machine Learning Algorithms: A Case Study on Post-Stroke Rehabilitation

Sofia Partilora^{(1)*}, Chiara Marzi^{(1)*}, Valeriano Videtta⁽¹⁾, Silvia Pancani⁽²⁾, Ilaria Pellegrini⁽²⁾, Jorge Navarro⁽³⁾, Lucia Falco⁽⁴⁾, Caterina Tramonti⁽⁵⁾, Ersilia Lucenteforte⁽¹⁾, Andrea Mannini^{(2)†}, Francesca Cecchi^{(2,6)†}

 Ersilia Lucenteforte 0000-0001-5608-5902

(1) Dipartimento di Statistica, Informatica, Applicazioni "G. Parenti", Università degli Studi di Firenze, Firenze, Italy

(2) IRCCS Fondazione Don Gnocchi ETS, Firenze, Firenze, Italy

(3) IRCCS Fondazione Don Gnocchi ETS, Milano, Italy

(4) Fondazione Don Gnocchi ETS, Parma, Italy

(5) Fondazione Don Gnocchi ETS, La Spezia, Italy

(6) Dipartimento di Medicina Sperimentale e Clinica, Università degli studi di Firenze, Firenze, Italy

* These authors contributed equally to this work

† These authors contributed equally to this work

CORRESPONDING AUTHOR: Ersilia Lucenteforte, Dipartimento di Statistica, Informatica, Applicazioni "G. Parenti", Università degli studi di Firenze, Firenze, Italy. E-Mail: ersilia.lucenteforte@unifi.it

INTRODUCTION

Sample size determination remains an important methodological challenge in the development of machine learning (ML) prediction models in clinical studies. Unlike classical statistical approaches such as regression-based methods, there is still no widely accepted procedure for estimating the amount of data required to develop reliable ML models, and current practice remains highly variable and largely based on informal or heterogeneous criteria.

OBJECTIVES

This study had two main aims: 1) to identify and critically compare the methods proposed for sample size determination in ML-based clinical prediction studies through a systematic review; and 2) to apply and compare the methods identified in the systematic review within a real-world case study based on the STRATEGY study, a multicentre Italian cohort originally including 743 post-stroke patients admitted to intensive inpatient rehabilitation.

METHODS

The systematic review was conducted using searches in three bibliographic databases: MEDLINE, the Cochrane Li-

brary, and EMBASE. In addition, a forward citation search was performed in Scopus to identify studies that applied the identified methodological approaches to justify sample size in their own ML-based analyses.

The identified methods were then applied to the STRATEGY dataset. Model development and evaluation were conducted using a nested validation framework. First, the dataset was split into a development set (80%) and an independent test set (20%). Within the development set, an inner 5-fold cross-validation procedure was used for models training (i.e. SVM, RF and XGBoost), hyperparameter tuning, and model selection. Before models fitting, missing data were imputed using the K-Nearest Neighbours algorithm, and synthetic data generation was performed within each training fold using the R synthpop package.

RESULTS

The literature search identified 35 studies meeting the inclusion criteria: 9 methodological papers and 26 empirical studies. The most used approaches for sample size determination for ML models are an a priori formula, an empirical learning curve method, and a parametric learning curve extrapolation. We applied them to the STRATEGY dataset, with the aim of predicting the modified Barthel Index (mBI) at discharge.

The three sample size determination approaches yielded concordant, but numerically different estimates. The a priori method produced the largest required sample size, consistent with its conservative analytical nature, followed by the empirical learning curve and the parametric extrapolation method. Notably, all three sample size estimates exceeded the sample size used by Campagnini et al. (2025, Scientific Reports) in a similar post-stroke rehabilitation population. Although not directly comparable, the lower predictive performance reported by Campagnini et al. is consistent with their smaller sample size, which was below estimates obtained in our post hoc assessment.

CONCLUSION

The sample size estimates obtained in this study should not be interpreted as fixed thresholds, but as complementary reference points that may support study planning. To the best of our knowledge, this is one of the first works to systematically map and compare the available methods for sample size determination in ML-based clinical prediction studies.

Target Trial Emulation of a Non-Pharmaceutical Intervention within the Stage Project: Smoking Cessation during Pregnancy as a Case Study

Alessandra Macciotta⁽¹⁾, Daniela Zugna⁽¹⁾, Lorenzo Richiardi⁽¹⁾

 Alessandra Macciotta 0000-0002-0287-7979

(1) Department of Medical Sciences, University of Turin and CPO Piemonte, Turin, Italy

CORRESPONDING AUTHOR: Alessandra Macciotta, Department of Medical Sciences, University of Turin and CPO Piemonte, Turin, Italy.
E-Mail: alessandra.macciotta@unito.it

INTRODUCTION

Randomized controlled trials (RCTs) are the gold standard for causal inference, but observational studies can also provide robust causal evidence. Target Trial Emulation (TTE) applies the principles of RCTs to observational data by explicitly specifying the protocol of a hypothetical trial before conducting the analysis. While this framework has been applied to pharmaceutical interventions, non-pharmaceutical interventions present additional challenges: exposures often correspond to behaviors rather than clearly defined interventions capable of inducing those behaviors.

OBJECTIVES

To describe the development and implementation of a TTE study within the European STAGE project investigating the effect of smoking cessation during pregnancy on maternal and child outcomes.

METHODS

We developed a target trial protocol specifying eligibility criteria, treatment strategies, assignment procedures, outcomes, and follow-up for both the hypothetical target trial and its observational emulation. The emulation used data from the NINFEA study, an internet-based birth cohort including approximately 7,500 mother-child pairs, with repeated questionnaires administered from pregnancy until 16 years after delivery.

RESULTS

The target trial includes adult pregnant women who smoked before pregnancy and compares two treatment strategies: smoking cessation (0 cigarettes/day) versus continued smoking (>0 cigarettes/day). Randomization is stratified by maternal age, educational level, parity, recruitment center, smoking history, and pre-pregnancy morbidities. Maternal outcomes include smoking habits after delivery and hospitalization within 18 months postpartum; child outcomes include birth weight and respiratory diseases. Follow-up begins at conception and ends at outcome occurrence, loss to- or end of follow-up. Both intention-to-treat and per-protocol effects are estimable.

The observational emulation includes all adult women in the NINFEA cohort who smoked before pregnancy. The availability of repeated smoking assessments allowed the emulation of two nested target trials, considering the assignment to treatment strategies at two time points: smoking cessation either at conception or during the first trimester. Randomization is emulated through adjustment for baseline covariates corresponding to the stratification variables. Outcomes are assessed using questionnaires administered 6 and 18 months, and 7 years after delivery.

Smoking cessation at conception substantially reduced the odds of postpartum smoking relapse at both 6 months (OR = 0.13; 95% CI: 0.10–0.18) and 18 months (OR = 0.21; 95% CI: 0.15–0.28), and was associated with a higher birth weight (β = 97.6 g; 95% CI: 42.5–152.7 g). Similar results were obtained with smoking cessation within the first trimester.

CONCLUSIONS

Our results showed positive effects of smoking cessation during pregnancy on both maternal and child health. The study, demonstrating the feasibility of emulating RCTs of non-pharmaceutical interventions using observational data, also highlights methodological considerations specific to behavioral interventions. Unlike pharmacological treatments, smoking cessation can be achieved through multiple interventions (e.g. counselling or financial incentives), which were not directly observed in the cohort. Consequently, the emulation focused on well-defined behavioral strategies using explicit cessation thresholds rather than the underlying intervention. Furthermore, identifying pregnancy onset as the trial time zero ensured alignment between the target trial and the observational data, while repeated exposure measurements enabled the assessment of adherence to the assigned treatment strategy over time.

Data-Driven Identification of Clinical Phenotypes Associated with Early Infection Risk after Liver Transplantation Using the Open-Source Framework PhenoCluster

Ettore Rocchi^(1,3), Cecilia Bonazzetti^(1,2), Martina Casarini⁽¹⁾, Matteo Rinaldi^(1,2), Cristiana Laici⁽⁴⁾, Antonio Siniscalchi⁽⁴⁾, Matteo Ravaoli⁽⁵⁾, Matteo Cescon⁽⁵⁾, Maria Cristina Morelli⁽⁶⁾, Gastone Castellani^(1,3), Pierluigi Viale^(1,2), Maddalena Giannella^(1,2)

 Ettore Rocchi 0000-0002-7612-2819

(1) Department of Medical and Surgical Sciences, Alma Mater Studiorum, University of Bologna, Bologna, Italy

(2) Infectious Diseases Unit, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy

(3) IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy

(4) Anesthesia and Intensive Care, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy

(5) Hepatobiliary and Transplant Surgery, Department of Medical and Surgical Sciences, University of Bologna, Bologna, Italy

(6) Internal Medicine Unit for the Treatment of Severe Organ Failure, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy

CORRESPONDING AUTHOR: Ettore Rocchi, Department of Medical and Surgical Sciences, Alma Mater Studiorum, University of Bologna, Bologna, Italy. E-Mail: ettore.rocchi3@unibo.it

INTRODUCTION

Early infections are a leading cause of morbidity and mortality after liver transplantation, yet risk stratification typically relies on single markers rather than on integrated donor-recipient-perioperative profiles. The data-driven identification of latent subgroups offers a complementary approach, although its clinical application is often limited by the lack of standardized, reproducible pipelines that are robust against data leakage.

OBJECTIVES

To identify, through latent class and latent profile analysis (LCA/LPA), clinically interpretable phenotypes associated with early infection risk in a liver transplantation cohort, while validating PhenoCluster, an open-source Python framework that integrates phenotype discovery and downstream inference into a single configurable pipeline.

METHODS

PhenoCluster (<https://github.com/EttoreRocchi/phenocluster>) implements an end-to-end configurable pipeline.

We analyzed 553 liver transplant recipients (2018-2024) described by 30 mixed donor, recipient, and perioperative variables. Phenotypes were identified through LCA/LPA with StepMix, handling missing data via full-information maximum likelihood (FIML) and selecting the number of classes using information criteria (BIC, AIC, ICL). To prevent data leakage, the train/test split (80/20) was performed before preprocessing. The analysis comprised three steps: (i) estimation of the latent class model and selection of the number of phenotypes on the training set; (ii) prediction of phenotypes on the test set, to assess robustness on unobserved patients; (iii) re-estimation on the full cohort for the final estimates. Classification quality was assessed using the average posterior probability (AvePP) and stability via consensus clustering. Associations with outcomes were estimated using logistic regression (OR, 95% CI, Benjamini-Hochberg corrected p-values) and survival models (Cox, Kaplan-Meier).

RESULTS

Three clinically coherent phenotypes emerged, along a gradient of liver disease severity. Phenotype A (n=251; median MELD 12) served as the reference (any infection at 30 days: 27%). Phenotype B (n=179; median MELD 27, Child-

Pugh B/C 90%, thrombocytopenia, hypoalbuminemia, and frequent CRE colonization) showed the highest infectious burden: an approximately threefold increase in any infection (54%; OR 3.12; 95% CI 2.01-4.84; HR 2.35; 95% CI 1.75-3.16), with excess bacterial (OR 2.70), viral (OR 3.10), and fungal (OR 4.39) risk. Phenotype C (n=123; median MELD 9, Child-Pugh A 84%) showed a selectively bacterial risk (OR 1.64), without excess viral or fungal risk; 30-day mortality did not differ across phenotypes. Classification quality was high (AvePP 0.921; 74.6% with a posterior probability above 0.90; AvePP 0.929 in the test set), and internal validation showed no overfitting (training vs test log-likelihood -23.46 vs -23.51; consensus clustering stability 0.699).

CONCLUSIONS

Latent class and latent profile analysis of integrated donor-recipient-perioperative profiles identifies reproducible and clinically interpretable phenotypes that stratify early infection risk after liver transplantation. The PhenoCluster framework, validated here, provides a standardized, reproducible approach that is domain-agnostic and applicable to heterogeneous clinical cohorts. Open-source code: <https://github.com/EttoreRocchi/phenocluster>.

Pretesting Count Time Series with Application to Epidemiology: A Fast-Double Bootstrap Approach

Riccardo Ievoli⁽¹⁾, Lucio Palazzo⁽²⁾, Laura Grisotto⁽¹⁾

 Riccardo Ievoli 0000-0001-9489-3564

(1) Department of Environmental and Prevention Sciences University of Ferrara, Ferrara, Italy

(2) Department of Human and Social Sciences University of Naples L'Orientale, Naples, Italy

CORRESPONDING AUTHOR: Riccardo Ievoli, Department of Environmental and Prevention Sciences University of Ferrara, Ferrara, Italy.
E-Mail: riccardo.ievoli@unife.it

INTRODUCTION

Integer Autoregressive Models (INAR) have demonstrated their versatility in applications involving epidemiological count data. Among the possible model specifications, the equidispersed first-order Poisson INAR (1) represents a straightforward and parsimonious solution for practitioners due to its simplicity of use and availability of bootstrap-based inference. However, two preliminary checks are relevant: testing for the presence of an INAR effect and assessing whether the arrival process can reasonably be modeled as equidispersed Poisson, which simplifies both estimation and forecasting.

OBJECTIVES

This work proposes a novel bootstrap procedure for testing the presence of an INAR(1) effect in count time series. The procedure adapts the fast-double bootstrap algorithm proposed in Davidson, et al by introducing an additional loop within each bootstrap replication. In addition, it provides diagnostic information on the equidispersion of the arrival process by extending the approach of Palazzo et al.

METHODS

To test the presence of the INAR effect, both the semi-parametric and the parametric bootstrap-based score statistics can be redesigned for the fast-double bootstrap algorithm with and without specifying a given distribution for the disturbances. Then, by exploiting the different behavior of the parametric and semi-parametric bootstrap distribution in the specific case of the Poisson score statistic under model misspecification, a procedure is implemented to diagnose the non-equidispersion of the disturbances. To show the applica-

bility of the procedure, a small-scale Monte Carlo simulation is implemented by employing four data-generating processes for the parametric arrivals to capture under-/over-dispersion and zero-inflation. Then, freely available surveillance low-count data regarding infectious diseases have been considered for an empirical assessment of the procedure. One of the data sources regards the weekly number of syphilis cases (dataset: syph) in the United States (51 states, T=209).

RESULTS

Simulation analysis confirms that the fast-double bootstrap algorithm can provide non-negligible improvements in terms of empirical size (under the null hypothesis of no INAR effect) and power (presence of INAR effect) due to the pivotality of both the bootstrapped and the asymptotic score statistics in many cases, especially when the time series are moderately small (T=50, 75). Moreover, the procedure fails when the parametric bootstrapped Poisson score statistic is used under model misspecification (presence of under-/over-dispersion or zero-inflation) due to the non-pivotality issue. In this sense, a further comparison between bootstrapped parametric and semi-parametric score statistics can represent a diagnostic tool to empirically verify equi-dispersion in the data. In terms of empirical applications, the syph data do not always show evidence of an INAR effect, whereas the diagnostic tool suggests avoiding the use of Poisson INAR for the majority of the series investigated. After the pretests, a summary of the applications can be provided through ex-post bootstrap-based confidence intervals.

CONCLUSIONS

The proposed fast-double bootstrap procedure represents a computationally feasible method to pretest count data for

the presence of a short-term mechanism. The diagnostic tool should be formalized in further research and also compared with the existing proposals in terms of performance achieved under different simulation scenarios.

The Joint Impact of Temperature and Fine Particles on Mortality in the Metropolitan Area of Florence (Italy): A Causal Inference Analysis Based on G-Computation

Chiara Marzi⁽¹⁾, Giovenale Moirano⁽²⁾, Daniela Nuvolone⁽³⁾, Rodolfo Saracci⁽⁴⁾, Michela Baccini⁽¹⁾

 Chiara Marzi 0000-0002-1791-3573

(1) Department of Statistics, Computer Science and Applications "G. Parenti," University of Florence, Florence (FI), Italy

(2) Cancer Epidemiology Unit, Department of Medical Sciences, University of Turin and CPO Piedmont, Turin, Italy

(3) Unit of Epidemiology, Regional Health Agency of Tuscany, Florence (FI), Italy

(4) European Educational Programme in Epidemiology, Florence, Italy

CORRESPONDING AUTHOR: Chiara Marzi, Department of Statistics, Computer Science and Applications "G. Parenti," University of Florence, Florence (FI), Italy. E-Mail: chiara.marzi@unifi.it

INTRODUCTION

Air pollution and extreme temperatures are among the leading environmental determinants of premature mortality. Although the short-term associations between particulate matter with an aerodynamic diameter smaller than 10 μm (PM₁₀), temperature, and mortality have been extensively documented, the nature of their joint effects remains largely unexplored.

OBJECTIVES

Conceptualizing both temperature and air pollution as modifiable treatments within a potential outcome (PO) approach to causal inference, we aimed to: (1) estimate the joint causal effect of particulate matter $\leq 10 \mu\text{g}/\text{m}^3$ (PM₁₀) and temperature on natural mortality in the metropolitan area of Florence (2008–2019) within the potential outcomes framework, and (2) quantify the attributable fraction among the exposed under a joint counterfactual reduction scenario.

METHODS

DATA

Daily time series data on air pollution and mortality were collected from 2008 to 2019 for an area including eight municipalities in Tuscany: Bagno a Ripoli, Calenzano, Campi Bisenzio, Florence, Lastra a Signa, Scandicci, Sesto Fiorenti-

no, and Signa (597740 inhabitants in 2019). Mortality data, sourced from mortality registers, encompassed all natural causes, excluding the external ones. Air quality data were obtained from the Tuscany Regional Environmental Protection Agency (ARPA), while meteorological data including temperature and relative humidity were provided by the Tuscany Regional Hydrological Service (SIR) and the Laboratory of Monitoring and Environmental Modelling for Sustainable Development (LaMMA Consortium).

STATISTICAL ANALYSIS

Daily time-series data on mortality, PM₁₀, and temperature were analyzed using g-computation under a bivariate treatment framework. POs were defined for joint exposure levels of PM₁₀ and temperature under assumptions of no unmeasured confounding, consistency, and positivity formulated for the two-dimensional exposure. A bi-dimensional average dose-response function (aDRF) was estimated to characterize how the joint variation in PM₁₀ and temperature influences mortality. Attributable deaths (AD) and attributable fractions among the exposed (AF) were computed by comparing observed exposure levels with a counterfactual scenario defined by PM₁₀ = 15 $\mu\text{g}/\text{m}^3$ (WHO 2021 annual guideline) and temperature = 25°C.

RESULTS

The estimated dose-response surface showed evidence of interaction between PM₁₀ and temperature, with amplified

mortality risks at higher temperatures. Under the joint reduction scenario, 919 deaths were attributable to the combined effect of elevated pollution and temperature, corresponding to an AF of 9.45%. The AF was substantially larger under concurrent high-temperature and high-PM₁₀ conditions compared with the reduction of one exposure only (AF = 3.66% for PM₁₀ alone and 5.53% for temperature alone), supporting the presence of synergistic effects.

CONCLUSIONS

This study provides causal evidence of a synergistic short-term effect of PM₁₀ and temperature on mortality. The estimated attributable fractions highlight the public health relevance of joint exposure reduction strategies, particularly in the context of climate change and increasing heat events.

Use of Large Language Models to Assist Risk of Bias Assessment Supported by an Implementation Document

Cinzia Del Giovane⁽¹⁾, Manuel Marques da Cruz⁽²⁾, Bernardo Sousa-Pinto⁽²⁾, Paweł Jemioło⁽³⁾, Rafael José Vieira⁽²⁾, Silvia Minozzi⁽⁴⁾

 Cinzia Del Giovane 0000-0002-5028-932X

(1) Department of Medical and Surgical Sciences for Children and Adults, Unit of Statistical and Methodological support, University-Hospital of Modena and Reggio Emilia, Modena, Italy; Cochrane Italy

(2) MEDCIDS - Department of Community Medicine, Information and Health Decision Sciences, University of Porto, Porto, Portugal; CINTE-SIS@RISE - Health Research Network, University of Porto, Porto, Portugal; Public Health Unit Marão e Douro Norte, Local Health Unit Trás-os-Montes e Alto Douro, Vila Real, Portugal

(3) AGH University of Krakow; Kraków, Poland

(4) Laboratory of Methodology of Systematic Reviews and Guidelines production; Istituto di Ricerche Farmacologiche Mario Negri IRCCS, Milan, Italy

CORRESPONDING AUTHOR: Cinzia Del Giovane, Department of Medical and Surgical Sciences for Children and Adults, Unit of Statistical and Methodological support, University-Hospital of Modena and Reggio Emilia, Modena, Italy. E-Mail: cinzia.delgiovane@unimore.it

INTRODUCTION

Recent studies showed poor performance of large language models (LLM) for assessing risk of bias (RoB) with the RoB2 tool. This is in line with the low reliability that humans have in assessing RoB. However, the use of an implementation document (ID) prepared by expert reviewers – i.e., a standardised document providing clear instructions on how to answer each signalling question – can increase inter-rater agreement, and may improve the performance of LLM in RoB-related tasks.

OBJECTIVES

To determine the best framework for using LLMs to support human RoB2 assessments, and to evaluate the value of an ID for LLM-assisted evaluation.

METHODS

We will include three outcomes from randomised controlled trials (RCTs) across three different systematic reviews in which RoB2 was manually assessed with the guidance of an ID. The final manual assessments will serve as the reference standard. We will prompt LLMs to assess RoB by providing them with (i) the full RoB2 guidance and (ii) the text and supplementary files of the included RCTs. The model output will

consist of quotes from the RCT or generated summaries for each signalling question, along with suggested judgements for each applicable signalling question, each domain and overall. Two human reviewers will assess the quality and usefulness of the LLM outputs in answering signalling questions. We will compare the performance of LLMs with and without additional ID guidance, focusing on information extraction quality, accuracy, and reliability of suggested answers and overall judgements. This study will be conducted using locally run LLMs to ensure data privacy and model stability.

RESULTS

This is an ongoing study. We anticipate developing a framework for LLM-supported RoB assessments through the RoB2 tool. We will also report on the perceived quality and usefulness of using LLMs to support these assessments. We expect the agreement between LLM-based outputs and human reference standard to be non-inferior to assessments performed by two humans using an ID.

CONCLUSIONS

This study will inform approaches to use LLMs to support RoB assessment. Future studies should focus on a prospective evaluation of this proposed framework to measure efficiency gains.

Patient-Reported Outcomes in Irritable Bowel Syndrome: The Role of Alexithymia in Health-Related Quality of Life

Angela Alibrandi⁽¹⁾, Liam Alexander MacKenzie Myles⁽²⁾ and Emanuele Maria Merlo⁽³⁾

 Angela Alibrandi 0000-0001-6695-151X

(1) Department of Human Pathology in Adulthood and Childhood Gaetano Barresi, University of Messina, 98124 Messina, Italy

(2) Berkshire Healthcare NHS Foundation Trust, Bracknell RG12 2UT, UK

(3) Department of Biomedical and Dental Sciences and Morphofunctional Imaging, University of Messina, 98124 Messina, Italy

CORRESPONDING AUTHOR: Angela Alibrandi, Department of Human Pathology in Adulthood and Childhood Gaetano Barresi, University of Messina, 98124 Messina, Italy. E-Mail: aalibrandi@unime.it

INTRODUCTION

Irritable bowel syndrome (IBS) is a disorder of gut–brain interaction, in which psychological factors, including alexithymia and psychosomatic syndromes, are thought to influence patients' perceived health and quality of life. Patient-Reported Outcome Measures (PROMs) provide valuable insights into these patient-centered outcomes.

OBJECTIVES

This study aimed to investigate the association between alexithymia, psychosomatic syndromes and patient-reported health status in individuals with IBS.

METHODS

A total of 143 patients diagnosed with IBS (81.8% female; age range 18–78 years, $M = 30.90$, $SD = 14.39$) were enrolled in this cross-sectional study. Participants completed a sociodemographic questionnaire, the Toronto Alexithymia Scale-20 (TAS-20), a self-report scale composed of 20 items based on a 5-point Likert scale ranging from “strongly disagree” to “Strongly agree” assessing alexithymia. The scale is composed of three factors: difficulty identifying feelings (DIF), difficulty describing feelings (DDF) and externally oriented thinking (EOT). It represents a widely recognized valid instrument for the assessment of alexithymia and related factors in a wide set of pathological and non-pathological domains, the Diagnostic Criteria for Psychosomatic Research–Structured Interview (DCPR-SI), and the SF-36 Health Survey as

a measure of patient-reported health-related quality of life. Generalized linear models were estimated with the TAS-20 (total score and single factors) as predictors and health status and psychosomatic variables as outcomes.

RESULTS

Higher levels of alexithymia were significantly associated with greater psychosomatic symptom burden and poorer patient-reported health-related quality of life, emerging as significant predictor of both psychosomatic syndrome severity and reduced scores across health status domains assessed by the SF-36. Specifically, higher levels of alexithymia were associated with lower scores on health status variables, including physical functioning, role limitations due to emotional problems, emotional well-being, social functioning, pain and general health. In terms of psychosomatic conditions, alexithymia was associated with higher levels of disease phobia, thanatophobia, Type A behaviour, irritable mood and demoralization.

Considering DIF, DDF and EOT as predictors, all significant associations between DIF and DDF and health status were negative, indicating a deterioration in health associated with alexithymia. Contrarywise, all significant associations between alexithymia factors and psychosomatic syndromes were positive, highlighting the role of alexithymia in the onset and maintenance of psychosomatic disorders. Significant negative associations involving DIF were observed with role limitations due to emotional problems, emotional well-being, social functioning, pain, and general health. Positive associations emerged with disease phobia, thanatophobia, illness denial, conversion symptoms, Type A behaviour, irritable

mood, and demoralization. A significant negative association was found between DDF and role limitations due to emotional problems, while significant negative associations were identified with conversion, Type A behaviour and clinically detected alexithymia. EOT was found associated with higher levels of role limitations due to emotional problems, energy/fatigue, and emotional well-being. Conversely, lower levels of Type A behaviour, irritable mood, and demoralization emerged.

CONCLUSIONS

Alexithymia is strongly associated with worse patient-reported outcomes in individuals with IBS. These findings support the integration of PROMs into the assessment of IBS and suggest that interventions targeting alexithymia may contribute to improving patients' perceived health and quality of life.

Prediction of Suspected Choledocholithiasis: A Comparison of Diagnostic Performance of ML Algorithms against Current Guidelines

Daniela Pacella⁽¹⁾, Fabrizio D'Acapito⁽²⁾, Adriano De Simone⁽³⁾

 Daniela Pacella 0000-0003-2343-5069

(1) Department of Public Health, University of Naples Federico II, Naples, Italy

(2) General and Oncologic Surgery, Morgagni-Pierantoni Hospital, AUSL Romagna, Forlì, Italy

(3) Department of Electric Engineering and Information Technologies, University of Naples Federico II, Naples, Italy

CORRESPONDING AUTHOR: Daniela Pacella, Department of Public Health, University of Naples Federico II, Naples, Italy. E-Mail: daniela.pacella@unina.it

INTRODUCTION

Choledocholithiasis or common bile duct stones (CBDS) is a frequent cause of hospitalization. Confirmed stones are usually removed by Endoscopic Retrograde Cholangiopancreatography (ERCP), but this procedure carries significant risks (infection, perforation, hemorrhage). Current guidelines from the European and American Societies for Gastrointestinal Endoscopy (ESGE and ASGE) suggest ERCP only in patients deemed at high-likelihood of CBDS. 2019 updated ESGE guidelines define high-likelihood as having cholangitis or suspected CBDS at ultrasound (US), while ASGE also requires the presence of both total bilirubin >4mg/dL and CBD dilation on US. However, these criteria do not show conclusive performance with all studies showing very low sensitivity and high false negative rate although showing high specificity (thus minimizing the risk of unnecessary ERCPs).

OBJECTIVES

The objective of the present study was twofold: 1) to evaluate the diagnostic performance of ASGE and ESGE 2019 criteria in confirming suspected CBDS on a novel single-center retrospective cohort; 2) to train different machine learning (ML) models and evaluate their performances in comparison with traditional guidelines.

METHODS

Data extracted from the Morgagni-Pierantoni Hospital database, included n=870 patients admitted between 2017

and 2022, with 13% later confirmed CBDS. The study included patients who underwent US and blood tests and had a minimum follow-up of 12 months. Variables used for ML training were those that define ESGE/ASGE likelihood plus age, sex, white blood cells (WBC) and C-reactive protein (CRP) and were distributed at baseline as follows: mean age 66 ± 17 , female sex 52%, cholangitis 12%, abnormal liver function tests 34%, CBDS on US 4.6%, CBD dilation on US 17%, mean total bilirubin 1.30 ± 1.73 mg/dL, mean WBC 10480 ± 6276 , mean CRP 76 ± 105 . 14% (ESGE) and 15% (ASGE) scored as high risk. Outcome definition was confirmed CBDS at intraoperative RX or ERCP. ML algorithms compared were LR, SVM, KNN, LightGBM, Random Forest, GradientBoosting, XGBoost and MLP. Training, validation and test split was 70%/15%/15%; 10-fold cross validation was used. AUC and Youden Index derived sensitivity, specificity, PPV, NPV and accuracy with 95% confidence intervals (CI) were used for evaluation.

RESULTS

The only model which achieved better results at all diagnostic measures was a deep feed-forward MLP architecture (layers size 32,16) trained on normalized predictor values (batch size=32, AdamW optimizer, cross-entropy loss). It showed good calibration (Brier score=0.14) and reached a test AUC of 0.85 [0.79-0.91]. At the best probability threshold (0.22) it reached better test accuracy (0.85 [0.80-0.88]) than ESGE (0.82 [0.79-0.84]) and ASGE (0.82 [0.80-0.85]). It also showed significantly higher test sensitivity (0.55 [0.39-0.71]) than ESGE (0.35 [0.27-0.44]) and ASGE (0.4

[0.32-0.49]). Similar specificity was observed (MLP 0.89 [0.87-0.92]), ESGE 0.89 [0.87-0.91], ASGE 0.89 [0.86-0.91]. Test PPV and NPV were also significantly better (MLP 0.42 [0.33-0.52] and 0.93 [0.91-0.95]; ESGE 0.32 [0.22-0.44] and 0.90 [0.87-0.91]; ASGE 0.34 [0.24-0.45] and 0.91 [0.89-0.91]).

CONCLUSIONS

We compared 8 ML models to ASGE and ESGE high-likelihood criteria for suspected CBDS and showed that MLP reached significantly better performance in terms of sensitivity, NPV and PPV and overall greater accuracy while keeping similar specificity.

Evaluating Global and Domain-Specific PCA in Exposome Research: A Diagnostic Framework

Vittoria Ercoli⁽¹⁾, Rachele De Giuseppe⁽²⁾, Gloria Bertoli⁽³⁾, Roberto Bergamaschi⁽⁴⁾, Eleonora Tavazzi⁽⁴⁾, Bruno Giovanni Galuzzi⁽³⁾, David Chierigato⁽⁵⁾, Marco Cremaschi⁽⁵⁾, Camilla Torriani⁽¹⁾, Cristina Montomoli⁽¹⁾, Maria Cristina Monti⁽¹⁾

 Vittoria Ercoli 0000-0002-0899-4885

(1) Unit of Biostatistics and Clinical Epidemiology, Department of Public Health, Experimental and Forensic Medicine, University of Pavia, Pavia, Italy

(2) Laboratory of Dietetics and Clinical Nutrition, Department of Public Health, Experimental and Forensic Medicine, University of Pavia, Pavia, Italy

(3) Istituto di Bioimmagini e Sistemi Biologici Complessi (IBSBC), National Research Council (CNR), Milan, Italy

(4) IRCCS Mondino Foundation, Pavia, Italy

(5) Department of Informatics, Systems and Communication DISCo, University of Milano Bicocca, Milan, Italy

CORRESPONDING AUTHOR: Vittoria Ercoli, Unit of Biostatistics and Clinical Epidemiology, Department of Public Health, Experimental and Forensic Medicine, University of Pavia, Pavia, Italy. E-Mail: vittoria.ercoli@unipv.it

INTRODUCTION

The exposome describes the totality of environmental, behavioural, occupational, social and biological exposures experienced across the life course. This framework is particularly relevant for complex chronic diseases, where multiple exposures may interact dynamically with biological susceptibility. Exposome studies generate heterogeneous high-dimensional datasets, including pollutants, lifestyle, diet, socioeconomic factors, clinical variables and molecular biomarkers. In this context, dimension-reduction techniques, particularly principal component analysis (PCA), are frequently used to summarize correlated exposures, but it is often applied with incomplete diagnostics and limited justification for accepting low explained variance. Thus, a dataset may be statistically suitable for PCA while still being poorly represented by a small, stable and interpretable set of components.

OBJECTIVES

We aimed to propose a diagnostic framework to distinguish applicability from dimension-reduction techniques usefulness in exposome research, with particular attention to the role of domain-specific PCA as a domain-informed unsupervised approach.

METHODS

The framework was applied to two datasets. The first was EXPOSITION, an in-house cross-sectional cohort of 150 Northern Italian patients with multiple sclerosis, including 68 variables across demographic, clinical, sociodemographic, dietary, lifestyle, NO₂ exposure and miRNA domains collected and managed through BIOMATRIX, a privacy-preserving digital platform for biodiversity management and advanced analysis. The second was iMSMS, an international gut microbiome dataset including 1,152 multiple sclerosis patients and paired household controls, with variables across demographic, clinical, sociodemographic, dietary, microbiome, metagenomic and metabolomic domains. For each dataset, we applied a structured pre-PCA diagnostic assessment to evaluate the statistical admissibility of component extraction. This included missingness evaluation, correlation matrix inspection, near-zero variance detection, outlier assessment, scaling sensitivity, Kaiser–Meyer–Olkin measure of sampling adequacy, Bartlett’s test of sphericity and variable-level adequacy. PCA usefulness was then evaluated through explained variance, number of retained components, loading stability, domain purity, interpretability and dimensionality-reduction efficiency. Global PCA was compared with domain-specific PCA, robust PCA, sparse PCA, factor analysis, FAMD and MFA.

RESULTS

The application to EXPOSITION and iMSMS highlighted that conventional pre-PCA diagnostics are necessary but not sufficient. Even when factorability diagnostics supported PCA application, global PCA could still require several components to explain an adequate proportion of variance, limiting its value as an efficient dimensionality-reduction strategy. This issue was particularly relevant when variables belonged to distinct conceptual domains or represented complex biological profiles. Domain-specific PCA, while remaining unsupervised because it does not use outcome information, incorporated a priori knowledge on exposure domains and provided more interpretable and epidemiologically coherent summaries than global PCA in domain-structured data.

CONCLUSIONS

PCA should not be applied as a default dimensionality-reduction tool in exposome research. Researchers should report both pre-PCA diagnostics and post-PCA usefulness criteria, especially explained variance, component stability and interpretability. Domain-specific PCA may represent a practical compromise between purely data-driven reduction and meaningful exposure summarization in heterogeneous multi-domain datasets.

Explainable Clustering as a Tool for Lung Neuroendocrine Tumors Classification

Valentina Veronesi⁽¹⁾, Giulia Orlando⁽²⁾, Eleonora Duregon⁽²⁾, Vanessa Zambelli⁽³⁾, Francesco Leo⁽⁴⁾, Elisa Carla Fontana⁽⁵⁾, Enrico Ruffini⁽⁵⁾, Luisella Righi⁽³⁾, Giuseppe Pelosi^(6,7), Marco Volante^(3,8), Marta Rigoni⁽⁹⁾, Mauro Papotti⁽²⁾

(1) IRCSS MultiMedica, Milan, Italy

(2) Division of Pathology, City of Health and Sciences, University Hospital, Department of Oncology, University of Turin, Turin, Italy

(3) Division of Pathology, Department of Oncology, University Hospital, University of Turin, Orbassano, Italy

(4) Division of Thoracic Surgery, San Luigi University Hospital, University of Turin, Orbassano, Italy

(5) Division of Thoracic Surgery, City of Health and Sciences, University Hospital, University of Turin, Turin, Italy

(6) Department of Oncology and Hemato-Oncology, University of Milan, Milan, Italy

(7) Division of Pathology, San Luigi University Hospital, Orbassano, Italy

(8) Inter-Hospital Division of Pathology, IRCSS MultiMedica, Milan, Italy

(9) Department of Biomedical, Surgical, and Dental Sciences, University of Milan, Milan, Italy

CORRESPONDING AUTHOR: Valentina Veronesi, IRCSS MultiMedica, Milan, Italy. E-Mail: valentina.veronesi@multimedica.it

INTRODUCTION

Grading and prognostic stratification are central tools in clinical practice: they synthesize complex biological and pathological information into categories, which can be used for risk assessment and decision making. In lung neuroendocrine tumors (NETs), the WHO classification traditionally distinguishes typical (TC) and atypical (AC) carcinoids on the basis of mitotic count (MC) and necrosis. This framework is about to evolve towards the integration of Ki-67, a proliferation index with recognized biological and prognostic relevance. However, Ki-67 role has been highly debated in literature, partly due to the lack of universally accepted cut-offs and inter-observer reproducibility issues. When tumor subgroups are hypothesized, clustering represents a useful tool for identifying homogeneous groups of observations. However, translating groups obtained via unsupervised methods into transparent rules that can be applied to the clinical decision-making process is not straightforward.

OBJECTIVES

The aim was threefold: (i) to explore an alternative data-driven classification of lung NETs, integrating Ki-67 to MC and necrosis, for identifying clinically meaningful tumor subgroups with added value in grading and prognosis determination with respect to the TC/AC dichotomy; (ii) to translate them into a set of interpretable, reproducible, and applicable decision rules; (iii) to quantify uncertainty associated to cut-

offs, providing information about variability and support the evaluation of cases that are close to the decision thresholds.

METHODS

We applied an explainable clustering strategy, transforming the data-driven groups into explicit classification criteria, to a retrospective cohort of 483 lung NETs. Observations were clustered using KAMILA, applied to necrosis, MC, and Ki-67. Partitions from 2 to 7 clusters were tested; the three-cluster solution was selected based on clinical interpretability and internal validation indices, while the two-cluster solution was compared against the TC/AC benchmark. Then, we fitted conditional inference trees to predict cluster labels from clustering variables, deriving a decision tree. Cut-off stability was assessed through repeated cross-validation, estimating an empirical confidence interval (ECI) for each selected threshold. Survival analyses were used to evaluate the prognostic value of the three-cluster solution for overall survival and relapse-free survival; transferability was explored in an external cohort.

RESULTS

Clustering identified three groups, named NET G1 to G3, characterized by progressively higher Ki-67, MC, and frequency of necrosis. The groups only partially overlapped with the TC/AC classification, and provided a better stratification of ACs. Ki-67 resulted as the main discriminating variable,

splitting at $\leq 6\%$ (95% ECI: 4-7%) and $> 16\%$ (95% ECI: 16-18%), for NET G1 and NET G3, respectively. MC contributed to classification in the intermediate group, while necrosis was not selected in the final system. The resulting groups showed prognostic stratification, particularly for relapse, and the derived decision tree showed transferability to the external cohort.

CONCLUSIONS

We developed an interpretable, practical, three-tier classification system, which redefines the grading of lung NETs. Results supported Ki-67 integration in lung NETs stratification, confirmed MC as a discriminating factor, and suggested that necrosis might carry lower discriminative information. More broadly, this work supports the wider use of explainable clustering in clinically oriented classification studies, especially when data-driven groups need to be translated into rules that can be inspected, communicated, and applied in practice. Incorporating ECI into the inferential decision tree enables quantification of measurement uncertainty, accounting for inter-observer variability.

AI-Assisted and Human-Driven Biostatistical Workflows: A Task-Based Comparison Using Finnish Health Registry Data

Camilla Torriani⁽¹⁾, Juha Karjalainen⁽²⁾, Maria Cristina Monti⁽¹⁾, Tuomo Hartonen⁽²⁾, Andrea Ganna⁽²⁾

 Camilla Torriani 0009-0006-5879-6795

(1) Department of Public Health, Experimental and Forensic Medicine, Unit of Biostatistics and Clinical Epidemiology, University of Pavia, Pavia, Italy

(2) Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland

CORRESPONDING AUTHOR: ICamilla Torriani, Department of Public Health, Experimental and Forensic Medicine, Unit of Biostatistics and Clinical Epidemiology, University of Pavia, Pavia, Italy. E-Mail: camilla.torriani01@universitadipavia.it

INTRODUCTION

Recent advances in generative and agentic artificial intelligence are rapidly changing how biomedical research workflows are designed and conducted. AI-based assistants increasingly support coding, workflow automation, data analysis and scientific reporting. These tools could reduce time for routine analytical tasks and help researchers generate reproducible statistical pipelines. However, in biomedical and population-health research, analytical outputs must be not only fast and executable, but also methodologically appropriate, reproducible, interpretable and safe. A key open question is whether agentic AI tools can support biostatisticians without compromising scientific validity. We will address this question within ML4Health at FIMM, University of Helsinki, which combines health data with other information from nearly the whole Finnish population, including data from approximately 7 million individuals. The research data include diagnoses, disease treatments, medications, laboratory values, medical notes, home and institutional care, personal information, residential and marital history, pregnancies and births, education, job position, social assistance, and dates and causes of death. In this setting, Goose will be used as the agentic workflow interface, while Gemini will represent the generative engine driving code generation and reasoning support.

OBJECTIVE

We aim to develop and apply a task-based comparative framework to quantify the methodological validity, reproducibility and efficiency of an AI-assisted biostatistical workflow using Finnish health registry data.

METHODS

We will compare the AI-assisted workflow with a human-driven biostatistical workflow. We will define approximately 20 biostatistical tasks relevant to ML4Health and classify them into domains of increasing complexity. For each task, we will prepare an a priori methodological assessment sheet specifying input information, analytical objective, expected output and key sub-components required for a valid response. An expert-validated reference solution will also be prepared. Both workflows will receive the same input data or metadata, task description and analytical objective. For the AI-assisted workflow, we will use a standardized prompt chain and document prompts, outputs, corrections and decisions. Each output will be compared with the reference solution using a predefined scoring grid. Each sub-component will be scored as 2 = correct and complete or equivalent to the reference solution, 1 = partially correct or requiring minor correction, and 0 = incorrect, incomplete or requiring major correction. For each task, we will calculate total and percentage performance scores. Outputs will be classified according to predefined validity thresholds, considering both percentage score and major validity-threatening errors. Errors will be categorized by type and severity, including data-handling, coding, recoding, modelling, visualization, interpretation, reporting and reproducibility errors. The time required to generate, review and correct each output will also be recorded. Results will be summarized overall and by analytical domain.

RESULTS

We expect to identify domains in which AI-assisted workflows can reduce time and resource use while maintaining acceptable analytical quality, and those requiring substantial expert supervision. Expected outputs include task-level quality scores, error profiles and time-effectiveness summaries.

CONCLUSIONS

This evaluation will clarify the potential role of agentic AI as a supervised support tool for biostatisticians and may inform practical recommendations for transparent, reproducible and efficient AI-assisted analyses in large-scale health data research.

Data-Driven Characterization of Heart Rate Dynamics during Sleep Using Functional Data Analysis

Samuele Minari⁽¹⁾, Serajeddin Ebrahimian^(2,3), Antonella Zamboni^(1,4), Carolina Lombardi^(5,6), Martino Pengo^(5,6)

 Samuele Minari 0009-0009-4102-8900

(1) Laboratory of Quantitative methods for Life, Health and Society (QmLHS-Lab), Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy

(2) Department of Technical Physics, University of Eastern Finland, Kuopio, Finland

(3) Clinical research Center, Kuopio University Hospital, Kuopio, Finland

(4) Biostatistics Unit, IRCCS Istituto Auxologico Italiano, Milano, Italy

(5) Istituto Auxologico Italiano, IRCCS, Department of Cardiology, Ospedale San Luca, Milano, Italy

(6) Department of Medicine and Surgery, University of Milano-Bicocca, Milan, Italy

CORRESPONDING AUTHOR: Samuele Minari, Laboratory of Quantitative methods for Life, Health and Society (QmLHS-Lab), Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy. E-Mail: s.minari@campus.unimib.it

INTRODUCTION

Obstructive Sleep Apnea (OSA) is a highly prevalent sleep disorder characterized by recurrent upper airway obstruction during sleep, leading to intermittent hypoxia and autonomic nervous system dysregulation. Heart Rate Variability (HRV), derived from RR interval series in ECG recordings, has been widely studied as a non-invasive marker of autonomic dysfunction in OSA. However, conventional HRV analyses rely on summary statistics that may fail to capture the full temporal structure of the signal.

OBJECTIVES

This pilot study explores the application of Functional Principal Component Analysis (FPCA) to RR interval series extracted from polysomnographic ECG recordings, with the aim of identifying functional features associated with OSA severity and respiratory event characterization.

METHODS

We retrospectively analyzed ECGs from full diagnostic polysomnographies from patients evaluated at the IRCCS Istituto Auxologico Italiano. Raw overnight ECGs were denoised and then segmented into non-overlapping event-locked epochs centered on individual respiratory events. Additionally,

non-event ECG segments were extracted in order to utilize them as a control group. For each segment, RR interval series were extracted using a validated delineation algorithm for R peaks detection. Each RR series was then modeled as a continuous function over time using Functional Data Analysis (FDA) techniques and Functional Principal Component Analysis (FPCA) was applied in each setting to extract the leading modes of variation in RR series. The resulting functional principal components (FPCs) were examined in relation to AHI values and respiratory event typology.

RESULTS

The event-locked analysis ($n = 281$) was conducted under two perspectives: discrimination among respiratory event typologies and comparison between event and non-event segments. Scatter plots of the four leading FPCs (which accounted for 48.8%, 12.5%, 9.5% and 7.3% of the total variance, respectively) revealed no clearly discriminative functional structure between event typologies. This is arguably interpretable, as both event types may elicit highly similar cardiac responses, producing largely overlapping functional signatures. Notably, the first functional principal component predominantly described fluctuations around the mean RR value, effectively capturing the overall sympathovagal balance, rather than event-specific modulation patterns. A similarly inconclusive picture emerged from the event vs non-event comparison, which is in this case more striking. Despite

the autonomic perturbation associated with respiratory events being expected to produce a detectable functional signature relative to baseline, the FPC scatter plots showed substantial overlap between the two groups. High intra- and inter-individual variability in RR dynamics appears to dominate the functional representation, masking any systematic difference attributable to the presence of a respiratory event.

CONCLUSION

These preliminary findings suggest that naive application of FPCA to raw RR series may be insufficient to extract OSA-related functional structure, due to the dominance of physiological variability unrelated to respiratory events. Rather than representing a negative outcome, these results provide methodological insight into the limitations of direct functional decomposition and motivate the exploration of individual-centered approaches to better isolate the autonomic signature of OSA from background variability. Further analyses are ongoing.

A Novel Framework for Conducting Meta-Analysis of Patient Preference Studies

Matteo Bracco⁽¹⁾, Emanuele Pietropaolo⁽¹⁾, Martina di Blasio⁽¹⁾, Ileana Baldi⁽²⁾, Alessia Visconti⁽¹⁾, Paola Berchiolla⁽¹⁾

 Matteo Braccio 0009-0001-2135-3070

(1) Center for Biostatistics, Epidemiology, and Public Health, Department of Clinical and Biological Science, University of Turin, Italy

(2) Department of Cardiac Thoracic Vascular Sciences and Public Health, University of Padua, Padua, Italy

CORRESPONDING AUTHOR: Matteo Bracco, Center for Biostatistics, Epidemiology, and Public Health, Department of Clinical and Biological Science, University of Turin, Italy. E-Mail: matteo.bracco@unito.it

INTRODUCTION

Information about patients' priorities and treatment preferences is essential to guide decision-making processes along medical product life cycles. Discrete choice experiments (DCEs) are one of the most common methods to elicit patient preferences. Through the use of a robust methodology rooted on random utility theory, DCEs estimate patients utility functions defined by a set of coefficients called Preference Weights (PWs). PWs are used to reconstruct complex preference patterns such as benefit-risk trade-offs and the relative importance of treatment attributes. DCEs are, however, expensive and time-consuming, highlighting the need for evidence synthesis via meta-analysis.

OBJECTIVES

Currently, meta-analytic approaches are limited to trade-offs parameters, as PWs estimated from different DCEs are not directly comparable due to methodological and design heterogeneity.

However, these require the identification of at least two trading attributes across pooled studies, hence limiting the possible therapeutic areas of application as well as the DCEs pool's size. Here we construct the first systematic framework to synthesize DCEs quantitative information at the higher level of PWs, addressing how to re-scale PWs from different DCEs to allow comparisons and aggregation.

METHODS

We develop a novel methodology stemming from network theory to rescale PWs across pooled DCEs. By anchoring

PWs through common attributes, we are able to define DCEs networks. Sampling spanning trees through an Markov-Chain MonteCarlo (MCMC) procedure, we can set common scales across PWs from different DCEs, hence producing a sample of re-scaled PWs.

RESULTS

From the sample or re-scaled PWs sets, we are able to produce flexible meta-analytic outputs as distribution of preference parameters. Importance rankings allow to create a comprehensive picture of patients priorities across pooled DCEs. Besides standard trade-offs parameters, our approach also allows constructing non-linear trade-off curves between two or multiple attributes, to better capture real patients' trade-off patterns.

CONCLUSIONS

Our framework represents a key step towards more robust and resource-efficient modelling of patients' preferences. The production of multiple aggregated preference parameters outputs defines a comprehensive picture of patients preferences and priorities across different therapeutic areas, to guide development of new pharmaceutical products and regulatory decision making processes to account for patients real needs.

Evaluating Therapeutic Outcomes in Childhood Interstitial Lung Disease through Privacy Preserving Bayesian Federated Inference

Pietro Fusco⁽¹⁾, Mohd Rashid Khan⁽²⁾, Domenico Iervolino⁽²⁾, Pasquale Dolce⁽¹⁾, Danila Azzolina^(1,2)

 Pietro Fusco 0009-0001-1560-6377

(1) Biostatistics and Clinical Trial Biometry, Clinical Research Center DEMeTra, Department of Translational Medicine, University of Naples Federico II, 80138, Naples, Italy

(2) Unit of Biostatistics, Epidemiology and Public Health, Department of Cardiac, Thoracic and Vascular Sciences, University of Padova, 35131, Padova, Italy

CORRESPONDING AUTHOR: Pietro Fusco, Biostatistics and Clinical Trial Biometry, Clinical Research Center DEMeTra, Department of Translational Medicine, University of Naples Federico II, 80138, Naples, Italy. E-Mail: pietro.fusco@unina.it

INTRODUCTION

Childhood interstitial lung disease (chILD) encompasses rare, diffuse lung disorders with high morbidity and mortality. No pharmacological therapy has received regulatory approval, leaving management dependent on off-label treatments. The extreme rarity of chILD prevents individual centers from gathering sufficient data for well-powered studies, while data governance regulations restrict traditional multicenter data pooling.

OBJECTIVES

This study aims to establish a privacy-preserving Bayesian Federated Inference (BFI) framework for treatment stratification in chILD across a multinational network. Specifically, it seeks to: (1) compare the effectiveness of off-label therapies on clinical trajectories without moving individual patient data; (2) implement local Explainable AI (XAI) modules to provide interpretable treatment-response stratification by chILD subtype; and (3) evaluate a fixed-noise differential privacy mechanism to protect shared posterior estimates while maintaining clinical utility.

METHODS

A retrospective multicenter cohort study will enroll patients aged 0–18 with a chILD diagnosis and a minimum of 12 months of follow-up. The BFI framework enables privacy-preserving analysis, where centers compute local Bayesian hi-

erarchical models and share only summary-level posterior parameter estimates. Federated aggregation combines these estimates into a global inference. Institutional heterogeneity and non-IID data distributions will be managed via BFI extensions. A fixed Local Differential Privacy mechanism will add calibrated noise to shared posteriors prior to aggregation. Local XAI post-hoc modules (SHAP) will generate interpretable explanations of treatment-response predictions using the global model, with only aggregated feature-importance patterns shared across the network.

RESULTS

A pilot proof-of-concept across European centers is expected to demonstrate the feasibility of collaborative therapeutic evaluation without centralizing sensitive data. The privacy-protected BFI models are projected to yield interpretable stratifications of treatment response, with XAI enhancing clinical trust and actionable insight.

CONCLUSION

Ultimately, this framework establishes a scalable paradigm for collaborative medical AI in rare diseases, overcoming data fragmentation to directly inform clinical practice while rigorously preserving patient privacy at every contributing institution.

Comparing Survival without vs after Heart Transplant: Non-Parametric Methods for Estimation and Testing of a Non-Markovian Process

Davide Paolo Bernasconi^(1,2), Lorenzo Del Castello⁽¹⁾, Paola Rebora^(1,3), Laura Antolini^(1,4)

 Davide Bernasconi 0000-0001-8771-5428

(1) Bicocca Bioinformatics Biostatistics and Bioimaging Centre - B4, Dipartimento di Medicina e Chirurgia, Università Milano-Bicocca, Monza, Italy

(2) S.C. Ricerca Innovazione e Trasferimento Tecnologico - Unità di Ricerca Clinica, ASST Grande Ospedale Metropolitano Niguarda, Milano, Italy

(3) S.C. Epidemiologia Clinica e Biostatistica, Fondazione IRCCS San Gerardo dei Tintori, Monza, Italy

(4) IRCCS Fondazione Don Carlo Gnocchi, Milano, Italy

CORRESPONDING AUTHOR: Davide Paolo Bernasconi, Bicocca Bioinformatics Biostatistics and Bioimaging Centre - B4, Dipartimento di Medicina e Chirurgia, Università Milano-Bicocca, Monza, Italy. E-Mail: davide.bernasconi@unimib.it

INTRODUCTION

The “illness-death model” describes a simple three-states process where subjects can move from the initial state to the final state (death) possibly transiting to an intermediate state (illness). This framework applies also when the intermediate state consists in a therapeutic intervention administered after waiting some time since the initial event, e.g. cardiopathic patients waiting for heart transplant.

Within such context, it could be of interest to compare the hazards of mortality without vs after treatment

administration. The problem could be tackled using multi-state hazard based models, often relying on the Markov assumption, considering adjustment for waiting time to treatment start. However, from a clinical perspective, it is often more interesting to answer counterfactual questions such as: what is the survival of a patient alive at a certain time after entry on list and that will never be transplanted compared to a patient alive and transplanted at that time? These estimands are different from state-occupation probabilities based on the joint distribution between survival time and treatment administration provided by multistate models.

OBJECTIVES

The aim of the work is to adapt and apply non-parametric methods to estimate and compare the survival of patients on the waiting list vs after being transplanted in a setting where

the underlying process is non-markovian. We propose an estimator of survival curves and a log-rank-type test.

METHODS

We show the importance of checking the validity of the Markov assumption in order to select the most appropriate time scale (clock-forward or clock-reset) guiding the mortality after transplant and to assess the role of the waiting time. We propose a non-parametric method to estimate survival in the presence of a time-dependent intervention, accounting for the impact of waiting time and a modification of the Mantel-Byar test suitable for semi-Markov and extended semi-Markov scenarios. All these procedures are based on the adoption of the “clock-reset” scale that allows the hazard of death for treated patients to depend on time since treatment administration.

Using theoretical arguments and simulations we show the validity of the proposed methods.

RESULTS

We considered 9 simulation scenarios combining three type of effects for treatment and waiting time (no effect, constant effect, time-varying effect). In all scenarios we obtained negligible bias of survival estimate and good coverage of the confidence interval. Regarding the test, in all scenarios under the null hypothesis (no effect of treatment) the fraction

of false rejections was close to the nominal level, while for scenarios under the alternative hypothesis (protective effect of treatment) the power of our test was higher than that of the Mantel-Byar test.

Finally, we analyzed data from a real cohort of nearly 1000 patients affected by severe cardiomyopathy included in a waiting list for heart transplant but with low priority where mortality rate after transplant tends to increase along with a longer waiting time to transplant.

CONCLUSIONS

We show an original approach for survival estimation and testing in the presence of a time-dependent treatment accounting for the waiting time until treatment switch in a non-markovian process.

Unveiling the Temporal Pathways Linking Psychological Distress, Disease Symptomatology, and Quality of Life in Inflammatory Bowel Disease

Martina Carnevale^(1,2), Clarissa Ferrari⁽³⁾, Viviana Gerardi⁽⁴⁾, Agostino Brugnera⁽⁵⁾, Paola Cesaro⁽⁴⁾, Stefano Calza⁽¹⁾

(1) Università degli Studi di Brescia, Brescia, Italy

(2) Fondazione Poliambulanza Istituto Ospedaliero, Brescia, Italy

(3) UO of Research and Clinical Trials, Fondazione Poliambulanza Istituto Ospedaliero, Brescia, Italy

(4) UO of Gastroenterology and Endoscopy, Fondazione Poliambulanza Istituto Ospedaliero, Brescia, Italy

(5) Department of Human and Social Sciences, University of Bergamo, Bergamo, Italy

CORRESPONDING AUTHOR: Martina Carnevale, Università degli Studi di Brescia, Brescia, Italy. E-Mail: martina.carnevale@unibs.it

INTRODUCTION

Inflammatory bowel disease (IBD) is a chronic relapsing condition associated with substantial psychological burden, including elevated anxiety, depression, and perceived stress. Psychological distress is consistently linked to poorer quality of life (QoL), yet the mechanisms underlying this relationship remain incompletely understood. In particular, whether distress affects QoL directly or indirectly through disease symptomatology is still debated. Most mediation studies assume a single temporal ordering, although alternative specifications are often equally plausible. This study compares temporal mediation models to evaluate how temporal assumptions influence conclusions on this relationship in IBD.

OBJECTIVES

To investigate the mediating role of disease symptomatology in the association between psychological distress and QoL in patients with IBD, comparing alternative temporal mediation models and evaluating the robustness of findings under different missing-data imputation strategies.

METHODS

A longitudinal cohort of 102 patients with IBD was assessed at baseline (T0) and four-month follow-up (T1). Psychological distress was measured with the Depression Anxiety Stress Scale-21 (DASS-21), disease symptomatology with the patient-reported IBD-Control symptom score, and QoL with

the EQ-5D-5L. Baron and Kenny's mediation framework was applied to three alternative specifications: (1) a fully cross-sectional model, all variables at T0; (2) a fully prospective model, distress at T0 and symptomatology/QoL at T1; and (3) a partially prospective model, distress and symptomatology at T0 and QoL at T1. Mediation paths (a, b, c, c') were estimated for each specification. About 20% of follow-up observations were missing; missing data were handled via predictive mean matching and missForest multiple imputation. As both approaches produced comparable estimates, missForest was retained for the final analyses due to its lower computational burden.

RESULTS

The fully cross-sectional model supported partial mediation of IBD symptoms on the relation distress-quality of life: path a) distress is significantly correlated with IBD symptoms (mediator); path b) IBD symptom significantly correlated with QoL, while path c') a direct effect of distress on QoL remained after accounting for the mediator. The fully prospective specification also identified significant indirect pathways; however, the direct effect of distress on QoL reversed direction, suggesting an unstable mediation structure under this temporal assumption. In contrast, the partially prospective model, linking baseline distress and symptomatology with follow-up QoL, showed significant effects for all mediation paths in the expected direction, providing the most coherent evidence for partial mediation. Results were highly consistent between complete-case and imputed datasets, indicating robustness to the handling of missing follow-up data.

CONCLUSIONS

Among the temporal specifications examined, the model linking baseline distress to concurrent symptomatology and subsequent QoL at four-month follow-up provided the most coherent mediation pattern. These findings suggest that psychological distress rapidly influences patients' perceived symptom burden, which subsequently contributes to poorer QoL over time, while a direct effect of distress on QoL also persists. More broadly, this study shows that mediation results may depend substantially on the temporal specification adopted. Explicitly comparing plausible temporal models should be considered good practice when applying the Baron and Kenny framework to longitudinal observational data, improving the robustness and interpretability of mediation analyses beyond the specific clinical context of IBD.

A Bayesian Model for Patient-Reported Age at FSHD Symptoms Onset

Riccardo Cuoghi Costantini⁽¹⁾, Andi Nuredini⁽²⁾, Sara Pini⁽²⁾, Noemi Albano⁽²⁾, Roberto D'Amico⁽¹⁾, Rossella Tupler⁽²⁾

 Riccardo Cuoghi Costantini 0000-0002-9116-5198

(1) Università di Modena e Reggio Emilia, Dipartimento di Scienze Mediche e Chirurgiche Materno Infantili e dell'Adulto, Modena, Italy

(2) Università di Modena e Reggio Emilia, Dipartimento di Scienze Biomediche, Metaboliche e Neuroscienze, Modena, Italy

CORRESPONDING AUTHOR: Riccardo Cuoghi Costantini, Università di Modena e Reggio Emilia, Dipartimento di Scienze Mediche e Chirurgiche Materno Infantili e dell'Adulto, Modena, Italy. E-Mail: riccardo.cuoghicostantini@unimore.it

INTRODUCTION

Facioscapulohumeral muscular dystrophy (FSHD) is one of the most common muscular dystrophies and is usually associated with reduction of the D4Z4 repeat array on the 4q subtelomere (DRA). Because FSHD is hereditary, risk prediction is central to genetic counselling. However, prediction of disease occurrence and age at onset remains challenging mainly because of the retrospective nature of symptom onset reporting. Age at first symptom onset is often patient-reported and may be affected by recall error, rounding or heaping at preferred ages. Treating such reports as exact event times may bias risk estimation.

OBJECTIVES

We developed an accelerated failure-time survival model to estimate the probability of FSHD onset and the expected age at disease onset in newborns of DRA-carriers, using only information from parents and accounting for censoring, familial clustering and uncertainty in reported onset age.

METHODS

This study used the genetic, clinical, and sociodemographic data from the Italian National Registry for FSHD. Clinical information included symptom presence and severity, and standardized phenotype descriptions obtained through the Comprehensive Clinical Evaluation Form. Eligible families included at least one child, one DRA-carrier parent, and, when available, additional relatives.

The model developed was a Bayesian hierarchical accelerated failure-time survival model with an observation

model for patient-reported onset age. The outcome was the unobserved age at first patient-reported FSHD symptom onset and was modelled through a Weibull distribution. A 3 levels hierarchical structure was introduced to model the correlation structure among subjects sharing the same family and the same parents. Weakly informative priors were used for all parameters involved. For affected individuals, the referred age at onset was treated as an error-prone report of the latent event time: non-rounded integer ages were assigned narrow Student-t reporting error distributions, whereas ages ending with 5 or 0 were assigned progressively wider error distributions to account for potential age heaping. In addition, clinically meaningful developmental codes (such as age 6 used to denote childhood onset) were modelled as interval-censored observations. Unaffected genetically at-risk individuals were considered censored at their last assessment age.

Predictive performance measures were computed using 10-fold family-level cross-validation, holding out entire family clusters to evaluate generalization to unseen families.

RESULTS

Overall, 355 children of FSHD-affected subjects from 205 families were analysed. Across the analysed age ranges, the model showed moderate to good discrimination. Family-level cross-validation time-dependent Area under the ROC curve (AUC) exhibited values ranging from approximately 0.70 to 0.80, with mean equal to 0.75. Overall, the model achieved a mean balanced accuracy equal to 0.69, but its performances were highly dependent on parents' characteristics, resulting in better results when analysing classical FSHD phenotype parents, and poorer results when applied to complex phenotype cases.

CONCLUSIONS

We developed a Bayesian model for analysing a patient-reported time-to-event outcome in FSHD. It was designed to account for the uncertainty surrounding the age at onset, that is a clinically important information but usually collected with imprecision. This method may represent also a useful tool to support clinicians during genetic counselling.

Machine Learning-Based Classification of Parkinson's Disease Using Clinical and Biomarker Data: A Leakage-Safe Benchmarking Study with Convergent Feature Ranking

Ambra Faraci⁽¹⁾, Alessia Venuto⁽¹⁾, Rosella Ciurleo⁽¹⁾, Chiara Sorbera⁽¹⁾, Giovanni Bellomo^(2,3), Giovanna Nardi^(3,4), Lorenzo Gaetani^(3,4), Lucilla Parnetti^(3,4), Lilla Bonanno⁽¹⁾

 Ambra Faraci 0009-0001-9806-6344

(1) IRCCS Centro Neurolesi "Bonino Pulejo", 98124 Messina, Italy

(2) Department of Life Science, Health, and Health Professions, Link Campus University, 00165 Rome, Italy

(3) Laboratory of Clinical Neurochemistry, Department of Medicine and Surgery, Section of Neurology, University of Perugia, 06132 Perugia, Italy

(4) Clinica Neurologica, Azienda Ospedaliera di Perugia, 06129 Perugia, Italy

CORRESPONDING AUTHOR: Ambra Faraci, IRCCS Centro Neurolesi "Bonino Pulejo", 98124 Messina, Italy. E-Mail: ambra.faraci@irccsme.it

BACKGROUND

Parkinson's Disease (PD) is a progressive neurodegenerative disorder characterized by nigrostriatal dopaminergic neuronal degeneration and α -synuclein aggregation. Motor symptoms emerge after substantial neurodegeneration, making early diagnosis challenging. A prodromal phase, characterized by non-motor manifestations such as REM sleep behaviour disorder and hyposmia, may precede clinical diagnosis by several years. Cerebrospinal fluid biomarkers and dopamine transporter imaging have shown potential for early risk identification, supporting the application of machine learning (ML) approaches to PD diagnosis and risk stratification.

OBJECTIVES

This study aimed to provide a transparent, reproducible and leakage-safe benchmark of supervised ML models for PD classification using clinical and biomarker data, and to identify the most informative features for discriminating PD patients from healthy controls (HC) through convergent feature ranking across multiple selection approaches.

METHODS

Data from 58 participants (40 PD, 18 HC) were analysed by integrating demographic, clinical, and biomarker variables,

including dichotomous biomarker-status indicators, core neurodegeneration markers (bNfL, bGFAP, bpTau217, bN42, bN33, NSE), vascular/endothelial/hemostasis markers (U-PAR, VEGFA, ICAM-2, VCAM-1, VWF, T-PA, P-selectin, PECAM-1, E-selectin, SELPLG), immune/inflammatory markers and spectroscopy-derived metabolites (NAA, Cho, Glu, Ins) measured in the substantia nigra and globus pallidus bilaterally. Leakage-safe preprocessing comprised feature selection and standardization, with balanced class weighting applied where supported. Eight supervised ML classifiers were evaluated: Logistic Regression, Linear and Radial Basis Function Support Vector Machines (SVM), Random Forest, Gradient Boosting, k-Nearest Neighbours, Decision Tree, and Multi-Layer Perceptron. Hyperparameters were optimized through grid-search with stratified ten-fold cross-validation, using balanced accuracy as the primary optimization criterion and receiver operating characteristic area under the curve (ROC-AUC) as a secondary metric. Analyses were repeated across eight biomarker subsets. Model interpretability was investigated using SHapley Additive exPlanations (SHAP), while feature relevance was quantified through convergent ranking across ANOVA F-test, Mutual Information, Random Forest importance, LASSO coefficients, and principal component analysis (PCA) weighted loadings.

RESULTS

Linear SVM achieved the highest cross-validation balanced accuracy (0.86 ± 0.19) and was selected as the

best-performing model. On the independent test set, it achieved a ROC-AUC of 0.88 and a balanced accuracy of 0.79, with F1-scores of 0.71 for HC and 0.82 for PD. SHAP analysis identified MoCA as the most influential predictor. Convergent feature ranking identified IOIT as the dominant predictor in the full-feature set, while MoCA ranked first after excluding dichotomous biomarkers, followed by MMSE and bNfL. Within biomarker domains, bNfL emerged as the leading neurodegeneration biomarker; U-PAR, VEGFA, and ICAM-2 as the most relevant vascular markers, IL2-RA, TNF-R1, IL17RA, and TNFSF13B as the leading immune markers and GP-DX-Glu and SN-DX-Ins as the top spectroscopy derived features. PCA loadings diverged from supervised rankings, indicating limited overlap with discriminative relevance.

CONCLUSIONS

Leakage-safe ML pipelines were promising for distinguishing PD patients from HC using multimodal clinical and biomarker data. Although limited by the small sample size, this transparent and reproducible framework supports the use of parsimonious feature sets, with MoCA emerging as a robust predictor. SHAP-based interpretability further enhanced model transparency and the identification of clinically meaningful markers.

Bias-Adjusted Augmented Functional Random Forests for Interpretable Classification of Functional Trajectories

Annamaria Porreca^(1,2), Fabrizio Maturo⁽³⁾, Stefano Bonassi^(1,2)

 Annamaria Porreca 0000-0003-3278-1561

(1) Department for the Promotion of Human Science and Quality of Life, San Raffaele University, Rome, Italy

(2) Unit of Clinical and Molecular Epidemiology, IRCCS San Raffaele Roma, Rome, Italy

(3) Department of Economics, Statistics and Business, Faculty of Technological and Innovation Sciences, Universitas Mercatorum, Rome, Italy

CORRESPONDING AUTHOR: Annamaria Porreca, Department for the Promotion of Human Science and Quality of Life, San Raffaele University, Rome, Italy. E-Mail: annamaria.porreca@uniroma5.it

INTRODUCTION

Functional data analysis treats time-indexed observations as curves rather than isolated measurements, allowing classification models to exploit signal level, local velocity and curvature. Random Forests provide accurate non-parametric classifiers for complex predictors. When functional trajectories are augmented with first and second derivatives and summarized through functional principal component scores, the resulting predictors may be highly correlated. In this setting, standard variable-importance measures may overstate the role of components that mainly act as proxies for other functional information.

OBJECTIVES

This study proposes a bias-adjusted extension of Augmented Functional Random Forests for interpretable classification of functional trajectories. The aim is not only to combine the original curve, its first derivative and its second derivative within an ensemble classifier, but also to obtain importance measures that are less affected by redundancy among functional components. The central methodological contribution is a conditional randomization scheme for augmented functional principal components, combined with a stability criterion for identifying components that are both informative and robust.

METHODS

Observed trajectories are smoothed into continuous functions and augmented through successive derivatives. A sep-

arate Functional Principal Component Analysis is performed for each functional block, namely the original function, the first derivative and the second derivative. The resulting scores are used as predictors in an augmented Random Forest. Because components extracted from different blocks may describe related aspects of the same dynamic process, variable importance is assessed through Conditional Permutation Importance rather than marginal permutation alone. The randomization of each component is performed conditionally on correlated components belonging to the same functional rank, following the logic of conditional variable importance for Random Forests. The conditional loss in predictive accuracy is interpreted as the component-specific contribution not already explained by alternative functional representations. This measure is then combined with a Conditional Importance Stability Index based on repeated refitting of the forest.

RESULTS

The proposed approach separates marginal relevance from conditional relevance in augmented functional components. Components that appear important under ordinary permutation may lose relevance after conditioning, suggesting an inflation effect due to functional redundancy. Conversely, components that preserve high conditional importance and stable behaviour across refitted forests provide stronger evidence of a genuine predictive role. The resulting rankings are more selective, less driven by proxy effects and more coherent with the dynamic interpretation of the trajectories. This distinction is especially relevant in biomedical applications, where signal level, velocity and curvature may contain overlapping but non-equivalent information.

CONCLUSIONS

Bias-adjusted Augmented Functional Random Forests preserve the predictive flexibility of ensemble learning while improving the interpretability of functional classification. Conditional randomization reduces the risk of attributing importance to redundant components, whereas the stability index favours predictors whose contribution is robust across repeated model fits. The framework therefore links classification accuracy with a more controlled assessment of variable relevance, supporting a clearer interpretation of the dynamic information contained in functional trajectories.

A Multidimensional Analysis Framework for the Study of Contaminated Sites: Construction of Composite Indexes and their Association with Mortality

Rossella Murtas⁽¹⁾, Daniela Vacca⁽²⁾, Vanda Costantino⁽²⁾, Silvana Garozzo⁽²⁾, Lucia Martina⁽²⁾, Debora Tronu⁽²⁾, Luigi Minerba⁽¹⁾

 Rossella Murtas 0000-0002-8989-7605

(1) University of Cagliari

(2) Italian National Institute of Statistics (ISTAT)

CORRESPONDING AUTHOR: Rossella Murtas, University of Cagliari. E-mail: rossella.murtas@unica.it

INTRODUCTION

Contaminated Sites of National interest (SINs) represent highly complex settings in which environmental, demographic, socioeconomic, and productive factors interact, influencing the health status of resident populations. In these areas, consistently with the One Health paradigm, health results from the interaction among determinants belonging to different and closely interconnected domains. Given this complexity, studying the health status of populations living in SINs requires multidimensional analytical approaches capable of integrating heterogeneous information and capturing the systemic nature of these territories. Consequently, there is a need to adopt statistical methodologies able to simultaneously represent the multiple dimensions characterizing these contexts, overcoming the limitations of analyses based on single indicators, which describe only isolated aspects of the phenomenon under study.

OBJECTIVES

This study proposes a methodological framework for the multidimensional characterization of SINs, based on the construction of composite indexes capable of integrating both the spatial component of the data and the multiple dimensions characterizing these contexts. The aim is to identify the main latent dimensions of the territory and assess their association with all-cause and cause-specific mortality, providing useful tools for epidemiological surveillance and public health decision-making.

METHODS

The approach is applied to the Sulcis-Iglesiente-Guspinese and Porto Torres SINs in Sardinia, using municipalities as the territorial unit of analysis. A set of indicators derived from institutional data sources describes demographic characteristics, environmental pressure, level of industrialization, socioeconomic vulnerability, and land use. Statistical techniques are employed to reduce the dimensionality of the dataset and construct composite indicators representing the main latent territorial dimensions. These synthetic indexes are subsequently incorporated into regression models to investigate their association with all-cause mortality and specific causes of death. The analyses are further complemented by spatial statistical techniques aimed at evaluating the geographical distribution of territorial components and the role of spatial dependence in interpreting relationships between territory and health.

RESULTS

The proposed approach makes it possible to reduce and summarize a large number of variables into a limited set of composite indicators capable of representing the main structural characteristics of the territories, while reducing problems related to information redundancy and multicollinearity. The integration of regression models and spatial analysis techniques provides deeper insight into the role of different territorial dimensions in the distribution of mortality and highlights potential geographical patterns of epidemiological interest.

CONCLUSIONS

The proposed framework integrates dimensionality reduction methods, spatial analysis, and environmental epidemiology into a unified approach for studying territorial determinants of health. The methodology is flexible and replicable in other contexts characterized by high environmental complexity, contributing to the development of tools that support epidemiological surveillance and the planning of public health interventions according to the One Health paradigm.

Eating by the Clock: A Three-Mode Component Analysis of Temporal Dietary Patterns in the Italian Adult Population from IV SCAI Nutrition Survey 2017-2020

Vanessa O Nwajei⁽¹⁾, Denise Pires-Marafon⁽¹⁾, Luigi Palla⁽¹⁾

 Luigi Palla 0000-0001-6944-1750

(1) Department of Public Health and Infectious Diseases, Sapienza University, Rome, Italy

CORRESPONDING AUTHOR: Luigi Palla, Department of Public Health and Infectious Diseases, Sapienza University, Rome, Italy. E-Mail: luigi.palla@uniroma1.it

INTRODUCTION

Italy, as a country of origin of the Mediterranean dietary tradition, embodies a culturally entrenched temporal meal structure that is chrono-nutritionally coherent but increasingly threatened by modern lifestyle pressures. Standard two-way statistical methods such as principal component analysis (PCA) cannot simultaneously model the three-dimensional structure of the dietary survey data (individuals, nutrients, and time) motivating the use of three-way component analysis.

OBJECTIVES

To apply Tucker3 tensor decomposition to the Fourth Italian National Food Consumption Survey (IV SCAI 2017-2020) to identify latent chrono-nutritional archetypes and characterise their demographic and lifestyle determinants.

METHODS

The IV SCAI surveyed 1,969 participant aged 3 months to 74 years using two non-consecutive 24-hour dietary recalls. Adults aged 18 and above were retained (N = 915). Four energy-yielding macronutrients (alcohol, carbohydrates, fat, protein) were considered. Three-way array (915 x 4 x 6 time slots) was constructed for each data representation. The Convex Hull heuristic selected model of dimension (2,2,4) for kcal (fit: 84.14%). Subject-mode scores were profiled against eleven demographic and lifestyle variables using correlation, ANOVA, t-test, and multiple regression (N = 877 with complete lifestyle data).

RESULTS

Variance partitioning of the kcal data attributed 3.65% of variance to SS_a (individuals – Mode A), 17.02% to SS_b (macronutrient – Mode B), 29.15% to SS_c (timeslot-Mode C), 3.69% to SS_ab (individual-nutrient interaction), 18.83% to SS_ac (individual-time interaction), 10.62% to SS_bc (nutrient-time interaction), and 16.67% to SS_abc (individual-nutrient-time interaction). The timeslot, nutrient and the individual-time interaction were the dominant variance sources, followed by the three-way interaction, confirming individual-nutrient-time heterogeneity and justifying the use of Tucker3. The rotated individuals (A) mode had two components: A2, captures the population average pattern, and A1, captures the individual deviation from the baseline pattern. The rotated macronutrient (B) had also two: B1, contrasting fat (0.61) and protein (0.24) against carbohydrate (-0.73), and B2, capturing an average sustained carbohydrate and fat intake (0.68 and 0.67 respectively). The rotated temporal (C) mode had four components: C1, late-night snacking ($h_6 = 0.96$); C2, both lunch and dinner eating ($h_3 = 0.67$, $h_5 = 0.69$); C3, morning eating ($h_1 = 0.78$, $h_2 = 0.48$); and C4, a contrast between lunch ($h_3 = 0.71$) and dinner ($h_5 = -0.69$). The dominant core interaction, $A2 \times B2 \times C2 = 20,222.76$, described the variation from average individual pattern of carbohydrate-fat intake concentrated at lunch and dinner. A carbohydrate-dominant, fat-depleted morning meal was captured by $A2 \times B1 \times C3 = -1,956.53$. The A1 subject component identified individuals who concentrate their caloric intake disproportionately at lunch ($A1 \times B2 \times C4 = 6,119.29$; $A1 \times B1 \times C4 = 1,350.24$). Demographic profiling indicated that the A1 pattern was significantly associated with South-

ern Italian residence ($p = 0.006$) and employment status ($p = 0.027$ for retirement), while body mass index showed no association.

CONCLUSION

Tucker3 three-way component analysis effectively decomposed the chrono-nutritional structure of the Italian adult diet into interpretable individual, macronutrient, and temporal components, showing that a large energy intake variation is attributable to time, nutrient main effects as well as their three way interactions. These findings support the application of Tucker3 analysis as a methodological approach for chrono-nutritional research.

Multidimensional Validation of Clinical Evidence: A Four-Domain Audit Framework

Francesco Masedu⁽¹⁾, Monica Mazza⁽¹⁾, Margherita Attanasio⁽¹⁾, Ilenia Le Donne⁽¹⁾, Nicole Covone⁽¹⁾, Marco Valenti⁽¹⁾

 Francesco Masedu 0000-0003-0290-5324

(1) Department of Biotechnological and Applied Clinical Sciences, University of L'Aquila, Italy

CORRESPONDING AUTHOR: Francesco Masedu, Department of Biotechnological and Applied Clinical Sciences, University of L'Aquila, L'Aquila, Italy. E-Mail: francesco.masedu@univaq.it

INTRODUCTION

The evaluation of clinical evidence continues to rely predominantly on measures of statistical significance and relative effect size. Although indispensable, these metrics alone may not adequately characterize evidence quality, particularly with respect to follow-up integrity, clinical relevance, predictive performance, and vulnerability to bias. As regulatory and clinical decision-making increasingly incorporates evidence from both randomized and real-world studies, complementary approaches for assessing the robustness of findings are needed.

OBJECTIVES

To present a four-domain audit framework for the multidimensional validation of clinical evidence and to demonstrate its application in a proof-of-concept survival dataset.

METHODS

The proposed framework comprises four sequential domains of evaluation: (1) Follow-up Integrity, assessed through inverse Kaplan–Meier analysis to identify informative censoring and differential attrition; (2) Clinical Relevance, evaluated using Restricted Mean Survival Time (RMST) to quantify benefit in clinically interpretable time units; (3) Predictive Robustness, assessed through discrimination (Harrell's C-index), calibration analysis, and bootstrap-based optimism correction; and (4) Inferential Resilience, evaluated using E-value methodology to estimate the susceptibility of observed findings to unmeasured influences arising during study conduct and follow-up. These domains were selected to represent complementary dimensions of evidence quality: data integrity, clinical relevance, predictive validity, and inferential

robustness. The framework was applied to the Mayo Clinic Primary Biliary Cirrhosis (PBC) dataset as a proof of concept.

RESULTS

Follow-up integrity was preserved, with no evidence of differential censoring between groups (inverse Kaplan–Meier log-rank $p=0.54$). However, subsequent analyses identified important limitations. RMST analysis showed a between-group difference of 51.98 days (95% CI -341.87 to 445.84 ; $p=0.796$), indicating limited clinical relevance. Predictive performance was modest, with an optimism-corrected C-index of 0.639 and a calibration slope of 0.933. Finally, the E-value was 1.98, indicating that relatively modest unmeasured influences arising during follow-up could potentially account for the observed association. Overall, findings that appeared methodologically acceptable under conventional statistical assessment did not consistently satisfy all dimensions of the proposed validation framework.

CONCLUSIONS

The proposed audit framework provides a structured approach for integrating assessments of follow-up integrity, clinical relevance, predictive performance, and inferential resilience within a unified methodological framework. In this proof-of-concept application, preserved data integrity coexisted with limited clinical benefit, modest predictive robustness, and vulnerability to follow-up-related biases. By making these dimensions explicit and reproducible, the framework may complement conventional statistical evaluation in evidence appraisal, peer review, and regulatory assessment. Further validation across randomized and real-world datasets is warranted.

A Bayesian Network Approach to Investigate Breakthrough Cancer Pain

Sergio Coluccia⁽¹⁾, Marco Cascella⁽¹⁾, Dalila Esposito⁽²⁾, Rosario De Feo⁽²⁾, Vittorio Santoriello⁽³⁾, Francesco Pio Maria Di Carlo⁽¹⁾, Paola Rocco⁽¹⁾, Anna Crispo⁽¹⁾

 Sergio Coluccia 0000-0003-4044-1217

(1) Istituto Nazionale Tumori-IRCCS "Fondazione G. Pascale", via M. Semmola 9, Naples, 80131, Italy

(2) Department of Medicine, Surgery and Dentistry "Scuola Medica Salernitana", University of Salerno, Baronissi, 84081, Italy, unisa.it

(3) Department of Electrical Engineering and Information Technology, University of Naples Federico II, Naples, Italy

CORRESPONDING AUTHOR: Sergio Coluccia, sergio.coluccia@istitutotumori.na.it

INTRODUCTION

Cancer pain is a complex, multidimensional syndrome affecting approximately 55% of patients with cancer, depending on tumor stage and treatment, and is associated with poorer clinical outcomes and reduced quality of life. Pain assessment and management become particularly challenging when patients experience transient exacerbations of pain despite otherwise controlled baseline pain with opioid therapy, a condition known as breakthrough cancer pain (BTcP), which affects up to 70% of patients with chronic cancer pain. Artificial intelligence (AI) and advanced statistical approaches offer new opportunities to investigate the complex relationships among clinical variables contributing to BTcP. The Ruggi Study (ClinicalTrials.gov: NCT07038434) is a single-center, non-profit, observational-interventional study designed to characterize chronic pain through multimodal AI-based assessments in adult oncologic and non-oncologic patients.

OBJECTIVES

This preliminary analysis aimed to investigate the causal relationship between opioid therapy and BTcP, evaluating the mediating role of chronic pain phenotype in a subcohort of 99 cancer patients assessed at their first follow-up visit.

METHODS

A Bayesian network was constructed including age, smoking status, body surface area (BSA), comorbidities, cardiovascular therapy, cancer severity (presence of metastases or bone metastases), ongoing anticancer treatment at enrollment, time from diagnosis, treatment duration, opioid therapy,

chronic pain phenotype, BTcP, and quality of life assessed by the Brief Pain Inventory (BPI, item 9g). The network structure was learned using the Bayesian Information Criterion, allowing the identification of statistical dependencies among variables. Missing data were preliminarily assessed using Little's test and subsequently imputed through a model-based approach.

RESULTS

Overall, 22% of the dataset contained missing values. Little's test excluded a missing-not-at-random mechanism ($p = 0.14$), supporting the imputation strategy, which mainly affected variables related to cancer therapies. The median patient age was 65 years, with a median BSA of 1.4 m²; approximately 98% had a sex-adjusted normal BSA. Most patients had comorbidities, 66% had a time from diagnosis exceeding 12 months, 17% presented with bone metastases, and 39% were receiving opioid therapy. Neuropathic pain was the most frequent chronic pain phenotype (>60%), while eight patients experienced BTcP. Twenty-six patients reported impaired quality of life (BPI item 9g >4). The Bayesian network did not identify a direct causal relationship between opioid therapy and BTcP. Instead, chronic pain phenotype emerged as a mediator: patients with nociceptive pain receiving opioids showed a higher probability of developing BTcP than those with neuropathic or mixed pain (25% vs 19%).

CONCLUSIONS

These preliminary findings suggest that the association between opioid therapy and BTcP is influenced by the underlying chronic pain phenotype rather than representing a direct

causal relationship. Bayesian network analysis provides a valuable framework for disentangling complex clinical interactions and identifying patient subgroups at increased risk of BTcP, potentially supporting more personalized pain management strategies. Larger prospective studies are warranted to validate these findings and assess their clinical applicability.

Index by Author

A

Agodi, Antonella 51, 83, 139
 Albanese, Nicolò Cosimo 103
 Albano, Noemi 179
 Aldigeri, Raffaella 69, 71
 Alibrandi, Angela 161
 Alicandro, Gianfranco 61, 100
 Amodio, Emanuele 75
 Antolini, Laura 175
 Arcangeli, Valentina 12
 Archibugi, Livia 12
 Arienti, Chiara 81
 Arshad, Mamoona 57
 Ascef, Bruna De Oliveira 42
 Attanasio, Margherita 189
 Attardi, Enrico 47
 Autiero, Martina 117
 Azzolina, Danila 115, 119, 121, 123, 127, 174

B

Baccini, Michela 89, 133, 135, 145, 158
 Balan, Denisa Gabriela 111
 Baldassari, Arianna 67
 Baldi, Enrico 44, 125
 Baldi, Ileana 173
 Baldini, Marco 35
 Barchitta, Martina 51, 139
 Barci, Francesco 29
 Bartolacci, Silvia 35
 Bartolomeo, Nicola 146
 Basileo, Teresa 117
 Battaglia, Salvatore 57
 Baviera, Marta 108

Belli, Annalisa 17
 Bellocco, Rino 87
 Bellomo, Giovanni 181
 Bellomo, Sofia 75
 Bencardino, Daniela 55
 Bendoni, Stefano 31
 Bendotti, Sara 44, 125
 Berchiolla, Paola 113, 173
 Bergamaschi, Roberto 165
 Bergui, Mauro 31
 Bernardinelli, Luisa 16
 Bernasconi, Davide Paolo 175
 Bertoli, Gloria 165
 Bertolotti, Marinella 85
 Besio, Roberta 141
 Biffi, Eleonora 111
 Bilò, Maria Beatrice 57
 Biscardi, Giulio 145
 Boccia, Stefania 40
 Bonanno, Lilla 137, 181
 Bonassi, Stefano 183
 Bonazzetti, Cecilia 154
 Bonifazi, Martina 67
 Bonomi, Alice 107
 Bono, Roberto 57
 Borgatti, Renato 59
 Borghi, Costanza 89
 Bortolotti, Marta 47
 Boschini, Antonio 146
 Bracco, Matteo 173
 Braggion, Riccardo 55
 Bravi, Francesca 39
 Brugnera, Agostino 177
 Bruno, Denise 63
 Buffo, Camilla 73
 Burbui, Giorgia 89

C

- Calabrese, Laura 44
 Calderaio, Claudia 117
 Caliri, Santina 129
 Caltagirone, Luca 85
 Calzari, Luciano 49
 Calza, Stefano 177
 Caminiti, Caterina 31
 Camparini, Luca 98
 Camussi, Elisa 42
 Cangiano, Lorenzo 107
 Cannella, Roberto 20
 Capaso, Angelo 123
 Cappellari, Manuel 31
 Caputo, Roberta 117
 Caranci, Nicola 11
 Carbone, Alessio 24, 25, 27
 Carbone, Erika 49
 Carinci, Fabrizio 69, 71
 Carle, Flavia 53, 65, 67, 94
 Carloni, Emanuela 47
 Carnevale, Martina 177
 Carrara, Silvia 12
 Casarini, Martina 154
 Cascella, Marco 190
 Casetta, Ilaria 31
 Castellani, Gastone 154
 Catalini, Alessandro 35
 Cattadori, Gaia 63
 Cattaneo, Andrea 14
 Cattaneo, Daniele 47
 Caudo, Danilo 129, 131
 Cecchi, Francesca 148, 150
 Centola, Carmela 94
 Cereda, Cristina 79
 Cesaro, Paola 177
 Cescon, Matteo 154
 Cheungpasitporn, Wisit 87
 Chiappini, Alessandra 35
 Chierigato, David 165
 Chiodini, Paolo 143
 Chirici, Gherardo 89
 Ciccirillo, Francesco 146
 Cinquepalmi, Alice 16
 Cirillo, Luisa 117
 Ciurleo, Rosella 181
 Cocco, Stefania 117
 Colombo, Giulia 79
 Coluccia, Sergio 190
 Conti, Michela 75
 Cooley, Victoria 37
 Corrao, Giovanni 47, 59
 Corsico, Angelo Guido 57
 Corsi, Leonardo 148
 Corsonello, Andrea 87
 Cosentino, Nicola 107
 Costantini, Riccardo Cuoghi 179
 Costantino, Maria Antonella 59
 Costantino, Vanda 185
 Covone, Nicole 189
 Cozzi, Ilaria 47
 Cremaschi, Marco 165
 Crispo, Anna 190
 Cristina, Maria Luisa 24, 25, 27
 Currao, Alessia 44, 125

D

- D'Acapito, Fabrizio 163
 da Cruz, Manuel Marques 160
 Dal Buono, Arianna 12
 D'Amico, Roberto 179

D'Angelo, Martina 119
 David, Alessio 42
 De Cola, Maria Cristina 129, 131
 De Feo, Rosario 190
 De Giuseppe, Rachele 165
 Dei Cas, Alessandra 69, 71
 De Laurentiis, Michelino 117
 Del Castello, Lorenzo 175
 Del Furia, Matteo Johann 81
 Del Giovane, Cinzia 160
 Della Franca, Chiara 17
 Della Porta, Matteo Giovanni 47
 De Luca, Raffaele 12
 De Martino, Maria 105
 De Silvestri, Annalisa 55
 De Simone, Adriano 163
 D'Esposito, Vittoria 119
 Di Bartolo, Paolo 69, 71
 Di Belgiojoso, Elisa Barbiano 65
 di Blasio, Martina 173
 Di Caprio, Giovanni 16
 Di Carlo, Francesco Pio Maria 190
 Di Gennaro, Piergiacomo 143
 Di Liberto, Erminia 139
 Di Mitri, Chiara 83
 Di Molfetta, Sergio 73
 D'ippoliti, Daniela 47
 Di Spazio, Lorenzo 57
 Di Stefano, Giovanni 20, 83
 Distefano, Salvatore Antonino 22
 Di Tolla, Michele Francesco 119, 121
 Dolce, Pasquale 115, 119, 121, 123, 127, 174
 Dolci, Alberto 79
 Drabczyk, Rebecca 46
 Dragani, Davide Fernando 49
 Dupont, Chiara 25, 27

Duregon, Eleonora 167

E

Ebrahimian, Serajeddin 171
 Enea, Marco 22, 51, 75, 83, 139
 Ercoli, Vittoria 165
 Esposito, Dalila 190
 Eusse Villa, Luisa Fernanda 11
 Evangelisti, Andrea 94

F

Fainardi, Enrico 31
 Falco, Lucia 148, 150
 Falconi, Massimo 12
 Fanti, Stefania 73
 Faraci, Ambra 137, 181
 Faragalli, Andrea 67
 Favara, Giuliana 139
 Favaretto, Francesco 73
 Fazia, Teresa 16
 Fazzi, Elisa 59
 Ferrante, Gianluigi 42
 Ferrari, Clarissa 177
 Ferraro, Simona 79
 Ferretti, Virginia Valeria 55
 Filonzi, Andrea 35
 Filosa, Asia 59
 Fiori, Federica 39
 Focati, Maria Elisena 85
 Fognani, Alice 14
 Fontana, Elisa Carla 167
 Fontanarosa, Alessandro 35
 Fordellone, Mario 143
 Forlino, Antonella 141
 Formisano, Pietro 119
 Fortino, Ida 108

Fortner, Renée Turzanski 37
 Fracchiolla, Nicola Stefano 47
 Franceschi, Roberto 73
 Francesconi, Paolo 46
 Franchi, Matteo 47, 59, 73, 107
 Franco Perez, Wendy 141
 Fresi, Eleonora 55
 Fulici, Gianmaria 29
 Fusco, Pietro 119, 174

G

Gaetani, Lorenzo 181
 Galli, Jessica 59
 Galuzzi, Bruno Giovanni 165
 Gambini, Giulia 55
 Ganna, Andrea 169
 Garozzo, Silvana 185
 Gasparrini, Antonio 108
 Gembillo, Guido 87
 Gemmi, Fabrizio 47
 Genovese, Dario 75
 Gentilini, Davide 49
 Gerardi, Viviana 177
 Germani, Sabrina 49
 Gesuita, Rosaria 53, 65, 67, 94
 Giacani, Julieta 148
 Giammanco, Giovanni 22
 Giannella, Maddalena 154
 Gibertoni, Dino 98
 Giometto, Sabrina 46
 Giorgianni, Gabriele 96
 Gobbi, Federico Giovanni 55
 Graceffa, Domenico 22
 Gregorini, Mariacristina 98
 Gregorio, Caterina 105
 Grieco, Vittorio 96

Grimaldi, Alessandro 137
 Grisotto, Laura 156
 Grzegorz, Bilo 92
 Guglielmelli, Paola 47
 Gurnari, Carmelo 47

H

Hartonen, Tuomo 169
 Huaman, Marisol Palomino 17

I

Iaconelli, Antonio 115
 Iervolino, Domenico 115, 119, 121, 123, 174
 Ievoli, Riccardo 156
 Iezzi, Elisa 69, 71
 Ippoliti, Anna 79
 Ippolito, Antonella 96
 Iraci, Nunzio 129, 131
 Isola, Miriam 105
 Iurlo, Alessandra 47

J

Jemioł, Paweł 160

K

Karjalainen, Juha 169
 Khan, Mohd Rashid 121, 123, 174
 Khoshroo, Seyedmohammad 69, 71
 Klersy, Catherine 55
 Koumantakis, Emanuele 113

L

Laici, Cristiana 154
 Lakhla, Ayman 65, 67
 Lanfranchi, Giuseppe 87
 Lantos, Tamás 57

La Vecchia, Carlo 39, 40, 61, 91, 100
 Lazzarini, Stefano Giuseppe 81
 Lazzarino, Laura 85
 Le Donne, Ilenia 189
 Lentini, Nicolò 40
 Leo, Francesco 167
 Leoni, Olivia 47, 59
 Ligi, Daniela 17
 Limosani, Ivan 59
 Locatelli, Francesca 11, 57
 Lo Cicero, Lorenzo 87
 Lodico, Manuela 20, 22
 Lombardi, Carolina 171
 Lombardi Fortino, Domenico 53
 Lorusso, Letizia 146
 Lucenteforte, Ersilia 46, 148, 150
 Lunetta, Christian 129, 131
 Lupi, Antonello 35

M

Macciotta, Alessandra 152
 MacKenzie Myles, Liam Alexander 161
 Maconi, Antonio 85
 Maffioli, Margherita 47
 Maglietta, Giuseppe 31, 111
 Maglio, Gabriele M. 63
 Magnano San Lio, Roberta 139
 Maines, Evelina 73
 Malo, Elvia 55
 Mameli, Chiara 79
 Manacorda, Tommaso 18, 33
 Mancini, Cristina 35
 Mancuso, Stefano 89
 Mandreoli, Marcora 98
 Mangiafico, Salvatore 31
 Maniscalco, Laura 20, 22, 83, 139
 Mannello, Ferdinando 17
 Mannini, Andrea 148, 150
 Mannocci, Alice 14
 Marchetti, Pierpaolo 11, 29, 57
 Marcon, Alessandro 11, 29, 57
 Marconi, Vittorio 35
 Marenzi, Giancarlo 107
 Maresca, Giuseppa 137
 Margotto, Ludovica 47
 Marra, Angela 137
 Marrella, Gabriele Biagio 75
 Martella, Donato 79
 Martina, Lucia 185
 Martino, Grazia Anna Rita 94
 Marzi, Chiara 77, 145, 148, 150, 158
 Masala, Daniele 14
 Masedu, Francesco 189
 Masiero, Mauro 11
 Masini, Stefano 135
 Matranga, Domenica 20, 22, 51, 75, 83, 139
 Mattei, Alessandra 145
 Mattei, Antonella 110
 Maturo, Fabrizio 183
 Maugeri, Andrea 51, 139
 Maugeri, Andrea Giuseppe 83
 Mazza, Monica 189
 Mazzucco, Riccardo 85
 Mencarelli, Martina 69, 71
 Menichelli, Lorenzo 73
 Mensi, Martina Maria 59
 Merlo, Emanuele Maria 161
 Merolle, Valentina 14
 Messeri, Lorenzo 77
 Messina, Rossella 69, 71
 Miceli, Silvana 20, 83
 Michelini, Massimo M. 69, 71

Mignozzi, Silvia 37, 61, 100
 Militi, Annalisa 137
 Minari, Samuele 171
 Minerba, Luigi 185
 Minozzi, Silvia 160
 Mion, Marco 44
 Moirano, Giovenale 158
 Monti, Maria Cristina 165, 169
 Montomoli, Cristina 59, 165
 Monzani, Dario 20
 Mora, Barbara 47
 Morelli, Maria Cristina 154
 Mori, Francesca 42
 Morini, Elisabetta 131
 Morsellino, Vera 24
 Murgia, Nicola 57
 Murtas, Rossella 185
 Musarò, Antonio 129, 131

N

Narcisi, Selene 85
 Nardi, Giovanna 181
 Natale, Arianna 39, 91
 Navarro, Jorge 148, 150
 Negri, Eva 61
 Negrotti, Anna 111
 Nencini, Patrizia 31
 Nobile, Federica 143
 Nova, Andrea 16
 Nuredini, Andi 179
 Nuvolone, Daniela 145, 158
 Nyári, Tibor András 57

O

Olivieri, Fabiola 63
 O Nwajei, Vanessa 187

Ojeda Fernandez, Luisa 108
 Orgiana, Arianna 27
 Orlando, Giulia 167
 Ornago, Alice Margherita 105
 Ostojić, Jelena 67
 Ottogalli, Paolo 146
 Ottria, Gianluca 25, 27

P

Pacella, Daniela 163
 Pacileo, Guglielmo 85
 Pagliuca, Martina 117
 Paiella, Salvatore 12
 Palazzo, Lucio 156
 Paletta, Pasquale 18, 33
 Palla, Luigi 187
 Pampaloni, Francesca 77
 Pancani, Silvia 150
 Panero, Benedetta 79
 Papa, Mattia 131
 Papotti, Mauro 167
 Parati, Gianfranco 92
 Parnetti, Lucilla 181
 Parola, Sara 117
 Parpinel, Maria 39
 Parsa, Neda 46
 Partilora, Sofia 148, 150
 Pascariello, Angelo 14
 Passamonti, Francesco 47
 Passantino, Andrea 63
 Pastorino, Roberta 40
 Patel, Linia 61
 Patuzzo, Eros 63
 Pedretti, Giulia 111
 Pedretti, Roberto 63
 Pellegrini, Ilaria 150

Pelosi, Giuseppe 167
 Pengo, Martino 171
 Peretti, Clara 29
 Perigord, Adrien 105
 Perna, Simone 16
 Pessano, Sara 24, 25, 27
 Pezzini, Alessandro 31
 Picasso, Erica 85
 Picconi, Barbara 14
 Pietro, Fusco 121
 Pietropaolo, Emanuele 173
 Piezzo, Michela 117
 Pini, Sara 179
 Pires-Marafon, Denise 187
 Pirina, Pietro 57
 Pizzato, Margherita 61
 Poli, Vanna 59
 Pollastro, Carla 143
 Ponzio, Michela 18, 33
 Porreca, Annamaria 183
 Pracucci, Giovanni 31
 Prattichizzo, Francesco 63
 Prima, AlessiaDi 96
 Primi, Roberto 44, 125
 Puliani, Natascia 53
 Puntoni, Matteo 31

Q

Quartarone, Angelo 137

R

Ragusa, Rosalia 96
 Rashid, Mohd 115
 Ravaioli, Matteo 154
 Razzanelli, Matilde 46
 Reale, Laura Alessia 59

Rebora, Paola 175
 Rebuffi, Chiara 24
 Remoundou, Konstantina 113
 Restuccia, Giovanni 137
 Richiardi, Lorenzo 152
 Righi, Luisella 167
 Rigoni, Marta 63, 167
 Rinaldi, Matteo 154
 Rocchi, Ettore 154
 Rocchi, Marco B. L. 17
 Rocco, Paola 190
 Romairone, Paolo 27
 Romano, Alessandra 96
 Rondanelli, Mariangela 16
 Rossi, Marta 39, 91
 Roussaki, Ioanna 113
 Ruffini, Enrico 167
 Ruotolo, Agostino 77
 Russo, Mariateresa 123, 127

S

Sacco, Davide 49
 Saia, Valentina 31
 Salerno, Sergio 20
 Saletti, Andrea 31
 Salivetto, Marco 18, 33
 Sallustio, Fabrizio 31
 Salvatori, Francesca 17
 Sanfilippo, Lorenza 55
 Santarelli, Luca 94
 Santelli, Enrica 47
 Santoriello, Vittorio 190
 Santoro, Antonio 98
 Santoro, Domenico 87
 Santos, João Vasco 108
 Santucci, Claudia 40, 91, 100

Saracci, Rodolfo 158
 Sartini, Marina 24, 25, 27
 Sattin, Davide 129, 131
 Savastano, Simone 44, 125
 Scapellato, Alessia 96
 Scarfi, Chiara 129, 131
 Schiattarella, Paola 143
 Schinca, Elisa 25, 27
 Schünemann Holger, Jens 81
 Sciascia, Savino 87
 Scisciola, Vittorio 119, 121
 Secchettin, Erica 12
 Seidita, Aurelio 75
 Seidita, Valeria 20
 Senore, Carlo 42
 Sguanci, Chiara 133
 Signoriello, Simona 143
 Silipigni, Salvatore 87
 Simeon, Vittorio 42, 143
 Siniscalchi, Antonio 154
 Sisti, Davide 17
 Skrami, Edlira 53, 65, 67, 94
 Smimmo, Annafrancesca 143
 Soraci, Luca 87
 Soranna, Davide 92
 Sorbera, Chiara 181
 Sousa-Pinto, Bernardo 160
 Spadaro, Savino 121
 Spatari, Gloria 85
 Speranza, Teresa 143
 Stafoggia, Massimo 143
 Stefania, Arena 42
 Stefania, Bertazzon 125
 Sturniolo, Alessandro 137
 Szasz, Claudia 47

T

Tagliabue, Elena 63
 Tagliaferro, Sofia 11
 Tassi, Rossana 31
 Tavazzi, Eleonora 165
 Terrin, Maria 12
 Teruzzi, Elisa 107
 Tettamanti, Mauro 108
 Tinelli, Ornella 117
 Tinervia, Giovanni 75
 Toni, Danilo 31
 Tonucci, Sonia 53, 94
 Torriani, Camilla 141, 165, 169
 Torrigiani, Ginevra 108
 Torrisi, Antonina 96
 Tramonti, Caterina 148, 150
 Trerotoli, Paolo 146
 Trewin-Nybråten, Cassia B. 37
 Triolo, Federico 105
 Tronu, Debora 185
 Tucci, Fabrizio 92
 Tupler, Rossella 179
 Turati, Federica 39

U

Ulmanova, Leila 44, 125
 Unim, Brigid 67

V

Vacca, Daniela 185
 Valenti, Marco 189
 Valleverdina, Anna Carrabs 53
 Vallone, Stefano 31
 Valsecchi, Paola Maria 111
 Vannucchi, Alessandro Maria 47

Vasse, Marc 17
Vella, Giuseppe 75
Venturelli, Francesco 42
Venuto, Alessia 137, 181
Verdi, Luca 29
Verlato, Giuseppe 57
Veronesi, Valentina 167
Vetrano, Davide Liborio 105
Viale, Pierluigi 154
Viarchi, Arianna 110
Vicentini, Massimo 42
Videtta, Valeriano 148, 150
Vieira, Rafael José 160
Villani, Simona 44, 125, 141
Virzi, Valeria 42
Viscardi, Cecilia 133, 135
Visconti, Alessia 113, 173
Vitiello, Fabiana Pia 129, 131
Volante, Calogero 85
Volante, Marco 167
Von Arx, Claudia 117
Voso, Maria Teresa 47
Vukić Dugac, Andrea 67

Z

Zambelli, Vanessa 167
Zambon, Antonella 79, 81, 92, 108, 171
Zanzi, Riccardo 100
Zini, Andrea 31
Ziqian, Zeng 92
Zuccotti, Gianvincenzo 79
Zugna, Daniela 152