Social Science & Medicine 327 (2023) 115955

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

Living longer in declining health: Factors driving healthcare costs among older people

Laia Maynou [a,b,c,*], Andrew Street [b], Anna García−Altés [d,e,f]

[a] Department of Economics, Econometrics and Applied Economics, Universitat de Barcelona, Avinguda Diagonal, 690, 08034, Barcelona, Spain
[b] Department of Health Policy, London School of Economics and Political Science, Houghton Street, London, WC2A 2AE, UK
[c] Center for Research in Health and Economics (CRES), Universitat Pompeu Fabra, Ramon Trias Fargas 25-27, 08005, Barcelona, Spain
[d] Direcció General de Planificació i Recerca en Salut, Departament de Salut, Generalitat de Catalunya, Barcelona, Spain
[e] CIBER de Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain
[f] Institut de Investigació Biomèdica (IIB Sant Pau), Barcelona, Spain

ARTICLE INFO

Handling Editor: Joanna Coast

Keywords:
Multimorbidity
Long term conditions
Adjusted morbidity groups
Proximity to death

ABSTRACT

Background: Developed countries are facing challenges in caring for people who are living longer but with a greater morbidity burden. Such people are likely to be regular users of healthcare.
Objectives: Our analytical aim is to identify factors that explain healthcare costs among: (1) people over 55 years old; (2) the top 5% and 1% high-cost users among this population; (3) those that transition into the top 5% and 1% from one year to the next; (4) those that appear in the top 5% and 1% over multiple years; and (5) those that remain in the top 5% and 1% over consecutive years.
Methods: The data covered 2011 to 2017 and comprised 1,485,170 observations for a random sample of 224,249 people aged over 55 years in the Catalan region of Spain. We analysed each person's annual healthcare costs across all public healthcare settings related to their age, gender, socio-economic status (SES), whether or not and when they died, and morbidity status, through Adjusted Morbidity Groups.
Results: After controlling for morbidity status, the oldest people did not have the highest costs and were less likely to be among the most costly patients. There was also only a modest impact on costs associated with SES and with dying. Healthcare costs were substantially higher for those with a neoplasm or four or more long term conditions (LTCs), costs rising with the complexity of their conditions. These morbidity indicators were also the most important factors associated with being and remaining in the top 5% or top 1% of costs.
Conclusion: Our results suggest that age and proximity to death are poor predictors of higher costs. Rather, healthcare costs are explained mainly by morbidity status, particularly whether someone has neoplasms or multiple LTCs. Morbidity measures should be included in future studies of healthcare costs.

1. Introduction

As populations get older, the burden of morbidity increases. This underpins a growing problem that all developed countries are trying to grapple with: how to care for an increasing proportion of people who are living longer but in declining health. Research in this area has been compromised by a lack of data about the morbidity burden of each individual member of the population (Werblow et al., 2008; Bilger and Chaze, 2008; Geue et al., 2014). This information deficit originally led researchers to consider other individual characteristics that might explain their healthcare costs. Early studies highlighted the positive association between age and costs, as critiqued by de Meijer et al. (de Meijer et al., 2013), but subsequent studies argued that costs in most healthcare settings (other than primary care (Atella and Conti, 2014)) are explained better by how close someone is to dying (their "proximity to death") rather than their age (Zweifel et al., 1999; Alemayehu and Warner, 2004; Murphy and Martikainen, 2013; Geue et al., 2014). However, this argument meant that researchers had to address the difficulty of dealing with the endogenous relationship between costs and proximity to death (spending increases as people are close to dying but increased spending postpones their death) (Salas and Raftery, 2001; Felder et al., 2010). The insight also served little practical purpose as, for

* Corresponding author. Department of Economics, Econometrics and Applied Economics, Universitat de Barcelona, Avinguda Diagonal, 690, 08034, Barcelona, Spain.
E-mail addresses: laia.maynou@ub.edu (L. Maynou), a.street@lse.ac.uk (A. Street), agarciaaltes@gencat.cat (A. García−Altés).

https://doi.org/10.1016/j.socscimed.2023.115955
Received 10 January 2023; Received in revised form 3 April 2023; Accepted 5 May 2023
Available online 10 May 2023

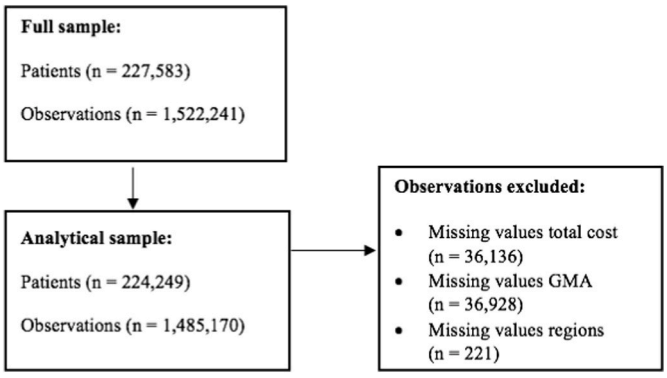


Fig. 1. Full sample to Analytical Sample.

most people, the timing of death cannot be accurately foretold (Colombier and Weber, 2011; Hazra et al., 2018; Howdon and Rice, 2018).

More recent studies have attempted to account for morbidity status in explaining healthcare costs. This has been a challenge because there is no standard way to measure morbidity, empirical practice dependent on data availability which varies geographically and by healthcare setting and dataset (Lehnert et al., 2011; Soley-Bori et al., 2021). Studies have used markers for causes of death (Polder et al., 2006; de Meijer et al., 2011; Luta et al., 2020), particular conditions (Seshamani and Gray, 2004; Violán et al., 2014), impairments and disability (Dormont et al., 2006; Hazra et al., 2018), chronic disease groups (Moore et al., 2017), Clinical Risk Groups (Carreras et al., 2018), Diagnosis and Drug Morbidity Groups (Longden et al., 2018) and Clinical Classification Software groups (Howdon and Rice, 2018). In studies that have included information about an individual's age, whether and when they died, and morbidity status, the latter proves the key determinant of costs, not age or proximity to death (de Meijer et al., 2011; Colombier and Weber, 2011; Howdon and Rice, 2018; Hazra et al., 2018; Carreras et al., 2018). We contribute to this literature by using Adjusted Morbidity Groups (AMGs) which are used in many regions of Spain to measure the form and complexity of the morbidity burden for each member of the population (Monterde et al., 2016; Caballer-Tarazona et al., 2019).

AMGs have three attractive features. First, they are built on the World Health Organisation's International Classification of Diseases (ICD) (WHO, 2019). This means that AMGs can be constructed in any context where ICD codes are used to code encounters with the health system. Second, there are only 36 AMGs, thus satisfying the call for simpler rather than overly complex systems to describe morbidity (Brilleman et al., 2014). Third, and most importantly, AMGs have substantial ability to explain variation among patients in their annual costs, as will be demonstrated by the analyses that follow.

We also contribute to a second strand of literature focusing on people with high needs and high costs (HNHC). These studies recognise that a relatively small proportion of people account for a large portion of health care costs. For example, in the US, the top 5% and top 1% of users account, respectively, for 50% and 22% of total health expenditure (Blumenthal et al., 2016). In a systematic review, Wammes et al. (2018) looking at the US, Canada, Germany, the Netherlands, Denmark and Taiwan found that the top 1% of users account for 23% of total health expenditure (ranging from 14% to 33%). In Catalunya, Vela et al. (2017) also show that the top 1% of users account for 23% of total health expenditure.

If it can be predicted who might join the HNHC group, interventions such as care coordination or case management might be developed to reduce avoidable costs and to ensure that services are available to meet their needs (Ronksley et al., 2015; Figueroa et al., 2021). A handful of studies have examined transitions into or persistence among the HNHC group, using panels of data spanning from two to five years (Anderson and Knickman, 1984; Monheit, 2003; Ronksley et al., 2015; Longden et al., 2018; Hwang et al., 2015). However, this literature has limitations. The paper by Ronksley et al. (2015) examined inpatient costs only for those admitted to a single hospital; Anderson and Knickman (1984)) and Longden et al. (2018) dropped from analysis those who died during the period (Longden et al. (2018) also dropped those with zero costs), limiting generalisability; and, with the exceptions of Monheit (2003) and Wammes et al. (2019), only summary statistics associated with transition or persistence are reported rather than multiple regression results, even if such analyses had been conducted (Hwang et al., 2015; Longden et al., 2018).

Building on these foundations, we analysed a panel of data covering seven years and annual costs associated with utilisation across all healthcare settings by older people. The aims were to assess the influence of individual characteristics in explaining annual healthcare costs among: (1) people over 55 years old; (2) the top 5% and 1% high-cost users among this older population; (3) those that transition into the top 5% and 1% from one year to the next; (4) those that appear in the top 5% and 1% frequently over multiple years; and (5) those that remain persistently in the top 5% and 1% over consecutive years. In the next sections we describe the data and our analytical approach, before summarising the results and drawing conclusions.

2. Methods

To analyse each individual's annual healthcare costs, we specified the following equation:

$$y_{it} = \alpha + \sum_{m=1}^{M} \beta_m X_{it} + \sum_{j=1}^{J} \gamma_j AMG_{it} + \delta_r + \lambda_t + \varepsilon_{it} \quad (1)$$

where $y_{it} = \{c_{it}, y_{it}^B, y_{it}^T\}$ is a set of different measures of cost for individual i in year t. The first, c_{it}, measures total annual costs across healthcare settings and is estimated as a generalised linear model (GLM), given that costs are highly skewed. The second, y_{it}^B, is a binary variable with $y_{it}^B = 1$ if the individual is among the top 5% (or 1%) in year t, and $y_{it}^B = 0$ otherwise, estimated using a logit model. A logit model is also used to assess factors associated with the transition into the top 5% (or 1%), y_{it}^T. Here, $y_{it}^T = 1$ if $y_{it_0}^B = 0$ and $y_{it_1}^B = 1$ where t_0 and t_1 are consecutive years; $y_{it}^T = 0$ otherwise.

X is a vector of variables capturing the individual's gender, age group in five year age bands, socio-economic status (SES), and year and month of death for those who died. *AMG* is a vector of dummy variables comprising 31 mutually exclusive AMGs to which each individual is allocated each year. We estimate equation (1) with and without the *AMG* vector to demonstrate the impact their inclusion has on the estimated effects of the X variables (only the results for c_{it} appear in the main paper, with the reduced form appearing in Appendix for y_{it}^B and y_{it}^T). δ_r is a set of seven health region dummies, λ_t is a set of year dummies (the first year in the series is omitted when estimating y_{it}^T), and ε_{it} is a classical

Table 1
Descriptive statistics: outcomes and covariates, 2011–2017.

	Analytical sample N = 1,485,170		Top 5% N = 74,261		Top 1% N = 14,854	
	Mean/Prop	Std	Mean/Prop	Std	Mean/Prop	Std
Annual cost (€)	1,936	4,359	15,867	11,507	32,562	16,543
Transition to top 5% cost	2.7%	16.1%				
Transition to top 1% cost	0.6%	7.6%				
Male	44.7%	49.7%	51.7%	49.9%	58.2%	49.3%
Female	55.3%	49.7%	48.3%	49.9%	41.8%	49.3%
Age - 55-64	22.4%	41.7%	11.7%	32.1%	12.7%	33.3%
Age - 65-74	36.4%	48.1%	27.7%	44.7%	31.5%	46.5%
Age - 75-84	25.9%	43.8%	35.3%	47.8%	35.6%	47.9%
Age - 85-94	14.7%	35.4%	24.7%	43.1%	19.7%	39.8%
Age≥95	0.6%	7.7%	0.7%	8.5%	0.4%	6.6%
SES <18,000	60.9%	48.8%	72.8%	44.5%	71.6%	45.1%
SES 18,000–100,000	34.5%	47.5%	21.9%	41.4%	22.9%	42.0%
SES >100,000	1.2%	11.0%	0.5%	7.0%	0.7%	8.2%
SES exempt	3.4%	18.2%	4.7%	21.2%	4.8%	21.4%
Alive	98.3%	12.8%	89.1%	31.1%	86.4%	34.3%
Died	1.7%	12.8%	10.9%	31.1%	13.6%	34.3%
Died in January	0.2%	4.0%	0.1%	3.6%	0.1%	3.6%
Died in February	0.2%	3.9%	0.3%	0.6%	0.1%	3.8%
Died in March	0.2%	4.0%	0.6%	0.8%	0.3%	5.7%
Died in April	0.1%	3.6%	0.7%	0.9%	0.4%	6.4%
Died in May	0.1%	3.6%	0.9%	0.9%	0.7%	8.3%
Died in June	0.1%	3.3%	0.9%	9.3%	0.9%	9.5%
Died in July	0.1%	3.5%	1.0%	9.8%	1.1%	10.5%
Died in August	0.1%	3.5%	1.0%	10.0%	1.3%	11.4%
Died in September	0.1%	3.5%	1.0%	10.3%	1.8%	13.1%
Died in October	0.1%	3.7%	1.3%	11.2%	2.0%	14.1%
Died in November	0.1%	3.8%	1.4%	11.9%	2.2%	14.8%
Died in December	0.2%	4.0%	1.6%	12.6%	2.5%	15.7%
Healthy	9.5%	29.3%	0.2%	3.9%	0.1%	3.5%
Acute disease c1	0.8%	9.0%	0.0%	1.5%	0.0%	1.6%
Acute disease c2	0.6%	7.6%	0.0%	1.3%	0.0%	1.4%
Acute disease c3	0.3%	5.5%	0.0%	0.9%	0.0%	1.4%
Acute disease c4	0.2%	4.2%	0.0%	1.1%	0.0%	1.2%
Acute disease c5	0.1%	3.2%	0.0%	1.6%	0.0%	1.4%
Neoplasm c1	2.4%	15.3%	2.4%	15.3%	2.6%	15.8%
Neoplasm c2	2.0%	14.2%	5.7%	23.3%	6.6%	24.9%
Neoplasm c3	1.2%	11.0%	7.1%	25.7%	8.5%	27.8%
Neoplasm c4	0.9%	9.5%	8.3%	27.6%	11.1%	31.4%
Neoplasm c5	0.7%	8.3%	9.8%	29.8%	16.6%	37.2%
1 LTC c1	2.3%	15.1%	0.1%	3.1%	0.1%	2.6%
1 LTC c2	4.8%	21.3%	0.1%	3.3%	0.1%	3.8%
1 LTC c3	2.1%	14.4%	0.2%	4.3%	0.2%	3.9%
1 LTC c4	1.6%	12.7%	0.4%	6.0%	0.2%	4.2%
1 LTC c5	0.9%	9.6%	0.4%	6.1%	0.2%	4.8%
2-3 LTC c1	5.3%	22.4%	0.2%	5.0%	0.2%	5.0%
2-3 LTC c2	9.3%	29.1%	0.9%	9.3%	0.4%	6.5%
2-3 LTC c3	6.4%	24.5%	1.3%	11.4%	0.6%	7.8%
2-3 LTC c4	5.3%	22.3%	1.7%	13.0%	0.9%	9.2%
2-3 LTC c5	3.4%	18.2%	3.8%	19.1%	2.7%	16.3%
4+ LTC c1	9.6%	29.5%	2.2%	14.6%	0.9%	9.3%
4+ LTC c2	14.0%	34.7%	7.4%	26.2%	3.9%	19.5%
4+ LTC c3	7.0%	25.4%	7.7%	26.6%	4.7%	21.1%
4+ LTC c4	5.9%	23.5%	14.1%	34.8%	10.2%	30.3%
4+ LTC c5	3.3%	17.8%	25.9%	43.8%	29.1%	45.4%

Notes: SES: Socioeconomic status; c1-c5: complexity levels; LTC: Long Term Condition.

Table 2
Descriptive statistics (patients): outcomes and covariates, 2011–2017.

	Analytical sample N = 224,249		Top 5% N = 43,372		Top 1% N = 10,195	
	Mean/Prop	Top 1%	Mean/Prop	Top 1%	Mean/Prop	Top 1%
Number of years in Top 5%	0.33	0.87	1.71	1.24	2.45	1.64
Number of consecutive years in Top 5%	0.29	0.78	1.52	1.12	2.18	1.57
Number of years in Top 1%	0.07	0.38	0.34	0.80	1.46	1.06
Number of consecutive years in Top 1%	0.06	0.36	0.32	0.76	1.38	0.99
Male	45.0%	49.7%	50.3%	49.9%	58.5%	49.3%
Female	55.0%	49.7%	49.7%	49.9%	41.5%	49.3%
Age - 55-64	21.7%	41.2%	10.8%	31.0%	12.6%	33.2%
Age - 65-74	35.5%	47.9%	26.8%	44.3%	30.7%	46.1%
Age - 75-84	26.0%	43.9%	34.9%	47.7%	35.3%	47.8%
Age - 85-94	15.9%	36.6%	26.6%	44.2%	20.9%	40.7%
Age≥95	0.8%	9.1%	0.9%	9.6%	0.6%	7.5%
SES <18,000	61.9%	48.6%	73.0%	44.4%	71.9%	44.9%
SES 18,000–100,000	33.5%	47.2%	22.2%	41.6%	22.9%	42.0%
SES >100,000	1.2%	10.8%	0.4%	6.6%	0.6%	7.5%
SES exempt	3.5%	18.3%	4.3%	20.2%	4.6%	21.0%
Alive for full period	89.0%	31.3%	69.2%	46.2%	57.1%	49.5%
Died 2011	1.1%	10.5%	2.2%	14.6%	2.7%	16.2%
Died 2012	1.5%	12.0%	3.9%	19.2%	5.6%	22.9%
Died 2013	1.5%	12.1%	4.3%	20.3%	5.7%	23.3%
Died 2014	1.5%	12.3%	4.5%	20.8%	6.8%	25.2%
Died 2015	1.7%	12.9%	5.0%	21.8%	6.9%	25.4%
Died 2016	1.8%	13.4%	5.4%	22.5%	7.1%	25.8%
Died 2017	1.9%	13.6%	5.6%	23.0%	8.0%	27.1%
Healthy	5.0%	21.7%	0.0%	1.6%	0.0%	1.7%
Acute disease c1	0.3%	5.8%	0.0%	0.5%	0.0%	0.0%
Acute disease c2	0.3%	5.6%	0.0%	0.5%	0.0%	0.0%
Acute disease c3	0.2%	4.4%	0.0%	0.7%	0.0%	0.0%
Acute disease c4	0.1%	3.7%	0.0%	0.7%	0.0%	0.0%
Acute disease c5	0.1%	2.7%	0.0%	0.8%	0.0%	0.0%
Neoplasm c1	1.6%	12.6%	1.3%	11.4%	1.5%	12.2%
Neoplasm c2	2.8%	16.4%	4.8%	21.4%	5.0%	21.7%
Neoplasm c3	2.4%	15.4%	6.7%	25.1%	7.7%	26.7%
Neoplasm c4	2.4%	15.2%	8.8%	28.4%	12.4%	33.0%
Neoplasm c5	2.2%	14.8%	10.3%	30.4%	19.5%	39.7%
1 LTC c1	1.0%	9.9%	0.0%	1.2%	0.0%	1.0%
1 LTC c2	1.7%	12.9%	0.0%	1.5%	0.0%	2.0%
1 LTC c3	1.2%	10.9%	0.0%	1.9%	0.0%	1.4%
1 LTC c4	1.0%	10.1%	0.1%	2.9%	0.0%	2.0%
1 LTC c5	0.7%	8.6%	0.2%	4.2%	0.1%	3.4%
2-3 LTC c1	2.6%	16.0%	0.1%	2.4%	0.0%	1.7%
2-3 LTC c2	5.7%	23.2%	0.3%	5.9%	0.3%	5.1%
2-3 LTC c3	4.9%	21.7%	0.6%	7.5%	0.4%	6.2%
2-3 LTC c4	4.7%	21.3%	0.9%	9.4%	0.5%	7.1%
2-3 LTC c5	3.4%	18.0%	2.2%	14.7%	1.5%	12.2%
4+ LTC c1	9.4%	29.1%	1.6%	12.4%	0.6%	7.9%
4+ LTC c2	18.2%	38.6%	7.6%	26.5%	3.4%	18.1%
4+ LTC c3	10.6%	30.8%	10.2%	30.3%	5.2%	22.2%
4+ LTC c4	10.0%	30.0%	18.0%	38.5%	11.7%	32.2%
4+ LTC c5	7.4%	26.1%	26.2%	44.0%	30.0%	45.8%

Notes: SES: Socioeconomic status; c1-c5: complexity levels; LTC: Long Term Condition. The Top 5% and Top 1% are the number of patients that have been into that categories at any point during the study period.

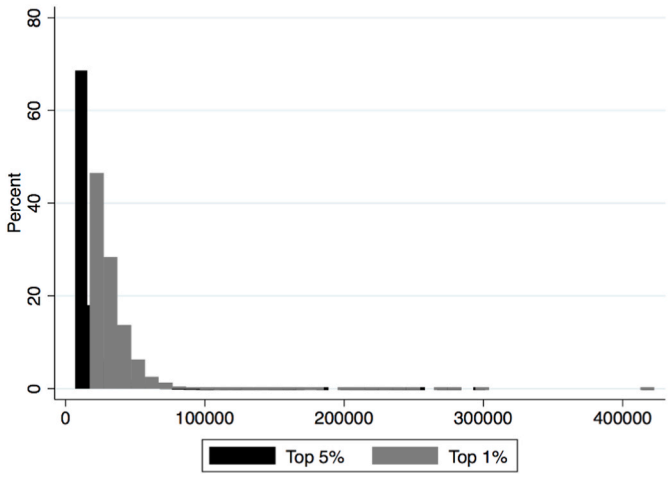


Fig. 2. Annual cost top 5% and top 1%.

error term. Standard errors are clustered at patient level and estimated coefficients from the GLM and logit models are reported as average marginal effects (Norton and Dowd, 2018).

We also examined factors associated with being frequently, y_i^F, and persistently, y_i^P, among high cost patients. Frequency is a count of the number of years that the individual appears among the top 5% (or 1%) across the seven year period. Persistence is measured by counting consecutive years that the individual appears among the top 5% (or 1%). In analysing both measures we applied an equation of the general form:

$$y_i = \alpha + \sum_{n=1}^{N} \beta_n X_i + \sum_{j=1}^{J} \gamma_j AMG_i + \delta_r + \varepsilon_i \quad (2)$$

Where $y_i = \{y_i^F, y_i^P\}$ and values for each X or AMG characteristic are those recorded either in 2017 or the year the individual died. Poisson models were used to estimate equation (2) for those in the top 5% but these models failed to converge for those with the top 1% of costs, so ordinary least squares was used instead. A robustness check was performed in which characteristics were set to 2011 rather than 2017 values (the results prove stable, as reported in Appendix).

Statistical analysis was conducted using Stata 15 (College Station, TX, USA).

3. Context and data

The Spanish public health system is funded via taxation giving all citizens access to all healthcare services free at point of use, with the exception of co-payments for dentistry, some types of medical equipment and pharmaceuticals (from which disabled people and those on selected state benefits are exempt) (OECD/European Observatory on Health Systems and Policies, 2021). In Catalunya in 2017 (the last year covered by our data), 70.5% of total healthcare expenditure was derived from taxation, while the remainder came from private spending (Fundación IDIS, 2022). In 2017, 28% of the Catalan population had additional private health insurance, giving them subsidised access to private health care, although this did not prevent them from using the public system (ESCA, 2017).

We analysed routine anonymised individual data about the use of publicly funded healthcare services collated by Agència de Qualitat i Avaluació Sanitàries de Catalunya (AQuAS) for the Catalan population of 7.5 m, of which 2 million are over 55 years old (Statistical Institute of Catalunya, Idescat, 2017). We took a 10% random sample from the Catalan Central Registry of Insured Persons of those aged 55+ on January 01, 2010, and analysed their utilisation of public healthcare services for seven years from 2011 to 2017, or to the date of death for those that died during the period. Every contact with the Catalan public healthcare system is recorded in administrative databases and the combined AQuAS dataset links pharmacy, primary care, hospital care, emergency department, long-term care, and mental health care data, as well as specialist care visits, non-emergency patient transportation, outpatient physical therapy, home-based oxygen therapy and dialysis.

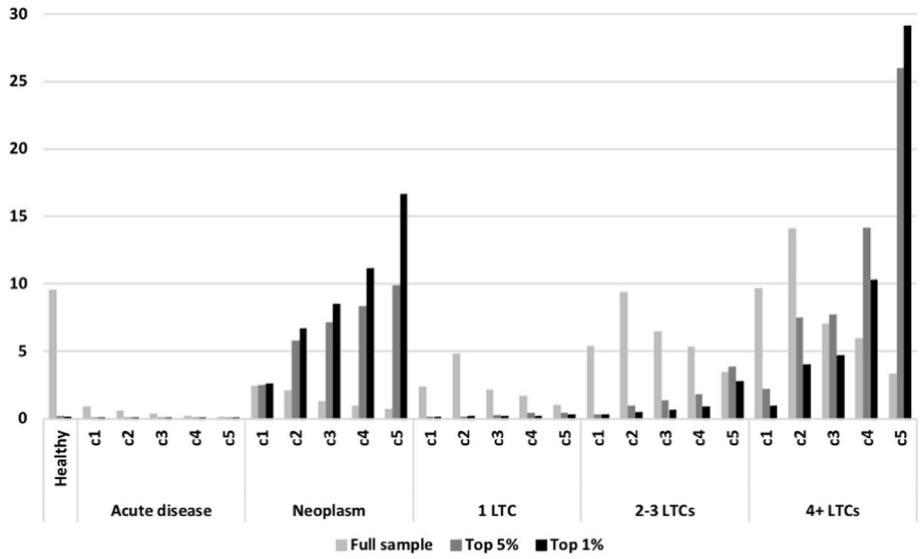


Fig. 3. Adjusted Morbidity Groups (% by full and subsamples).

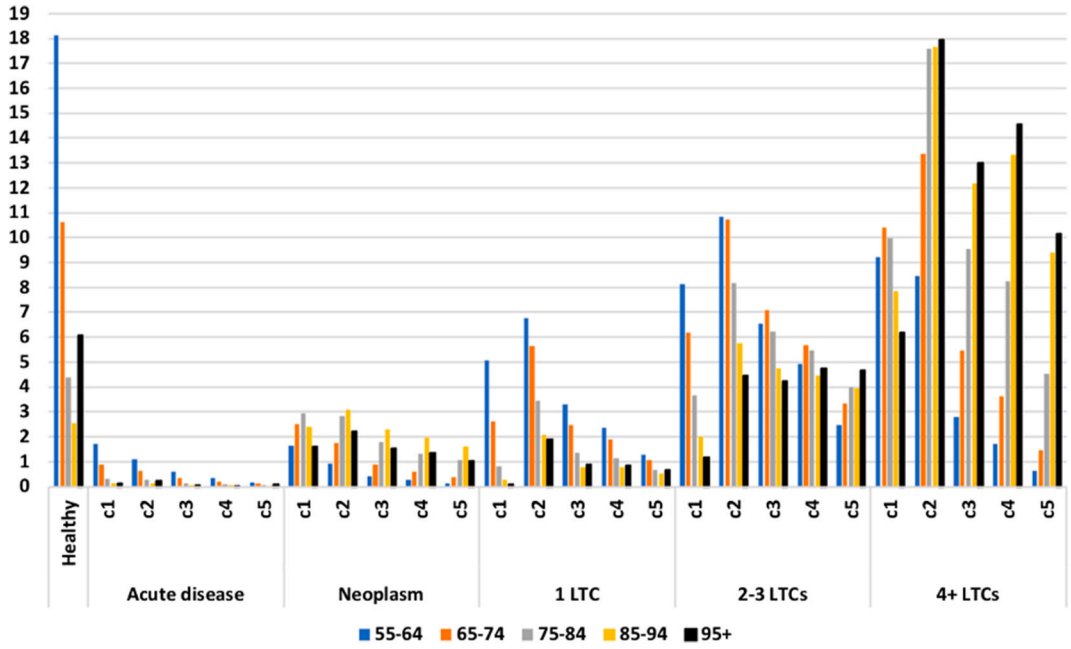


Fig. 4. Adjusted Morbidity Groups (% by age group).

The data also included the cost of each service delivered in each of these settings, based on reimbursement tariffs (Vela et al., 2018). The data indicated only the year not the actual date when services were used but this allowed us to calculate total annual health care expenditure (re-based to 2017 prices) across all settings for each individual.

The AQuAS dataset contains patient characteristics, including gender, age, and date of death (where applicable). We constructed dummy variables indicating the month of death, thereby accounting for the likelihood that individuals who died early in the year would have lower annual costs to those that died near the year end.

SES is proxied according to broad income brackets, based on the income reported in the previous year's tax return. This information is used to calculate pharmaceutical co-payments, such that there are four SES categories: Exempt (those with higher levels of disability, and people receiving selected state benefits or on minimum income policies); Low Income (<18,000€ a year); Medium Income (18,000€–100,000€ a year); and High Income (>100,000€ a year) (García-Altés et al., 2018; Carrilero et al., 2021; Servei Català de la Salut Generalitat de Catalunya, 2022a).

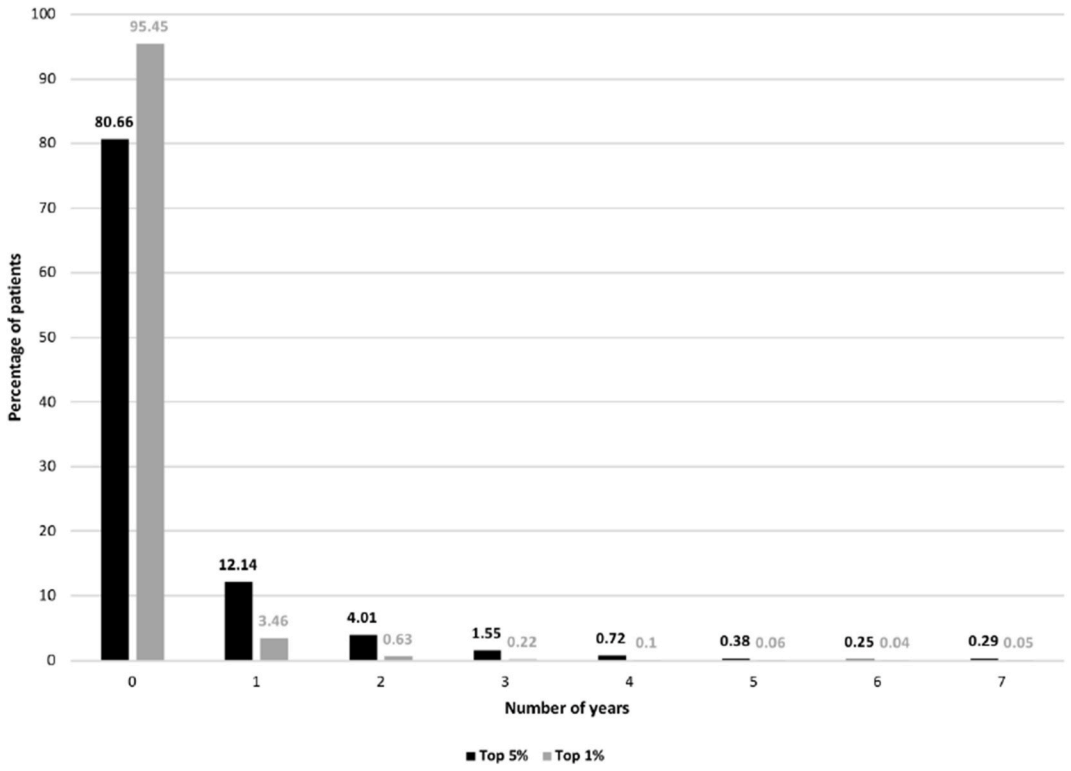
The dataset also includes information about each person's morbidity status, defined using AMGs which are based on a risk score tool validated for the Catalan population (Monterde et al., 2016, 2020). Individuals are classified each year to a particular AMG on the basis of diagnostic information recorded during the year using ICD10 codes (WHO, 2019). AMGs are structured in a hierarchy of six broad morbidity categories: healthy; pregnancy and childbirth (not relevant for our sample); acute disease; neoplasms (cancer); 1 LTC; 2–3 LTCs; and 4+ LTCs. The calculation and allocation of this grouping is explained in Appendix. As explained in Appendix, some of those in the healthy category did have positive health costs (notably pharmaceuticals), implying that they were not completely free of health problems, but their diagnoses were not used for allocation to the other five morbidity categories. With the exception of the healthy category, each of the other categories is divided into five levels of clinical complexity, generating 31 mutually exclusive AMG groups (and another five for pregnancy and childbirth). AMGs have been validated in the Catalan population in several studies (Monterde et al., 2016, 2020; Cleries et al., 2020; Vela et al., 2021) and have better explanatory power than the Charlson Comorbidity Index, counts of chronic conditions and Clinical Risk Groups when explaining healthcare resource use or costs.

4. Results

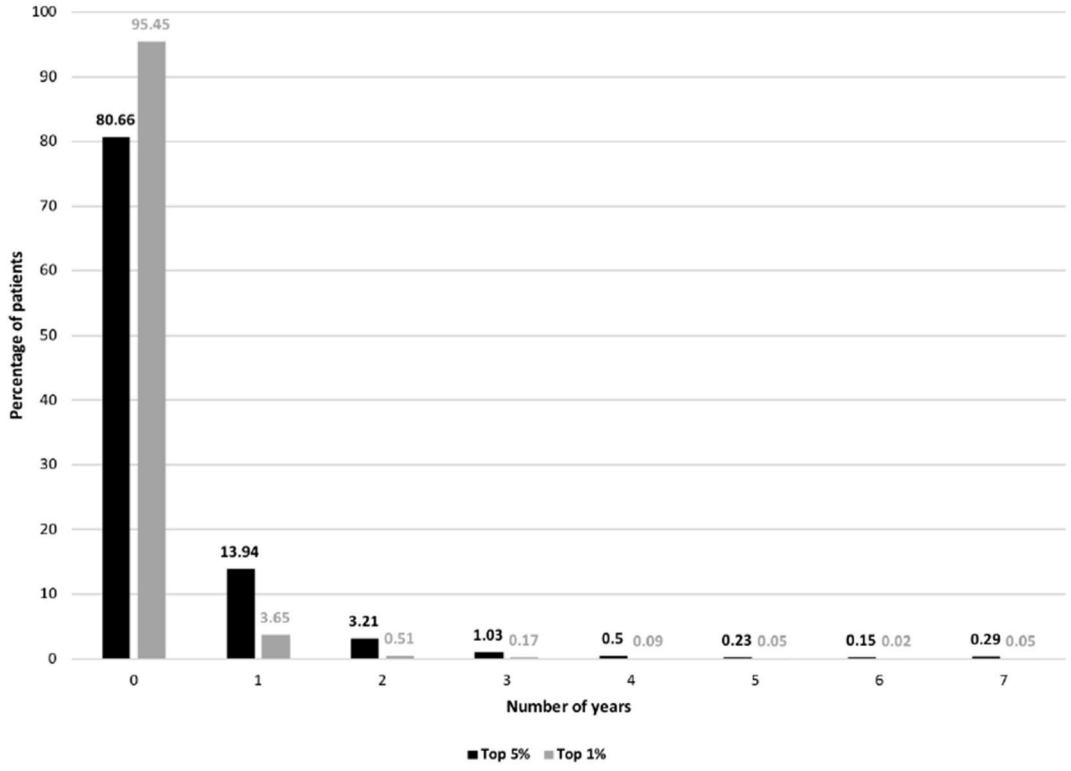
4.1. Characteristics of study subjects

As depicted in Fig. 1, the starting sample consisted of 1,522,583 observations for 227,583 patients, reduced to an analytical sample of 1,485,170 observations for 224,249 patients after accounting for missing values. Descriptive statistics are presented in Table 1 and Table 2 for the full set of observations and for the patients respectively and for those among the 5% and 1% with the highest costs each year. The average annual cost (in 2017 prices) across the full sample was 1,936€, rising to 15,867€ for those in the top 5% (who accounted for 41% of total costs) and to 32,562€ for those in top 1% (who accounted for 17% of total costs). Fig. 2 shows the skewed upper end of the cost distribution. 2.7% of the sample made the transition into the top 5% and 0.6% into the top 1% during the period.

The AMG distributions, by category and complexity level, are shown



(I) - Frequency: number of years



(II) - Persistence: consecutive years

Fig. 5. Frequency and persistence among the top 5% and top 1% (I) - Frequency: number of years (II) - Persistence: consecutive years.

Table 3
Results Annual cost, Top 5% and Top 1%.

	(1) GLM (AME) Annual cost	(2) GLM (AME) Annual cost	(3) Logit (AME) Top 5%	(4) Logit (AME) Top 1%	(5) Logit (AME) Trans. to top 5%	(6) Logit (AME) Trans. to top 1%
Female	−399	40**	−0.002***	−0.001***	−0.0003	−0.001***
	(16)	(16)	(0.001)	(0.0002)	(0.0003)	(0.0001)
Age - 65-74	503***	90***	−0.013***	−0.005***	−0.006***	−0.004***
	(17)	(25)	(0.001)	(0.001)	(0.001)	(0.0004)
Age - 75-84	1339***	175***	−0.022***	−0.009***	−0.009***	−0.006***
	(19)	(24)	(0.001)	(0.001)	(0.001)	(0.0004)
Age - 85-94	1708***	−45	−0.034***	−0.013***	−0.015***	−0.009***
	(21)	(30)	(0.001)	(0.001)	(0.001)	(0.0004)
Age≥95	1076***	−397***	−0.046***	−0.015***	−0.020***	−0.010***
	(60)	(72)	(0.002)	(0.001)	(0.001)	(0.001)
SES 18,000–100,000	−707***	−185***	−0.003***	0.0002	−0.001***	0.0003*
	(15)	(17)	(0.001)	(0.0003)	(0.0003)	(0.0002)
SES >100,000	−1201***	47	0.008*	0.007***	0.001	0.003**
	(63)	(219)	(0.004)	(0.003)	(0.002)	(0.001)
SES exempt	570***	118***	0.003***	0.0003	0.0004	5.18e-05
	(53)	(34)	(0.001)	(0.001)	(0.001)	(0.0003)
Died in January	−7.266	−1280***	−0.043***	−0.009***	−0.028***	−0.006***
	(65)	(28)	(0.001)	(0.0003)	(0.001)	(0.0002)
Died in February	1173***	−891***	−0.033***	−0.009***	−0.024***	−0.006***
	(82)	(46)	(0.001)	(0.0003)	(0.001)	(0.0003)
Died in March	2204***	−637***	−0.021***	−0.007***	−0.018***	−0.005***
	(106)	(35)	(0.001)	(0.0004)	(0.001)	(0.0004)
Died in April	3143***	−356***	−0.005**	−0.006***	−0.011***	−0.004***
	(141)	(61)	(0.002)	(0.001)	(0.001)	(0.0004)
Died in May	4128***	−227***	0.002	−0.004***	−0.005***	−0.004***
	(241)	(50)	(0.002)	(0.001)	(0.002)	(0.001)
Died in June	4965***	39	0.017***	−0.001	−0.001	−0.001
	(201)	(53)	(0.003)	(0.001)	(0.002)	(0.001)
Died in July	5044***	135**	0.019***	0.001	0.003*	0.0004
	(219)	(58)	(0.003)	(0.001)	(0.002)	(0.001)
Died in August	5783***	241***	0.019***	0.002**	0.005***	−0.001
	(218)	(60)	(0.003)	(0.001)	(0.002)	(0.001)
Died in September	6806***	447***	0.031***	0.007***	0.010***	0.004***
	(268)	(66)	(0.003)	(0.001)	(0.002)	(0.001)
Died in October	6985***	521***	0.035***	0.007***	0.014***	0.004***
	(249)	(68)	(0.003)	(0.001)	(0.002)	(0.001)
Died in November	7540***	644***	0.045***	0.008***	0.016***	0.006***
	(274)	(62)	(0.003)	(0.001)	(0.002)	(0.001)
Died in December	7415***	750***	0.048***	0.009***	0.025***	0.007***
	(244)	(66)	(0.003)	(0.001)	(0.002)	(0.001)
Acute disease c1		218***	0.0003	0.0001	0.0002	0.0002
		(7)	(0.0002)	(0.0001)	(0.0002)	(0.0001)
Acute disease c2		329***	0.001	0.0002	0.001*	6.00e-05
		(10)	(0.0003)	(0.0001)	(0.0003)	(9.78e-05)
Acute disease c3		459***	0.001	0.0004	0.001	0.0004
		(20)	(0.001)	(0.0003)	(0.0004)	(0.0003)
Acute disease c4		655***	0.0024**	0.001	0.002**	0.0003
		(23)	(0.001)	(0.0004)	(0.001)	(0.0003)
Acute disease c5		1270***	0.011***	0.002*	0.010***	0.001
		(57)	(0.003)	(0.001)	(0.003)	(0.0001)
Neoplasm c1		2107***	0.046***	0.010***	0.031***	0.007***
		(31)	(0.001)	(0.001)	(0.001)	(0.001)
Neoplasm c2		4878***	0.163***	0.041***	0.097***	0.027***
		(57)	(0.003)	(0.001)	(0.002)	(0.001)
Neoplasm c3		8225***	0.341***	0.094***	0.204***	0.063***
		(98)	(0.004)	(0.003)	(0.003)	(0.002)
Neoplasm c4		12,164***	0.528***	0.172***	0.311***	0.119***
		(145)	(0.005)	(0.004)	(0.005)	(0.004)
Neoplasm c5		19,157***	0.781***	0.343***	0.439***	0.245***
		(221)	(0.004)	(0.007)	(0.006)	(0.006)
1 LTC c1		115***	0.001***	9.89e-05	0.0004***	6.05e-05
		(6)	(0.0003)	(7.86e-05)	(0.0001)	(5.09e-05)
1 LTC c2		230***	0.0002	0.0001*	0.0003***	6.72e-05*
		(6)	(0.0001)	(7.30e-05)	(9.26e-05)	(3.97e-05)
1 LTC c3		427***	0.003***	0.001***	0.001***	0.0003***
		(9)	(0.0004)	(0.0002)	(0.0002)	(0.0001)
1 LTC c4		690***	0.009***	0.001***	0.002***	0.0003***
		(15)	(0.001)	(0.0002)	(0.0003)	(0.0001)
1 LTC c5		1305***	0.019***	0.002***	0.012***	0.001***
		(26)	(0.001)	(0.0004)	(0.001)	(0.0003)
2-3 LTC c1		309***	0.001***	0.0003***	0.001***	0.0001***
		(5)	(0.0002)	(9.60e-05)	(0.0001)	(4.99e-05)

(*continued on next page*)

Table 3 (*continued*)

	(1) GLM (AME) Annual cost	(2) GLM (AME) Annual cost	(3) Logit (AME) Top 5%	(4) Logit (AME) Top 1%	(5) Logit (AME) Trans. to top 5%	(6) Logit (AME) Trans. to top 1%
2-3 LTC c2		586***	0.004***	0.0003***	0.002***	0.0002***
		(6)	(0.0002)	(7.01e-05)	(0.0001)	(4.53e-05)
2-3 LTC c3		924***	0.009***	0.001***	0.004***	0.001***
		(9)	(0.0004)	(0.0001)	(0.0002)	(8.26e-05)
2-3 LTC c4		1324***	0.017***	0.002***	0.009***	0.001***
		(12)	(0.001)	(0.0002)	(0.0004)	(0.0001)
2-3 LTC c5		2690***	0.063***	0.009***	0.045***	0.006***
		(26)	(0.001)	(0.001)	(0.001)	(0.0004)
4+ LTC c1		1032***	0.012***	0.001***	0.005***	0.0004***
		(10)	(0.0004)	(0.0001)	(0.0002)	(6.51e-05)
4+ LTC c2		1981***	0.034***	0.004***	0.017***	0.002***
		(14)	(0.001)	(0.0002)	(0.0003)	(0.0001)
4+ LTC c3		3247***	0.078***	0.010***	0.042***	0.005***
		(24)	(0.001)	(0.001)	(0.001)	(0.0003)
4+ LTC c4		5128***	0.173***	0.029***	0.099***	0.017***
		(38)	(0.002)	(0.001)	(0.001)	(0.001)
4+ LTC c5		11,740***	0.513***	0.155***	0.297***	0.096***
		(106)	(0.003)	(0.003)	(0.003)	(0.002)
N	1,485,170	1,485,170	1,485,170	1,485,170	1,264,675	1,264,675
AIC or Pseudo R2	16.922	15.920	0.351	0.323	0.284	0.308
Region fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
SE cluster	patient-level	patient-level	patient-level	patient-level	patient-level	patient-level
Years	2011–2017	2011–2017	2011–2017	2011–2017	2012–2017	2012–2017

Notes: Significance levels: ***p < 0.01, **p < 0.05, *p < 0.1. SES: Socioeconomic status; c1-c5: complexity level; LTC: Long Term Condition. Average Marginal Effects (AME) reported. Reference category: male, age 55–64, SES <18,000, alive, healthy.

in Fig. 3. Those with neoplasms of higher complexity levels were more likely to appear among those with the highest costs. The same was true of people with 4+ LTCs.

The associations between age and the presence of AMGs are shown in Fig. 4. Higher proportions of relatively younger (55–75 years) and older (>95 years) people were allocated to the healthy group than were those aged 75–94 years. Those in the older age categories with 2–3 LTCs and with 4+ LTCs were more likely to have conditions of greater complexity.

Fig. 5-I shows the frequency that someone was among the 5% and 1% of those with the highest costs each year across the seven year period. 80.7% (95.5%) of the full sample never appear among the top 5% (1%) in any year, while 12.1% (3.5%) were in the top 5% (1%) for a single year, 4.0% (0.6%) appeared in two years, and the remaining 3.2% (0.5%) appeared in three or more years.

Fig. 5-II reports persistence among the top 5% and top 1%. 3.2% (0.5%) remained in the top 5% (1%) for two consecutive years, 1.0% (0.2%) for three consecutive years, and 0.3% (0.05%) were in the top 5% (1%) for the full seven years.

4.2. Factors influencing costs

Table 3 presents the results of estimating equation (1) for the three samples of observations. Columns (1) and (2) report the results of the analysis of annual costs c_{it} without and with the set of AMG variables. The inclusion of AMG dramatically attenuates the estimated effects of age, gender, SES and of dying, as shown by comparing the estimates from columns (1) and (2) depicted as grey and black dots respectively in Fig. 6-I.

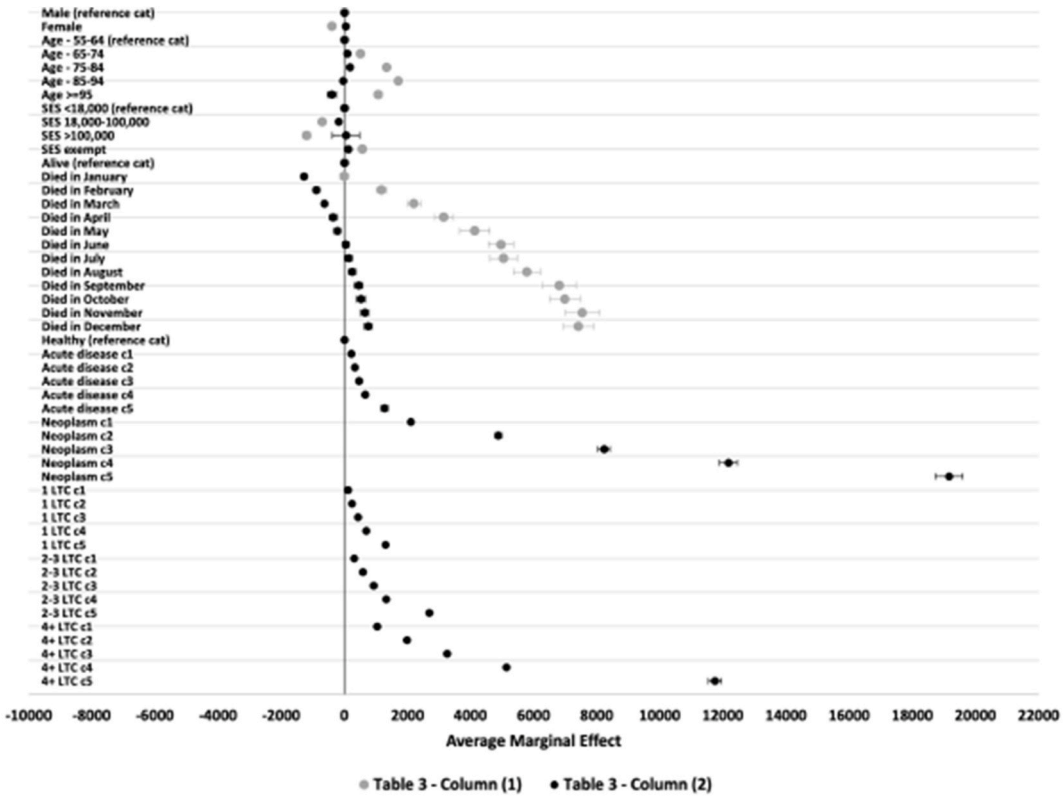
As shown in column (2) annual costs were 40€ higher for women than men, and those aged 95+ had lower costs (−397€) than those aged 55–64. Compared to those who didn't die, annual costs were −891€ lower for those who died in January, increasing monotonically month-by-month up to 644€ higher for those that died in December.

Fig. 6-I clearly indicates the dominant influence of neoplasms and 4+ LTCs on annual costs. Compared to people allocated to the healthy group, costs for someone with a neoplasm increased from 2107€ at complexity level 1 to 19,157€ at complexity level 5. For someone with 4+ LTCs, costs increased from 1032€ at complexity level 1 to 11,740€ at complexity level 5.

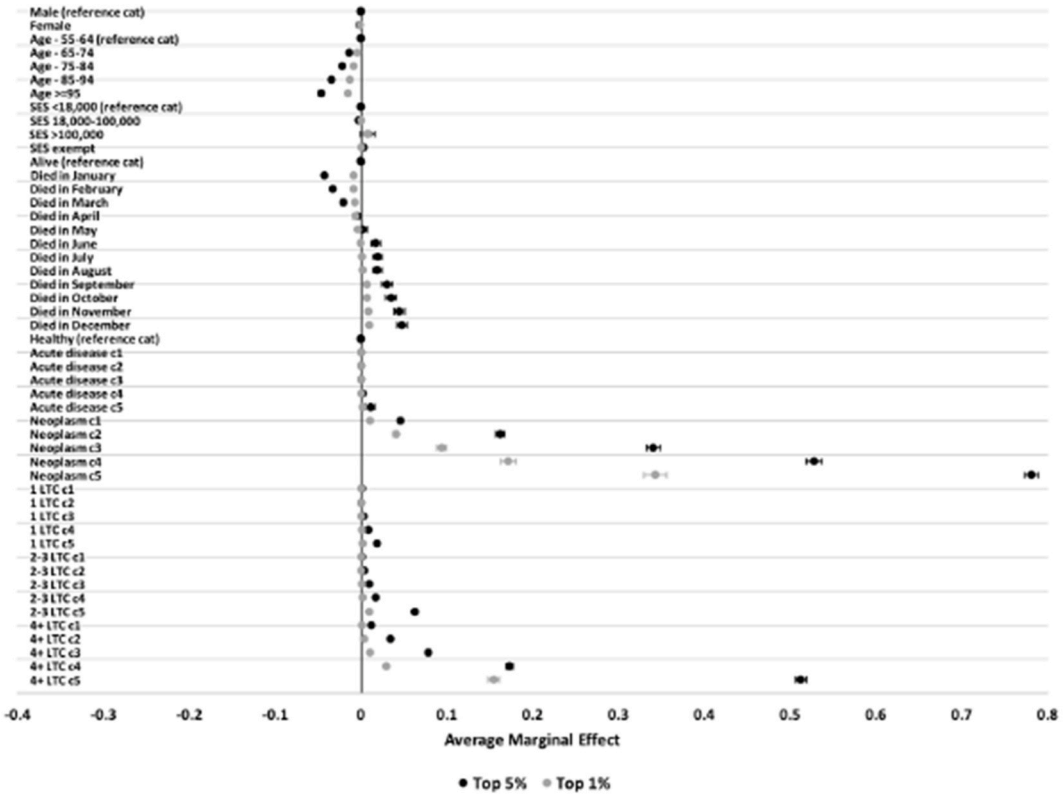
Columns (3) and (4) in Table 3 and Fig. 6-II report the factors associated with being among the most expensive 5% or 1% in any given year. The likelihoods were higher for males and for those with Low Income or Exempt from pharmaceutical co-payments status, and progressively lower for those in older age groups. Compared to those who were alive for the full data period, those who died during the first (latter) six months of any year were less (more) likely to be among the top 5% or 1% (effect sizes were very small for the top 1%).

All these effects, however, are dwarfed by those associated with the AMG variables. Most notably, those with complex neoplasms were far more likely than those without to be among the top 5% and top 1%. Those with 4+ LTCs of higher levels of complexity were also much more likely than those with no LTCs to be among the top 5% and top 1%.

Columns (5) and (6) in Table 3 and Fig. 6-III show the results from the logit models estimating the factors associated with transition into the high cost groups. These effects generally had a similar direction to those reported in Columns (3) and (4). Compared to survivors, those who died in the first (latter) five months of the year were less (more) likely to make the transition into the top 5% or 1%. Having a neoplasm



(I) - Annual Cost



(II) - Probability of being in the Top 5% and in the Top 1%

Fig. 6. Results: Annual Costs, Top 5% and Top 1% (I) - Annual Cost (II) - Probability of being in the Top 5% and in the Top 1% (III) - Probability of transition into the Top 5% and in the Top 1%.

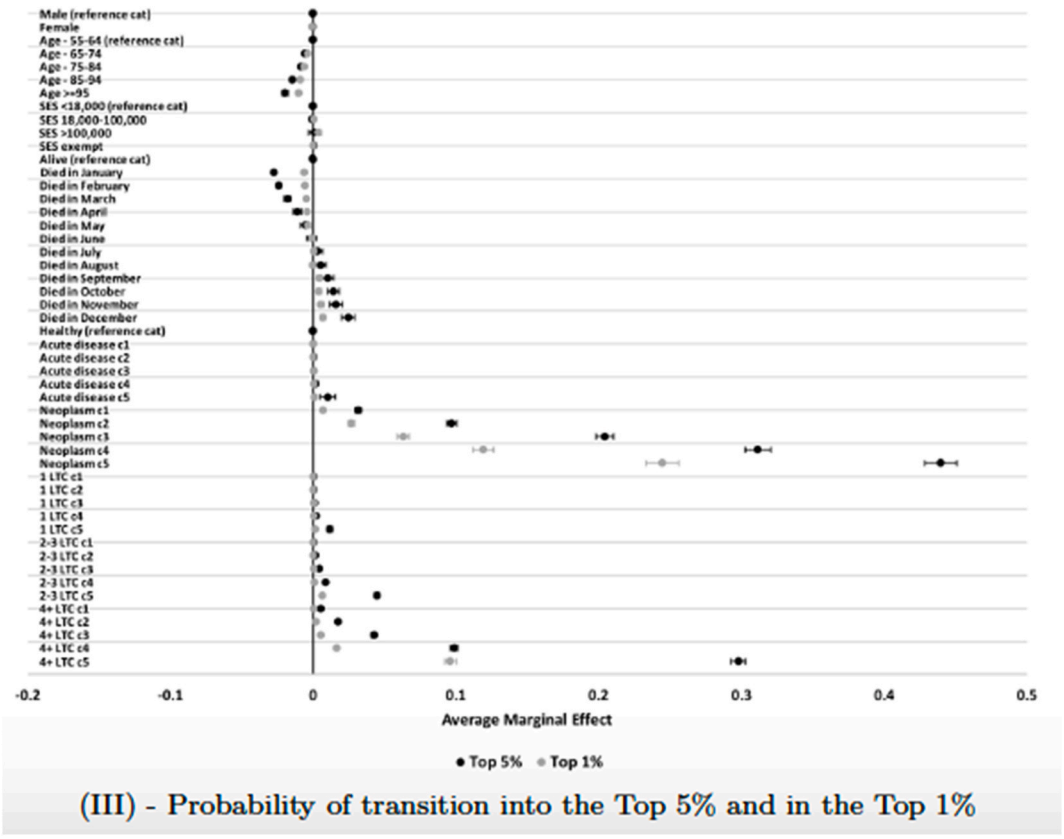


Fig. 6. (*continued*).

or 4+ LTCs were the most powerful predictors of transition, their importance rising in the level of complexity.

The results of estimating equation (2) are reported in Table 4 and Fig. 7. Columns (1) and (2) show that men, those in younger age groups, those with lower SES, those who died later in the period, those with complex neoplasms, and those with more LTCs of higher complexity were all associated with appearing more frequently among the high cost groups. These factors had a similar effect when considering persistence among those with high costs, as reported in columns (3) and (4) respectively.

5. Discussion

Our analyses demonstrate that morbidity status, not age or proximity to death, is the key determinant of annual healthcare costs. Other studies have found that, having accounted for morbidity status, the age effect collapses, sometimes to zero (Dormont et al., 2006; Moore et al., 2017; Howdon and Rice, 2018; Carreras et al., 2018). Our analyses not only indicate that older age does not drive higher costs but indeed suggests quite the reverse: when analysing annual costs for the full sample of those aged 55+, those in the oldest age group (>95) were estimated to have *lower* costs than those in the other age categories. This is consistent with the negative age effect found by Wammes et al. (2019) in their analysis of costs for those with chronic heart failure. There is an even more striking message when focusing on those with the HNHC group: there is a *negative* age gradient associated with being among, transitioning into and remaining in the top 5% and top 1%. The most likely explanation for this finding is that people that live longer are likely to be healthier, this probably being a factor in their longevity (Hazra et al., 2018). If studies of healthcare costs fail to account for morbidity status they will give a misleading impression of the impact of age on costs.

Consistent with the literature examining proximity to death on costs, we found associations between annual costs and whether and when people died. Those that died earlier (later) in the year had lower (higher) annual costs than those who survived, and were less (more) likely to be among those with the highest 5% and 1% of costs. But all of these effects, though generally statistically significant, were of small magnitude.

In contrast to age and proximity to death, morbidity status was found to be the key determinant both of annual costs and of membership among HNHC group. Of the set of AMGs, the dominant predictors of costs were whether people had a neoplasm or 4+ LTCs, with costs rising in relation to the complexity of these problems. Cancer stands out as a main driver of costs for two main reasons. First, cancer is the main mortality cause in Catalunya and generates a lot of related healthcare

Table 4
Frequency and persistence among the Top 5% and Top 1%.

	(1) OLS Frequency top 5%	(2) OLS Frequency top 1%	(3) OLS Persistence top 5%	(4) OLS Persistence top 1%
Female	−0.027***	−0.011***	−0.026***	−0.010***
	(0.003)	(0.002)	(0.003)	(0.0015)
Age - 65–74	−0.028***	−0.005***	−0.026***	−0.005***
	(0.004)	(0.002)	(0.004)	(0.002)
Age - 75–84	−0.063***	−0.027***	−0.063***	−0.0257***
	(0.005)	(0.002)	(0.004)	(0.002)
Age - 85–94	−0.212***	−0.091***	−0.197***	−0.085***
	(0.006)	(0.003)	(0.006)	(0.003)
Age≥95	−0.348***	−0.129***	−0.330***	−0.123***
	(0.016)	(0.006)	(0.015)	(0.006)
SES 18,000–100,000	−0.034***	−0.003**	−0.029***	−0.003*
	(0.003)	(0.002)	(0.003)	(0.002)
SES >100,000	−0.002	0.013**	−0.0002	0.012*
	(0.012)	(0.006)	(0.011)	(0.006)
SES exempt	0.050***	0.009*	0.040***	0.008*
	(0.010)	(0.005)	(0.009)	(0.005)
Died 2011	−0.249***	0.002	−0.161***	0.008
	(0.021)	(0.012)	(0.020)	(0.012)
Died 2012	−0.007	0.083***	0.091***	0.089***
	(0.022)	(0.013)	(0.020)	(0.013)
Died 2013	0.107***	0.081***	0.200***	0.086***
	(0.023)	(0.014)	(0.022)	(0.013)
Died 2014	0.041	0.082***	0.119***	0.083***
	(0.025)	(0.016)	(0.024)	(0.015)
Died 2015	0.088***	0.068***	0.140***	0.071***
	(0.026)	(0.015)	(0.024)	(0.014)
Died 2016	0.194***	0.089***	0.191***	0.0766***
	(0.026)	(0.016)	(0.024)	(0.015)
Died 2017	0.144***	0.090***	0.126***	0.074***
	(0.027)	(0.016)	(0.025)	(0.015)
Acute disease c1	−0.005	−0.003***	−0.002	−0.003***
	(0.010)	(0.001)	(0.010)	(0.001)
Acute disease c2	−0.008***	−0.002**	−0.008***	−0.002**
	(0.003)	(0.001)	(0.003)	(0.001)
Acute disease c3	−0.007*	−0.002**	−0.005	−0.002**
	(0.004)	(0.001)	(0.004)	(0.001)
Acute disease c4	−0.010*	−0.003**	−0.007	−0.003**
	(0.005)	(0.001)	(0.005)	(0.001)
Acute disease c5	0.002	−0.003*	0.002	−0.003*
	(0.012)	(0.002)	(0.011)	(0.002)
Neoplasm c1	0.226***	0.066***	0.215***	0.065***
	(0.011)	(0.006)	(0.010)	(0.006)
Neoplasm c2	0.501***	0.126***	0.463***	0.119***
	(0.011)	(0.006)	(0.010)	(0.006)
Neoplasm c3	0.855***	0.202***	0.769***	0.192***
	(0.014)	(0.008)	(0.013)	(0.008)
Neoplasm c4	1.275***	0.330***	1.118***	0.312***
	(0.018)	(0.011)	(0.016)	(0.010)
Neoplasm c5	1.935***	0.589***	1.638***	0.545***
	(0.022)	(0.014)	(0.020)	(0.013)
1 LTC c1	−0.013***	−0.004***	−0.010***	−0.004***
	(0.004)	(0.001)	(0.004)	(0.001)
1 LTC c2	−0.004	0.001	−0.003	0.001
	(0.003)	(0.002)	(0.002)	(0.002)
1 LTC c3	0.003	0.0003	0.004	0.0003
	(0.004)	(0.001)	(0.005)	(0.001)
1 LTC c4	0.027***	0.001	0.027***	0.001
	(0.007)	(0.001)	(0.007)	(0.001)
1 LTC c5	0.062***	0.009***	0.059***	0.008***
	(0.009)	(0.002)	(0.009)	(0.002)
2–3	−0.004	−0.001**	−0.002	−0.001*
LTC c1	(0.003)	(0.001)	(0.002)	(0.001)
2–3	0.015***	0.004***	0.017***	0.004***
LTC c2	(0.003)	(0.001)	(0.003)	(0.001)
2–3	0.041***	0.008***	0.041***	0.008***
LTC c3	(0.003)	(0.001)	(0.004)	(0.001)
2–3	0.061***	0.012***	0.059***	0.012***
LTC c4	(0.004)	(0.001)	(0.004)	(0.001)
2–3	0.192***	0.035***	0.180***	0.034***
LTC c5	(0.007)	(0.003)	(0.007)	(0.003)
4+	0.052***	0.009***	0.051***	0.008***
LTC c1	(0.003)	(0.001)	(0.003)	(0.001)

(*continued on next page*)

Table 4 (*continued*)

	(1) OLS Frequency top 5%	(2) OLS Frequency top 1%	(3) OLS Persistence top 5%	(4) OLS Persistence top 1%
4+	0.140***	0.024***	0.134***	0.023***
LTC c2	(0.003)	(0.001)	(0.003)	(0.001)
4+	0.304***	0.052***	0.285***	0.050***
LTC c3	(0.005)	(0.002)	(0.005)	(0.002)
4+	0.571***	0.100***	0.515***	0.095***
LTC c4	(0.007)	(0.003)	(0.007)	(0.003)
4+	1.422***	0.312***	1.195***	0.293***
LTC c5	(0.012)	(0.007)	(0.011)	(0.006)
N	224,249	224,249	224,249	224,249
R^2	0.291	0.108	0.270	0.106
Region fixed-effects	Yes	Yes	Yes	Yes
SE cluster	patient-level	patient-level	patient-level	patient-level
Years	2011–2017	2011–2017	2011–2017	2011–2017

Notes: Significance levels: ***p < 0.01, **p < 0.05, *p < 0.1. Average Marginal Effects (AME) reported. Reference category: male, age 55–64, SES <18,000, alive, healthy. The month when the patient died is also included in the models as dummy variables.

activity across all settings (Departament de Salut Generalitat de Catalunya, 2022). Second, expenditure on medicines accounts for 10% of the Catalan healthcare budget, and cancer drugs are among the most expensive of these medicines (Servei Català de la Salut Generalitat de Catalunya, 2022b).

The AMGs also proved powerful explanations of whether someone made the transition into and remained among those with the top 5% or top 1% of annual costs; and of the number of years spent among the top 5% or top 1%. Previous studies have attempted to develop general prediction models for those who might join the HNHC group (Chechulin et al., 2014; Rosella et al., 2018; Yang et al., 2018; Cohen et al., 2022) but the adoption of these models has been limited by "a very large number of predictor variables and the heavy data requirements to run the model" (Chechulin et al., 2014). The AMG system overcomes this complexity by reducing morbidity status to just 36 categories (including those for pregnancy and childbirth), thereby offering the potential to make prediction models more tractable.

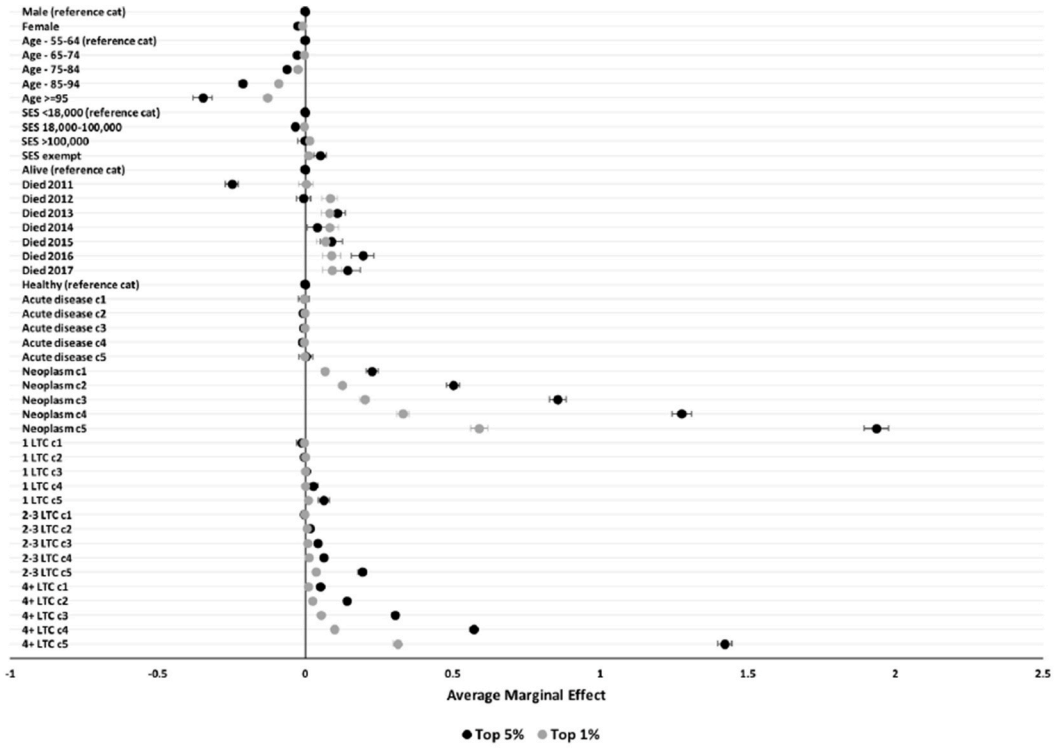
Our study demonstrates that valuable insights can be obtained by exploiting routine administrative data. A key attribute of the Catalan data is that morbidity markers are assigned according to the AMG classification system to all health care users each year (Monterde et al., 2016). The AMG system is not perfect, however, notably because a non-trivial proportion of people classified as "healthy" had positive health expenditure during the year. Indeed 0.08% of those in "healthy" AMG category appeared among those with the highest 5% of annual costs. This suggests that these people might not be completely healthy, but might have unrecorded health problems or that their problems are not used by the grouping algorithm to allocate them to other AMG categories. This merits further exploration.

There are drawbacks to using administrative data, of course, with three standing out for the dataset we employed. First, it is possible that the older that someone is, the more likely that health problems have been detected. If so, morbidity status would be more accurately coded as people grow older. This bias in coding accuracy might lead to upward (downward) bias of the estimated age effects for those in younger (older) age groups.

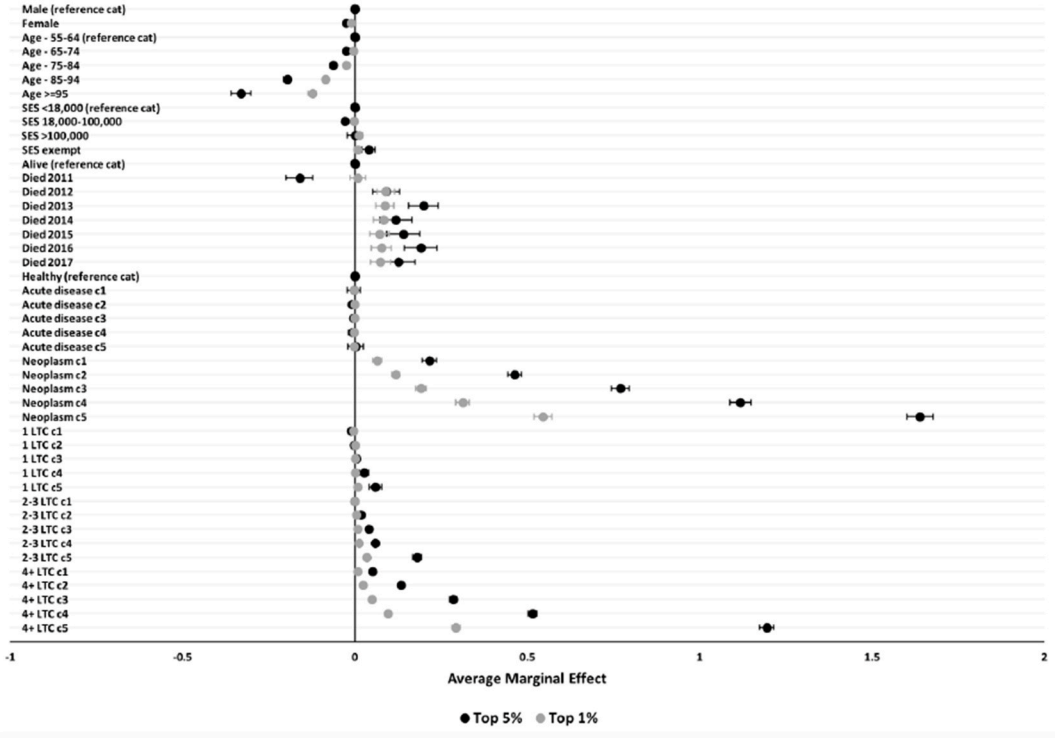
Second, we have had to rely on the income brackets used to determine pharmaceutical co-payments as our measure of SES. The low and medium SES categories are not finely differentiated, containing 62% and 34% of observations respectively. This imposes limits on our ability to assess the extent to which healthcare costs vary within these large SES groups. Nevertheless, it is reassuring that these SES groups explained very little of the variation in costs, as would be expected in a public health system where access is based on need rather than income.

Third, the data capture utilisation of publicly funded healthcare services only. In Catalunya, 28% of the population has additional private health insurance, giving them subsidised access to private healthcare. Information about the use of private healthcare is not collected by the government, so we were unable to analyse the totality of healthcare utilisation or total costs. It is worth emphasising also that, while we are able to analyse the characteristics of those used healthcare, we are not able to ascertain the extent to which people were under-served, and hence not receiving healthcare in accordance with their need. Compared to countries that lack universal health coverage, however, this may be less of a problem in Spain where there is a national health system free at the point of use. Nevertheless access to services may still be discriminatory. For instance, gender inequalities have been shown to be present in the use of cancer services (Cheung et al., 2020; Gill et al., 2019) and to persist during the end-of-life trajectory (Fernandez et al., 1999). As Dalmau-Bueno et al. (Dalmau-Bueno et al., 2021) point out "the mechanisms underlying gender inequalities are complex and may overlap with other factors such as longevity (women in our population tended to be older), socioeconomic status (women tended to receive lower annual incomes), and social environment (living alone was twice more frequent among women than men), among others." To address discrimination, the Catalan government is incorporating an explicit gender perspective in all phases of healthcare planning (WHO, 2020).

In all developed countries, people are living longer but in declining health. This study has demonstrated that age and proximity to death are poor predictors of higher healthcare costs, which are explained mainly by morbidity status, particularly whether someone has neoplasms or multiple LTCs. These conditions are also the key reasons why people appear among or are likely to transition into the group with the highest costs. This makes intuitive sense: people don't use heath care because



(I) - Frequency: number of years in the Top 5% and Top 1%



(II) - Persistence: consecutive years in the Top 5% and Top 1%

Fig. 7. Results: Frequency and persistence in the Top 5% and Top 1% (I) - Frequency: number of years in the Top 5% and Top 1% (II) - Persistence: consecutive years in the Top 5% and Top 1%.

they are old or approaching death but because they are unwell. Identifying each individual's morbidity status is key to ensuring that they can receive appropriate care at the right time and in the right settings. Such information would also enhance health expenditure forecasting models that typically rely only on age and gender structure, and occasionally proximity-to-death, when describing demographic pressures (Marino et al., 2017). The AMG system looks like a valuable way to characterise morbidity status that can be used in future studies of healthcare costs and membership among the HNHC group.

Funding

LM is funded by the Spanish Ministry of Science, Innovation and Universities (PID2019104319RB-I00). AS received no funding for this project. AGA was employed at AQuAS when the project was underway.

Ethics approval

Ethics Approval granted by the Institutional Commitee for Ethics Review of Projects (CIREP) at Universitat Pompeu Fabra, number 180 on 9th of February 2021.

Credit author statement

LM: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing – review & editing, Visualization. AS: Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, Visualization, Supervision, Project administration. AGA: Conceptualization, Writing – review & editing.

Declaration of competing interest

LM and AS have nothing to disclose. AGA was employed at AQuAS when the project was underway.

Data availability

The authors do not have permission to share data.

Acknowledgements

We thanks researchers from Agència de Qualitat i Avaluació Sanitàries de Catalunya (AQuAS) and, in particular, Elisenda Martínez Carbonell. This study used anonymised data provided by AQuAS, within the Public Data Analysis for Health Research and Innovation Program (PADRIS). We thank participants at the AES and EuHEA conferences and at the AQUAS and Department of Health of Catalonia internal seminars in 2022 for their comments, in particular, the valuable suggestions from Ed Kendall.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.socscimed.2023.115955.

References

Alemayehu, B., Warner, K., 2004. Ageing of population and health care expenditure: a red herring? Health Econ. 39 (3), 627–642.
Anderson, G., Knickman, J.R., 1984. Patterns of expenditures among high utilizers of medical care services: the experience of Medicare beneficiaries from 1974 to 1977. Med Care. 22 (2), 143–149.
Atella, V., Conti, V., 2014. The effect of age and time to death on primary care costs: the Italian experience. Soc. Sci. Med. 114, 10–17.
Bilger, M., Chaze, J., 2008. What drives individual health expenditure in Switzerland? Swiss Journal of Economics and Statistics 144, 337–358.
Blumenthal, D., Chernof, B., Fulmer, T., Lumpkin, J., Selberg, J., 2016. Caring for high-need, high- cost patients - an urgent priority. NEJM 375 (10), 909–911.
Brilleman, S.L., Gravelle, H., Hollinghurst, S., Purdy, S., Salisbury, C., Windmeijer, F., 2014. Keep it simple? Predicting primary health care costs with clinical morbidity measures. J. Health Econ. 35 (100), 109–122.
Caballer-Tarazona, V., Guadalajara-Olmeda, N., Vivas-Consuelo, D., 2019. Predicting healthcare expenditure by multimorbidity groups. Health Pol. 123 (4), 427–434.
Carreras, M., Ibern, P., Inoriza, J., 2018. Ageing and healthcare expenditures: exploring the role of individual health status. Health Econ. 27 (5), 865–876.
Carrilero, N., Dalmau-Bueno, A., García-Altés, A., 2021. Socioeconomic inequalities in 29 child- hood diseases: evidence from a 1,500,000 children population retrospective study. BMC Publ. Health 21, 1150.
Chechulin, Y., Nazerian, A., Rais, S., Malikov, K., 2014. Predicting patients with high risk of becoming high-cost healthcare users in Ontario (Canada). Healthc Policy 9 (3), 68–79.
Cheung, M., Croxford, R., Earle, C., Singh, S., 2020. Days spent at home in the last 6 months of life: a quality indicator of end of life care in patients with hematologic malignancies. Leuk. Lymphoma 61 (1), 146–155.
Cleries, M., Monterde, D., Vela, E., Guarga, À., García-Eroles, L., Pérez-Sust, P., 2020. Grupo de validación. Validación clínica de 2 agrupadores de morbilidad en el ámbito de atención primaria. Atención Primaria 52 (2), 96–103.
Cohen, A.G., Vinker, S., Isaacson, A., Avramovich, E., Merzon, E., 2022. A Practical Model for Early Identification of Prospective High Need High Cost Patients. medRxiv, 2022.02.04.22270056. Available from: https://www.medrxiv.org/content/medrxiv/early/2022/02/06/2022.02.04.22270056.full.pdf, 2022.02.04.22270056. Available from:
Colombier, C., Weber, W., 2011. Projecting health-care expenditure for Switzerland: further evi- dence against the 'red-herring' hypothesis. Int. J. Health Plann. Manag. 26 (3), 246–263.
Dalmau-Bueno, A., García-Altés, A, Vela, E, Cleries, M., Perez, C, Argimon, J., 2021. Frequency of health-care service use and severity of illness in undocumented migrants in Catalonia, Spain: a population-based, cross-sectional study. Lancet Planet Health 5 (5), e286-e296.
de Meijer, C., Koopmanschap, M., Bago d' Uva, T., van Doorslaer, E., 2011. Determinants of long- term care spending: age, time to death or disability? J. Health Econ. 30 (2), 425–438.
de Meijer, C., Wouterse, B., Polder, J., Koopmanschap, M., 2013. The effect of population aging on health expenditure growth: a critical review. Eur. J. Ageing 10 (4), 353–361.
Departament de Salut Generalitat de Catalunya, 2022. Avanç de resultats de les causes de mort a Catalunya,2021. Available from: https://scientiasalut.gencat.cat/bitstream/handle/11351/9245/avanc_resultats_causes_mort_catalunya_2021_%20resultats_provisionals_2023.pdf?sequence=1&isAllowed=y.
Dormont, B., Grignon, M., Huber, H., 2006. Health expenditure growth: reassessing the threat of ageing. Health Econ. 15 (9), 947–963.
ESCA, 2017. Enquesta de salut de Catalunya. L'estat de salut, els comporta- ments relacionats amb la salut i l'ús de serveis sanitaris a Catalunya. Resultats principals de l'ESCA 2017. Resum executiu. Available from: https://salutweb.gencat.cat/web/.content/_departament/estadistiques-sanitaries/enquestes/Enquesta-de-salut-de-Catalunya/Resultats-de-lenquesta-de-salut-de-Catalunya/documents/resum-executiu_%20esca_2017.pdf
Felder, S., Werblow, A., Zweifel, P., 2010. Do red herrings swim in circles? Controlling for the endogeneity of time to death. J. Health Econ. 29 (2), 205–212.
Fernandez, E., Schiaffino, A., Rajmil, L., Badia, X., Segura, A., 1999. Gender inequalities in health and health care services use in Catalonia (Spain). J. Epidemiol. Community Health 53 (4), 218–222.
Figueroa, J.F., Horneffer, K.E., Riley, K., Abiona, O., Arvin, M., Atsma, F., et al., 2021. A methodol- ogy for identifying high-need, high-cost patient personas for international comparisons. Health Serv. Res. 56 (S3), 1302–1316.
Fundación IDIS, 2022. Informe Sanidad Privada Aportando Valor 2022. Informe 12. Available from: https://www.fundacionidis.com/uploads/informes/informe_sanidad_privada_aportando_valor_2022.pdf.
García-Altés, A., Ruiz-Muñoz, D., Colls, C., Mias, M., Martin-Bassols, N., 2018. Socioeconomic inequalities in health and the use of healthcare services in Catalonia: analysis of the individual data of 7.5 million residents. J. Epidemiol. Community Health 72 (10), 871–879.
Geue, C., Briggs, A., Lewsey, J., Lorgelly, P., 2014. Population ageing and healthcare expendi- ture projections: new evidence from a time to death approach. Eur. J. Health Econ. 15 (8), 885–896.
Gill, T., Gahbauer, E., Leo-Summers, L., Murphy, T., Han, L., 2019. Days spent at home in the last six months of life among community-living older persons. Am. J. Med. 132 (2), 234–239.
Hazra, N., Rudisill, C., Gulliford, M., 2018. Determinants of health care costs in the senior elderly: age, comorbidity, impairment, or proximity to death? Eur. J. Health Econ. 19 (6), 831–842.
Howdon, D., Rice, N., 2018. Health care expenditures, age, proximity to death and morbidity: implications for an ageing population. J. Health Econ. 57, 60–74.
Hwang, W., LaClair, M., Camacho, F., Paz, H., 2015. Persistent high utilization in a privately insured population. Am. J. Manag. Care 21 (4), 309–316.
Lehnert, T., Heider, D., Leicht, H., Heinrich, S., Corrieri, S., Luppa, M., et al., 2011. Review: health care utilization and costs of elderly persons with multiple chronic conditions. Med. Care Res. Rev. 68 (4), 387–420.
Longden, T., Wongb, C.Y., Haywooda, P., Halla, J., van Goola, K., 2018. The prevalence of persistence and related health status: an analysis of persistently high healthcare costs in the short term and medium term. Soc. Sci. Med. 211, 147–156.
Luta, X., Diernberger, K., Bowden, J., Droney, J., Howdon, D., Schmidlin, K., et al., 2020. Healthcare Trajectories and Costs in the Last Year of Life: a Retrospective Primary Care and Hospital Analysis. BMJ Supportive & Palliative Care.

Marino, A., Morgan, D., Lorenzoni, L., James, C., 2017. Future trends in health care expenditure: a modelling framework for cross-country forecasts. OECD Health Working Papers 95.
Monheit, A.C., 2003. Persistence in Health Expenditures in the Short Run: Prevalence and Consequences. Med. Care 41 (7), III53-III64..
Monterde, D., Vela, E., Clèries, M., 2016. Los grupos de morbilidad ajustados: nuevo agrupador de morbilidad poblacional de utilidad en el ámbito de la atención primaria. Atención Primaria 48 (10), 674–682.
Monterde, D., Vela, E., Cleries, M., Garcia-Eroles, L., Roca, J., Pérez-Sust, P., 2020. Multimorbidity as a predictor of health service utilization in primary care: a registry-based study of the Catalan population. BMC Fam. Pract. 21 (1), 39.
Moore, P., Bennett, K., Normand, C., 2017. Counting the time lived, the time left or illness? Age, proximity to death, morbidity and prescribing expenditures. Soc. Sci. Med. 184, 1–14.
Murphy, M., Martikainen, P., 2013. Use of hospital and long-term institutional care services in relation to proximity to death among older people in Finland. Soc. Sci. Med. 88, 39–47.
Norton, E.C., Dowd, B.E., 2018. Log odds and the interpretation of logit models. Health Serv. Res. 53 (2), 859–878.
OECD/European Observatory on Health Systems and Policies, 2021. State of Health in the EU. Spain: Country Health Profile 2021. Available from: https://eurohealthobservatory.who.int/publications/m/spain-country-health-profile-2021.
Polder, J., Barendregt, J., van Oers, H., 2006. Health care costs in the last year of life—the Dutch experience. Soc. Sci. Med. 63 (7), 1720–1731.
Ronksley, P.E., McKay, J.A., Kobewka, D.M., Mulpuru, S., Forster, A.J., 2015. Patterns of health care use in a high-cost inpatient population in Ottawa, Ontario: a retrospective observational study. CMAJ Open 3 (1), E111–E118.
Rosella, L.C., Kornas, K., Yao, Z., Manuel, D.G., Bornbaum, C., Fransoo, R., et al., 2018. Predicting high health care resource utilization in a single-payer public health care system: development and validation of the high resource user population risk tool. Med. Care 56 (10), e61–e69.
Salas, C., Raftery, JP., 2001. Econometric issues in testing the age neutrality of health care expenditure. Health Econ 10 (7), 669–671.
Servei Català de la Salut Generalitat de Catalunya, 2022a. El model de copagament farma- cèutic. Available from: https://catsalut.gencat.cat/ca/serveis-sanitaris/atencio-farmacèutica/financament-public-medicaments/model-copagament/.
Servei Català de la Salut Generalitat de Catalunya, 2022b. Memòria 2021. Servei Català de la Salut. Available from: https://memoria.catsalut.gencat.cat/wp-content/uploads/2022/10/Memoria2021_CatSalut.pdf.
Seshamani, M., Gray, A., 2004. A longitudinal study of the effects of age and time to death on hospital costs. J. Health Econ. 23 (2), 217–235.
Soley-Bori, M., Ashworth, M., Bisquera, A., Dodhia, H., Lynch, R., Wang, Y., et al., 2021. Impact of multimorbidity on healthcare costs and utilisation: a systematic review of the UK literature. Br. J. Gen. Pract. 71 (702), e39–e46.
Vela, E., Cleries, M., Vella, V., Adroher, C., García-Altés, A., 2017. Análisis poblacional del gasto en servicios sanitarios en Cataluña (España): ¿qué y quién consume más recursos? Gac. Sanit. 33 (1), 24–31.
Vela, E., Tenyi, A., Cano, I., Monterde, D., Cleries, M., García-Altés, A., et al., 2018. Population- based analysis of patients with COPD in Catalonia: a cohort study with implications for clinical management. BMJ Open 8, e017283.
Vela, E., Cleries, M., Monterde, D., Carot-Sans, G., Coca, M., Valero-Bover, D., et al., 2021. Per- formance of quantitative measures of multimorbidity: a population-based retrospective analysis. BMC Publ. Health 21 (1), 1881.
Violán, C., Foguet-Boreu, Q., Roso-Llorach, A., Rodriguez-Blanco, T., Pons-Vigués, M., Pujol- Ribera, E., et al., 2014. Burden of multimorbidity, socioeconomic status and use of health services across stages of life in urban areas: a cross-sectional study. BMC Publ. Health 14 (530).
Wammes, J., Van der Wees, P., Tanke, M., Gert, P., Westert, G., Jeurissen, P., 2018. Systematic review of high-cost patients' characteristics and healthcare utilisation. BMJ Open 8, e023113.
Wammes, J.J.G., Auener, S., van der Wees, P.J., Tanke, M.A.C., Bellersen, L., Westert, G. P., et al., 2019. Characteristics and health care utilization among patients with chronic heart failure: a longitudinal claim database analysis. ESC Heart Fail 6 (6), 1243–1251.
Werblow, A., Felder, S., Zweifel, P., 2008. Population ageing and health care expenditure: a school of 'red herrings. Health Econ. 16 (10), 1109–1126.
WHO, 2019. International Statistical Classification of Diseases and Related Health Problems 10th Revision. Available from: https://icd.who.int/browse10/2019/en#/.
WHO, 2020. Thirty-year Retrospective of Catalan Health Planning. Available from: https://www.euro.who.int/en/health-topics/Health-systems/health-systems-financing/publications/2020/thirty-year-retrospective-of-catalan-health-planning-2020. Available from:
Yang, C., Delcher, C., Shenkman, E., Ranka, S., 2018. Machine learning approaches for predicting high cost high need patient expenditures in health care. Biomed. Eng. Online 17 (1), 131.
Zweifel, P., Felder, S., Meiers, M., 1999. Ageing of population and health care expenditure: a red herring? Health Econ. 8 (6), 485–496.