Application of simultaneous equation models to temporary disability prescriptions in primary healthcare centres

Natividad Guadalajara Olmeda and Isabel Barrachina Martínez

Economy and Social Sciences Department, Universitat Politècnica de València, Valencia, Spain

Abstract

Non-medical characteristics of primary healthcare centres (PHCs) influence the prescription of temporary disability leaves (incidence and absence rates) due to anxiety, as studied in the Valencian Community (E Spain; 5,111,706 inhabitants, 2009), where 485 centres (66%) were analysed. A structural two-equation model was used to explain which centres’ factors impact prescriptions more.

This model determined the influence of PHCs’ factors (location, delay in specialised care, sick leave duration, etc.) on the incidence and absence rates, and on the interdependence between both rates.

The results suggest the need to improve centres’ management (clinical guidelines) and labour market regulations, and to control the disability benefits paid.

1. Introduction

In most western European countries, prescribing temporary disability (TD) for mental disorders has increased more than for other diseases [1], as have the costs incurred by TD for mental disorders [2]. This might be due to the high prevalence of mental illnesses in patients who visit primary healthcare centres (PHCs), which ranges from 20.0% to 34.4% depending on the country [1,3-6]. Therefore, mental illness is not only one of the main
causes of TD [2,7], but also one of the illnesses with the greatest degree of recurrence [8],
disability [9], and duration [10-15].

Anxiety, and dissociative and somatoform disorders with diagnosis code 300 (cod300), are the second highest ranking disorders for which TD leave days are taken in Spain [16]. However, little is known about the factors which allow TD for mental disorders to be assessed (gender, age, etc.) [2,6,15,17-19]. Furthermore, no consensus among authors has been reached. For example, some authors have observed more women prescribed TD than men for all age groups, unlike others who have concluded that this does not apply to all age groups.

Regression techniques [11,13,20] and their variants; Cox regression [10,12,21,22], Poisson regression [14], logistic regression [5,23,24] and Block-wise linear regression [25]; were used to estimate the influence of: gender; patient’s level of education; presence of chronic diseases (including depression and anxiety); marital status; socio-economic status; type and duration of employment, and salary level.

In Spain, primary care physicians (PCPs) determine whether patients are prescribed TD after examining them, and they apply a therapeutic plan in accordance with the prognosis [26]. TD should be prescribed based on patients’ factors (age, gender, marital status, disease, the legal employee regime (self-employed or general worker), etc.). However, this study raises the possibility of there also being factors relating to PHCs that influence prescribing TD leave, such as PHCs’ geographic location, the age and training of prescribing doctors [12], among others [27,28].

Therefore, this study aims to analyse which factors of the PHCs and Health Districts (HDs) impact prescribing TD for those diseases included under cod300 of the ICD-9-CM, and their modelling in 2009.
2. Material and Methods

2.1 Participants

The entire population of the Valencian Community (VC) in 2009 was analysed, which consisted in 5,117,190 inhabitants, with a working population of 2,395,598. This population was classified into the 739 PHCs, grouped into 23 HDs. The population assigned to each PHC ranges from 97 to 27,712 inhabitants, the mean being 5,554 (median of 2,708), while the average working population assigned to each PHC was 3,242 inhabitants. Currently, the number of primary care physicians per PHC ranges from 1 to 24, with a mean of 4.56 (median of 2) primary care physicians, who attend to a mean of 1,204 patients (114 to 3,008).

In these PHCs, 480,755 TDs were prescribed for all diagnoses made in 2009, 25,859 of which corresponded to cod300, and the related total leave days amounted to 2,347,394. The number of leaves finalised in 2009 for cod 300 diseases, regardless of the year in which they started, was 20,194. This represents an increase of 5,665 leaves as compared to 2008. In other words, on average, there were almost eight more leaves prescribed per PHC in 2009 than in 2008, thus evidencing the magnitude of the problem.

One of the novelties of the present study is the unit of analysis used, formed by 739 PHCs, whereas the unit of analysis employed in previous studies was either the individuals themselves or patients.

2.2 Information sources

The data relating to TD (population assigned to PHCs, TD leave duration, number of leaves started and finalised in 2009) were obtained from the Ambulatory Information System (AIS) integrated into [29] the ABUCASIS Information System of the Regional Valencian Government's Department of Health. The AIS includes several management modules, such as electronic medical records, in which all the healthcare information (TD leaves
prescribed, active and passive diagnostics, reasons for visits, etc.) about patients' visits to PHCs and specialist centres is recorded. The population-based data (age and gender) were collected from the Population Information System (PIS) of the Valencian Healthcare Agency.

PHCs-related information (accreditation for healthcare training, age of healthcare professionals, diagnoses of hospitalised patients, use of electronic medical records, etc.) was taken from the PHC Record of Accreditation for specialised healthcare training and from the Cyrus Information System for human resource management and the Management Agreement Information System of the Regional Valencian Health Department.

Information on hospital stays was taken from the Valencian Region's Minimum Basic Data Set (MBDS) of hospital admission services.

2.3. Mathematical model
In all the previous studies analysed in the literature review on TD modelling, uniequational regression models were used (models with a single dependent variable), in which the cause-effect relationship was unidirectional. However, this is indeed not often the case, therefore it is preferable to use simultaneous equation models which identify a set of variables simultaneously through a remaining set of variables. In this way, model parameters are estimated by taking other equations into account. Simultaneous equation models have been used in health areas, such as the patient's relationship of trust in PCPs and its influence on the responsibility they take for a follow-up of treatment [30].

In general, a model of m dependent or endogenous variables \((Y_1, Y_2 \ldots Y_m)\) and n prescribing or exogenous variables \((X_1, X_2 \ldots X_n)\) is formulated as follows:

\[
\begin{align*}
Y_{1i} &= \beta_{10} + 0 \ast Y_{1i} + \beta_{12} \ast Y_{2i} + \cdots + \beta_{1m} \ast Y_{mi} + \alpha_{11} \ast X_{1i} + \alpha_{12} \ast X_{2i} + \cdots + \alpha_{1n} \ast X_{ni} + \mu_{1i} \\
Y_{2i} &= \beta_{20} + \beta_{21} \ast Y_{1i} + 0 \ast Y_{2i} + \beta_{23} \ast Y_{3i} + \cdots + \beta_{2m} \ast Y_{mi} + \alpha_{21} \ast X_{1i} + \alpha_{22} \ast X_{2i} + \cdots + \alpha_{2n} \ast X_{ni} + \mu_{2i} \\
\vdots &\quad \vdots \\
Y_{mi} &= \beta_{m0} + \beta_{m1} \ast Y_{1i} + \beta_{m2} \ast Y_{2i} + \cdots + \beta_{mm-1} \ast Y_{m-1i} + 0 \ast Y_{mi} + \alpha_{m1} \ast X_{1i} + \cdots + \alpha_{mn} \ast X_{ni} + \mu_{mi}
\end{align*}
\]
In order to estimate the parameters or structural coefficients of the model (1), either systems methods or complete information methods should be used. In other words, the equations should be estimated simultaneously by considering all the restrictions imposed on the various system equations. However, in practice, these methods are not used as they are computationally intensive, provide solutions whose parameters are highly non-linear, and the potential resulting specification errors are transferred to the rest of the system [31].

Therefore, uniequational methods are utilised: ordinary least squares (OLS), indirect least squares (ILS) and two-stage least squares (2SLS). The OLS method can be implemented in the case of recursive, triangular or causal models. However, most simultaneous equation models do not have a unilateral cause and effect relationship, thus rendering it necessary to use the other two methods. The ILS method is implemented when equations are exactly identified, whereas the 2SLS method should be used if an equation is over-identified. The 2SLS method involves two successive stages [32,33]:

Stage 1: Regression equations are obtained for the reduced form of each endogenous variable in all default or exogenous variables in the complete system to eliminate the probable correlation among all endogenous variables $Y$ and the perturbation equations of the remaining endogenous variables.

\[
Y_{1i} = \hat{c}_{10} + \hat{c}_{11} * X_{1i} + \hat{c}_{12} * X_{2i} + \cdots + \hat{c}_{1n} * X_{ni} + \hat{s}_{1i}
\]

\[
Y_{2i} = \hat{c}_{20} + \hat{c}_{21} * X_{1i} + \hat{c}_{22} * X_{2i} + \cdots + \hat{c}_{2n} * X_{ni} + \hat{s}_{2i}
\]

\[
Y_{mi} = \hat{c}_{m0} + \hat{c}_{m1} * X_{1i} + \hat{c}_{m2} * X_{2i} + \cdots + \hat{c}_{mn} * X_{ni} + \hat{s}_{mi}
\]  

where $\hat{s}$ are the traditional OLS residuals. If the estimates of the average conditional $Y$ values giving a fixed $X$ are called $\hat{Y}$, then Equation (2) can be expressed as follows:

\[
Y_{1i} = \hat{Y}_{1i} + \hat{s}_{1i}
\]

\[
Y_{2i} = \hat{Y}_{2i} + \hat{s}_{2i}
\]
\[ Y_{mi} = \hat{Y}_{mi} + \hat{s}_{mi} \quad (3) \]

Stage 2: The Y values in the original structural equation (1) are replaced with the estimated \( \hat{Y} \) (3) and the OLS method is then applied to the over-identified structural equations as follows:

\[
Y_{i1} = \beta_{10} + 0 \ast (\hat{Y}_{i1} + \hat{s}_{i1}) + \beta_{12} \ast (\hat{Y}_{i2} + \hat{s}_{i2}) + \cdots + \beta_{1m} \ast (\hat{Y}_{mi} + \hat{s}_{mi}) + \alpha_{11} \ast X_{i1} + \alpha_{12} \ast X_{i2} + \cdots + \alpha_{1n} \ast X_{ni} + \mu_{i1} \\
Y_{i2} = \beta_{20} + \beta_{21} \ast (\hat{Y}_{i1} + \hat{s}_{i1}) + 0 \ast (\hat{Y}_{i2} + \hat{s}_{i2}) + \beta_{23} \ast (\hat{Y}_{i3} + \hat{s}_{i3}) + \cdots + \beta_{2m} \ast (\hat{Y}_{mi} + \hat{s}_{mi}) + \alpha_{21} \ast X_{i1} + \alpha_{22} \ast X_{i2} + \cdots + \alpha_{2n} \ast X_{ni} + \mu_{i2} \\
Y_{mi} = \beta_{m0} + \beta_{m1} \ast (\hat{Y}_{i1} + \hat{s}_{i1}) + \cdots + \beta_{m(m-1)} \ast (\hat{Y}_{m-1i} + \hat{s}_{m-1i}) + 0 \ast (\hat{Y}_{mi} + \hat{s}_{mi}) + \alpha_{m1} \ast X_{i1} + \alpha_{m2} \ast X_{i2} + \cdots + \alpha_{mn} \ast X_{ni} + \mu_{mi} \quad (4) \]

The estimators obtained in this manner are consistent.

2.4. Explanatory variables and variables to be explained
Two dependent or endogenous variables for the year 2009 were used:

1. The PHCs’ incidence rate (IR) for cod300 [18] measured as:
   \[ IR = \frac{NP}{RP} \times 100 \quad (5) \]

2. The PHC’s absence rate (AR) for cod300 measured as:
   \[ AR = \frac{DA}{(RP \times 365)} \times 100 \quad (6) \]

where: NP: number of patients in a PHC who were prescribed a TD for cod300. DA is the total number of leave days due to cod300 in a PHC. RP are the number of persons susceptible to having a TD in the PHC, and because they have a labour contract on 1 January, which is considered constant throughout the year. In other words, it is the population aged 16-65 years who is either employed or in a situation of contributory unemployment, and is likely to be prescribed sick leave.

The PHC variables considered were: total population attended to by the PHC, and their gender and age, province and a PHC on a coastline; average age of PHCs’ physicians; accreditation for health training and use of electronic medical records; average TD duration
of sick leaves for cod300 and for all diagnoses, the absence rate for all diagnoses, psychiatric services, psychiatric hospital stays and days of delay in the HD for both primary health care and specialised care (Table 1).

3. Results

Of the 739 PHCs, there were 163 with 0 TDs and 1 with 10 TDs. They were eliminated, which left the descriptive analysis with 575 PHCs serving a working population of 2,338,852 inhabitants (98%) (Table 2). The 163 eliminated PHCs were small in size with an average working population of 346 inhabitants per PHC.

The NP ranged from 1 to 313, with a mean of 44.97 TDs initiated by PHCs. Therefore, the average TD for cod300 was 1.079, meaning that 1.079 TDs were initiated for cod300 diseases in 2009 per 100 working individuals or individuals in a contributory unemployment situation in each PHC.

Each TD leave for cod300 which finalised in 2009 lasted an average of 120.40 days, which gave rise to an average of 4,082.37 total sick leave days per PHC. Having distributed these days over all the days of the working population, an average of 0.25 leave days per 100 individuals for such diseases in 2009 was obtained.

For all the TDs prescribed for all diagnoses, the AR increased by 10-fold from 0.25 to 3.036, but the average TD duration lowered from 120.40 to 70.46; i.e., by almost 50%, thus evidencing the social and economic impact of diagnosing cod3000 illnesses as compared to other diagnoses. The total number of psychiatric hospital stays for the entire working population was 203 days, which is an average of 0.354 days per PHC. On average, 90% of PHCs used electronic medical records, and the average number of days of delay for a first specialised care visit was 60.64 days.

To obtain a simultaneous equation model, crossing the various sources of information allowed the assessment of 485 PHCs of the initial 739 PHCs (i.e., 66%); this
was because information relating to some variables was lacking. The analysed PHCs served a total working population of 2,163,456 (90%), thus ensuring the representativeness of the data analysed.

The normality of distribution of the dependent IR variable was achieved by log transformation. Firstly, according to the rates obtained between the dependent variables and the remaining explanatory variables, the following system (7) based on two simultaneous equations was adopted:

\[
\begin{align*}
\ln (IR + 1) &= \beta_{10} + \beta_{12} \times AR + \alpha_{11} \times \text{PhcCol} + \alpha_{12} \times \text{PhcPsiqSt} + \mu_1 \\
AR &= \beta_{20} + \alpha_{21} \times \text{PhcCoL} + \alpha_{23} \times \text{PhcUseEMR} + \alpha_{24} \times \text{PhcAR} + \alpha_{25} \times \text{PhcAvgDur} + \alpha_{26} \times \text{HdDel} + \\
&\quad \alpha_{27} \times \text{Prov} + \mu_2
\end{align*}
\]

(7)

It can be observed that the two equations were over-identified; therefore, the 2SLS method was applied. Table 3 contains the coefficients obtained with the equation system using 2SLS.

In the ANOVA analysis, it can be noted that the significance of F was less than 0.000. Therefore, all the independent variables significantly related to the dependent variable. The t statistic and its significance below 0.05 indicate that the variables significantly contributed to improve the quality of the regression model.

The AR for cod300 in PHCs, the coastline location of PHCs and the number of psychiatric stays in hospitals of the PHC population allowed almost 60% of the TDs initiated by PHCs for cod300 to be explained. In the PHCs close to coastlines, the number of TDs was larger due to a higher population density, while more hospital stays for psychiatric disorders reduced TDs, possibly due to the improved treatment of the disease, although its influence was very slight. Clearly, the larger the number of TDs, the higher the AR became. The constant for the dependent variable model leaves the AR non-significant,
indicating that the result would not be valid in those cases when all the explanatory variables took values close to 0.

This AR for cod300 was also due to the fact that the population was located closer to the coastline, and the increase became even greater if PHCs fell within the province of Valencia, the most populated area in the VC. In the PHCs with a higher AR for all diagnoses, the AR for cod300 was also higher, demonstrating that the higher or lower number of TDs prescribed was linked to PHCs’ operation. However, an inverse ratio was also found between the TD duration for all the diagnoses in a PHC and its AR for cod300, evidencing that behaviour differed in terms of the duration of the TDs for cod300 as compared to all the other diseases observed in the descriptive analysis.

Higher quality management by PHC physicians, indicated by a more extensive use of electronic medical records, led to a higher AR for cod300, most probably due to the fact that use of electronic medical records leads to better diagnosed diseases.

Finally, although less explanatory than the remaining variables, the number of days that a first specialised care visit was delayed in all the PHCs of a HD increased the AR for cod300. This was expected since a delay in providing specialised care to patients increases the number of leave days for TD.

4. Discussion

As in other Western European countries [16], an increase in TD for mental disorders is observed, evidencing that an occupational health problem exists. An average of 1.0794% of the working population initiated a TD for cod300, implying 2,347,365 leave days.

The average duration of TDs for cod300 is also longer than for all the other diagnoses made, as other studies have also concluded [10-15]. The introduction of changes into sick leave benefits, economic market positions and company policies may help cut this rate, which occurred in The Netherlands in the years 2004 to 2007 [2].
The variables related to the PHC population have very little impact on TD variability, and do not appear in the simultaneous equations model, unlike what happened in previous studies [2, 6, 15, 18, 19], where they were displaced by the variables relating to PHC and HDs. This demonstrates that it is necessary to carry out a more extensive longitudinal cross-sectional study into the PHCs’ characteristics in order to take measures to reduce the number of TDs, such as increased physicians’ attendance in the PHCs on the coastline and in the province of Valencia, and quicker specialised care services. The reasons why an increased use of electronic medical records leads to a higher AR for cod300 should be extensively explored. Although it is true that diseases may be better diagnosed if such records are employed, it is also true that this might entail a propensity to lengthen leaves in the event of substitutions in these PHCs.

5. Limitations

It was impossible to include the mental health care resource indicators for the Valencian Healthcare System since the available information did not suit the purposes of this study.

In addition, ABUCASIS AIS allowed information on all the TDs for cod300 in 2009 (not previous years), to be collected because it was impossible to obtain data on all the patients with this same diagnosis, regardless of them being prescribed TD leave. Therefore, it was not possible to obtain prevalence and TD measurements for both cod300 and the remaining diagnosis codes for all the patients diagnosed with a previous cod300 disease.

Consequently, it was impossible to compare differences between patients diagnosed with a TD for cod300 and a TD prescribed to patients with other diseases using the MBDS information system.
Additionally, it was not feasible to study the recurrence of TD or its relationship with the study population’s labour and cultural characteristics because information was lacking.

Finally, information was also lacking on the TDs prescribed to the patients included in the Insurance Agency (special scheme for civil servants, the armed forces and court administrative officials -MUFACE, MUGEJU and ISFAS) since their registration, and confirmation and sick leaves were not recorded in the official forms used by the National Healthcare System, but on those forms supplied by these entities. Nonetheless, it is not relevant because this group accounts for a very small percentage (2.73%) of the working population studied.

Acknowledgments: The translation of this paper was funded by the Universitat Politècnica de València, Spain. We are grateful to Joaquín Andani for supplying the database.

6. References


Table 1: Variables used in the model

<table>
<thead>
<tr>
<th>Variables</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prov</td>
<td>Province of the VC takes value 1 if PHC is located in Valencia and 0 otherwise (Castellón or Alicante, these being the other two provinces making up the VC)</td>
</tr>
<tr>
<td>AvgDur</td>
<td>Average duration of all TD leaves for cod300 illnesses finalised</td>
</tr>
<tr>
<td>PhcCoL</td>
<td>Coastal or not coastal location of PHC. It takes value 1 if the PHC has coastline and 0 otherwise</td>
</tr>
<tr>
<td>PhcAR</td>
<td>Number of days of absence due to TD for all diagnoses divides by the total number of days corresponding to the working population</td>
</tr>
<tr>
<td>PhcAvgDur</td>
<td>Average duration of all TD leaves for all illnesses finalized in 2009</td>
</tr>
<tr>
<td>PhcPsiSt</td>
<td>Number of hospital stays resulting from a primary diagnosis included in group 5 of the tabular list of ICD9MC</td>
</tr>
<tr>
<td>PhcUseEMR</td>
<td>Percentage of use of electronic medical records in the PHC</td>
</tr>
<tr>
<td>HdDel</td>
<td>Total number of days a first visit was delayed in specialized health care, in all of the PHCs under each HD.</td>
</tr>
</tbody>
</table>

Table 2. Descriptive statistics for the 575 PHCs and information sources.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP</td>
<td>44.97</td>
<td>18</td>
<td>1</td>
<td>313</td>
<td>AIS</td>
</tr>
<tr>
<td>TD</td>
<td>1.079</td>
<td>0.9887</td>
<td>0.03</td>
<td>3.7556</td>
<td>AID and PIS</td>
</tr>
<tr>
<td>DA</td>
<td>4,082.37</td>
<td>1,627</td>
<td>2</td>
<td>30,866</td>
<td>AIS</td>
</tr>
<tr>
<td>AR</td>
<td>0.25</td>
<td>0.24</td>
<td>0.00</td>
<td>0.92</td>
<td>AID and PIS</td>
</tr>
<tr>
<td>AvgDur</td>
<td>120.40</td>
<td>115.87</td>
<td>2</td>
<td>425.5</td>
<td>AIS</td>
</tr>
<tr>
<td>PhcAR</td>
<td>3.036</td>
<td>3.07</td>
<td>0.09</td>
<td>6.08</td>
<td>AIS</td>
</tr>
<tr>
<td>PhcAvgDur</td>
<td>70.46</td>
<td>66.88</td>
<td>29.33</td>
<td>186.5</td>
<td>AIS</td>
</tr>
<tr>
<td>PhcPsiSt</td>
<td>0.354</td>
<td>0</td>
<td>0</td>
<td>24.138</td>
<td>MBDS</td>
</tr>
<tr>
<td>PhcUseEMR</td>
<td>90.00</td>
<td>91.61</td>
<td>13.36</td>
<td>99.62</td>
<td>*</td>
</tr>
<tr>
<td>HdDel</td>
<td>60.64</td>
<td>62.35</td>
<td>7.7</td>
<td>134.65</td>
<td>*</td>
</tr>
</tbody>
</table>

* Valencia Health Care Agency's Management Agreement Information System for 2009
Table 3. Estimates derived from the equation system.

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Ln (IR +1)</th>
<th>IR+1</th>
<th>AR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>Coefficient</td>
<td>T Statistic</td>
</tr>
<tr>
<td>Constant</td>
<td>0.223</td>
<td>1.250</td>
<td>11.845</td>
</tr>
<tr>
<td>AR</td>
<td>1.687</td>
<td>5.406</td>
<td>24.729</td>
</tr>
<tr>
<td>PhcCoL</td>
<td>0.037</td>
<td>1.038</td>
<td>2.852</td>
</tr>
<tr>
<td>PhcPsiSt</td>
<td>-0.010</td>
<td>0.990</td>
<td>-2.629</td>
</tr>
<tr>
<td>PhcUseEMR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PhcAR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PhcAvgDur</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HdDel</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prov</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.592</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>235.278</td>
<td>0.000</td>
<td>44.309</td>
</tr>
</tbody>
</table>

Source: Compiled by author.