Use of Real World Data for development of an algorithm to evaluate the complexity of people with chronic diseases controlling for socioeconomic and environmental inequalities

Álvaro Franquet[1,2,3], Pere Plaja[1], Montse Saguer[1], Inma Quintana[4], Marc Saez[2,3], Maria A. Barceló[2,3]

[1] Fundació Salut Empordà, Figueres, Spain
[2] Research Group on Statistics, Econometrics and Health (GRECS), Universitat de Girona, Girona, Spain
[3] CIBER in Epidemiology and Public Health (CIBERESP), Spain
[4] Consell Comarcal de l'Alt Empordà, Figureres, Spain

INTRODUCTION

- For more than ten years, various coordinated strategies for the care of people with chronic diseases have been implemented in Spain.
- Stratification, based on predictive models, has become the most important axis on which these strategies gravitate[1,2].
  - Level 0: Healthy people with or without risk factors.
  - Level 1: Presence of 1-2 chronic diseases. Low risk patients, with conditions still in incipient states.
  - Level 2 (complex chronic patients, **PCC**): Presence of multiple diseases (3-4 chronic diseases). High risk patients but of less complexity.
  - Level 3 (advanced chronic patients, **MACA**): Patients with multiple chronic diseases and greater burden of fragility. 5 or more chronic diseases.
- There is evidence that there are socioeconomic inequalities in health[3-5]. A non-negligible proportion of them is caused by environmental problems[3].
- In general, environmental conditions can contribute to socioeconomic inequalities in health through *differential exposure* and through *differential susceptibility*.
- Only some predictive model of stratification of the population with chronic diseases adjusts for socioeconomic inequalities and none for environmental inequality[6].
- As a consequence, which could have important implications regarding the main strategy of care, the use of specialized hospital care and, therefore, in the allocation of available resources.
- Our objective is to assess the survival of the different levels of patient prioritization, taking and not taking into account socioeconomic and environmental inequalities.

METHODS

- We used a retrospective population-based cohort, covering the region 'La Selva Interior', Catalonia, composed of adult population 15 years or more), followed between 2005 and 2012 (in total 21038 subjects were followed during that period, of which 598 died).
- *Variables obtained from other sources*: Deprivation index. Source: Spanish Population and Housing Census 2011; Analyses: Index built with the IneqCities methodology[7]; Unemployment rate: number of men and women 16-64 years old and >64 years old, all at the census tract level. Source: Spanish Population and Housing Census 2011. Air pollutants. Source: Catalan Network for Pollution Control and Prevention, 2005-2011. Analyses:INLA-SPDE[8].
- Random effects were included to control for unobserved confounders, individual heterogeneity and spatial adjustment.

Estimation of the prevalences

- We used a two-part econometric model (Hurdle model)[9]. In the first part we estimated the probability that the subject was present in the sample using a generalized linear mixed model with binomial link (i.e. logistic regression).

$$\mu_{1i} = Prob(y_i = 1|X_1, \beta_1)$$
$$\log\left(\frac{\mu_{1i}}{1-\mu_{1i}}\right) = X_1'\beta_1$$
$$Var(y_i|X_1, \beta_1) = \mu_{1i}(1-\mu_{1i})$$

- In the second part we estimated the number of PCCs and MACAs cases in each census tract by means of a truncated negative binomial, reweighted by the probabilities estimated in the first part.

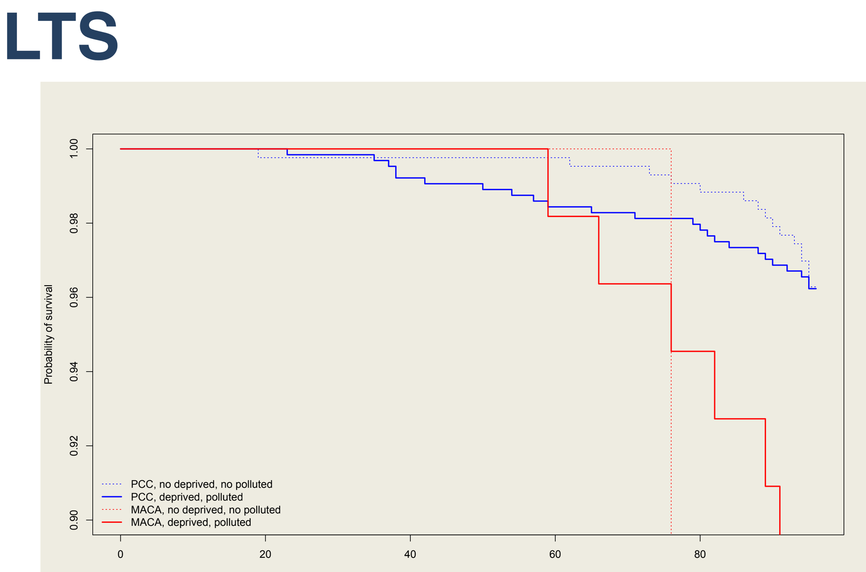
$$f_2(y_i|X_2, y_i > 0, \beta_2) = \frac{\Gamma(y_i+\psi_{2i})}{\Gamma(\psi_{2i})\Gamma(y_i+1)} \times \left(\frac{\mu_{2i}}{\mu_{2i}+\psi_{2i}}\right)^{y_i}\left[\left(\frac{\mu_{2i}+\psi_{2i}}{\psi_{2i}}\right)^{\psi_{2i}} - 1\right]^{-1}$$

- Inferences, made simultaneously, were performed following the Integrated Nested Laplace Approximation (INLA) approach, within a Bayesian framework.

Survival analysis

- Andersen-Gill models allowing delayed entry and time-varying covariables[10,11].

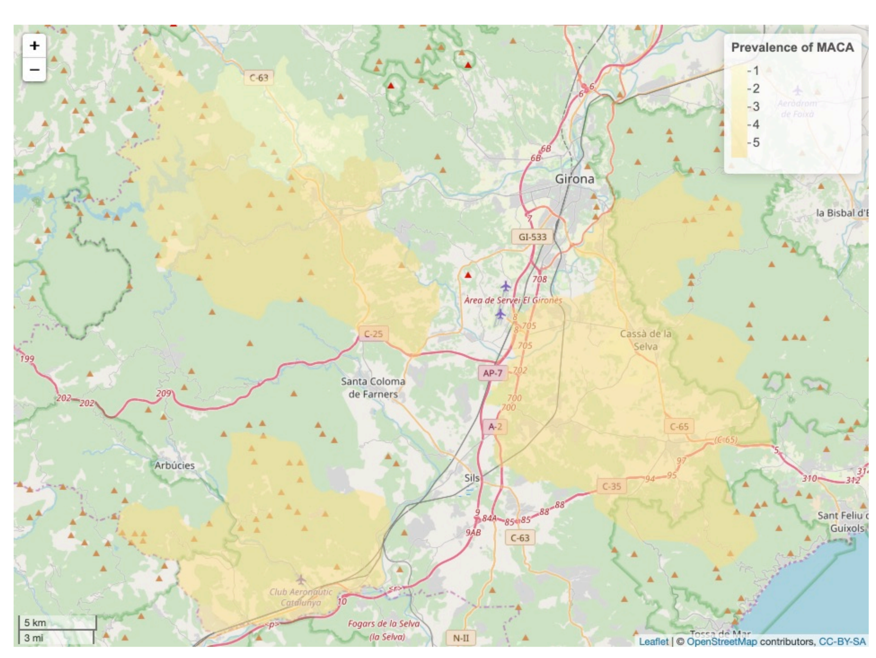
RESULTS



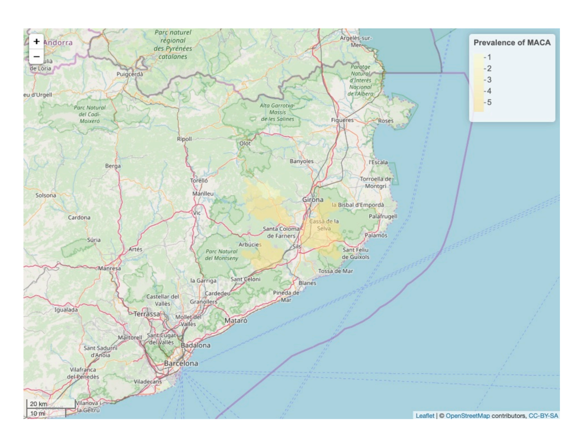
Stratification	Hazard ratio (95% confidence interval)	
	Without taking into account SES and environmental inequalities	Taking into account SES and environmental inequalities
Level 1	1.145 (0.843-1.554)	1.211 (0.811-6.029)
PCC	1.604 (1.163-2.211)	1.992 (0.675-5.881)
MACA	2.847 (1.837-4.413)	4.449 (1.423-6.597)

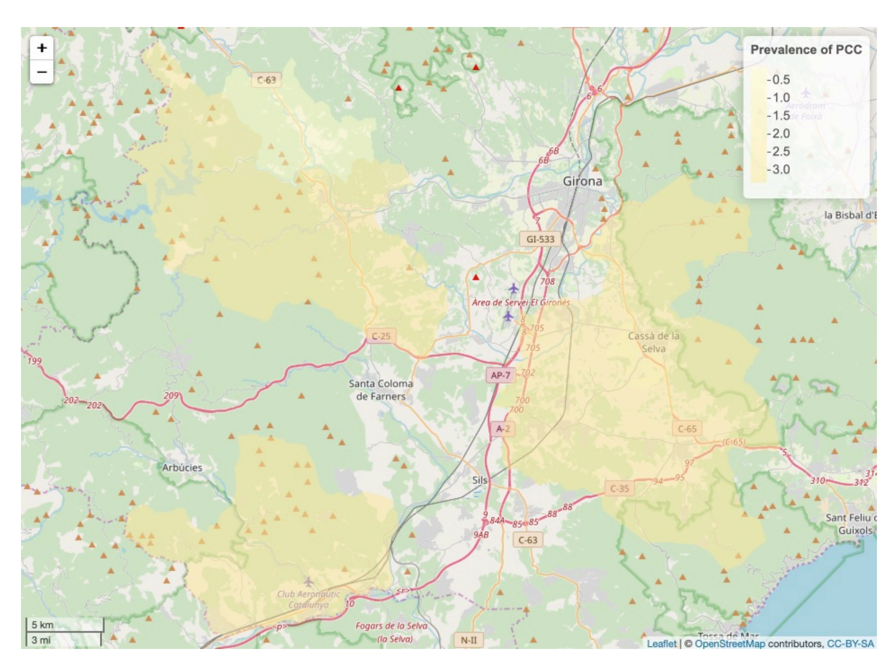
Adjusted by sex, age (15 5-year groups), diagnosis of cancer, smoking and enolic status (No, Yes, ex), obesity.

Reference: Healthy people with or without risk factors.



Prevalence of MACA (per 1000 habs)





Prevalence of PCC (per 100 habs)

CONCLUSIONS

- Using a method free of selection bias we have estimated the prevalence of MACAs and PCCs: 0.663 (95% CI:0.453-0.927) and 10.439 (95% CI:8.709-12.347) per 100 inhabitants of 15 years and more, respectively.
- These prevalences, however, present a considerable spatial heterogeneity, with a higher prevalence in census sections that are more polluted and more economically deprived.
- Take into account socio-economic and environmental inequalities, practically double the risk of dying of MACA subjects and eliminate the risk difference of PCCs in relation to healthy subjects.

References

1.- Blay C, Limón E (coords). *Bases para un modelo catalán de atención a las personas con necesidades complejas. Conceptualización e introducción a los elementos operativos*, 2017
2.- Estupiñán-Ramírez M, Tristancho-Ajamil R, Company-Sancho MC, Sánchez-Janáriz H. Comparación de modelos predictivos para la selección de pacientes de alta complejidad. *Gac Sanit.* 2019; 33(1):60-65.
3.- Acheson D. *Independent Inquiry into Inequalities in Health Report*. London, England: The Stationary Office, 1998 [Available at:: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/265503/ih.pdf, last accessed on January 31, 2019].
4.- Marmot M. Social determinants of health inequalities. *Lancet* 2005; 365(9464):1099-1104.
5.- Kunst AE. Describing socioeconomic inequalities in health in European countries: an overview of recent studies. *Revue d'Epidemiologie et de Sante Publique* 2007;55:3–11.
6.- Deguen S, Zmirou-Navier D. Social inequalities resulting from health risks related to ambient air quality – A European review. *European Journal of Public Health* 2010; 28(1):27-35.
7.- Hoffmann R, Borsboom G, Saez M, Mari Dell'Olmo M, Burström B, Corman D, Costa D, Deboosere P, Domínguez-Berjón MF, Dzúrová D, Gandarillas A, Gotsens M, Kovács K, Mackenbach J, Martikainenn P, Maynou L, Morrison J, Palència L, Pérez G, Pikhart H, Rodríguez-Sanz M, Santana P, Saurina C, Tarkiainen L, Borrell C. Social differences in avoidable mortality between small areas of 15 European cities: an ecological study. *Int. J Health Geogr.* 2014; 13:8.
8.- Lindgren F, Rue H, Lindström J. An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach (with discussion).
9.- Saez M, Barceló MA, Coll-de-Tuero G. A selection-bias free method to estimate the prevalence of hypertension from an administrative primary health care database in the Girona Health Region, Spaiin. *Computer Methods and Programs in Biomedicine* 2009; 93:228-240-
10.- Andersen PK, Gill RD. Cox's regression model for counting processes: a large sample study. *Ann Stat* 1982; 10:1100-1120.
11.- Barceló MA. Marginal and conditional models in multivariate survival analysis. *Gac Sanit.* 2002; 16(52):59-68.

This work was partly funded by one of the Industrial Doctorates 2018 of the AGAUR, Catalonian Government, and the projects MPCUdG2016/69, GDRCompetUdG2017 and TRANSFER2016/24 of the University of Girona.