Factors associated with hospitalisations in chronic conditions deemed avoidable: ecological study in the Spanish healthcare system

Ester Angulo-Pueyo,1,2 Manuel Ridao-López,1,2 Natalia Martínez-Lizaga,1,2 Sandra García-Arnesto,1,2 Salvador Peiró,3 Enrique Bernal-Delgado1,2

ABSTRACT

Objectives: Potentially avoidable hospitalisations have been used as a proxy for primary care quality. We aimed to analyse the ecological association between contextual and systemic factors featured in the Spanish healthcare system and the variation in potentially avoidable hospitalisations for a number of chronic conditions.

Methods: A cross-section ecological study based on the linkage of administrative data sources from virtually all healthcare areas (n=202) and autonomous communities (n=16) composing the Spanish National Health System was performed. Potentially avoidable hospitalisations in chronic conditions were defined using the Spanish validation of the Agency for Health Research and Quality (AHRQ) preventable quality indicators. Using 2012 data, the ecological association between potentially avoidable hospitalisations and factors featuring healthcare areas and autonomous communities was tested using multilevel negative binomial regression.

Results: In 2012, 151,468 admissions were flagged as potentially avoidable in Spain. After adjusting for differences in age, sex and burden of disease, the only variable associated with the outcome was hospitalisation intensity for any cause in previous years (incidence risk ratio 1.19 (95% CI 1.13 to 1.26)). The autonomous community of residence explained a negligible part of the residual unexplained variation (variance 0.01 (SE 0.008)). Primary care supply and activity did not show any association.

Conclusions: The findings suggest that the variation in potentially avoidable hospitalisations in chronic conditions at the healthcare area level is a reflection of how intensively hospitals are used in a healthcare area for any cause, rather than of primary care characteristics. Whether other non-studied features at the healthcare area level or primary care level could explain the observed variation remains uncertain.

INTRODUCTION

The Spanish healthcare system shows low levels of potentially avoidable hospitalisations (PAH) in chronic care conditions compared with other countries with similar institutional features.1 However, the variation within the country is vast, whether looking at composite measures or focusing on specific conditions.2 3

Generally, PAH have been used as a proxy for quality of ambulatory care (primary care, in the Spanish context).4 5 Therefore, monitoring PAH variation and trends over time could be a powerful tool to improve service performance. Indeed, the Spanish National Health Service (SNHS) has adopted PAH to evaluate the effectiveness of the National Strategy for Chronic Patients.6

However, numerous studies have shown that, rather than being associated with quality, PAH might be associated with other factors, only some of which are related to ambulatory care. Geographical barriers or effective access to healthcare facilities,7 8 care continuity across levels9 and availability of primary care professionals are examples of ambulatory care-related factors.10–12

Strengths and limitations of this study

The study builds on virtually all hospital admissions discharged in 2012 and analyses nearly all healthcare areas and autonomous communities composing the Spanish National Health System.

The random effect multilevel modelling approach enables a better understanding of the relative influence of the relevant decision units composing the healthcare system in Spain (healthcare areas and autonomous communities).

Beyond the intrinsic limitations of a cross-section ecological study, some predictors found relevant in the literature were not explored, particularly at the primary care level, and the number of healthcare areas (n=202) and autonomous communities (n=16) in the study might be insufficient to find ecological associations if they exist.
care-related factors with which PAH have been associated are supply of acute beds,13 existence and type of long-term care services14–16 and differences in insurance schemes.17–18 Socioeconomic gradient has also been observed to have an influence beyond epidemiological differences between geographical areas.5 19–23

In Spain, some studies have partially addressed the question of the factors that explain PAH variation beyond differences in age and sex. They have mainly focused on the impact of socioeconomic status of the population, distance to a hospital and primary care supply.5 24–27 However, the findings were not always consistent, each study focused on a specific autonomous community (AC), the units of analysis were different across studies and the topic was broader than chronic conditions. The only nationwide study, carried out in the context of an international comparison initiative, barely explored contextual or systemic factors.1

The question as to whether PAH might be useful as a proxy for primary care performance is still current in Spain. In this paper we have explored the association of a broad number of contextual and systemic factors, using nearly all potentially avoidable hospitalizations in chronic conditions produced in a year in the SNHS.

**METHODS**

**Study design**

An ecological study of administrative data was performed, analysing the association between PAH for chronic conditions occurring in people aged ≥40 years and systemic and contextual factors that characterise healthcare areas (HCAs) in Spain.

**Population and setting**

The SNHS, a quasi-federal decentralised system,8 is composed of 17 ACs with full responsibility for policy making, planning and financing at the regional level; in turn, each region is administratively organised into 203 HCAs, the locus for hospital and primary care provision. This hierarchy (ie, HCAs within ACs) implies that health outcomes or performance measures might be influenced by factors affecting both loci (see general characteristics of ACs and HCAs in Spain in online supplementary material appendix 1). Given that multilevel analysis requires the existence of several groups within the upper level of analysis, we excluded one AC composed of a single HCA. Consequently, in this study we analysed virtually all PAH discharges produced in 202 HCAs and 16 ACs in 2012. Data correspond to all publicly funded hospital activity, discharged from either private or public hospitals.

**Variables**

The outcome variable was the number of PAH observed in each HCA. PAH were defined as those unplanned admissions with a primary diagnosis of congestive heart failure (CHF), chronic obstructive pulmonary disease (COPD), adult asthma, second episode of angina, diabetic coma or dehydration in patients aged ≥40 years. The definitions used in this study stem from the validation of the Agency for Health Research and Quality (AHRQ) preventable quality indicators affecting chronic conditions28 in the Spanish context29 (see ICD-9-CM codes in online supplementary material appendix 2).

The independent variables included in this study were as follows:

1. Age, defined as the ratio between the population aged 65–79 years and the population aged ≥80 years in each HCA; the ratio would capture the exponential effect of age in the appearance of chronic conditions.
2. Sex, as the percentage of men in each HCA.
3. Burden of disease, defined as the cumulative total of hospitalisations for hip fracture, acute myocardial infarction, ischaemic stroke and cancer of the colon, lung or breast treated surgically occurring in the population aged ≥40 years per 10 000 inhabitants aged ≥40 years; these hospitalisations very likely reflect differences in health population across HCAs and not differences in supply-side factors.30
4. Mix of PAH conditions: as PAHs are a composite of six conditions with a potentially different basal risk of hospitalisation, this variable represented the mixture of conditions as the proportion of all PAH admissions which were for the two most common conditions (COPD and CHF).
5. Intensity of hospitalisation, defined as the rate of discharges for any cause (except avoidable hospitalisations) in patients aged ≥40 years per 10 000 inhabitants aged ≥40 years in the 3 years prior to the year of analysis (2009, 2010 and 2011).
6. Distance to a hospital, as the percentage of people living less than 30 min distance time to the closest hospital.
7. Hospital supply factor, essentially characterised by bed supply, non-surgical admissions, and physician and nurse hospital workforce.
8. Socioeconomic status factor, represented by unemployment level.
9. Social care supply factor, which mainly clustered beds in social care institutions, social workers and physiotherapists.
10. Primary care factor, mainly gathering the total number of primary care centres, with or without emergency wards, and general practitioner and nurse consultations. As the literature has proposed primary care as a main driver in the reduction of PAH, this factor was considered as the main independent factor in the models.
11. Specialised ambulatory care factor, mainly represented by cardiology and ophthalmology consultations in specialised ambulatory centres.
12. The AC of residence, considered as a random effect to explain unobservable variables that could homogeneously affect HCAs within an AC.
Variables 7–11 were extracted from an ad hoc factor analysis. The Madrid AC was excluded in this specific analysis as it lacks primary care data disaggregated at the HCA level. Online supplementary appendix 3 describes the list of independent variables used in this study as well as the value distribution across HCAs. Online supplementary appendix 4 describes the factor analysis and results.

Sources of data and linkage

All data used in this study were extracted from the linkage of secondary sources. PAH discharges were obtained from routinely collected administrative data integrated, harmonised, linked and exploited in the context of the Atlas VPM Project (http://www.atlasvpm.org), a research initiative that has systematically studied unwarranted variations in healthcare performance in Spain since 2002. This dataset is constructed on the 17 Minimum Basic Hospital Datasets cleaned and consolidated in the AC, and collects individual information from virtually all publicly funded episodes produced in Spanish hospitals since 2002 (roughly 5 million per year). This study used PAH, age groups, sex, hospitalisation intensity and burden of disease retrieved from this data source.

Atlas VPM also collates various other data from sources maintained by the Spanish National Statistics Office or Health Authorities on their behalf, as well as from La Caixa Foundation, a private initiative that regularly collects socioeconomic data from different official sources. In this study, population counts stratified by sex and age were extracted from the 2012 Spanish National Institute of Statistics’ Municipal Register of Inhabitants (ie, annual update of the National Census). Hospital supply data were obtained from the 2012 Annual Hospital Survey, primary care activity was extracted from the 2012 Primary Care Information System, primary care supply was retrieved from the 2012 Primary Care Centres catalogue and supply of social care institutions was acquired from the 2009 Social Care Survey, the four datasets are held by the Ministry of Health and Social Policy. Socioeconomic status was retrieved from the 2013 Economic Yearbook for Spain, published by La Caixa Foundation. Finally, distance to hospital was extracted from one of the datasets developed by Atlas VPM as part of their master.

With regard to linkage, each PAH episode was deterministically allocated to the place of residence of the patient, making it possible to aggregate the episodes produced in a particular geographical site at the HCA and AC. Given that the Minimum Basic Hospital Dataset covers virtually all hospitalisations produced in a year, the allocation of patients to HCA was successful in more than 98% in 2012, and that the remaining sources contain information for 100% of the HCA level. The deterministic linkage across datasets is complete except for the primary care factor in the region of Madrid (n=11 HCAs).

Statistical analysis

The statistical analysis sought to examine the association between counts of PAH admissions in each HCA and variables and composite factors extracted from the factor analysis. To decide which variables and factors should be modelled in the regression, a bivariate analysis was performed using type 1 error <10% as the entrance threshold. Four models were then fitted using a population offset to adjust for differences in the density of the population aged ≥40 years.

Because composite factors are normalised (Kaiser method), single variables were also normalised to avoid scale effects in the estimation of coefficients.

We used negative binomial regressions (given that the dependent variable exhibited overdispersion) with clustered robust estimation for SEs. The magnitude of the association was assessed in terms of the incidence rate ratio (IRR) and its 95% CIs. For the particular case of model 4, a two-level mixed effects negative binomial regression was fitted, specifying AC as a random effect. The goodness of fit of the models was estimated by the likelihood ratio test. We used the variance of the intercept and its CI as the indicator for residual second-level variation. We then explored whether the HCA level variables retained in the model had heterogeneous behaviour across ACs (ie, multiplicative effect), including them as random slopes in the second level. All the analyses were performed with Stata V.13.

Ethics statement

This study, observational in design, used retrospective anonymised non-identifiable and non-traceable data and was conducted in accordance with the amended Helsinki Declaration, the International Guidelines for Ethical Review of Epidemiological Studies and Spanish laws on data protection and patients’ rights. This study implies the use of anonymised individual data with double dissociation, in the original source and once stored in the database, which impedes patient traceability. Spanish legislation does not require informed consent or the approval of an Ethics Committee in observational studies, where data come from secondary sources and datasets do not include personal information as a consequence of the de-anonymisation and pseudo-anonymisation procedures over the original raw information.

RESULTS

In 2012, 151 468 admissions for chronic conditions in patients aged ≥40 years were flagged as potentially avoidable. Of these, the reason for admission was CHF in 56 092 patients, COPD in 64 383 patients, adult asthma in 56 78 cases, angina in 16 142 patients, diabetic coma in 2587 patients and dehydration in 6596 patients. COPD and CHF represented 79% of the cases. Across areas, the proportion of the two conditions ranged from 47% to 91%.
With regard to the distribution of PAH across the 202 areas considered in the analyses, the variation was fairly high with a coefficient of variation (ie, ratio of SD to mean) of 33.1% and a 3.4-fold difference between the areas in the 95th and 5th percentiles of the distribution of PAH rates (the map in online supplementary appendix 1 provides additional insight into the distribution of PAH rates across HCAs and ACs). Once PAH were adjusted for age and sex, only four factors showed an association with PAH: burden of disease (IRR 1.12 (95% CI 1.05 to 1.20)), intensity of hospitalisation over the previous 5 years (IRR 1.17 (95% CI 1.08 to 1.27)), distance to a hospital (IRR 1.08 (95% CI 1.03 to 1.14)) and primary care supply (IRR 0.94 (95% CI 0.88 to 1)). The remaining factors did not show a statistical association with PAH (table 1).

Four multivariate models were fitted and the results are presented in table 2. Model 1, the basic model, included the main confounders (age, sex and burden of disease) in the association between HCA and AC factors and PAH. Those areas with relatively younger people were less likely to exhibit higher PAH (IRR 0.84); in contrast, areas with a higher burden of disease had higher PAH (IRR 1.12). In model 2 the primary care supply factor was added, hypothesising that, beyond differences in population features, the higher the primary care supply (resources and activity), the lower would be the toll of PAH in the area. However, on the contrary, this factor lost the weak significance observed in table 1 with an IRR of 0.96 and was excluded in subsequent models. Model 3 tested whether, beyond differences in population characteristics, intensity of hospitalisation and distance to a hospital (the variables that were found to be associated in table 1) were associated with PAH. Both factors upheld their positive association (IRR 1.15 and IRR 1.07, respectively). Finally, in model 4, AC of residence was included as a random effect with a view to capturing unobserved factors at that level. As observed in table 2, hospitalisation intensity in previous years remained associated with PAH (IRR 1.19) whereas distance to a hospital lost significance. The residual variance explained by ACs (variance not explained by the differences in health status and demography and hospital intensity across HCAs) was statistically significant but negligible in magnitude (variance 0.01 (SE 0.008)). None of the random slope models provided an improvement in model fit (see online supplementary appendix 5).

### DISCUSSION

PAH in chronic patients were associated with demography and health of the HCA population (age, sex and burden of disease distribution) and with the intensity of hospitalisations for any cause in previous years. The AC of residence explained a negligible fraction of the residual variation.

Hospitalisation intensity was the main factor associated with PAH differences across HCAs. This is consistent with previous evidence observed in different healthcare systems.1,13 In our opinion, once the potential effect of legitimate reasons (ie, differences in the health status of the populations) has been extensively adjusted, we can hypothesise that the association between hospital intensity and PAH stems from systemic factors such as planning or patient management. With regard to planning, this argument was proposed for Medicare in the mid 1990s, where a high correlation between PAH discharges and medical discharges was observed.13 In that case, the

### Table 1

<table>
<thead>
<tr>
<th>Incidence rate ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burden of disease</td>
<td>1.12</td>
</tr>
<tr>
<td>Intensity of hospitalisation</td>
<td>1.17</td>
</tr>
<tr>
<td>Mix of PAH conditions</td>
<td>1.06</td>
</tr>
<tr>
<td>Distance to a hospital</td>
<td>1.08</td>
</tr>
<tr>
<td>Hospital supply</td>
<td>1.14</td>
</tr>
<tr>
<td>Socioeconomic status</td>
<td>0.96</td>
</tr>
<tr>
<td>Social care supply</td>
<td>0.99</td>
</tr>
<tr>
<td>Primary care supply</td>
<td>0.94</td>
</tr>
<tr>
<td>Specialised ambulatory care</td>
<td>0.96</td>
</tr>
<tr>
<td><strong>PAH, potentially avoidable hospitalisations</strong></td>
<td></td>
</tr>
</tbody>
</table>

### Table 2

<table>
<thead>
<tr>
<th>HCA level</th>
<th>Model 1 (IRR (95% CI))</th>
<th>Model 2 (IRR (95% CI))</th>
<th>Model 3 (IRR (95% CI))</th>
<th>Model 4 (IRR (95% CI))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>0.98 (0.90 to 1.06)</td>
<td>0.98 (0.90 to 1.06)</td>
<td>0.99 (0.95 to 1.04)</td>
<td>1.00 (0.95 to 1.04)</td>
</tr>
<tr>
<td>Age group</td>
<td>0.84 (0.75 to 0.95)</td>
<td>0.84 (0.75 to 0.95)</td>
<td>0.89 (0.83 to 0.96)</td>
<td>0.93 (0.89 to 0.98)</td>
</tr>
<tr>
<td>Burden of disease</td>
<td>1.12 (1.05 to 1.20)</td>
<td>1.12 (1.05 to 1.18)</td>
<td>1.07 (1.03 to 1.13)</td>
<td>1.06 (1.01 to 1.10)</td>
</tr>
<tr>
<td>Primary care supply</td>
<td>0.96 (0.91 to 1.02)</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Intensity of hospitalisation</td>
<td>1.15 (1.08 to 1.22)</td>
<td>1.19 (1.13 to 1.26)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Distance to a hospital</td>
<td>1.07 (1.04 to 1.10)</td>
<td>1.03 (0.98 to 1.08)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>AC level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Int</td>
<td>intercept SE</td>
<td>0.01</td>
<td>0.008</td>
<td></td>
</tr>
<tr>
<td>LRT $\chi^2$ (p value)</td>
<td>1.97 (p=0.16)</td>
<td>37.99 (p&lt;0.0001)</td>
<td>15.50 (p&lt;0.0001)</td>
<td></td>
</tr>
</tbody>
</table>

Model 2, with primary care supply factor, included 191 HCAs given the lack of information in the Madrid AC. AC, autonomous community; HCA, healthcare area; IRR, incidence rate ratio; LRT, likelihood ratio test.
uneven bed supply distribution (ie, more beds in urban, most affluent areas) played an important role. In our study, the hospital supply factor was also found to be associated with hospital intensity (IRR 1.09; 95% CI 1.05 to 1.15). Regarding patient management, a plausible explanation is related to the way chronic patients are managed in hospital emergency wards. In a previous study, the decision to admit a chronic patient or treat the relapse episode in a day-case ward was hypothesised as the reason for the difference in hospitalisation rates across areas. The lack of association between primary care features or socioeconomic status and PAH is inconsistent with most of the literature, in which PAH rates are observed to fall with higher primary care supply or higher socioeconomic status. \textsuperscript{16–19, 23, 27}

With regard to primary care features, two main arguments could explain the lack of association. On the one hand, the variables contained in the factor might not be sensitive enough to detect differences in PAH. Indeed, given data availability, the only variables that were included in the factor were the number of primary care centres, centres with emergency wards and primary care visits. Unfortunately, we have not been able to collect and analyse variables that have been shown to be predictive in the literature (eg, continuity of care \textsuperscript{5}) or variables expected to be sensitive to differences and that could better represent the quality of performance in the clinical setting (eg, treatment appropriateness, early detection of disease relapse or adherence to treatment). On the other hand, this study focuses on HCAs and ACs; however, HCAs are also composed of a number of Primary Care Areas (PCAs) that act as the basic units of assistance for a defined population. We could therefore postulate that the primary care variables in this study should have been used at the PCA level of analysis rather than at the HCA level to be sensitive to the detection of differences in PAH. In fact, some of the studies that found an association between primary care and PAH in Spain were developed at PCAs.\textsuperscript{27}

As far as socioeconomic status is concerned, the lack of association may be explained by some characteristics of the SNHS such as mandatory insurance, universal coverage, extensive primary care coverage and small financial barriers (particularly in the period of study). However, there are two other factors that might explain the findings according to other studies developed in the context of the SNHS. First, unemployment might not be sensitive enough to detect differences in PAH at the HCA level. Indeed, a study carried out in the context of a specific AC in Spain found that unemployment rates were associated with PAH at the primary care level.\textsuperscript{24} Second, unemployment as the single variable explaining socioeconomic status at the population level might not capture socioeconomic differences well. Two other studies on PAH in two different ACs found an association when using a compound of deprivation measures.\textsuperscript{5, 27}

Along a different line, although the distance to a hospital was found to be associated in model 3, a finding consistent with previous research, \textsuperscript{7, 8, 25, 26, 36} the association disappeared when the AGs in the second level were included. This finding is compatible with the fact that distance to a hospital is a good proxy of rurality in Spain, and the distribution of rural versus urban areas in the country is markedly associated with the AC of residence.

Finally, this is the first study in the SNHS context that explores residual variance at the AC level. Given that the AGs are the loci for policy, planning and financing decisions (eg, the implementation of the National Strategy on Chronic Care is deployed at the AC level), we could expect a higher effect. However, the effect was seen to be negligible, which implies that HCAs (ie, provision decisions) are responsible for the vast majority of the observed variation in PAH.

**Study limitations**

Beyond the intrinsic limitations of a cross-section ecological study (ie, limited capacity to establish causal associations and the risk of ecological fallacy if interpretations are not confined to the HCAs and ACs), three main caveats should be considered when interpreting our study results. First, we were not able to explore some predictors found relevant in the literature (eg, continuity of care) because of the lack of data. In addition, HCAs have been considered as a homogenous body, although primary care centres within an HCA behave heterogeneously when treating chronic conditions. Finally, the number of HCAs (n=202) and ACs (n=16) might be insufficient to find an association if it exists, which could partially explain the lack of association in some factors.

Other limitations worth mentioning are related to residual confounding and misclassification bias. With regard to the former, because the outcome variable is defined as a composite measure, a potential reason for the differences in PAH across HCAs could stem from the existence of variability in the mix of conditions included in the composite. As observed in table 1, the mix of PAH conditions was not associated with differences in PAH, ruling this out as a confounding variable. Regarding misclassification, three sources of potential bias could be distinguished: (1) differences in coding intensity across HCAs could play a role in the identification of PAH and burden of disease cases but, fortunately, coding intensity is not expected to affect diagnosis at admission since primary diagnoses are missed in less than 0.7% of the 4.9 million discharges; (2) different coding practices across HCAs could play a role in the misclassification of those variables based on administrative data; to avoid such a threat, we used variables formally validated for the SNHS (construct, face and empirical validation) within the context of the Atlas VPM project;\textsuperscript{29, 37–40} and (3) the lack of studies on the
quality of the data sources used to build the factors poses concerns about their reliability.

In addition, although the study focused on virtually all the publicly funded admissions in the country discharged from both private and public hospitals, hospitalisations funded privately were excluded. Whether this might imply some limitation in terms of internal validity... a longitudinal examination of the... coding need to be more specific to aid policy makers and clinicians. Public Health... 2009;123:169–71.


