Electronic medical records in primary care: management of duplicate records and a contribution to epidemiological studies

Luiz Felipe Pinto (https://orcid.org/0000-0002-9888-606X)¹
Leda Jung dos Santos (https://orcid.org/0000-0002-9554-1621)²

Abstract Primary health care electronic medical records were analyzed in Rio de Janeiro for two chronic diseases, namely, hypertension and diabetes, in a population-based study with a cross-sectional epidemiological design that considered the Rio de Janeiro population enrolled in Family Health Teams. Calculation of the prevalence rate was stratified by gender and age group, and the condition of the disease was measured by family doctors in their visits using the ICD-10. Except for the last two age groups (75-79 years and 80 years and over), with apparent under-registration of the diagnosed cases, a positive association was found between prevalence rates and age in both genders. The generation of objective and reliable statistical information is fundamental for local management, allowing the evaluation of demographic dynamics and the peculiarities of each territory, and assisting in the planning and monitoring of the quality of Rio de Janeiro people’s records registered in each family health unit. Thus, the regular management of duplicate records in the registered user roster is essential to minimize the over-registration of clinical cases reported in the electronic medical records.

Key words Electronic medical record, Hypertension, Diabetes mellitus, Primary health care, Brazil

¹Departamento de Medicina em Atenção Primária à Saúde, Faculdade de Medicina, Universidade Federal do Rio de Janeiro. R. Laura de Araújo 36/2º parte, Cidade Nova. 20211-170 Rio de Janeiro RJ Brasil. felipepinto.rio@medicina.ufrj.br
²Curso de Graduação em Medicina, Centro Universitário Serra dos Órgãos. Teresópolis RJ Brasil.
Introduction: contextualizing the use of electronic records in the health area

Health information technology – the hardware, software, and infrastructure needed to collect, store, and exchange patient information in clinical practice has been changing healthcare in the world. Its implementation involves a complex process with the participation of several technical, human, individual and organizational dimensions. The first step in using the software consists of recording registration and administrative data of all health care providers – units, health professionals of the services employing medical records. Its aspects of usability must also be planned, and should be maximized for gains in adherence, efficiency, and quality of user registration. We are referring to the issue of the need for previous elaboration and validation by managers, doctors, and users of a data dictionary, an important component that must be developed when building an information system containing the description of the layout of variables.

Hendy et al. highlight the long period required to implement electronic health records in the UK experience. In 1998, it was estimated at 10 years. However, after four years, only 3% of the trusts had met the target. The reasons given were funding decentralization to the local level, and the lack of standardized health information technologies. Rozenblum et al., on the other hand, describe a nationwide Canadian project started in 2001 that aimed to implement a unique information system containing the description of the layout of variables.

In this paper, we considered electronic health records to describe a set of integrated, interoperable and customizable modules in the daily care of primary health care services, covering administrative, registration and clinical bases, with individual actions and group activities, procedures, medical consultations with records as per the International Classification of Diseases (ICD-10), International Classification of Primary Care (ICPC), immunization, home visits, laboratory and exam results, medication and problem lists, administrative, registration and clinical bases, with individual actions and group activities, procedures, medical consultations with records as per the International Classification of Diseases (ICD-10), International Classification of Primary Care (ICPC), immunization, home visits, laboratory and exam results, medication and problem lists, accessed by all team professionals that are part of the services. These records can be linked with other databases and using a unique identification code such as the Individual Taxpayer Registration Number (CPF). Furthermore, geographically-wise, a country’s postal code and Census data can be used to obtain additional information on social determinants in health, including for academic research.

This paper aims to analyze the electronic medical records of the Family Health Teams in the city of Rio de Janeiro, presenting the process of managing duplicate records, which allowed calculating the prevalence rates of two of the
most important chronic diseases to be monitored in PHC, namely, hypertension and diabetes.

Methods

This is a population-based, cross-sectional epidemiological study that considered (i) the administrative bases related to the registration data of the Rio de Janeiro’s population with “Family Health Teams” in Municipal Centers or Family Clinics, (ii) the demographic and clinical records of the population of a certain area of the city of Rio de Janeiro, which started the implementation and use of electronic medical records in PHC in the mid-2010s. This area was chosen because it was the first area of the city to universalize the use of electronic medical records, as well as the existence of a Family and Community Medicine Residency Program19.

The municipality of Rio de Janeiro is the second most populous in Brazil, the 1229th in geographical area, and 18th in population density. Its strategic location in the country and its importance from the historical, political, economic and social viewpoint has been attracting the attention of researchers worldwide, concerning the PHC Reform, implemented since 2009, as highlighted by the then President of WONCA, Professor Amanda Howe20. In 2016, the city held the Twenty-First World Conference of Family Doctors (WONCA)21 and presented its “Primary Health Care Reform” 22, in which more than four million Rio de Janeiro dwellers started to have 1,244 own teams of primary health care in operation until December 2016 (including Family Clinics launched in the last quarter of 2016, and which still did not contain records in the National Registry of Health Establishments (CNES) in December of the same year), representing around 66% of real coverage between its residents – calculated from the registries of electronic medical records, after management of duplicate records.

From the perspective of the Ministry of Health, the city was, in 2016, the second municipality in Brazil with the most significant number of people potentially covered by Family Health Teams in that period (“potential coverage”, whose numerator is calculated from the number of ESF x average number of 3,450 people registered per ESF). However, after 2017, a significant overall reduction was observed in the number of Rio de Janeiro’s teams and the number of community workers per team and, consequently, the potential coverage of ESF in September 2019, dropped to 45.3% (882 teams23 x average number of 3,450 people, divided by 6,718,903 inhabitants, estimated population for 2019, by IBGE).

The units that develop the Primary Health Care actions in the municipality of Rio de Janeiro recorded in the National Registry of Health Establishments (CNES) of the Ministry of Health are the Municipal Health Centers and Family Clinics existing in ten health planning areas (AP), the city’s “health districts”. They are classified into ‘type A’ units, where only Family Health Teams (with one family doctor, one nurse, one nursing technician, six community health workers, and one health surveillance worker) exist, and ‘type B’ units, in which the Family Health Teams coexist with medical professionals from other specialties and other health professionals. All Family Clinics have been implemented as ‘type A’ units, and we can find ‘type A’ and ‘type B’ Municipal Health Centers.

We selected two chronic diseases, namely, diabetes and arterial hypertension to calculate the prevalence rate or point prevalence, as defined by Pereira24 (considering the 12-month period – July 2014–June 2015). These were stratified by gender and age group. Initially, we are interested in presenting the “duplicate records” management process, which is essential to keep the database of unique identification of users clean and with quality to perform linkages between other clinical databases and individual data. Next, we aimed to evaluate the quality of medical records’ coding in the visits performed by them, computing the ICD-10.

This study was approved by the Ethics Committee of the Municipal Health Secretariat of Rio de Janeiro within the scope of the Laboratory of Epidemiology and Health Statistics (LEES) Project for Primary Care and Health Surveillance in the city of Rio de Janeiro.

The management of duplicate records in the Municipal Health Department from 2013 to 2016: enabling the calculation of more reliable epidemiological indicators

The cadastral baseline with sociodemographic data was consolidated in the first half of 2013, with 171 decentralized face-to-face meetings involving a total of 2,560 community health workers (ACS), directors/managers of the Municipal Centers and Family Clinics. These workshops were developed to qualify the definitive registrations (called “Sheet A”) of the population enrolled in the entire municipality of Rio de Ja-
of ‘duplicate’ records
• Health care and the management
electronic medical records in Primary
population-based indicators, considering the res-
over six months, under the supervision of the
were updated in the electronic medical records
one, hundreds of thousands of people’s records
and starting the process of reviewing their list
each community worker’s receipt of a notebook
and binding on A3 paper of the list of variables
Sheet A, by micro area and ESF, facilitating
workshops developed in 2013, it was possible to
in the areas covered by the ESF. With this set of
people – allowing the construction of a baseline
micro area and ESF, facilitating
each community worker’s receipt of a notebook
and starting the process of reviewing their list
of registered and duplicated users. Thus, one by
hundreds of thousands of people’s records
were updated in the electronic medical records
over six months, under the supervision of the
Health Observatory Network Stations (Rede OT-
over six months, under the supervision of the
Health Observatory Network Stations (Rede OT-
people with duplicate records; that is, 13.7% of the
approximately 340,000 records of homonymous
ics-Rio/SMS-RJ).
• Unique identification number of each person.
people below 16 years of age as the
unique identification number of each person.
16 years of age as the
unique identification number of each person.
16 years of age as the
unique identification number of each person.
• CPF for people over 16 and the number of the Live Birth Certificates
people – allowing the construction of a baseline
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identification number of each person.
unique identifi
Table 1. Population registered by the Family Health Teams before x after the “Sheet A” Qualification Workshops in the electronic medical records according to planning areas (AP) - Municipality of Rio de Janeiro - July / 2013.

<table>
<thead>
<tr>
<th>Projected duplicate records among the registered population</th>
<th>Projected registered population without duplicate CPF/DNV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population registered in electronic medical records</td>
<td>Duplicate population (for example, homonyms)</td>
</tr>
<tr>
<td>Total</td>
<td>A</td>
</tr>
<tr>
<td>1.0</td>
<td>123,456</td>
</tr>
<tr>
<td>2.1</td>
<td>157,819</td>
</tr>
<tr>
<td>2.2</td>
<td>57,736</td>
</tr>
<tr>
<td>3.1</td>
<td>375,544</td>
</tr>
<tr>
<td>3.2</td>
<td>225,945</td>
</tr>
<tr>
<td>3.3</td>
<td>393,579</td>
</tr>
<tr>
<td>4.0</td>
<td>125,821</td>
</tr>
<tr>
<td>5.1</td>
<td>332,569</td>
</tr>
<tr>
<td>5.2</td>
<td>355,289</td>
</tr>
<tr>
<td>5.3</td>
<td>329,618</td>
</tr>
<tr>
<td>Total</td>
<td>2,477,376</td>
</tr>
</tbody>
</table>


Note: Only definitive records of people linked to the micro area. The data of five units that used electronic medical records were not informed since the respective databases were not received. A total of 11 units with ESF did not use electronic medical records by May 2013.

Captions: CPF = Individual Taxpayer Registration Number. DNV = Live Birth Certificate. (A) Number of definitive active records in the micro areas of community health workers, registered in the electronic medical records. (C) = (B/A) (%). (D) Total duplicate records of homonymous people. (E) Total duplicate records of homonymous persons with valid CPF/DNV. (F) Total duplicate records of homonymous people with the same CPF/DNV. (G) = Proportion of duplicate records among records of homonymous persons with the same CPF/DNV compared to the total records of homonymous persons. (H) = Projected total number of duplicate records among the records of homonymous people with the same CPF/DNV. (I) = Projected duplicate records for the total registered population among the records of homonymous people with the same CPF/DNV.
up working with aggregated (ecological) data. Several European countries (Portugal, Spain, the Netherlands) have routines/algorithms for managing duplicate records, and annually update the list of each doctor/family health team, also using computer and probabilistic algorithms to identify duplicate cases and join records.

Final considerations

For the first time in the history of primary health care of the SUS in the city of Rio de Janeiro, this study used registers of electronic medical records to calculate the prevalence of chronic diseases based on population in a city planning area, showing the need for monitoring periodic and regular management of “duplicate user records” in primary health care for the calculation of more reliable health indicators. The results found are relevant to what was expected in the literature, namely, the older the age, the higher the prevalence of hypertension or diabetes, with differences between men and women, thus showing the quality of clinical records.

The generation of objective and reliable statistical information is essential for the management of SUS from micro to macro level, allowing the assessment of demographic dynamics and the particularities of each territory, and assisting in the planning and monitoring of Brazilians registered in each family health unit, with continuous analysis of the list of users to remove citizens with more than one record.

In 2019, the Brazilian Ministry of Health defined some innovations to improve the management of teams, one of which provides that some monitoring indicators will be followed-up in the new federal funding for primary care. Thus, we believe it is essential to prepare management reports/algorithms that allow returning to each Family Health Team their list of active users in the year, and the list of registered users who have not used the services (non-users/regulars) and other information. Only in this way will it be possible to calculate the most appropriate clinical and epidemiological indicators, from the micro (family health team) to the macro (total in the municipality). Thus, finally, after 30 years of implantation of the Unified Health System (SUS)

### Table 2. Prevalence of systemic arterial hypertension and diabetes mellitus diagnosed by family doctors between consultations by gender and age group of users Selected area of the Municipality of Rio de Janeiro (N = 115,280) - 2015 (*)

<table>
<thead>
<tr>
<th>Age group (years)</th>
<th>Systemic Arterial Hypertension (SAH)</th>
<th>Diabetes (SAH and Diabetes)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>18-24</td>
<td>0.5%</td>
<td>0.5%</td>
</tr>
<tr>
<td>25-29</td>
<td>1.1%</td>
<td>1.2%</td>
</tr>
<tr>
<td>30-34</td>
<td>2.1%</td>
<td>3.7%</td>
</tr>
<tr>
<td>35-39</td>
<td>4.2%</td>
<td>7.8%</td>
</tr>
<tr>
<td>40-44</td>
<td>8.0%</td>
<td>12.9%</td>
</tr>
<tr>
<td>45-49</td>
<td>13.6%</td>
<td>21.2%</td>
</tr>
<tr>
<td>50-54</td>
<td>20.7%</td>
<td>28.2%</td>
</tr>
<tr>
<td>55-59</td>
<td>28.8%</td>
<td>35.9%</td>
</tr>
<tr>
<td>60-64</td>
<td>37.8%</td>
<td>42.9%</td>
</tr>
<tr>
<td>65-69</td>
<td>41.2%</td>
<td>47.0%</td>
</tr>
<tr>
<td>70-74</td>
<td>42.7%</td>
<td>51.2%</td>
</tr>
<tr>
<td>75-79</td>
<td>46.6%</td>
<td>49.7%</td>
</tr>
<tr>
<td>80 and over</td>
<td>39.5%</td>
<td>41.2%</td>
</tr>
<tr>
<td>Total</td>
<td>13.0%</td>
<td>19.3%</td>
</tr>
</tbody>
</table>

Source: Special tabulation of aggregated medical visits, based on microdata from validated electronic records, SMS-RJ, December 2016.

(*) The data consider a specific area of city planning, representative of the municipality and one of the first to implement the use of electronic medical records in all units with family health teams.
and after 25 years of the Family Health Strategy, the Federal Government defined that the CPF would be the unique identification number of Brazilian citizens in dozens of public services\(^2\).

We believe that, given the regular administration by the Ministry of Economy of the CPF database (active, inactive, and invalid), this process may progressively favor a significant reduction in the millions of duplicate records of users of the Family Health Teams. After all, if the (conservative) results of the projection observed for the municipality of Rio de Janeiro (13%) are used for the whole country (which must be much worse, given the difficulty of small municipalities in managing the duplicate records), we will find about 20 million duplicate entries in the Family Health Strategy through the CPF alone. If we add to this the duplicates by name (homonyms; linkage of name, mother’s name and date of birth) and those registered in more than one health unit or more than one municipality, another million records will be found in this situation.

The redirection of the Ministry of Health towards the qualification of individual records is a virtuous point in this direction, as long as it is accompanied by continuous monitoring of electronic records and support from the State Health Secretariats for small municipalities, with less management capacity.

Expanding access is also enhancing the management of the lists of Brazilian citizens, uniquely identifying them, so that one can know the real statistical denominators for calculating the relevant indicators in each case, starting with the actual coverage of the Family Health Strategy.

**Collaborations**

LF Pinto and LJ Santos participated equally in all stages of drafting this paper.

**References**