An Analysis of the Costs of Treating Schizophrenia in Spain: a Hierarchical Bayesian Approach

Francisco-Jose Vázquez-Polo,1* Miguel Negrín,1 Juan M. Cabasés,2 Eduardo Sánchez,2 Joseph M. Haro,3 Luis Salvador-Carulla4

1 Department of Quantitative Methods, University of Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, Spain
2 Department of Economics, Public University of Navarre, Pamplona, Spain
3 Research and Development Unit, Sant Joan de Déu-SSM, Sant Boi, Barcelona, Spain
4 Department of Psychiatry, University of Cádiz, Cádiz, Spain

Abstract

Background: Health care decisions should incorporate cost of illness and treatment data, particularly for disorders such as schizophrenia with a high morbidity rate and a disproportionately low allocation of resources. Previous cost of illness analyses may have disregarded geographical aspects relevant for resource consumption and unit cost calculation.

Aims: To compare the utilisation of resources and the care costs of schizophrenic patients in four mental-health districts in Spain (in Madrid, Catalonia, Navarra and Andalusia), and to analyse factors that determine the costs and the differences between areas.

Methods: A treated prevalence bottom-up three year follow-up design was used for obtaining data concerning socio-demography, clinical evolution and the utilisation of services. 1997 reference prices were updated for years 1998-2000 in euros. We propose two different scenarios, varying in the prices applied. In the first (Scenario 0) the reference prices are those obtained for a single geographic area, and so the cost variations are only due to differences in the use of resources. In the second situation (Scenario 1), we analyse the variations in resource utilisation at different levels, using the prices applicable to each healthcare area. Bayesian hierarchical models are used to discuss the factors that determine such costs and the differences between geographic areas.

Results: In scenario 0, the estimated mean cost was 4918.948 euros for the first year. In scenario 1 the highest cost was in Gava (Catalonia) and the lowest in Loja (Andalusia). Mean costs were respectively 4547.24 and 2473.98 euros. With respect to the evolution of costs over time, we observed an increase during the second year and a reduction during the third year. Geographical differences appeared in follow-up costs. The variables related to lower treatment costs were: residence in the family household, higher patient age and being in work. On the contrary, the number of relapses is directly related to higher treatment costs. No differences were observed between health areas concerning resource utilisation.

Discussion: Calculating the costs of a given disease involves two principal factors: the resource utilisation and the prices. In most studies, emphasis is placed on the analysis of resource utilisation. Other evaluations, however, have recognized the implications of incorporating different prices into the final results. In this study we show both scenarios. The factors that determine the cost of schizophrenia for the Spanish case are similar to the factors encountered in studies carried out in other countries.

Implications for Health Policies: Treatment costs may be reduced by the prevention of psychotic symptoms and relapse.

Implications for Future Research: The use of the same price data in multicentre studies may not be realistic. More effort should be made to obtain price data from all the centres or countries participating in a study. In the present study, only direct healthcare and social costs have been included. Future research should consider informal and indirect costs.

Received 30 June 2004; accepted 13 August 2005

Introduction

Clinical decisions cannot be taken without a prior analysis of the cost of treatment, particularly for disorders with a high morbidity rate, such as schizophrenia. According to the 1993 World Bank report, although neuropsychiatry disorders constitute the second most prevalent non-infectious disease, they receive a disproportionately small allocation of resources in countries with a consolidated market economy.

In the international context, in a study of the cost of treating mental illness in the U.S.A., schizophrenia was found to be the most costly individualized disease, with 25.8% of direct costs, 17% of indirect costs by morbidity and 11% of indirect costs by mortality. With respect to other costs (crime, benefit payments, imprisonment and care in the family), schizophrenia represents 54.5% of the subtotal. In...
In Spain, the Psicost group has contributed to developing a methodology to evaluate services and the costs arising from chronic disease in Spain.\textsuperscript{5-7}

Most studies of the factors determining such costs have been developed at hospital level.\textsuperscript{8-10} However, the growing decentralization of social and healthcare services and emphasis on out-patient care (such as rehabilitation units, day-care services and home-based treatment) has reduced the attention paid to studying hospital-aggregated costs. In general, patients are admitted when the disease is at an advanced stage, and so the consideration of hospital costs alone is not appropriate. This situation is more acute in the case of diseases like schizophrenia, in which the relative weight of hospital costs is less than that of other expenses such as medicines and informal care.\textsuperscript{11-13} In the present study, we use micro-units (“patients”) to analyse the factors determining treatment costs in Spain.

There are considerable differences between different regions in Spain concerning the provision of social and healthcare services,\textsuperscript{14} a fact to be taken into consideration when costs are evaluated. For this reason, we propose a statistical technique, hierarchical modelling, that is suitable for examining the effects of regional variations on treatment costs. Hierarchical models enable us to analyse data that present a hierarchical structure comprising multiple “micro” units nested within “macro” units. Various authors have proposed applying such models to different areas of healthcare economics.\textsuperscript{15-17}

In the present study, we present the Bayesian estimation of hierarchical models, using Markov Chain Monte Carlo (MCMC) simulation methods.\textsuperscript{18,19} The Bayesian perspective enables us to derive a natural interpretation of the results in terms of probability, and also to incorporate prior information into the analysis.\textsuperscript{20}

The objectives of this study are to analyse costs of schizophrenia in different healthcare regions and to determine the relevant variables underlying these costs. For these purposes, we used Bayesian hierarchical analysis, applied to four healthcare regions in Spain.

We propose two different scenarios, varying in the prices applied. In the first (Scenario 0) the reference prices are those obtained for a single geographic area, and so the cost variations are only due to differences in the use of resources. In the second situation (Scenario 1), we analyse the variations in resource utilisation at different levels, using the prices applicable to each healthcare area.

### Methods

#### Data

This project is a treated prevalence based bottom-up prospective study on the cost of schizophrenia in Spain, which provides complementary information to a previous incidence based study on this topic.\textsuperscript{21} Both societal and health decision making perspectives have been addressed.

Data were obtained from four small areas, which were selected as being representative of different socio-economic contexts and of different kinds of organization and availability of services. The four areas or health districts analysed in the study were in the regions of Catalonia (Barcelona: Gava health district), Andalusia (Granada: Loja health district), Madrid (Salamanca health district) and Navarre (Burlada health district).\textsuperscript{14} Further detail about the characteristics of the Spanish mental health system and the selected areas is published elsewhere.\textsuperscript{21} In Spain, public mental health care delivery is organized in sectors. Each sector has a mental health care community centre and other related resources that in some cases, for example hospitals, may be shared among several sectors.

For each centre, we selected from the register with outpatient visit data a representative sample of cases of schizophrenia determined by the prevalence of cases treated. Criteria for inclusion in the study were: diagnosis of schizophrenia (DSM-IV diagnosis), aged 18-65 years and having been in contact with the mental health treatment services in one of the selected areas within the six-month period designated for inclusion. After excluding patients with a primary diagnosis of neurological disorder or mental handicap, a sample of 356 patients was obtained (see Table 1).

The patients were evaluated at three points in time: at the beginning of the study, after one year and after two years (1998-2000). The following measuring techniques were used to evaluate the patient’s clinical record and use of treatment services:

- Socio-demographic questionnaire and clinical record.
- Positive and Negative Symptoms Scale (PANSS), Spanish version.\textsuperscript{22,23} PANSS measures 30 items of symptoms, using a semi-structured interview format. Symptoms are

#### Table 1. Healthcare Area Data

<table>
<thead>
<tr>
<th>Healthcare Area</th>
<th>Province</th>
<th>Inhabitants</th>
<th>Nº of patients</th>
<th>Source data record</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gavá</td>
<td>Barcelona</td>
<td>135,000</td>
<td>86</td>
<td>Gavá mental health centre</td>
</tr>
<tr>
<td>Loja</td>
<td>Granada</td>
<td>63,490</td>
<td>73</td>
<td>Schizophrenia cases in the South Granada area</td>
</tr>
<tr>
<td>Salamanca</td>
<td>Madrid</td>
<td>142,001</td>
<td>105</td>
<td>Psychiatric cases in the Madrid region</td>
</tr>
<tr>
<td>Burlada</td>
<td>Navarre</td>
<td>65,000</td>
<td>92</td>
<td>Navarre health information system (SISNA)</td>
</tr>
<tr>
<td>Total</td>
<td>–</td>
<td>405,491</td>
<td>356</td>
<td>–</td>
</tr>
</tbody>
</table>
classified into three subscales: positive, negative and general psychopathology. PANSS positive subscale is a subset of items in the PANSS that rates seven positive symptoms of schizophrenia (delusions, conceptual disorganization, hallucinations behavior, excitement, grandiosity, suspiciousness/persecution, and hostility). The PANSS negative subscale is a subset of items in the PANSS that rates seven negative symptoms of schizophrenia (blunted affect, emotional withdrawal, poor rapport, passive apathetic withdrawal, difficulty in abstract thinking, lack of spontaneity/flow of conversation, stereotyped thinking).

- Global Assessment of Functioning (GAF-general), Spanish version. This scale forms part of the V axis of the DSM-IV diagnosis, and measures the patient’s overall functioning on a scale of 0-100, with higher scores representing better functioning. We used the Goldman et al.24 modified version of the Global Assessment of Functioning Scale, separating the measures of social and occupational functioning (GAF-social or SOFAS) from the measures of symptoms and psychological functioning (GAF-clinical). It permitted the measure of social functioning and mental impairments in separate axis, as shown in previous studies.5

- Disability Assessment Schedule, short version.25 The DASsv is a measure of disability based on the International Classification of Diseases (ICD-10), and evaluates 4 areas of disability: self-care, family and work role fulfilment and social adaptation. Values of 0-5 are assigned, with higher scores expressing a greater degree of disability.

- EuroQol-5 Dimensions (EQ-5D), a measure of health-associated generic life-quality. This measure constitutes a visual analogue scale (VAS) by which the patient’s health is evaluated on a scale of 0-100. It has been adapted for use in Spanish.26,27

- The Schizophrenia Cost Evaluation Questionnaire (SCEC). This instrument comprises an inventory of the utilisation of healthcare and social services, together with information on indirect costs.4 It is based on the Client Service Receipt Inventory (CSRI).28 The instrument collects information about service used using four sources of information: interviews with patients, family members and clinicians and review of clinical records. All the instruments are administered by a trained external researcher.

We did not use a previously available data base or claim data set to identify service use, but rather, a bottom-up approach in which the external evaluator determined service use through interviews with the patient and the family. Additionally, data bases and clinical charts that contained information about the services the patient received were used to complete the information.

Table 2 summarises descriptive statistics of the sample for the first operational year of the study.

Calculating the costs of a given disease involves two principal factors: resource utilisation and prices. In most studies, emphasis is placed on the analysis of resource utilisation. Thus, for example, in the context of multinational

| Table 2. Sampling Descriptions. Mean (Standard Deviation) or Percentage (%), 1997 |
|-------------------------------|------------------|----------------|-------------------|-------------------|-------------------|
|                             | Gava | Loja | Salamanca | Burlada | Total |
| Age                          | 37.53 (11.67) | 38.12 (9.07) | 37.68 (9.03) | 40.4 (10.91) | 38.44 (10.25) |
| Female                       | 31.40 | 21.92 | 33.33 | 36.96 | 31.46 |
| Years ill                    | 15.01 (10.68) | 15.01 (7.29) | 13.87 (8.65) | 14.25 (9.09) | 14.47 (9.04) |
| Partner                      | 20.93 | 9.59 | 10.48 | 14.13 | 13.76 |
| Employment                   | 12.79 | 17.81 | 24.76 | 25.00 | 20.51 |
| Household                    |       |       |       |       |       |
| Family                       | 87.21 | 87.67 | 79.05 | 84.78 | 84.27 |
| Alone                        | 8.14 | 4.11 | 12.38 | 6.52 | 8.15 |
| Sheltered accom./inst.       | 4.65 | 0.00 | 1.90 | 4.35 | 2.81 |
| N° relapses                  | 0.8 (0.97) | 0.7 (1.01) | 0.68 (0.78) | 0.17 (0.41) | 0.58 (0.85) |
| N° suicide attempts          | 0.05 (0.26) | 0.11 (0.43) | 0.08 (0.3) | 0.02 (0.15) | 0.06 (0.3) |
| GAF-clinical                 | 46.59 (14.94) | 49 (11.85) | 55.37 (19.43) | 50.96 (13.97) | 51.36 (16.2) |
| GAF-social                   | 48.63 (14.59) | 42.11 (13.48) | 51.91 (20.55) | 47.46 (13.45) | 48.18 (16.8) |
| GAF-general                  | 44.14 (14.58) | 43 (13.05) | 52.88 (20.01) | 49.44 (12.78) | 48.44 (16.47) |
| DAS-personal                 | 0.96 (1.15) | 0.95 (1.11) | 1.24 (1.29) | 0.69 (0.98) | 0.98 (1.17) |
| DAS-occupational             | 3.96 (1.77) | 4.7 (1) | 2.96 (1.9) | 3.74 (1.95) | 3.69 (1.85) |
| DAS-family                   | 2.08 (1.26) | 2.33 (1.94) | 1.96 (1.4) | 1.74 (1.66) | 2 (1.58) |
| DAS-others                   | 3.06 (1.45) | 2.18 (1.15) | 2.44 (1.47) | 2.12 (1.36) | 2.41 (1.41) |
| PANSS-positive               | 16.12 (6.26) | 15.03 (5.75) | 12.36 (6.27) | 11.99 (5.47) | 13.85 (6.19) |
| PANSS-negative               | 22.2 (9.33) | 19.8 (7.19) | 17.01 (8.95) | 20.85 (7.38) | 20.04 (8.5) |
| PANSS-general                | 35.2 (11.6) | 32.65 (8.42) | 33.08 (11.71) | 31.49 (8.64) | 33.15 (10.34) |
| VAS                          | 63.35 (24.1) | 57.18 (22.7) | 55.08 (23.81) | 57.82 (17.59) | 58.4 (22.17) |
| N                            | 86 | 73 | 105 | 92 | 356 |
The range of prices applicable within a certain country is used to evaluate the use of resources in all the countries forming part of the study. Other evaluations, however, have recognized the implications of incorporating different prices into the final results. In Spain, there is no complete database of prices, although partial information is provided by some (for example, the SOIKOS data base of Spanish health care costs).

In this study we have attempted to obtain prices for each of the areas considered, based on accounting data for each healthcare unit. The list of prices obtained is given in Table 3. It was not possible to derive prices for each and every one of the resources considered, except for the Loja district in Granada. The prices for which data were unavailable were simulated under the assumption that the distribution of costs among resources, for each of the three blocks considered (out-patient health care, in-patient attention, and intermediate aid and social services) is the same in all areas, taking that of Loja as a reference. We show, as an example, the simulation of the price of the casualty department in the Salamanca district (Madrid). In Loja, our area of reference, the total price for the first block is 373.12, of which 13.61% corresponds to the casualty department. The total price for Salamanca district is estimated as the ratio between the sum of the price data available and the percentage of those services in the distribution of prices in the Loja reference area

\[
\left( \frac{41.49 + 38.28 + 30.96 + 30.90}{0.1361 + 0.1029 + 0.0676 + 0.65} \right) = 381.08.
\]

Thus, the simulated price for the casualty department is 381.08 / (58.24/373.12) = 59.49. The same process is followed for the other simulated prices. The price of the casualty department for the area of Gavà (Barcelona) was not used in the simulation process because it incorporates services that are not included in other areas. The mean percentages of the total cost obtained with simulated prices are 11.61%, 19.97% and 5.08% for the areas of Gavà (Barcelona), Salamanca (Madrid) and Burlada (Navarra), respectively. These simulated prices are shown in bold print in Table 3.

We propose two alternative scenarios for the analysis of disease-related costs. Scenario 0 is that in which the prices for a single area (Loja) are applied to evaluate the use of resources in the four study areas. Thus, cost differences between patients are due solely to the differing utilisation of resources. In Scenario 1, on the other hand, the prices for each of the areas listed in Table 3 are applied.

The costs for subsequent years were obtained by applying the interannual rate of inflation, provided by the Spanish Statistical Institute (I.N.E., Spanish initials), to the relevant costs for the year 1997. The prices of medicines were the same for the four areas, and were obtained from the official drug-price catalogue of the public health system.

Model

Bayesian multilevel models are used in the present study. Bayesian statistics are widely used in the specific context of the economic evaluation of healthcare activities. Cost data are individualized by patients and are obtained from various centres, and so it is desirable to obtain probabilistic models with realistic error structures.

Multilevel models, also known as hierarchical linear models or as random parameter models, have been applied in numerous areas of the social sciences. Hierarchical models, designed for the analysis of individual data grouped into hierarchies or levels, are especially suitable for the present study, as they combine effects that are peculiar to the individual with others that refer to the study area in which the individual resides.
The area can influence the cost in various ways: one important reason for variation in costs between areas is the variation in resource use which, in turn, arises from other factors such as clinical practice variations. Another possible reason for the variation is the price of resources. Other factors that might result in variability in costs between areas are heterogeneity in training, education or capital expenditure in previous years.

According to the literature, the most commonly used models to analyse the area effect are ordinary least-squares (OLS) models. These models include binary variables referring to hierarchies, but do not take into account the hierarchical structure of the data. OLS models assume that observations across areas are independent and have a common variance. Furthermore, area-level variables included in an OLS model are considered as if they were measured at the patient level, thus spuriously inflating the amount of information supplied.

A number of efficient algorithms are available for obtaining maximum likelihood (ML) estimates of a multilevel model, for example the iterative generalized least squares procedure (IGLS) or restricted maximum likelihood estimates (RIGLS). Nevertheless, Bayesian methods can implement multilevel models without statistical limitations. Bayesian MCMC methods yield inferences based upon samples from the full posterior distribution and allow exact inference in cases where, as mentioned above, the likelihood based methods yield approximations. They are easily implemented with the software package WinBUGS. Browne and Draper made a detailed comparison of classical and Bayesian estimation for this type of model.

The three levels, namely year, patient and geographic area, form a natural hierarchic structure for the simultaneous analysis of each individual level. By means of this model, we can evaluate the proportion of the variability of the dependent variable, the cost \( c_{ijk} \), that is due to individual characteristics, and the proportion that is due to the membership group.

The individual cost of each patient, per year, is clustered by patients, who in turn are clustered by geographic areas. In order to model this hierarchy, we include a random term to correspond to each level. This perturbation term incorporates the effects of the patient and of the area on the dependent variable. Furthermore, costs often present a high degree of skewness, and thus the assumption of normality would not be justified. Log-normal distribution fits such an asymmetry well, and so the logarithmic transformation of costs is used as the dependent variable. Figure 1 shows the histogram of the costs and the logarithmic transformation of costs for each scenario.

As explicative variables we use those listed in Table 2, taking care to distinguish between first and second-level covariates. First-level covariates, \( x_{1ijk}, x_{2ijk}, \ldots, x_{hijk} \), vary for each year, while covariates at the second level, the patient level, are constant in time \( x_{(h+1)jk}, x_{(h+2)jk}, \ldots, x_{(h+p)jk} \). In social sciences, the explicative variables in regressions are normally centred with respect to the mean, mainly in order to allow direct interpretation of the constant, that is, the value of the variable to be explained when all the explicative variables are assigned a value of zero. In the present study, it was decided to use variables that were centred with respect to the mean, except in the case of qualitative variables such as sex or employment situation. The use of centred variables not only provides advantages in interpreting the results but also reduces the convergence time in simulation processes.

Thus, the model described can be expressed as:

\[
\ln c_{ijk} = \beta_0 h_{ik} + \beta_1 x_{1ijk} + \beta_2 x_{2ijk} + \ldots + \beta_h x_{hijk} + \mu_{ijk} \\
\beta_0 h_{ik} = \beta_0 + \beta_{h+1} x_{(h+1)jk} + \ldots + \beta_{h+p} x_{(h+p)jk} + \epsilon_{ijk} \\
\mu_{ijk} = \mu_{ik} + \nu_k \\
\mu_{ik} \sim N(0, \sigma^2_{\mu}), \quad \epsilon_{ijk} \sim N(0, \sigma^2_{\epsilon}), \quad \nu_k \sim N(0, \sigma^2_{\nu})
\]

\[
\text{cov}(\mu_{ijk}, \epsilon_{ijk}) = 0, \quad \text{cov}(\mu_{ijk}, \nu_k) = 0, \quad \text{cov}(\epsilon_{ijk}, \nu_k) = 0
\]

The subindices \((i = 1, \ldots, n_1), (j = 1, \ldots, n_2)\) and \((k = 1, \ldots, n_3)\) refer to each of the levels considered, that is, year, patient and geographic area. The study comprised 4 such areas \(n_3\) and 356 patients \(n_2\), while the value of the year data was 1068 \(n_1 = 3 \times 356 = 1068\).

For the sake of simplicity, we have assumed a level effect only for the constant term, and have not analysed the level effect for each of the explicative variables. Thus, each patient and each geographic area possess their own cost levels.

The characteristics for the first level include:

- The patient’s age. Age can be interpreted as the variable tendency of the model. We assumed a model presenting linear growth, due to the time horizon being too short to permit the observation of polynomial growth. We expect a negative effect of the age, confirming that the need for cost-intensive treatment interventions is maximum at the early stages of the illness.
- The employment situation. The use of centred variables not only provides advantages in interpreting the results but also reduces the convergence time in simulation processes.
- The area can influence the cost in various ways: one possible reason for variation in costs between areas is the price of resources. Other factors that might result in variability in costs between areas are heterogeneity in training, education or capital expenditure in previous years.
- The employment situation. The use of centred variables not only provides advantages in interpreting the results but also reduces the convergence time in simulation processes.
- The area can influence the cost in various ways: one possible reason for variation in costs between areas is the price of resources. Other factors that might result in variability in costs between areas are heterogeneity in training, education or capital expenditure in previous years.
a positive effect of this variable.
- PANSS. The PANSS scale measures clinical symptoms, divided into three subscales: positive (delusions, hallucinations, for example), negative (social withdrawal, poverty of speech) and general symptoms. We would expect an increase in costs with higher levels of positive and negative symptoms, especially with the first ones.
- VAS. As a measure of quality of life, we expect a negative effect on costs of this variable, although this is not confirmed in the literature.

With respect to gender, the literature reports lower costs for women, and how many years he/she had been affected by the disease prior to the study period, as a proxy of the stage of the illness, with the variables corresponding to the second level, namely, the patient.

The Bayesian approach allows us to incorporate information previous to the study, by means of prior distributions. In hierarchical models, one of the underlying themes is the choice of prior distributions to be used. In the present study, we utilise the conjugate conditional model, which enables Gibbs sampling to be applied. For a detailed discussion of the choice of prior distributions, see Browne.19

\[
\begin{align*}
(\beta_1, \beta_2, \ldots, \beta_h, \ldots, \beta_{h+p}) & \sim \text{NMV}(\beta_0, \Sigma_0), \\
\sigma^2_\mu & \sim \text{IG}(a_\mu, b_\mu), \\
\sigma^2_\varepsilon & \sim \text{IG}(a_\varepsilon, b_\varepsilon), \\
\sigma^2_\nu & \sim \text{IG}(a_\nu, b_\nu)
\end{align*}
\]

where \(\beta_0\) and \(\Sigma_0\) are the mean and the prior matrix of the variances-covariances of the vector \(\beta = (\beta_1, \beta_2, \ldots, \beta_h, \ldots, \beta_{h+p})\); \(a\) and \(b\) refer to the shape and scale parameters of the gamma distribution.

In the present study we examine, for the first time in Spain, the determinants of treatment costs for schizophrenic patients. For this reason we assume a lack of initial
knowledge, expressed as non-informative prior distributions. Such a prior structure is described by the following parameters of the prior distribution:

\[
\beta_0 = (0, 0, \ldots, 0),
\]

\[
\Sigma^{-1} = \begin{pmatrix}
0.00001 & 0 & \cdots & 0 \\
0 & 0.00001 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0.00001
\end{pmatrix}
\]

\[
a_0 = 0.001 \quad \text{and} \quad b_0 = 0.001
\]

Results

Table 4 shows descriptive statistics of the estimated costs for each healthcare area, for each of the years comprising the study.

The costs showed in Table 4 are similar to that calculated by Duggan\textsuperscript{43} for the same period of time in U.S.A. The average inpatient and outpatient costs for patients with schizophrenia in that country were 3834$ in 1997, 4250$ in 1998 and 4504$ in 1999. It is important to point out that these costs are adjusted to 2001 dollars. In that year the exchange rate was 1$ = 1.117€.

For our comparison purpose, we remark that the characteristics of both samples are also similar. The average age for the Spanish sample is 38.44, while the average age in U.S.A. is 43.8. The percentage of population male is 68.54% in Spain and 54.9% in U.S.A. If we analyse some measure of use of services, the percentage of people with any hospitalization in the first year of study (1997) is 25.42% in Spain, versus the 32.6% of U.S.A. However, the average days in hospital are 10.91 in Spain and 7.3 in U.S.A.
Scenario 0

The first scenario represents an analysis of resource utilisation, as the same prices were used for each of the geographic areas. Table 4 shows that the estimated costs according to Scenario 0, with respect to Scenario 1, are lower for the Gavà and Burlada areas, and higher for the Salamanca area, due to the higher prices in Gavà and Burlada than those corresponding to Loja.

The first column of Table 5 gives the results of the model in which only the constant is included. This preliminary information is useful in that it informs us how the total cost variance is distributed by levels. The estimated mean variance per level (patient years nested within patients) was 0.9026. The estimated mean variance between patients is 1.106, there existing a 95% probability that such a variance lies between 1.004 and 1.215. Concerning the variance between geographic areas, a mean a posteriori value of 0.4491 units was estimated, with an associated probability interval of (0.1321, 1.334). As was to be expected, most of the variation in costs was situated at the patient level (45.00% of total variance). The geographic area was responsible for 18.27% of the total data variation, although the Bayesian interval is very wide. The estimated mean cost for all the data considered was 4918.948 euros. These values refer only to variations in the constant, and can be used as a reference for the complete model, which incorporates characteristics to predict the costs at the various levels.

The second column in Table 5 shows the results for the hierarchical model specified in the ‘Model’ section. The variance in the data that is not explained by the covariates is now estimated at 0.5969 euros. The variance between patients (level 2) is also reduced when their individual characteristics are controlled, and is evaluated at 0.8012 units. Finally, the mean variance between healthcare areas is calculated to be 0.2559 units.

As the dependent variable has been transformed into logarithmic form, the coefficients cannot be interpreted directly, but require exponential transformation. Thus, the exponential of the coefficients can be interpreted as the proportional change in costs arising from a unit change in the independent variable.

On average, the women in the study presented costs that were 12.10% lower than those corresponding to the men, although the very large probability interval obtained prevents us from claiming the latter variable to be statistically relevant. We also observed that higher patient age was directly associated with lower treatment costs. In fact, it is estimated that one year older is associated with a reduction about 2.58% in cost. The negative effect of increased age on treatment costs has been observed in many other studies of cost analysis concerning schizophrenic patients, thus confirming the hypothesis that the use of high-cost resources such as hospital admissions is more prevalent during the early stages of the disease, and decreases over time. Being in work is associated with greater patient independence and, probably, with lower levels of symptoms, resulting in lower cost levels incurred, estimated at 24.87%, with a 95% probability of their being between 0.21% and 44.77%. The years during which the disease had been suffered and the existence or otherwise of a stable relationship were not found to be determinant factors in treatment costs.

A factor that was found to be relevant to treatment costs was the type of household in which the patient lived. Those who lived in a family environment, whether with their own children or with their parents, incurred costs that were 75.92% lower than those corresponding to patients living in sheltered housing or in institutions. Patients who lived alone also presented costs that were 67.95% lower than those of the reference patients. Finally, the patients who lived in other circumstances incurred costs that were 69.63% lower than those living in sheltered housing or in public institutions.

The number of psychotic relapses was also found to be a cost-determining variable. The patients with a higher number of relapses incurred 49.8% higher costs than those of the reference patients, with a 95% probability that this increase was between 35.1% and 65.4%. The number of suicide attempts made was not a relevant factor concerning treatment costs.

The GAF measures the patient’s overall level of functioning, with higher scores representing better functioning. Higher GAF scores imply lower treatment costs. However, the inclusion of the value 1 in the probability intervals of this measure means that it is not statistically relevant.

The DAS scale measures the patient’s level of disability during a given year, with higher scores reflecting greater disability and, presumably, higher treatment costs. The family DAS is the measure that has greatest impact on treatment costs. A unit increase in the family DAS corresponds to a cost increase of between 0.1% and 14.3%, with a probability of 95%. As regards self-care, occupational functioning and functioning in other activities do not show any positive effect with respect to treatment costs.

Neither the positive, nor the negative nor the general PANSS were found to be relevant regarding the estimation of treatment costs. The same was true for the visual analogue scale, which was found to have virtually no relation with costs.

Figure 2 illustrates the error terms of the third level, analysed by healthcare areas, which enables us to determine whether there exist relevant differences between such areas in the utilisation of resources, after having made appropriate corrections for the explicative variables.

The overlapping of the Bayesian probability intervals does not allow us to determine whether there are relevant differences between geographic areas in the use of resources, although we observed that the Salamanca and Gavà areas presented higher levels of utilisation than did Loja and Burlada.

The absence of differences between areas justified the inclusion of an alternative model, where the third level (geographic area) is not considered as a determinant of costs. This model is shown in the third column of Table 5. The variance in the data that is not explained by the covariates is now estimated at 0.5073 euros. The variance between patients is evaluated at 0.636 units. There are no differences in the estimation of the parameters of the model.
Scenario 1

A second scenario proposed in this study is that of a model with heterogeneous prices between geographic areas. Again we must note the limitations of the results presented, resulting from the absence of some prices, as noted above (Table 3). Nevertheless, in this scenario it is possible to carry out a more realistic cost analysis than that proposed in the above section, one that is of use as a first approach to the problem.

As shown in Table 4, in this scenario the geographic areas with highest mean costs during the first year were Gavá, Salamanca and Burlada, with mean costs of 4547.24, 4515.69 and 4225.13 euros, respectively. The lowest costs (2473.98 euros) were measured in Loja. The differences between areas are thus seen to be considerable. With respect to the evolution of costs over time, we observed an increase during the second year of the study (in total terms, from 4002.11 to 4210.60 euros) and a reduction during the third year, to 3989.20 euros. This tendency was observed in Gavá and in Loja, while in Salamanca the costs fell during succeeding years, and, on the contrary, in Burlada, costs increased every year.

The fourth column of Table 5 shows the results of the model when only the constant is included. The estimated mean variance in the first level is 1.012 units. The variation between patients, level two, is 1.101 units, with a probability interval of (0.8857, 1.351). The mean variance at the level of geographic area was calculated to be 0.5851 units. The coefficient corresponding to the constant was a mean value of 6.973 units. The estimated mean cost in this scenario was 4111.58 euros.

The fifth column in Table 5 shows the model, incorporating explicative variables for the patients. With regard to the predictive variables in the model, a difference of one year of age is associated with lower costs (2.42%), subjects in employment are associated with lower costs (26.92%), living in a family environment is associated with 85.69% lower costs with respect to those incurred by patients living in sheltered housing or in an institution, the number of relapses is associated with increased costs (46.7%), and high personal DAS values are linked with increased treatment costs (14.9%).

When we evaluated random effects, the greatest variance (0.8218) was observed in the first level. The variance corresponding to the geographic area was higher than that in Scenario 0, with an average increase from 0.2559 to 0.411. We see, thus, that the use of different prices between areas increases the differences between communities.

Figure 3 shows the error component in the third level by geographic area. Again we see there is an overlap between the different probability intervals, and so we cannot be categorical about differences between the levels of prices by areas, although certain differences are evident. Thus, in Loja levels are lower than in the other areas. This fact is due to the joint effects of the lower utilisation of services in Loja (see Figure 1) and of lower prices (see Table 3). The area of Burlada presents levels of resource use that are lower than those found in Gavá and Salamanca. However, when we include the high prices of these resources, the total costs are found to be similar to those of the latter two areas.

Due to the lack of differences between areas we have added the results of an alternative model where the third
Table 5. Bayesian Hierarchical Models. Results of Parameter Estimation

| Scenario 0 | | Scenario 1 | |
|-----------|---|---|---|---|---|---|---|---|---|
|           | Exp ($\beta$) | BI-95% | Exp ($\beta$) | BI-95% | Exp ($\beta$) | BI-95% | Exp ($\beta$) | BI-95% | Exp ($\beta$) | BI-95% |
| Constant $\beta$ | 7.272 (6.718, 7.871) | 8.191 (7.58, 8.768) | 8.273 (7.889, 8.652) | 6.973 (6.328, 7.513) | 8.359 (7.449, 9.444) | 8.52 (8.144, 8.891) |
| Age | 0.9742 (0.956, 0.9927) | 0.9799 (0.964, 0.9961) | 0.9758 (0.958, 0.9936) | 0.9841 (0.9683, 1) |
| Female | 0.879 (0.664, 1.136) | 0.9183 (0.716, 1.163) | 0.8716 (0.6684, 1.117) | 0.9428 (0.7321, 1.192) |
| Years ill | 1.004 (0.9843, 1.024) | 1.005 (0.9877, 1.022) | 1.005 (0.9859, 1.024) | 1.008 (0.991, 1.025) |
| Partner | 1.063 (0.7414, 1.478) | 1.065 (0.7695, 1.443) | 1.069 (0.7509, 1.482) | 1.114 (0.8094, 1.504) |
| Employment | 0.7513 (0.5523, 0.998) | 0.714 (0.541, 0.9258) | 0.7308 (0.5241, 0.989) | 0.7037 (0.5373, 0.907) |
| Family | 0.2408 (0.1197, 0.435) | 0.2443 (0.164, 0.486) | 0.1431 (0.0683, 0.259) | 0.1253 (0.0646, 0.252) |
| Alone | 0.3205 (0.1534 0.590) | 0.2862 (0.1426, 0.588) | 0.1946 (0.091, 0.367) | 0.1897 (0.0957, 0.345) |
| Other | 0.3037 (0.1259, 0.618) | 0.3105 (0.1245, 0.599) | 0.2202 (0.0899, 0.461) | 0.2012 (0.0888, 0.452) |
| N° relapses | 1.498 (1.351, 1.654) | 1.429 (1.303, 1.562) | 1.467 (1.31, 1.634) | 1.334 (1.221, 1.453) |
| N° suicide attempts | 1.238 (0.7984, 1.841) | 1.095 (0.7347, 1.573) | 1.352 (0.8239, 2.092) | 1.043 (0.7116, 1.476) |
| GAF-clinical | 0.9921 (0.9857, 1.073) | 0.9917 (0.9793, 1.004) | 0.9923 (0.9762, 1.009) | 0.9921 (0.9802, 1.004) |
| GAF-social | 0.9946 (0.9849, 1.004) | 0.9901 (0.9764, 1.004) | 1.006 (0.9952, 1.018) | 0.993 (0.9798, 1.006) |
| GAF-general | 0.9939 (0.98, 1.008) | 1.013 (0.9933, 1.032) | 0.9977 (0.982, 1.014) | 1.017 (0.9917, 1.029) |
| DAS-personal | 1.033 (0.9414, 1.13) | 1.055 (0.9722, 1.143) | 1.149 (1.04, 1.265) | 1.047 (0.9684, 1.131) |
| DAS-occupational | 0.955 (0.8915, 1.022) | 0.9458 (0.8891, 1.004) | 0.9708 (0.8994, 1.045) | 0.9454 (0.8909, 1.002) |
| DAS-family | 1.067 (1.001, 1.143) | 1.059 (0.9944, 1.127) | 1.062 (0.9851, 1.147) | 1.054 (0.9923, 1.119) |
| DAS-others | 0.9719 (0.8837, 1.067) | 0.9962 (0.9143, 1.083) | 0.9359 (0.8421, 1.036) | 0.9905 (0.912, 1.074) |
| PANSS-positive | 0.9869 (0.9674, 1.007) | 0.9905 (0.9727, 1.009) | 0.999 (0.9781, 1.021) | 0.9916 (0.9744, 1.009) |
| PANSS-negative | 1.003 (0.9864, 1.019) | 0.999 (0.9843, 1.014) | 1.006 (0.9892, 1.024) | 1.004 (0.9901, 1.019) |
| PANSS-general | 1.01 (0.997, 1.023) | 1.014 (1.003, 1.026) | 1.008 (0.9942, 1.022) | 1.011 (0.9996, 1.022) |
level (geographic area) is not considered. The results are shown in the final column of the Table 5. The variance of the first level error term is estimated to be 0.4563 euros. The variance between patients is evaluated at 0.6442 units.

**Discussion**

The results obtained in this study represent a calculation of the cost of treating schizophrenia in different healthcare areas in Spain, based on information on costs arising from the disease, and on measures of clinical evolution and of resource utilisation. A bottom-up approach was used for the calculation of costs. The samples selected were representative of the prevalence samples of individuals with schizophrenia who receive treatment from the public healthcare services of the participating catchments areas. Estimates of prices were based on available accounting data for healthcare resources. The procedure applied is probably not the most appropriate, and there do exist methodologies to ascertain costs, such as Activity Based Costing (ABC) which would almost certainly provide a closer approximation of the real value of prices.

In this study, two scenarios are used, with Scenario 1 representing the analysis of real costs and therefore the ideal objective. Due to the above-commented problems in obtaining such data, we also examine Scenario 0, to analyse resource utilisation, and present Scenario 1 as an analysis of sensitivity.

![Figure 3. Error Component of Level Three. Mean and 95% Bayesian Interval (Scenario 1).](image-url)
The indicators of social functioning and social integration are powerful predictors of costs. First, patients living in an institution have the higher costs, which are related to accommodation costs. Patient living with their family have lower costs than patients living alone, which could indicate that family members do provide informal care that reduces service needs. Finally, patients who are currently working have less cost.

The results show that greater age of the patient is associated with lower costs, which probably indicates that older individuals have less service needs. On the contrary, more years since onset is associated to higher overall costs, which can be interpreted as patients with younger age of onset having more severe forms of the disorder and thus requiring more care.

Social functioning is a more powerful predictor of costs than clinical symptoms. The effects of symptoms measured by the PANSS score, the effects of social and occupational functioning as measured by the GAF, and the effects of disability measured by the DAS are not relevant in our models. Only the disability scale in the family and self-care area shows some association with the costs. A recent paper concludes that the effects of these measures seem to be pure within effects which can be attributed to the intrapersonal variance of these variables over time. Roughly speaking, our model is focused on the analysis of the interpersonal differences and this might be the reason for the poor relevance of these measures as determinants of the costs.

We were unable to detect significant differences in the utilisation of resources or in healthcare costs between the areas analysed, although differences between the areas were found in the mean values of both variables. This means that differences in the areas could be accounted for patient differences and implies that further studies should adjust for patient characteristics before reaching conclusions when comparing aggregate data between areas or regions.

An important limitation of this work is that only direct healthcare and social costs have been included, but informal care and indirect costs have not been evaluated. We know that informal care and indirect costs may be greater than direct costs.

The costs of informal care situations have been estimated in a separate study and were found to vary greatly between the two areas analysed (Gavà and Burlada), which might have led to bias if they had been included in the analyses performed in the present study.

Other limitations should be mentioned. First, sample size is relatively small. Second, although the authors tried to include all clinical variables that could affect cost, we cannot rule out the existence of unobserved characteristics, for example sociodemographic characteristics or differences in service availability, which could be different among the different areas and thus affect costs. It should be noted, though, that the sample is representative of every catchments area included in the analysis. This is particularly relevant in Spain, where mental health care is provided by sectors. Furthermore, every case included in the study was reassessed by an external researcher previously standardised in the assessment battery used in this study.

Further work could be conducted to analyse whether the effect of each of the explicative variables on costs varies between geographic areas. This possibility has not been addressed in the present study, as our analysis focuses on comparing geographic areas at the level of the constant, with a more extended analysis proposed for future studies.

Acknowledgements

The SIG-RIRAG group, a multidisciplinary network funded by Instituto de Salud Carlos III, Ministerio de Salud y Consumo (Spain), is acknowledged for its collaboration to the study.

References

20. O’Hagan A, Stevens JW. Assessing and comparing costs: how robust are