

# Geolocation of hospitalizations registered on the Brazilian National Health System's Hospital Information System: a solution based on the R Statistical Software

doi: 10.5123/S1679-49742018000400016

Thiago Augusto Hernandes Rocha<sup>1</sup> –  orcid.org/0000-0002-6262-3276

Núbia Cristina da Silva<sup>1</sup>

Pedro Vasconcelos Maia Amaral<sup>2</sup>

Allan Claudius Queiroz Barbosa<sup>1</sup>

João Ricardo Nickenig Vissoci<sup>3</sup>

Erika Bárbara Abreu Fonseca Thomaz<sup>4</sup>

Rejane Christine de Sousa Queiroz<sup>4</sup>

Matthew Harris<sup>5</sup>

Luiz Augusto Facchini<sup>6</sup>

<sup>1</sup>Universidade Federal de Minas Gerais, Faculdade de Ciências Econômicas, Belo Horizonte, MG, Brasil

<sup>2</sup>Universidade Federal de Minas Gerais, Centro para o Desenvolvimento e Planejamento Regional, Belo Horizonte, MG, Brasil

<sup>3</sup>Duke University, Divisão de Medicina de Emergência, Carolina do Norte, Estados Unidos

<sup>4</sup>Universidade Federal do Maranhão, Departamento de Saúde Pública, São Luís, MA, Brasil

<sup>5</sup>Imperial College London, Faculdade de Medicina, Londres, Reino Unido

<sup>6</sup>Universidade Federal de Pelotas, Faculdade de Medicina, Pelotas, RS, Brasil

## Abstract

**Objective:** to describe a solution enabling geolocation of hospital admissions (AIH), processed on the Brazilian National Health System's Hospital Information System. **Methods:** in order to spatialize AIHs an R language script was written, based on the microdatasus and CepR packages; the script was applied to identify all AIHs in Goiás state in the year 2015; after downloading and pre-processing the data, the procedure for AIH spatialization was detailed. **Results:** of the 361,213 AIHs processed, we were able to retrieve 24,220 different ZIP codes (CEPs); from this set of ZIP codes, 23,910 (98.7%) were geolocated; these geolocated ZIP codes enabled spatialization of 97.7% of AIHs processed for the state of Goiás. **Conclusion:** it is possible to spatialize AIHs with a high success rate; the method detailed in this paper opens a new range of possibilities for the design of evaluation studies, formulation of policies and planning of health care actions.

**Keywords:** Hospitalization; Information Systems; Automatic Data Processing; Spatial Analysis; Geographic Information Systems.

## Correspondence:

**Thiago Augusto Hernandes Rocha** – Organização Pan-Americana da Saúde, Setor de Embaixadas Norte, lote No. 19, Brasília, DF, Brazil. Postcode: 70800-400  
E-mail: rochahernandes3@gmail.com

## Introduction

Brazil is a privileged space for discussions on health service evaluation thanks to diverse information on the service provision system being made publicly available. Standing out in this sense is the role played by the Brazilian National Health System's Information Technology Department (DATASUS),<sup>1</sup> responsible for making access available to its diverse health information systems. Despite the relevance of this initiative, difficulties persist as to the documentation of information publicized, the possibility of data disaggregation and the large number of files to be handled.<sup>2,3</sup>

*Despite the potentialities associated with SIH/SUS, most of the data publicized is disaggregated to the maximum as far as the municipal level. As such, conducting analyses to examine characteristics within municipalities is a challenging task.*

Worthy of highlight among the universe of information systems that provide the data made available by DATASUS is the Brazilian National Health System (SUS) Hospital Information System (SIH/SUS). The SIH/SUS system is responsible for processing information taken from hospital admission authorization forms (AIHs).<sup>4</sup> The same system monitors payments made to each SUS hospital, the codes defined by the 10<sup>th</sup> International Statistical Classification of Diseases and Related Health Problems (ICD-10) associated with hospitalizations, average length of patient stay, postcode (CEP) of each AIH, among other important information. This system has been widely used to inform health evaluation studies in Brazil.<sup>5-9</sup>

Despite the potentialities associated with SIH/SUS, most of the data publicized is disaggregated to the maximum as far as the municipal level.<sup>4</sup> As such, conducting analyses to examine characteristics within municipalities is a challenging task which implies the need to collect primary data of a more granular nature than the data usually publicized by DATASUS. Moreover, the need to manipulate a large volume of data that has a low level of documentation means that researchers are required to use tools with limited flexibilization capacity, permeated by access difficulties and limited from the point of view of information generation, as is the

case of the TabWin and TabNet applications.<sup>10</sup> Although some efforts have been made to improve documentation relating to SIH/SUS,<sup>11</sup> challenges still remain.

Analyzing events that occur on the municipal level is crucial for conducting health evaluation studies. It is a mistake to consider geographical territory as an expression of uniform characteristics, especially with regard to health events. Cities such as Belo Horizonte, for instance, have neighbourhoods with human development indices (HDIs) identical to those found in Switzerland, whilst at the same having regions with HDIs equivalent to those found in Sub-Saharan Africa.<sup>12</sup> When analyzing municipal indicators, both these extremes are considered in an aggregated manner. This has the effect of homogenizing differences and catalyzing mistakes associated with ecological fallacy.

Reflecting on social and geographical determinants is fundamental for improving understanding of the health-disease dyad.<sup>13</sup> Bearing in mind the importance of studies based on the geographic location of health events<sup>14-16</sup> and the relative incipency of studies supported by this type of methodology in the Brazilian context,<sup>13</sup> there is an urgent need to develop more effective strategies for health event spatialization.

The importance of developing solutions capable of enabling the examination of intra-municipal characteristics, together with the limitations and difficulties inherent to using TabWin and TabNet, contributed to the design of this study, the objective of which was to present a solution capable of enabling the geolocation of hospitalizations processed on SIH/SUS.

## Solution development

This is a methodological study.<sup>17</sup> In order to enable spatialization of SIH/SUS hospitalizations, the postcodes (CEP) informed on each hospital admission authorization form (AIH) were geolocated with the aid of R language.

Certain challenges needed to be overcome in order to be able to spatialize the hospitalizations processed on SIH/SUS, in particular: the need to handle multiple 'RD\*.dbc', files through which the data is publicized; retrieving the postcode data for each AIH, as well as its geolocation. Once these stages have been completed, the latitude and longitude coordinates for each AIH can be obtained, with precision limited to the street of origin of the patient, thus enabling more granular analyses to be performed than those carried out using only aggregated municipal indicators.

R statistical software was used to structure the solution. This software allows the inclusion of packages intended to carry out specific functions, such as handling DATASUS data.

The first stage of the AIH geolocation process involves downloading the publicized data from SIH/SUS. The information on hospitalizations is made available at the following web address: [ftp://ftp.datasus.gov.br/dissemin/publicos/SIHSUS/200801\\_/](ftp://ftp.datasus.gov.br/dissemin/publicos/SIHSUS/200801_/); the files are named following the RDUFAAMM standard, where RD is the abbreviation of 'reduced', UF is the Federative Unit, AA is the year and MM is the month. As such, for each Federative Unit there is a reduced monthly file containing the data relating to hospitalizations at the health establishments in the respective UF. Santos<sup>11</sup> has described the information contained in each variable publicized in the SIH/SUS 'RD\*.dbc' files. In order to analyze just one year,<sup>12</sup> files need to be handled – one per month – for each of the 27 UFs. Furthermore, each of these files needs to be decompressed, as they are publicized in \*.dbc format and need to be converted to \*.dbf format. In order to address the challenges related to this onerous process, Saldanha<sup>18</sup> developed the microdatasus package.

The R microdatasus package has functions for downloading DATASUS microdata files (\*.dbc format), reading them using the Read.dbc package and pre-processing them for use.<sup>18</sup> The \*.dbc format files are decompressed to \*.dbf format automatically using the Read.dbc package.<sup>3</sup> These packages can be installed using Github, via the devtools package, directly via the front-end/GUI, using R or RStudio,<sup>19</sup> by giving the following commands:

```
install.packages("devtools") # Install the package
devtools
devtools::install_github("rfsaldanha/microdatasus")
install.packages("read.dbc")
```

The microdatasus package only has two functions: `fetch_datasus` and `process_sih`. The former is for automating file downloads from DATASUS, while also decompressing and aggregating the individual files into a single file. By performing this task, this function substantially reduces the number of operations needed to create a processable database from the data available on DATASUS. The second function is for pre-processing the data downloaded from DATASUS, attributing labels to the

raw variables obtained by using the `fetch_datasus` function.

The `fetch_datasus` function has seven arguments:

- `year_start` : initial year
- `month_start` : initial month
- `year_end` : final year
- `month_end` : final month
- `UF` : Federative Units
- `information_system` : information system
- `vars` : variables of interest

The arguments beginning with 'year' or 'month' are used to define the start and end months and years of the dataset to be downloaded, respectively. The UF argument defines which states are to be downloaded. The `information_system` argument defines from which systems the data are to be downloaded. Although SIH/SUS data was used in this study, the microdatasus package is able to perform automated downloads from other systems, such as the Mortality Information System (SIM), the Live Births Information System (SINASC) and the National Registry of Health Establishments (CNES). In the future the SUS Outpatient Information System (SIA/SUS) will also be included.<sup>18</sup> Finally, the `vars` argument defines which variables will be downloaded.

The `process_sih` function is responsible for pre-processing the data downloaded from Datasus by the `fetch_datasus` function.<sup>18</sup> The `process_sih` function has three functions:

- `data` : file created by the function `fetch_datasus`
- `information_system` : sub system of SIH/SUS
- `municipality_data` : hometown data of each patient

The `data` argument defines the file that received the result of the `fetch_datasus` function without any modifications. The `process_sih` function should be used straightaway after the `fetch_datasus` function. The `information_system` argument details the source of the data stored in the file to be processed. Finally, the `municipality_data` argument adds information to the file about the municipality of residence, such as full municipality name, latitude and longitude.<sup>18</sup> In order for the functions to operate satisfactorily, the user needs to be connected to the internet and the DATASUS FTP needs to be working. More detailed information on the package can be obtained from Wikipedia: <https://github.com/rfsaldanha/microdatasus/wiki>

What follows is a detailed practical example of the commands used to load the microdatasus package in the R workspace and download the SIH/SUS data for the state of Goiás in 2015:

```
library(microdatasus) # load the microdatasus
package in R environment.

internacoes <- fetch_datasus(year_start = 2015,
month_st = 1, year_end = 2015, month_end =
12, uf = "GO", information_system = "SIH-RD")

str(internacoes)
```

These commands generate a file with 361,213 hospitalizations (internacoes) for the year 2015 in the state of Goiás. In all, 113 variables associated with each hospitalization were brought together. The complete description of the variables can be found in the work of Santos.<sup>11</sup> The structure of the database in relation to the first six variables follows below.

```
> str(internacoes)

'data.frame':   361213 obs. of  113 variables:

 $ UF_ZI:Factor w/ 172 levels
"520000","520013": 1 1 1

 $ ANO_CMPT: Factor w/ 1 level "2015": 1 1 1

 $ MES_CMPT: Factor w/ 12 levels "01","02","03": 1 1 1

 $ ESPEC : Factor w/ 12 levels "01","02","03",...: 1 1 1

 $ CGC_HOSP : Factor w/ 106 levels
"00029827000128": NA NA NA

 $ CEP: Factor w/ 24220 levels
"03145010","03728210": 6464 6447 6458
```

It must be emphasized that hospitalizations do not allow unique patient identification, since the same patient may be hospitalized more than once. Solutions exist for probabilistic linkage of hospitalizations, such as that performed by

OpenRecLink. However, the application of this type of technique is beyond the scope of this study.<sup>20</sup> The process\_sih function was applied to the 'hospitalizations' object generated in order to label the variables downloaded from DATASUS. The variables recorded on the AIH form include the postcode (CEP) of the residence of the patient admitted to the health establishment, this information being obligatory for the AIH to be registered. By using the CepR package, each AIH can be geolocated based on the postcode provided by the patient.

The Brazilian postcode database is monopolized by the Brazilian Post and Telegraph Company and its use is conditioned to payment of a license. Even so, the paid version does not contain the latitude and longitude parameters associated with all of Brazil's postcodes. In order to overcome these difficulties, the CEP Aberto (open postcode) project was developed collaboratively.<sup>21</sup> This project aims both to provide free access and also to collaboratively build a database containing all of Brazil's geolocated postcodes. A total of 980,955 postcodes currently exist in Brazil. The CEP Aberto project also developed a free of charge application programming interface (API), which enables data searches using a postcode to be done for data such as: state (UF), municipality, telephone code, neighbourhood, street or equivalent, latitude, longitude and altitude. In order to use the API, the user needs to register via this web address <http://cepaberto.com/>. After registration an access token is provided. As the project is maintained collaboratively, the volume of searches is limited to one search every three seconds and a maximum of ten thousand searches per token per day. Considering the functionalities of the CEP Aberto project, Robert Myles developed an R package, called CepR, that searches for CEP Aberto postcode data directly from R.<sup>22</sup> A download, pre-processing and AIH geolocation script was developed which integrates the CepR and microdatasus solution.

What follows are the steps for integrating both of the R package solutions. Firstly, a summary needs to be made of the postcodes obtained from SIH/SUS using the fetch\_datasus function in order to minimize the number of searches on the CEP Aberto API. This reduces the time needed for the geolocation process, given that a postcode is generally repeated several times on the AIH records.

As such, the first step is to create a single postcode list with no duplicated codes. The following commands exemplify how to do this, based on a file resulting from the `fetch_datusus` function.

The same hospitalization file defined previously is used. The user needs to iterate over a list of unique postcodes (`cep_unicos`) in order to perform the searches. To do this operation a loop needs to be configured. For each postcode entry listed on the `cep_unicos` vector, this command will search for the respective geographic location information. The following steps define how to create a file to receive the data for each postcode search done and how to set the loop parameters:

```
install.packages("cepR") # Install teh CepR package
#Creates a temporary dataframe that will be used
to fill during each zip code iteration.
geo_coded<- data.frame(estado=character(),cid
ade=character
bairro=character(),cep=character(),logradouro
=character(),latitude=character(),longitude=ch
aracter(),altitude=character(),ddd=character()
,cod_IBGE=character(),quality = logical(),cep_
buscado = character(),stringsAsFactors=FALSE)

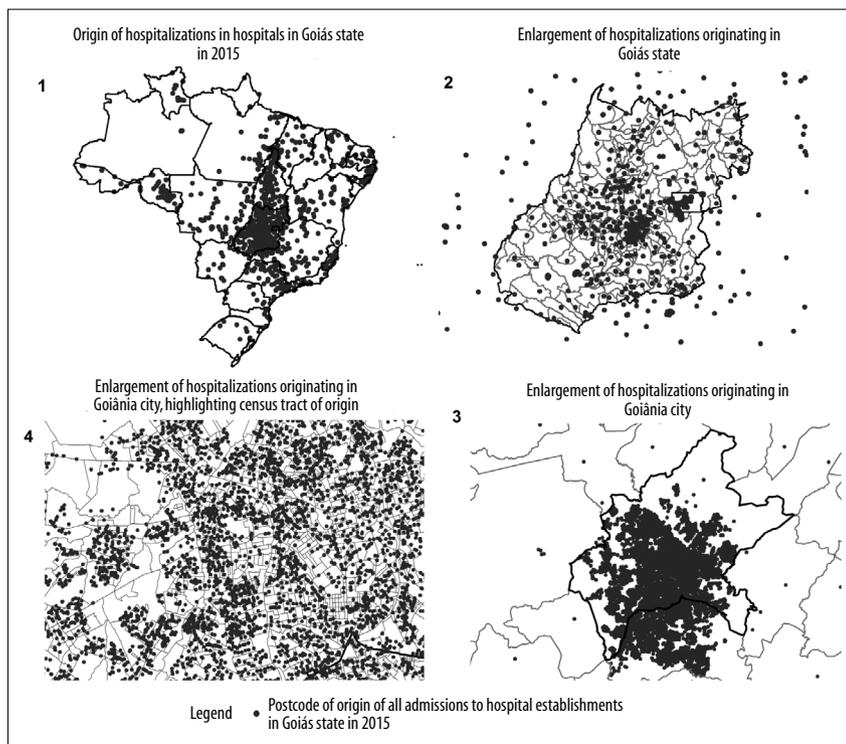
cep_unico <- as.character(unique(internacoes$C
EP)) # cria um vetor com os CEPS sem repetição
para minimizar a realização das consultas na API
do CEP Aberto
#Loop de consulta de cada CEP listado no vetor:
cep_unico
for (i in 1:10000) {
  sys1 <- Sys.time()
  consulta <- busca_cep(cep=(cep_unico[[i]]),
token= 'seu token')
  consulta$quality <- anyNA(c(consulta$latitude,
consulta$longitude))
  consulta$cep_buscado <- cep_unicos[[i]]
  geo_coded <- rbind(geo_coded, consulta,
make.row.names=FALSE)
  if(Sys.time()-sys1 <=4.0) Sys.sleep(4.0-(Sys.
time()-sys1))}
```

The loop in question takes into account the minimum time interval for performing searches

on the CEP aberto API. Only the initial parameters in bold letters in the section (i in 1:10000) need to be modified in order for the indexes of up to ten thousand postcodes per day to be provided. In addition the 'seu token' (your token) field needs to be replaced by the token identification provided when registering on the CEP Aberto project website. The search code keeps a record in the quality column of the success or not of the search. A TRUE value indicates success in obtaining the latitude and longitude coordinates of the checked postcode. When the search process has been finished, the `geo_coded` file with the following structure is generated:

```
str(geo_coded)
Classes 'tbl_df', 'tbl' and 'data.frame':
24220 obs. of 12 variables:
 $ estado : chr "GO" "GO" "GO"
 $ cidade : chr "São Simão" "Mineiros"
"Quirinópolis"
 $ bairro : chr NA NA NA
 $ cep : chr "75890000" "75830000"
"75860000"
 $ logradouro : chr "São Simão" "Mineiros"
"Quirinópolis" "AC Jataí, Avenida Dorival de
Carvalho, 1007"
 $ latitude : chr "-18.9964906" "-17.5624415"
"-18.4476442"
 $ longitude : chr "-50.547432" "-52.5489206"
"-50.4551598"
 $ altitude : chr "478.400000" "787.900000"
"512.900000"
 $ ddd : chr NA NA NA "64" ...
 $ cod_IBGE : chr "5220405" "5213103"
"5218508"
 $ quality : logi TRUE TRUE TRUE
 $ cep_buscado: chr "75890000" "75830000"
"75860000"
```

The steps detailed above were applied to the AIH data for the state of Goiás for the year 2015. The results of the AIH spatialization process were plotted using ARCMAP.<sup>23</sup> Although we opted to use ARCMAP, any GIS software can be used for the same purpose, such as the QGIS, GEODA and TerraView free solutions.



**Figure 1 – Distribution of hospital admission authorizations given in Goiânia and Goiás, by patients' place of residence, 2015**

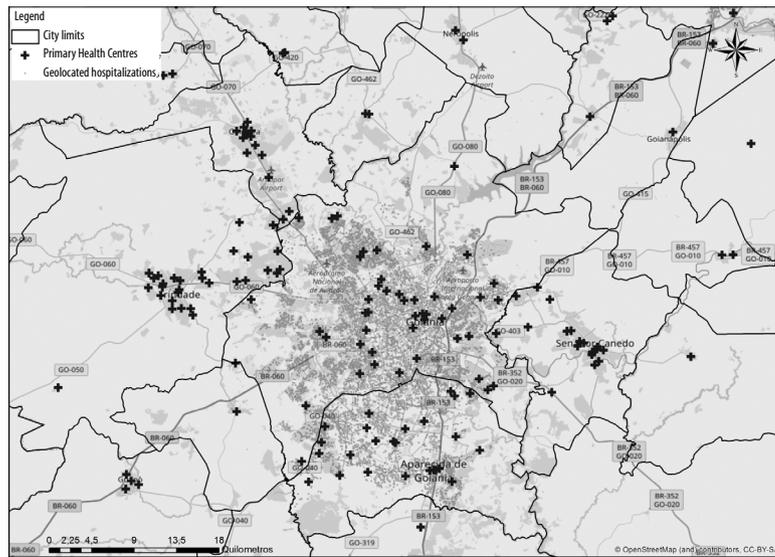
## Results

23,910 (98.7%) of the 24,220 unique postcodes were geolocated, corresponding to 353,004 AIHs out of 361,213 possible hospitalizations (97.7%). No postcodes were retrieved from SIH/SUS with fewer than eight digits. Non-geolocated postcodes corresponded to generic postcodes for entire cities or large regions and also to postcodes that had not yet been input to the CEP Aberto project databases. With regard to the spatialization of the points relating to the AIHs, we found that were patients from all of the Brazilian states who were admitted to one hospital or another in the state of Goiás during 2015 (Figure 1); each point indicates the place of residence of the patient admitted to a Goiás health establishment in 2015. Each item of Figure 1 displays the geolocated AIHs, with different levels of visualization granularity.

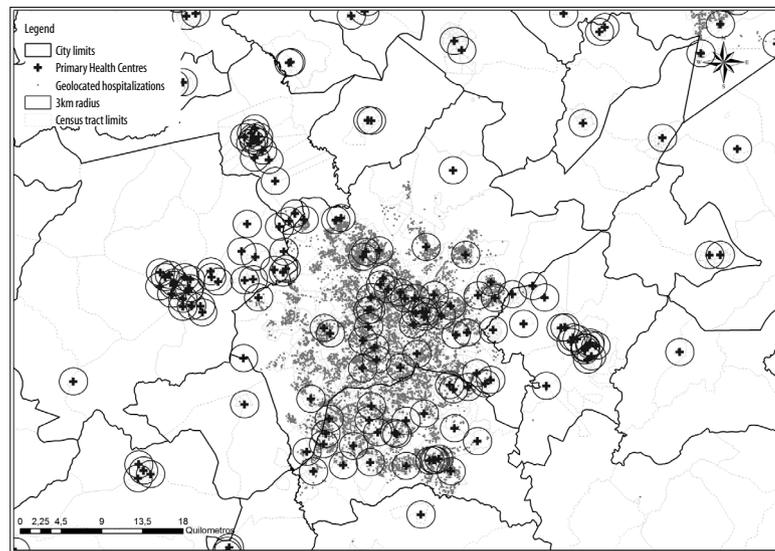
Figure 2 presents grey points corresponding to each geolocated AIH, solely for the city of Goiânia (capital of the state of Goiás) and surrounding region, in relation to primary health care centres, street layout

and river courses. By superimposing different layers of geographical information, using geoprocessing and surveillance techniques, it is possible to analyze probable relationships between hospitalization patterns and linkage with geographical elements. As SIH/SUS provides CID-10 codes, spatial dependence of specific diseases and geographical elements can be analyzed. Moreover, as SIH/SUS holds data with effect from 2008 onwards, time series analyses can be developed.

Figure 3 demonstrates how social determinants of health, health care provided by primary care teams and volume of hospitalizations in a specific region, can be analyzed together, with a degree of granularity not reported thus far in the literature. The circles around each cross drawn on Figure 3 correspond to a distance of 3km from each primary health care centre. Taking capture techniques into consideration, it is possible to attribute a hospitalization burden to Primary Health Care teams, weighted by sociodemographic information linked to census tracts, for example. This hospitalization burden can be used to design quasi-experimental evaluation research.



**Figure 2 – Geolocated hospital admission authorizations and their relationship with other geographical elements in Goiânia and surrounding area, Goiás, 2015**



**Figure 3 – Relationship between primary health care centres, hospitalizations and census tracts in Goiânia and surrounding area, Goiás, 2015**

## Discussion

Applying health geography methods in order to solve methodological problems associated with evaluation research is, nevertheless, an unfinished agenda. This study aimed to provide a small contribution to expanding analytical possibilities related to SIH/SUS. Our findings are promising given the potentiality for developing higher quality studies, proposing public

policies informed by local evidence and qualifying the health care supply planning process.

The design we proposed opens the possibility of better investigation of a myriad of questions. Analyses of social determinants of health can be better informed. AIH geolocation enables concomitant analysis of disease incidence profiles and other social elements, such as resident population profile, socio-economic status, degree of urbanization and other

published data per census tract, for instance. With regard to environmental variables, pollution, street layout and river course databases, among others, can be compared with the occurrence of diseases in specific regions.

Health care network analysis can be benefitted by the method proposed here. It will be possible to analyze a patient's place of residence and compare this information with the location of the health establishment that provided care to them. This data can provide a powerful insight into the real flows of ill people in looking for care. Regular and constant flows of patients with given health conditions, as well as accessibility flows, can provide health professionals with information about regions dealing with service shortages.

Analyses of equity in access to services can be better detailed and service catchment areas can be defined based on geoprocessing techniques. Quality of care provided by hospitals, primary health care centres and urgent care services can be evaluated according to geographical proximity and potential influence area parameters. Probabilistic linkage techniques can be optimized through the inclusion of geographic location parameters, increasing their degree of precision. Hospitalization burdens can be linked to specific health services, such as in the case of Primary Health Care. Health service supply planning actions can be done in more detail, given the possibility of analyzing the epidemiological profile of a given population allocated to a geographic region. The availability of data with greater granularity opens up a range of options to health service managers, health workers, public policy formulators and scholars.

Despite the progress and the potentiality of the method developed, there are limits to be addressed. The first of them relates to the existence of a degree of imprecision in locating the origin of the patient by using the postcode. As geolocation is done based on the street layout code, there exists a degree of uncertainty that has the power to affect more sensitive analyses, generating imprecision regarding problems occurring in small areas. Another limitation of geolocation to be considered is that owing to limitations as to the volume of API searches, the process developed is slow to perform, so that large volumes of data can consume a substantial amount of time before being geolocated. In the case of generic postcodes, that cover entire cities

and large regions and end in '000', the technique is incapable of providing a precise location. Finally, we highlight the dependence on a collaborative solution without regular funding. The CEP Aberto project is fundamental for the operationalization of the solution proposed here. Moreover, the CEP Aberto project databases do not contain all Brazilian postcodes, thus contributing to some level of imprecision which needs to be analyzed case by case. Without the CEP Aberto project, the method defined here loses its propulsive capacity.

Notwithstanding the limitations presented, the solution this paper presents has more potentialities than limitations. Applying the technique that is described in detail here is shown to be capable of fomenting new research using SIH/SUS data. The historical perspective of the Brazilian National Health System's Hospital Information System and the new degree of granularity obtained by using the geolocation method can serve to inform the design of quasi-experimental Public Health research. This type of design is capable of providing more robust evidence regarding health policies and programmes and should be encouraged.

### Acknowledgements

Our thanks to Fúlvio Nedel, Danieal Petruzalek, Daniel Saldanha and Robert Myles, for giving their time, effort and dedication to developing the solutions aimed at furthering enhancements to the Brazilian National Health System (SUS). We thank CAPES, for granting a Ph.D. sandwich course scholarship to the first author of this paper. Our thanks to the Pan American Health Organization.

### Authors' contributions

Rocha TAH took part in the conception, scheduling and writing, literature review, data analysis, discussion of the textual elements and critical review of the manuscript. Silva NC, Amaral PVM, Barbosa ACQ and Vissoci JRN took part in writing, literature review and discussion of the textual elements of the manuscript. Thomaz EBAF, Queiroz RCS, Harris M and Facchini LA took part in the discussion of the textual elements and critical review of the manuscript. All the authors approved the final version of the manuscript and are responsible for all aspects of the study.

## References

1. Ministério da Saúde (BR). Departamento de Informática do SUS – Datasus. Informações de saúde (TabNet) [Internet]. 2015 [citado 208 nov 8]. Disponível em: <http://www2.datasus.gov.br/DATASUS/index.php?area=02>
2. Rocha TAH, Rocha J, Silva N, Amaral P, Facchini L, Thumé E, et al. Cadastro nacional de estabelecimentos de saúde: evidências sobre a confiabilidade dos dados. *Ciêns Saúde Coletiva*. 2017 jan;23(1):229-40. doi: 10.1590/1413-81232018231.16672015.
3. Petruzalek D. READ.dbc – um pacote para importação de dados do Datasus na linguagem R [Internet]. In: Anais do XV Congresso Brasileiro de Informática em Saúde; 2016 27 nov – 30 nov [citado 2018 nov 8]; Goiânia, Brasil. Disponível em: [http://docs.bvsalud.org/biblioref/2018/07/906543/anais\\_cbis\\_2016\\_artigos\\_completos-601-606.pdf](http://docs.bvsalud.org/biblioref/2018/07/906543/anais_cbis_2016_artigos_completos-601-606.pdf)
4. Ministério da Saúde (BR). Departamento Nacional de Auditoria do SUS. Coordenação-Geral de Desenvolvimento Normatização e Cooperação Técnica. Auditoria no SUS: noções básicas sobre sistemas de informação [Internet]. Brasília: Ministério da Saúde; 2004 [citado 2018 nov 8]. 94 p. Disponível em: [http://bvsms.saude.gov.br/bvs/publicacoes/auditoria\\_sus.pdf](http://bvsms.saude.gov.br/bvs/publicacoes/auditoria_sus.pdf)
5. Gerhardt TE, Pinto JM, Riquinho DL, Roese A, Santos DL, Lima MCR. Utilização de serviços de saúde de atenção básica em municípios da metade sul do Rio Grande do Sul: análise baseada em sistemas de informação. *Ciêns Saúde Coletiva*. 2011; 16(suppl 1):1221-32. doi: 10.1590/S1413-81232011000700054.
6. Bittencourt SA, Camacho LAB, Leal MC. O Sistema de informação hospitalar e sua aplicação na saúde coletiva hospital. *Cad Saúde Pública*. 2006 jan;22(1):19-30. doi: 10.1590/S0102-311X2006000100003.
7. Loyola Filho AI, Leite Matos D, Giatti L, Afradique ME, Viana Peixoto S, Lima-Costa ME. Causas de internações hospitalares entre idosos brasileiros no âmbito do Sistema Único de Saúde. *Epidemiol Serv Saúde*. 2004 dez;13(4):229-38. doi: 10.5123/S1679-49742004000400005.
8. Escosteguy CC, Portela MC, Medronho RA, Vasconcellos MT. The Brazilian hospital information system and the acute myocardial infarction hospital care. *Rev Saúde Pública*. 2002 Aug;36(4):491-9.
9. Schramm JM, Szwarcwald CL. Sistema hospitalar como fonte de informações para estimar a mortalidade neonatal e a natimortalidade. *Rev Saúde Pública*. 2000 jun;34(3):272-9. doi: 10.1590/S0034-89102000000300010.
10. Silva NP. A utilização dos programas TabWin e TabNet como ferramentas de apoio a disseminação das informações em saúde [dissertação]. Rio de Janeiro (RJ): Fundação Oswaldo Cruz; 2009.
11. Santos AC. Sistema de informações hospitalares do Sistema Único de Saúde: documentação do sistema para auxiliar o uso das suas informações [dissertação]. Rio de Janeiro (RJ): Fundação Oswaldo Cruz; 2009.
12. Fonseca B, Silva K. Atribuição de IDH aos bairros de Belo Horizonte [Internet]. *Rev Transite*. 2017 [cited 2017 set 5]. Disponível em: <http://transite.fafich.ufmg.br/idh-bairros-de-belo-horizonte/>
13. Guimarães RB. Geografia e saúde coletiva no Brasil. *Saúde e Soc*. 2016 out-dez;25(4):869-79. doi: 10.1590/s0104-12902016167769.
14. Kearns R, Moon G. From medical to health geography: novelty, place and theory after a decade of change. *Prog Hum Geogr*. 2002 Oct;26(5):605-25. doi: 10.1191/0309132502ph389oa.
15. Macintyre S, Ellaway A, Cummins S. Place effects on health: How can we conceptualise, operationalise and measure them? *Soc Sci Med*. 2002 Jul;55(1):125-39.
16. Dummer TJB. Health geography: supporting public health policy and planning. *CMAJ*. 2008 Apr;178(9):1177-80. doi: 10.1503/cmaj.071783.
17. Lima DVM. Research design: a contribution to the author. *Online Brazilian J Nurs*. 2011;10(2):1-18. doi: 10.17665/1676-4285.2011v10n2.
18. Github. Download de dados do DataSUS e pré-processamento no R [Internet]. 2017 [citado 2018 nov 7]. Disponível em: <https://github.com/rfsaldanha/downloadDataSUS>
19. Nedel FB. csapAIH: uma função para a classificação das condições sensíveis à atenção primária no programa estatístico R\*. *Epidemiol Serv Saúde*. 2017 jan-mar;26(1):199-209. doi: 10.5123/S1679-49742017000100021.
20. Camargo Júnior KR, Coeli CM. Going open source: some lessons learned from the development of OpenReLink. *Cad Saúde Pública*. 2015 Feb;31(2):257-63.

21. CEP Aberto. O programa CEP aberto [Internet]. 2017 [citado 2017 set 5]. Disponível em: <http://cepaberto.com/>
22. Github. Um pacote R para buscar informações sobre CEPs, endereços, bairros e cidades. (An R package for accessing Brazilian postal code data). [Internet]. 2017 [citado 2018 nov 7]. Disponível em: <https://github.com/RobertMyles/cepR>
23. Environmental Systems Research Institute. ArcGIS desktop: release 10.3 [Internet]. 2014 [cited 2018 Nov 8]. Available in: <https://www.esri.com/esri-news/releases/15-1qtr/arcgis-10-3-and-arcgis-pro-modernize-gis-for-organizations-and-enterprises>

Received on 22/02/2018  
Approved on 01/11/2018