Risk factors for the ill-defined causes of death in the Brazilian states: a multilevel analysis

Fatores de risco associados aos óbitos por causas mal definidas nos estados do Brasil: uma análise multinível

Abstract This study describes the spatial-temporal changes of the proportion of ill-defined causes of death in Brazil (1998-2012) and investigates which demographic and socioeconomic factors affect this proportion. We collected information of the proportion of ill-defined causes of death by age (15-59 years), sex, period, locality, and socioeconomic data. We used a multilevel Poisson model to investigate which factors affect the risk of ill-defined causes of death. Unlike states located in the South and Midwest, we detected clusters with high proportional levels of these deaths in states in the North and Northeast regions. A greater proportion occurred in 1998-2002 (0.09), in the North and Northeast (0.14 and 0.12, respectively), in older age groups (0.09), and in places with poor socioeconomic conditions. The adjusted analysis showed differences in proportion according to the region, age, period, schooling, social inequality, and income. The results indicate that the lower the age group and the better the socioeconomic situation, the lower the risk to register the cause of death as ill-defined. Although over the past years, the quality of Brazil’s mortality data has gradually increased, investments towards improving mortality registries cannot be discontinued.

Key words Mortality, Cause of death, Vital statistics, Mortality registries, Multilevel analysis

Resumo Este estudo descreve as mudanças espaço-temporais da proporção de causas mal definidas no Brasil (1998-2012) e seus fatores associados. Coletamos informações da proporção de óbitos por causas mal definidas por idade (15-59 anos), sexo, período, local de residência, além de fatores socioeconômicos. Utilizou-se modelo multinível de Poisson para investigar os fatores associados às causas mal definidas dos óbitos. Ao contrário dos estados do sul e centro-oeste, identificou-se clusters com elevados níveis proporcionais destes óbitos nos estados do norte, nordeste e sudeste. A maior proporção de óbitos mal definidos ocorreu em 1998-2002 (0.09), no norte e nordeste (0.14 e 0.12, respectivamente), nos grupos etários mais velhos (0.09) e nos locais com condições socioeconômicas desfavoráveis. A análise ajustada indicou diferença dos níveis de causas mal definidas de acordo com a região, a idade, o período, a escolaridade, a desigualdade social e a renda. Nosso resultados sugerem que quanto menor a faixa etária e melhores as condições socioeconômicas, menor o risco de óbitos mal definidos. Apesar de nos últimos anos a qualidade dos dados de mortalidade no Brasil ter aumentado, os investimentos na melhoria do registro dos óbitos não podem ser descontinuados.

Palavras-chave Mortalidade, Causas de morte, Estatísticas vitais, Registros de mortalidade, Análise multinível
Introduction

Improving vital registrations is a priority for the World Health Organization (WHO)\(^1\) In 2003, only one quarter of the WHO member nations had high quality of vital statistics, while most countries (51%), including Brazil, had intermediate quality, covering between 70% and 90% of the country’s deaths or showing percentage of ill-defined causes between 10% and 20%\(^2\).

In the 2008-2011 period, 30% of WHO member nations had high-quality death registrations. The United Kingdom, United States of America and Chile figured among these countries, but Brazil still showed 76% of average usability of mortality data for this period\(^3\). In Brazil, the quality of the Mortality Information System (SIM) is slowly improving\(^4,5\). Some authors, which investigated the underlying cause of death related with neoplastic diseases, showed high reliability of the registered information\(^6,7\). Causes classified as ill-defined refer to deaths without medical care and those, which, although had been provided medical care, the cause of mortality could not be determined. Deaths classified as ill-defined affect the quality of data, reducing its reliability. The lack of accuracy of the mortality data makes the adequate allocation of health resources difficult\(^8,9\).

The main hypothesis of the present research is the relationship between the proportion of ill-defined causes of death and the sociodemographic characteristics. Sociodemographic determinants have been cited in much of the research related to mortality registrations\(^10-12\). Using the Brazilian federative units (states), we evaluate the proportion of ill-defined causes of death along the 1998-2012 period, through descriptive techniques and statistical modeling. To identify the spatial profile of the proportion of ill-defined causes of death and evaluate possible spatial changes over time, we also apply spatial analysis techniques. Lima e Queiróz\(^13\), who performed statistical techniques in order to investigate the distribution of ill-defined causes of death in Minas Gerais (Southeast of Brazil), identified spatial clusters in the North and Northeast of the state from 1980 to 2007.

Considering the importance of reliable government-provided mortality information to support appropriate sectoral actions and the high social inequality in the Brazilian states, this study describes the spatial and temporal changes of the proportion of ill-defined causes of death in Brazil and investigates which demographic and socioeconomic factors affect this proportion over the 1998-2012 period.

Methods

This study evaluates Brazilian data from 1998 to 2012 using the country’s member states as units of analysis.

Brazil is the largest country of South America with significant regional socioeconomic differences. While the Human Development Index in the Southeast, South and Midwest regions of the country varies around 0.75-0.76, in the North and Northeast, it ranges from 0.66-0.67\(^14\).

The South region includes Paraná, Santa Catarina, Rio Grande do Sul states; the Southeast region, São Paulo, Minas Gerais, Rio de Janeiro, and Espírito Santo; the Midwest region, Mato Grosso do Sul, Mato Grosso, Goiás, and the Federal Distrito; the North region, Acre, Rondônia, Amazonas, Roraima, Pará and Amapá; the Northeast region, Bahia, Piauí, Maranhão, Ceará, Rio Grande do Norte, Paraíba, Pernambuco, Alagoas, and Sergipe states (Figure 1).

Data was collected from information on death records of those aged 15 to 59 years from the Mortality Information System (SIM) based on death certificates entered into the system. To avoid confounding factors due the distinct diseases, which prevail in each stage of life, we selected only the population aged 15-59. The classification of ill-defined causes of death followed the definition of the second volume of 10th Revision (ICD-10), which includes: “symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified” (codes R00-R99, except R-95), “ill-defined circulatory disorders” (codes 146.1, 146.9, 195.9, 199), “ill-defined respiratory disorders” (codes J96.0, J96.9), and “ill-defined conditions originating in the perinatal period” (code P28.5)\(^15\).

Using the number of ill-defined causes of death as the numerator and the total number of deaths as the denominator, we calculated the proportion of ill-defined causes of death by sex, age group and period. Information about Gross Domestic Product (GDP) per capita/1,000 dollar (1 real = 0.27 dollars), Gini index, and illiteracy rate in each state were also collected from DATA-SUS website for each period\(^16,17\).

Statistical analysis

In this paper, we apply spatial analysis techniques. We used the georeferenced mesh of states of Brazil (shapefile file), available from the “Brazilian Institute of Geography and Statistic” website\(^18\).
To verify if there is spatial dependence, we used Global Moran I and pseudo-significance test for asymmetric data, which tested whether the proportion of ill-defined causes of death showed greater similarity in contiguous Brazilian states than would be expected by chance.

The Global Moran index can be defined by the formula below:

$$I = \frac{1}{n \Sigma (y_i - \bar{y})^2} \sum_{i} \sum_{j} w_{ij} (y_i - \bar{y})(y_j - \bar{y})$$

where, $w_{ij}$ is an element of contiguity matrix $\omega$; $y_i$ is the proportion of state $i$; $y_j$ is the proportion of the state $j$; $\bar{y}$ is the sample mean and $n$ is the total number of states.

To evaluate the presence of potentially distinct spatial correlation regimes in the proportion of ill-defined causes of death in different Brazilian states, we plotted cluster maps using Local Indicators of Spatial Association (LISA) and used a Queen matrix to compute the spatial statistics.

The Local Moran index can be defined by the formula below:

$$I_i = \frac{n}{\Sigma (y_i - \bar{y})^2} \sum_{j} w_{ij} (y_i - \bar{y})(y_j - \bar{y})$$

where, $z_i = y_i - \bar{y}$; $z_j = y_j - \bar{y}$; $\omega_{ij}$ is an element of contiguity matrix $\omega$; $y_i$ is the proportion of state $i$; $y_j$ is the proportion of the state $j$; $\bar{y}$ is the sample mean and $n$ is the total number of states.

For each period, we plotted Moran scatter graphs to evaluate the spatial association of the proportion of ill-defined causes of death in each state with the average proportion of its neighbors. To perform this analysis, the normalized values of the proportion of ill-defined causes of death in each state are compared to the mean values of its neighbors, which are divided in four quadrants that provide a classification of four types of spatial autocorrelation: direct spatial autocorrelations (Q1: high-high; Q2: low-low) and inverse spatial autocorrelations (Q3: high-low; Q4: low-high). The points of Q3 and Q4 can be seen as units that do not follow the same spatial dependence process of the area.

We used statistical tests to access the association between ill-defined causes of death and the following variables: sex, age group, period, region, GDP per capita, Gini index, and illiteracy rate. We used a multilevel Poisson regression model to explain the occurrence of ill-defined causes of death (response variable).

The model in the first level included the following factors: sex, age group, period, region, GDP (logarithm), Gini Index and illiteracy rate, and, as the “offset”, the total deaths (logarithm) of each specific population. In the second level, we included the Brazilian states. Poisson regression is frequently used to model count data, but it is often inadequate for over-dispersion situations. Therefore, we applied the multilevel Poisson regression to circumvent over-dispersion in the study.

The selection of the variables included in the model was based on previous scientific knowledge and the availability of the indicators on government databases. All analyses were performed with TerraView (version 4.2.2.), GeoDa (version 1.4.6) and R-Project (version 3.2.4) softwares.

This study did not require submission to the Ethics and Research Committee (i.e. an Institutional Review Board), as it uses public access...

Results

From 1998 to 2012, there was a reduction in the proportion of deaths classified as ill-defined in Brazil. The proportion of deaths with ill-defined causes in the periods 1998-2002, 2003-2007 and 2008-2012, were 0.09, 0.08 and 0.07, respectively. Both men and women show the same proportion over the period (0.08). The highest proportion of ill-defined causes of death occurred in the North and Northeast regions, 0.14 and 0.12, respectively, while the lowest occurred in the South and Midwest, 0.5 and 0.6, respectively (Table 1).

Age strongly correlates with the proportion of ill-defined causes of death. The proportion of the older group (50-59 years) was greater than double compared to the younger age groups (15-19 and 20-29 years). The states with GDP per capita US$ 8,100 or greater showed almost half of the share of ill-defined causes of death compared to those with lower per capita GDP. The states with Gini index 0.62 or greater had 50% higher proportional levels of ill-defined causes of death than those with lower levels of inequality. The states with illiteracy rates 11.60 or greater also showed proportional levels of ill-defined causes of death 50% higher than the others (Table 1).

The highest Global Moran’s I index was observed in the 1998-2002 period (0.28, p-value < 0.04). In the other periods, the Moran’s I index was not statistically significant. As expected, the scatterplots showed that most Brazilian states are located in Q1 (high-high) and Q2 (low-low) quadrants (Figure 2). These points represent units with low (Q2) or high (Q1) incidence rates following the same spatial dependence process of the area.

The LISA Cluster maps, Figure 3, highlight three low-low areas comprised of one state in the South (Paraná) in 1998-2002 and two others in the Midwest (Mato Grosso and Goiás states) in 2003-2007 (i.e. units with low incidence rates following the same spatial dependence process of the area). We found the most high-high areas in the Northeast and North regions (i.e. units with high incidence rates following the same spatial dependence process of the area). However, the map of 2008-2012 highlights one important Brazilian state in the Southeast (São Paulo) as a significant high-high area.

We could detect only four significant low-high areas. Although these areas are close to areas with high proportions, they have lower proportions of ill-defined causes of death (Figure 3).

In Table 2, we investigate which factors affect the risk of ill-defined causes of death. There was not an association between sex and the occurrence of ill-defined causes of death. The adjusted analysis also did not show a difference between the proportion of ill-defined causes of death between the Midwest (reference category) and South regions. However, the risks of ill-defined causes of death in the Southeast, North and Northeast regions were 1.87, 2.05 and 2.08 times greater, respectively, than that observed in the Midwest region. The adjusted analysis showed an increasing trend in the risk of ill-defined causes of death as age increases. The proportion of ill-defined causes of death was about two times greater in the older group than that of the younger. The adjusted analysis suggests a decreasing trend in the proportion of ill-defined causes of death...
death over the entire period under study. The risk of ill-defined causes of death was 37% lower in the last period than that of the first one.

In the adjusted analysis, both income indicators, GDP and Gini index, were associated with the levels of ill-defined causes of death. The greater the GDP of a state, the lower the risk of ill-defined causes of death. The greater the Gini index, the greater the risk of ill-defined causes of death. Finally, there is a direct association between the illiteracy rate and the risk of ill-defined causes of death. The greater the illiteracy rate, the greater the proportion of ill-defined causes of death (Table 2).

Figure 2. Local Moran’s I scatter plot of the proportion of ill-defined causes of death in Brazilian states by period.

The quadrants Q1, Q2, Q3 and Q4 represent the four types of spatial autocorrelation: high-high (Q1), low-low (Q2), high-low (Q3) and low-high (Q4).

Figure 3. Lisa maps of the average of the proportion of ill-defined causes of death in Brazilian states by period.
Table 2. Adjusted relative risk of ill-defined causes of deaths according to the period and demographic and socioeconomic factors.

<table>
<thead>
<tr>
<th>Categories</th>
<th>RR*</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Midwest</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>North</td>
<td>2.05</td>
<td>0.0003</td>
</tr>
<tr>
<td>Northeast</td>
<td>2.08</td>
<td>0.0001</td>
</tr>
<tr>
<td>Southeast</td>
<td>1.86</td>
<td>0.005</td>
</tr>
<tr>
<td>South</td>
<td>1.29</td>
<td>0.26</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0.99</td>
<td>0.56</td>
</tr>
<tr>
<td>Age group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15-19</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>20-29</td>
<td>1.10</td>
<td>0.00001</td>
</tr>
<tr>
<td>30-39</td>
<td>1.70</td>
<td>0.00001</td>
</tr>
<tr>
<td>40-49</td>
<td>2.00</td>
<td>0.00001</td>
</tr>
<tr>
<td>50-59</td>
<td>2.04</td>
<td>0.00001</td>
</tr>
<tr>
<td>Period</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1998-2002</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>2003-2007</td>
<td>0.80</td>
<td>0.00001</td>
</tr>
<tr>
<td>2008-2012</td>
<td>0.62</td>
<td>0.00001</td>
</tr>
<tr>
<td>GDP1</td>
<td>0.86</td>
<td>0.00001</td>
</tr>
<tr>
<td>Gini Index</td>
<td>11.57</td>
<td>0.00001</td>
</tr>
<tr>
<td>Illiteracy rate</td>
<td>≤ 11.60</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>≥ 11.60</td>
<td>1.06</td>
</tr>
</tbody>
</table>

RR = Relative Risk; GDP = Gross Domestic Product (the conversion real to dollar considered: 1 real = 0.27 dollars); 1Logarithm of GDP. *We used a multilevel Poisson regression model to calculate risk of ill-defined causes of death. The model included in the first level: sex, age group, period, region, logged GDP, Gini Index, illiteracy rate and, as the “offset”, the logarithm of the total deaths of each specific population. In the second level, we included the state.

Discussion

The findings indicated that ill-defined causes of death in Brazil decreased gradually in recent years. Other Brazilian investigations also showed reductions in the proportion of ill-defined causes of death over the same period. Costa and Marcopito highlighted a reduction of 25% in the percentage of ill-defined causes of death in Brazil from 1979 to 2002. Some of the explanations provided by the authors for this reduction was the proportional increase of external causes of death and higher determination of the natural causes of death. De Lima and Queiroz (2014) reported that the percentage of ill-defined causes of death in Brazil fell 53% over 1991-2010 period. The authors reported that this reduction probably occurred due to the successful efforts of the Brazilian government to improve data quality.

Our results indicated an association between the region where the death occurred and the proportion of ill-defined causes of death. The South region showed the lowest proportion of ill-defined causes of death of the country. This probably occurred because this region presents favorable socioeconomic conditions. In the South, we found the lowest illiteracy rates (5.9) and Gini index (0.56) for the period. Except for the South, all other regions showed a higher proportional risk of ill-defined causes of death compared to the Midwest. We also detected clusters of high level of ill-defined causes of death in North and Northeast until 2007 compared to the South and Midwest where we found low proportion clusters. Martins Junior et al. also highlighted a reduction of about 80% in ill-defined proportion causes of death in the Northeast from 1979 to 2009. Similarly, other researchers pointed that the highest percentages of ill-defined causes of death are located in the North and Northeast regions. Costa and Marcopito found that, from 1979 to 2002, the percentages of ill-defined causes of death in several Brazilian regions fell significantly. The highest levels of ill-defined causes of death were detected in some states in the Northeast (37-58% in 2002). According to the authors, the variation of the proportion of ill-defined causes of death along the country is due to the distinct socioeconomic development stages of the different Brazilian areas.

In last period, the proportion of ill-defined causes of death in Sao Paulo state was higher compared to its Southern and Midwest neighbors. However, Sao Paulo showed a lower proportion of ill-defined causes of death compared to its neighbors in the southeast region (Rio de Janeiro and Minas Gerais states). In fact, Minas Gerais showed the highest proportion from 2007-2012 (0.11). Although Minas Gerais improved its coverage of registering mortality, the proportion of ill-defined causes of death increased substantially from 1980 to 2007.

Except in the South, in other Brazilian regions we detect some states with low level of ill-defined causes of death close to others with high level, classified as low-high areas. No one of these states became significant high-high areas during this period.

The proportion of ill-defined causes of death in Brazil was similar in both sexes. In Sao Paulo, the State Health Secretary reported in 2010 that the proportion of ill-defined causes of death was almost the same in both sexes (around 0.05 to 0.06).

Our findings detected that proportional values of ill-defined causes of death were directly associated with age. Other Brazilian studies also showed this association. Considering that
the higher the age group, the greater the proportion of ill-defined causes of death, non-communicable chronic diseases, which are more prevalent in older ages, seem to have a significant role on the proportional levels of ill-defined causes of death32,33.

The results also indicated that the higher education level of the state’s population, the lower the proportion of ill-defined causes of death. Some authors corroborated these findings36,39. Teixeira et al.39 indicated that the proportion of ill-defined causes of death for those who do not complete high school or graduate is twice those who do.

The state based income levels also can increase the proportion of ill-defined causes of death. As we observed in this study, both low per capita GDP contributed to increase this proportion, as well as high inequality (represented by Gini index). Kanso et al.38 also showed a declining trend in the proportion of ill-defined causes of death as GDP increases.

Our study contributes to our understanding of Brazilian public health indicators; however, the results must be analyzed with caution, considering the variation of coverage in the quality of data within each state and region. The method used in this study assumes that the population under the study are demographically stable inside each region, which does not reflect the complexity of our real world. As is discussed in the scientific literature, some local factors can affect the amount of ill-defined causes of death. The most violent places where there are more external causes of death, for example, probably have a greater proportion of ill-defined causes of death than others40. According to official data, the most violent Brazilian states are located in the Northeast region in the following states, Alagoas, Ceará, Rio Grande do Norte, Sergipe and Bahia. In 2013, the homicide rates in these states were higher than 50 deaths per 100,000 inhabitants46.

Although the completeness of death records in Brazil have improved in recent years, there are still wide regional variations in the distribution of ill-defined causes of death43; with significant differences across states and regions over time41. The greatest percentage of incomplete records is observed in the North and Northeast regions, while the most complete is in the South42. In 2010, the vital registration data system covered less than 80% of the deaths in the following states, Tocantins (North) and Maranhão (Northeast), and more than 90% of the deaths in Distrito Federal (Midwest), Paraná (South), Rio Grande do Sul (South), Rio de Janeiro (Southeast), São Paulo (Southeast) and Espírito Santo (Southeast)42.

Many efforts have been made to improve the quality of Brazilian data. Among these tools, we should highlight the implementation of investigation programs of ill-defined causes of death in each city, which corrects and redistributes data by researching deaths. Another important tool is to link databases in order to correct ill-defined mortality registers4.

The use of the state as the unit of analysis also represents a limitation. Others studies using smaller areas as units of analysis could help to identify the main locales where governmental data improvement programs should be focused. Although several small municipalities have already improved their data systems for tabulating death registrations, there are still several gaps like the high turnover of hired professionals and the lack of professional qualified to perform systematic registrations of local death data43. Besides, we must consider the ecological fallacy, which makes impossible deriving any inferences in individual level problematic. Another limiting factor of this study was the official information which was available through government sources. Other socioeconomic information in individual level could also have had contributed to explaining the levels of ill-defined causes of death. Despite of these limitations, the present study makes important contributions about the quality of data in Brazil in recent years and indicates that mortality data has gradually improved in recent decades.

Although managers, physicians, and other professionals have made a positive impact in this improvement, there is still a lot of work to do. Nevertheless, some efforts – investigations of the ill-defined deaths and the use of linkages techniques to compare different databases and correct possible inconsistencies among the data – are already performed in order to gradually decrease the proportion of ill-defined causes of death throughout the country. If these initiatives are maintained, Brazil’s vital registration system should soon reach adequate quality, according to international standards1.

Conclusion

The results of the present study indicated that the highest proportions of ill-defined causes of death are found in the North and Northeast regions of the country. However, these values are falling over the time. Our findings suggest that the lower
the age group and the better the socioeconomic condition, the lower the risk of registering the cause of death as ill-defined.

Although over the past years, there is better quality of mortality data in Brazil, investments targeting improvements in mortality registration cannot be discontinued, especially in Brazil's North and Northeast regions, where data quality remains subpar.

Collaborations

NCP Rodrigues conceived the study, participated in its design and coordination, helped to draft the manuscript, performed the statistical analysis, read and approved the final manuscript. AS Almeida helped in the statistical analysis, read and approved the final manuscript. RP Daumas, G O'Dwyer, MKN Andrade, MB Flynn and VTS Lino helped to draft the manuscript, read and approved the final manuscript.
References