ABSTRACT: Objectives: To estimate the prevalences of hypertension and diabetes for small areas in Belo Horizonte, according to the Health Vulnerability Index (HVI). Methods: Ecological study with data from the Surveillance of Risk and Protection Factors for Chronic Diseases by Telephone Survey (Vigitel) conducted in Belo Horizonte, from 2006 to 2013. The self-reported diagnosis of diabetes and hypertension were evaluated. The estimates of prevalence and the 95% confidence interval (95%CI) were calculated using the direct and indirect method by HVI grouped into four categories: low, medium, high and very high health risk. Results: During the period evaluated, 26% (95%CI 25.2 – 26.8) and 6.1% (95%CI 6.7 – 6.5) of the adult population from Belo Horizonte reported being hypertensive and diabetic, respectively. According to the indirect method to obtain estimates of hypertension and diabetes prevalences per HVI, it was found that areas of very high risk had a higher prevalence of adults with hypertension (38.6%; 95%CI 34.8 – 42.4) and diabetes (16.2%; 95%CI 13.1 – 19.3) when compared to the low risk (28.2%; 95%CI 27.0 – 29.4 and 6%; 95%CI 5.4 – 6.7, respectively). Conclusion: The adult population living in areas at high risk for health had a higher prevalence of hypertension and diabetes compared to those with a lower risk. Keywords: Diabetes mellitus. Hypertension. Small-area analysis. Logistic models. Health surveys.
INTRODUCTION

Noncommunicable diseases (NCD) are the main causes of morbidity and mortality worldwide, in addition to causing premature deaths, disabilities and high costs and financial burdens for individuals, societies, and health systems. In Brazil, NCD accounted for 75% of the causes of death in 2015, with four major groups of diseases standing out: cardiovascular, cancer, diabetes, and chronic respiratory diseases.

Studies point out that these diseases are multifactorial and emphasize the importance of social determinants in their occurrence, with worse indicators in the most vulnerable and socially marginalized population. An example is the fact that NCD more frequently affect low-income populations, as they are more vulnerable, exposed to risks, and have less access to health services and health promotion and disease prevention practices.

NCD and their risk factors have been monitored in the country by large national surveys, which are essential to aggregate evidence on individual and collective health; however, national surveys do not allow estimates by municipalities or for smaller areas within the sampled cities. Thus, it is necessary to advance the monitoring of these diseases in small geographic areas, in order to identify the existing inequalities, which are often masked by national or state average statistics.

Estimates for small areas can help managers in public health in planning, defining intervention priorities and allocating resources, in addition to identifying possible decentralized solutions to the diagnosed problems, which may include social, cultural, legal, political or health-related approaches in small areas. The use of composite indicators in the health...
field, such as the Health Vulnerability Index (HVI), makes it possible to identify health inequalities and helps in the redesign of health care network and promotion of population development in various geographic scales of action, contributing to guide public policies in the definition of priorities for the allocation of resources\textsuperscript{9,10}.

The present study investigated arterial hypertension (AH), as it is considered the most prevalent risk factor for cardiovascular diseases, being responsible, in Brazil, for 2,283.48 (95\%CI 2,050.77 – 2,496.19) disability-adjusted life years (DALYs)/100 thousand, in 2019\textsuperscript{11}; and diabetes mellitus (DM), for being responsible for 1,076.51 (95\%CI 915.36 – 1,256.42) DALYs/100 thousand, in 2019\textsuperscript{11}, in addition to both presenting high prevalence in the Brazilian population\textsuperscript{6,7}. Thus, the study aimed to estimate the prevalences of AH and DM in small areas of Belo Horizonte, according to the HVI.

**METHODS**

This is an ecological study that used data from the Surveillance of Risk and Protection Factors for Chronic Diseases by Telephone Survey (Vigilância de Fatores de Risco e Proteção para Doenças Crônicas por Inquérito Telefônico – Vigitel) system carried out in Belo Horizonte, from 2006 to 2013. More detailed information on data collection and Vigitel samples are available in a previous publication\textsuperscript{12}.

The present study evaluated the self-reported medical diagnosis of DM and AH through positive replies to the questions:

- Has any doctor ever told you that you have diabetes?
- Has any doctor ever told you that you have hypertension?

The HVI was considered in this study as the “small areas”. It was developed in 1998 and updated in 2012 with data from the 2010 census, by the Municipal Health Secretariat of Belo Horizonte to point out priority areas for intervention and allocation of resources, in addition to allowing the analysis of characteristics of population groups residing in census tracts\textsuperscript{10}. It is a composite indicator that combines socioeconomic (residents per household, percentage of illiterate people, percentage of private households with a per capita income of up to half the minimum wage, average nominal income of the head of the household, percentage of mixed and black people, and indigenous) and environmental variables (sewage, water supply, and solid waste destination).

HVI ranged from 0 to 1 — values close to 1 indicate high social vulnerability and values close to 0 indicate low or nonexistent social vulnerability — and is categorized according to the following cut-off points:

- low: $\text{HVI} \leq 0.1957$
- medium: $0.1957 < \text{HVI} \leq 0.2865$
- high: $0.2865 < \text{HVI} \leq 0.3782$
- very high: $\text{HVI} > 0.3782$. 

3

REV BRAS EPIDEMIOL 2021, 24: E210015.SUPL.1
In 2012, the 3,830 sectors of Belo Horizonte were grouped into the following categories:
- low risk: 1,330 – 34.7%;
- medium risk: 1,460 – 38.1%;
- high risk: 737 – 19.2%;
- very high risk: 303 – 7.9%.

In order to include the census tracts in the Vigil databases, a linkage was made with the National Register of Addresses for Statistical Purposes of the 2010 census by Postal Codes. Then, the HVI information by census tract was inserted. This procedure was performed in a data center with a high level of security — physical and virtual.

The prevalence of DM and AH and the respective 95% confidence intervals (95%CI) were estimated, according to the HVI, using the direct and indirect estimation methods for small areas.

The direct estimation method consists of using the sample design variables to obtain the estimates in a smaller area. For joint analysis of Vigil data, it was necessary to calculate post-stratification weights, using the rake method, to adjust the distribution of the Vigil sample to the 2010 census population by HVI. The weights were calculated in the STATA program using the SURVWGT package, requiring the sample weight information to execute the package:

\[
\text{weight} = \frac{\text{number of adults in the household}}{\text{number of telephones in the household}}
\]

The indirect estimation method consists of using statistical models to input the variable dichotomous response (Y), in the set of census tracts without any Vigil interview. Of the total of 3,830 census tracts in Belo Horizonte, 513 (13.3%) had no Vigil interview. Therefore, imputations of the outcome variables in these sectors were necessary.

For the construction of the logistic regression model, the census tracts with a single interview were selected (535 sectors). This criterion is due to the similarity in the distribution of sectors according to HVI in the group with one and no Vigil interview. The sample of 535 sectors was divided into two subsamples of equal sizes, the first (sample 1) being used for the development of the logistic regression model and the second (sample 2) for its validation, in order to ensure that the model obtained in the first sample was robust.

The following covariables from the census tract, extracted from the 2010 census, were used to build the models: percentage of households without water supply, percentage of households without a bathroom, percentage of households with literate people, percentage of households with females, percentage of households with residents over the age of 55, percentage of households with non-white residents, percentage of households with income above one minimum wage, percentage of households with six or seven residents, percentage of households with eight or more residents.

For the imputation of the outcome variable in the set of sectors without any Vigil interview, a cutoff point of 0.260 was considered for AH and 0.062 for DM. The adult resident in
the sector with a predicted probability greater than or equal to the cutoff point was classified as 1 or 0, otherwise. This cut-off point was obtained in the group of census tracts with a single Vigitel interview.

To evaluate the model performance, the 2 × 2 classification matrix was used, with the proportion of correct classification given as: the answer is AH diagnosis and the model classified the subject as hypertensive, so it was classified as true positive (TP); the subject is not hypertensive and the model classified them as non-hypertensive, so they were classified as true negative (TN). On the other hand, the proportion of incorrect classification was given as: the response category is hypertensive and the model classified the individual as non-hypertensive, that is, as a false negative (FN); and the response category is non-hypertensive and the model classified it as hypertensive, called false positive (FP). The sensitivity of the model was defined by \( \frac{TP}{TP + FN} \), specificity by \( \frac{TN}{TN + FP} \), and accuracy by \( \frac{TP + TN}{TP + FN + TP + FN} \).

After imputing data on outcomes in the census tracts without interviews, the prevalence of outcomes by HVI was estimated. Post-stratification weights were calculated to adjust the sample distribution for the 2010 census population by HVI, using the rake method. These weights were calculated in the STATA version 14.0 program using the SURVWGT package, requiring the sample weight information to execute the package. In this study, data from the N1, N2, and N populations extracted from the 2010 Belo Horizonte census were considered for calculating the weight of the group of sectors with and without Vigitel interviews — \( \text{weight} = \frac{N}{N_1} \) and \( \text{weight} = \frac{N}{N_2} \), respectively —, with N being the total adult population aged 18 years old and older, \( N_1 \) the total of adults in the sectors with Vigitel interviews, and \( N_2 \) the total of adults in the sectors without Vigitel interviews.

More detailed information on the direct and indirect methods of estimation for small areas employed can be obtained in another publication.

To assess the differences between the prevalence of AH or DM by IVS, the Student’s t-test was used, with a significance level of 5%.

The present study was developed according to Resolution No. 466/2012, of the National Health Council, and is integrated with the research project entitled “Inequalities in small geographical areas of the indicators of noncommunicable diseases, violence, and its risk factors (Desigualdades em pequenas áreas geográficas dos indicadores de doenças crônicas não transmissíveis, violências e seus fatores de risco), approved by the Research Ethics Committee of Universidade Federal de Minas Gerais.

RESULTS

Between 2006 and 2013, Vigitel interviewed 15,833 adults living in the city of Belo Horizonte, with 14,174 (90%) geocoded interviews.

During the study period, 26% (95%CI 25.2 – 26.8) and 6.1% (95%CI 6.7 – 6.5) of the adult population in Belo Horizonte reported being hypertensive and diabetic, respectively. The prevalences over the years are shown in Table 1.
According to the adjusted logistic regression model, the chance of an adult being classified as hypertensive is 0.006 (exponential -5.103). Three out of a total of nine variables showed associations with AH. The percentage of households with eight or more residents and the percentage of households with females were associated with an increased likelihood that an adult in the census tract would be classified as hypertensive. In contrast, the variable percentage of households with residents over the age of 55 decreased the likelihood of an adult in the census tract being classified as hypertensive (Table 2).

The chance of an adult being classified as diabetic is 0.008 (exponential of -4.745). Three variables were associated with DM — two (percentage of households with eight or more residents and percentage of households with non-white residents) increased and one (percentage of households with six or seven residents) decreased the likelihood of an adult from the census tract to be classified as diabetic (Table 2).

In assessing the adjusted model for AH, the cut-off point of 26% was used to classify the census tracts without an interview as hypertensive or non-hypertensive. In sample 1, sensitivity was 64%, specificity 53%, and accuracy 56%. In sample 2, the sensitivity, specificity, and accuracy values were 51, 55, and 54%, respectively. In assessing the accuracy of the adjusted model, it was noted that the results of the two samples are very close. This indicates that the model is consistent.
In sample 1, the model performed better in the hypertensive than in the non-hypertensive category, while in sample 2 the performance in the non-hypertensive category was greater when compared to the result of sample 1. On the other hand, in sample 2, it was observed that the performance in the hypertensive category was lower than that found in sample 1. This variation was expected due to the use of one of the samples to build the model. In general, the performance of the model’s correctness in the validation sample is lower when compared to the result of the training sample. These results suggest that the adjusted logistic regression model is consistent and reasonably accurate, 56 and 54%, for samples 1 and 2, respectively (Table 3).

For DM, the cut-off point of 6.2% was used to classify census tracts without interview as diabetic or non-diabetic. In sample 1, the sensitivity was equal to 62%, with specificity of 65% and accuracy of 64.8%. In sample 2, sensitivity, specificity, and accuracy were 53, 67, and 66.4%, respectively. In assessing the accuracy of the adjusted model, it was noted that the results of the two samples are very close. This indicates that the model is consistent.

In sample 1, the model performed well in the non-diabetic and diabetic classification, while in sample 2 the performance in the non-diabetic category was higher when compared to the result of sample 1. On the other hand, it was observed that the performance in the diabetic category reduced in sample 2. This variation was expected by the use of one of the samples to build the model. In general, the performance of the model’s correctness in the validation sample is lower when compared to the result of the training sample. These results

Table 2. Analysis of logistic regression for hypertension and diabetes mellitus, Belo Horizonte, MG, Brazil. Vigil, 2006 to 2013.

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Coefficient</th>
<th>SE</th>
<th>p-value</th>
<th>Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hypertension</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage of households with eight or more residents</td>
<td>0.119</td>
<td>0.071</td>
<td>0.093</td>
<td>1.13</td>
</tr>
<tr>
<td>Percentage of households with females</td>
<td>8.436</td>
<td>4.811</td>
<td>0.079</td>
<td>4.610.0</td>
</tr>
<tr>
<td>Percentage of households with residents over the age of 55</td>
<td>-0.008</td>
<td>0.003</td>
<td>0.013</td>
<td>0.99</td>
</tr>
<tr>
<td>Constant</td>
<td>-5.103</td>
<td>2.574</td>
<td>0.047</td>
<td>0.006</td>
</tr>
<tr>
<td><strong>Diabetes mellitus</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage of households with six or seven residents</td>
<td>-0.178</td>
<td>0.083</td>
<td>0.032</td>
<td>0.83</td>
</tr>
<tr>
<td>Percentage of households with eight or more residents</td>
<td>0.238</td>
<td>0.152</td>
<td>0.119</td>
<td>1.26</td>
</tr>
<tr>
<td>Percentage of households with non-white residents</td>
<td>4.027</td>
<td>2.042</td>
<td>0.048</td>
<td>56.1</td>
</tr>
<tr>
<td>Constant</td>
<td>-4.745</td>
<td>1.237</td>
<td>0.000</td>
<td>0.008</td>
</tr>
</tbody>
</table>

SE: standard error; Vigil: Surveillance of Risk and Protection Factors for Chronic Diseases by Telephone Survey (Vigilância de Fatores de Risco e Proteção para Doenças Crônicas por Inquérito Telefônico).
indicate that the adjusted logistic regression model is consistent and has good accuracy, 64.8 and 66.4% for samples 1 and 2, respectively (Table 3).

Tables 4 and 5 compare the prevalence estimates for AH and DM according to the HVI, using the direct and indirect methods of estimation in small areas. According to the indirect method, in the 2006 to 2013 period, the areas of very high risk had a higher prevalence of hypertensive adults (38.6%; 95%CI 34.8 – 42.4) when compared to the low (28.2%; 95%CI 27.0 – 29.4) and medium risk ones (33.4%; 95%CI 32.1 – 34.6). In turn, the areas of very high risk had a higher prevalence (16.2%; 95%CI 13.1 – 19.3) of DM when compared to the medium (9.7%; 95%CI 8.9 – 10.5) and low risk areas (6%; 95%CI 5.4 – 6.7). In the direct method the prevalences did not differ statistically among the HVI strata, and it was observed that the prevalence of AH ranged from 26% (95%CI 24 – 27) for low risk to 30% (95%CI 28 – 32) for high risk, and DM ranged from 5.9% (95%CI 5.2 – 6.6) to 7.7% (95%CI 6.5 – 9.0).
The present study estimated the prevalence of AH and DM, according to the HVI, in Belo Horizonte. The direct and indirect methods of estimation in small areas were compared. It was found that areas with high risk had a higher prevalence of AH and DM compared to areas with low risk. The table below shows the estimated prevalence of hypertension and diabetes mellitus according to the health vulnerability index using the direct and indirect estimation methods in small areas, Belo Horizonte, MG, Brazil. Vigil, 2006 to 2013.

Table 4. Estimated prevalence of hypertension and diabetes mellitus according to the health vulnerability index using the direct and indirect estimation methods in small areas, Belo Horizonte, MG, Brazil. Vigil, 2006 to 2013.

<table>
<thead>
<tr>
<th>HVI</th>
<th>Hypertension</th>
<th></th>
<th>Diabetes mellitus</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Indirect method</td>
<td>Direct method</td>
<td>Indirect method</td>
<td>Direct method</td>
</tr>
<tr>
<td>Low</td>
<td>28.2 (27 – 29.4)</td>
<td>26 (24 – 27)</td>
<td>6 (5.4 – 6.7)</td>
<td>5.9 (5.2 – 6.6)</td>
</tr>
<tr>
<td>Medium</td>
<td>33.4 (32.1 – 34.6)</td>
<td>30 (29 – 32)</td>
<td>9.7 (8.9 – 10.5)</td>
<td>7.7 (7.0 – 8.5)</td>
</tr>
<tr>
<td>High</td>
<td>35.8 (33.7 – 37.8)</td>
<td>30 (28 – 32)</td>
<td>12.8 (11.3 – 14.3)</td>
<td>7.7 (6.5 – 9)</td>
</tr>
<tr>
<td>Very high</td>
<td>38.6 (34.8 – 42.4)</td>
<td>29 (25 – 33)</td>
<td>16.2 (13.1 – 19.3)</td>
<td>7.6 (5.3 – 10)</td>
</tr>
</tbody>
</table>

HVI: health vulnerability index; 95%CI: 95% confidence interval; Vigil: Surveillance of Risk and Protection Factors for Chronic Diseases by Telephone Survey (Vigilância de Fatores de Risco e Proteção para Doenças Crônicas por Inquérito Telefônico).

Table 5. p-value of the differences between the prevalence of hypertension and diabetes mellitus between the categories of the health vulnerability index, Belo Horizonte, MG, Minas Gerais. Vigil, 2006 to 2013.

<table>
<thead>
<tr>
<th>HVI</th>
<th>Hypertension</th>
<th></th>
<th>Diabetes mellitus</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Indirect method p-value*</td>
<td>Direct method p-value*</td>
<td>Indirect method p-value*</td>
<td>Direct method p-value*</td>
</tr>
<tr>
<td>Very high/Low</td>
<td>&lt; 0.001</td>
<td>0.125</td>
<td>&lt; 0.001</td>
<td>0.171</td>
</tr>
<tr>
<td>Very high/Medium</td>
<td>0.010</td>
<td>0.544</td>
<td>&lt; 0.001</td>
<td>0.923</td>
</tr>
<tr>
<td>Very high/High</td>
<td>0.191</td>
<td>0.647</td>
<td>0.054</td>
<td>0.926</td>
</tr>
<tr>
<td>High and Low</td>
<td>&lt; 0.001</td>
<td>0.846</td>
<td>&lt; 0.001</td>
<td>0.010</td>
</tr>
<tr>
<td>High and Medium</td>
<td>0.048</td>
<td>0.001</td>
<td>0.001</td>
<td>0.996</td>
</tr>
<tr>
<td>Medium and Low</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

HVI: health vulnerability index; *p lower than 0.05 indicates a significant difference; Vigil: Surveillance of Risk and Protection Factors for Chronic Diseases by Telephone Survey (Vigilância de Fatores de Risco e Proteção para Doenças Crônicas por Inquérito Telefônico).

**DISCUSSION**

The present study estimated the prevalence of AH and DM, according to the HVI, in Belo Horizonte. The direct and indirect methods of estimation in small areas were compared. It was found that areas with high risk had a higher prevalence of AH and DM compared to
those with low risk, according to the indirect method of estimation in small areas. The direct method showed no prevalence differences among the areas.

Previous studies have shown that less economically privileged areas have worse health conditions when compared to more privileged areas both between countries and within the same country, revealing the unequal geographical distribution of diseases.

Inequalities in the health status between groups with different socioeconomic status are configured as a public health problem in several locations around the world. It is known that worse socioeconomic conditions are associated with unhealthy behaviors, lack of knowledge about self-care, barriers to healthy eating and unsafe spaces that perpetuate structural inequalities, which could explain the higher occurrence of NCD in these areas. Previous studies also point out that the unfavorable socioeconomic situation and the regional deprivation of individuals with DM and AH are related to the poorer quality of care, which directly interferes with the results of the control of these diseases, leading to a higher incidence of complications.

Research that identifies regional health disparities is essential for promoting social programs and policies aimed at socially vulnerable groups. In addition to individual factors, environmental factors are directly associated with lifestyle and the occurrence of NCD. Thus, analyses that consider the context in which the individual is inserted are fundamental for the recognition of social contexts in the determination and control of diseases, contributing to the planning of prevention and treatment actions.

The indirect method of estimation for small areas used in this study is highlighted, which used the imputation of missing data in areas that do not have Vigitel telephone interviews, avoiding the under or overestimation of the prevalences of chronic diseases by HVI obtained by the direct method, which considers only the sectors with Vigitel interviews. An example is the presence of AH and DM, which is strongly influenced by age, and in the areas of greater health vulnerability there is a smaller number of aged people compared to the most economically privileged locations.

The HVI for estimating disease prevalence in small areas also deserves to be highlighted. This index takes into account the census tracts, which are the smallest territorial division adopted by the Brazilian Institute of Geography and Statistics. As it is composed of socioeconomic and environmental variables, the HVI allows the identification of inequities in census tracts and the prioritization of areas in the planning of health actions by managers.

In addition to this, primary health care (PHC) emerges as a facilitator in monitoring individuals and reducing regional inequalities, enabling the performance of health professionals in the territories in different contexts. The territories are subdivided into coverage areas composed of contiguous census tracts, allowing Family Health Strategy teams to hold information on the population and propose local interventions.

This study has some limitations. Firstly, the exclusion of 10.5% of Vigitel’s interviews due to the non-identification of census tracts by linkage, which could lead to a selection bias, however the use of post-stratification weights, according to the 2010 census population by HVI, minimizes potential bias. Secondly, the accuracy of the adjusted model for imputing
the variable hypertensive or non-hypertensive adults in the set of sectors without any Vigitel interview was between 50 and 60%, which is considered acceptable, but subject to under or overestimation of the outcome. Thirdly, the information on DM and AH is self-reported by the participants, which can overestimate or underestimate the prevalence of the outcomes, but direct diagnostic measures, obtained by laboratory and clinical tests, are difficult to be performed in large populations, in addition to being costly, and studies of validation comparing self-reported and clinical measures indicate good reliability results\(^{28}\). Fourthly, to estimate the prevalence of the indicators, data from 2006 to 2013 were aggregated, which reflects the mean in the period and, consequently, the loss of the trend over the years.

Finally, the estimates obtained by the indirect method require external validation of the results found in this article. It is emphasized that Vigitel’s external validation study, compared to the household survey, has already been carried out for the outcome of tobacco use by HVI and has identified acceptable values\(^{8}\). Thus, it is also recommended to proceed with an external validation study to estimate the prevalence of chronic diseases for small areas using Vigitel data.

It was concluded that high health risk areas have higher AH and DM prevalences than low risk areas. The results demonstrate that investigations carried out for all municipalities, without taking into account regional differences, can bring biased estimates of the health situation in vulnerable areas and may not identify inequalities in the population. Producing reliable estimates for small regions allows to know and locate differences in the distribution of health events in the territory. Thereby, it contributes to the understanding of local realities, being an important strategy for the identification of areas of greater vulnerability and for the redesign of a health care network and social promotion, at different geographical scales, considering the territory, a principle that governs PHC, in addition to assisting managers in planning public policies according to the community needs.

REFERENCES


Received on: 10/20/2020
Revised on: 12/02/2020
Accepted on: 12/12/2020
Preprint on: 12/15/2020

Authors’ contribution: DCM designed the study. RTIB developed the method of management and statistical analysis of the data. CSG, LSMC, and ADM helped in the analysis of the data, the interpretation of the results and the preparation of the preliminary version of the manuscript. DCM, RTIB, and ALPR critically reviewed the manuscript and contributed to the interpretation of the results. All authors read, contributed to, and approved the final version of the manuscript.