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

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
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 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.

Analyzing the rate of implementation of electronic medical records by doctors in the provinces of Canada, Chang & Gupta highlight, first of all, the regional gap in the level of their implementation. They point out that funding in this country is not the main barrier, and the existence of a “super-user” clinical physician, a flexible regional leadership for the implementation, and local support are factors that have contributed to its success.

The literature highlights Denmark’s role in successfully achieving this goal. In a comparative report on the health systems of Germany, the United Kingdom, France, the Netherlands, and Denmark, Freudmann & Studer highlight that the latter country was the one that achieved the best results in the use of electronic medical records that reduced, on average, 50 minutes per day, the completion of paper forms, and increased the number of serviced patients by 10%.

Some authors point out that these records are a broad, low-cost data source, which allows longitudinal comparisons, cohort studies, quality assessment, and clinical trials with large sample sizes, which can be used for health surveillance, particularly for chronic diseases such as diabetes, hypertension, and cardiovascular diseases, contributing in no small part to the morbimortality of the Brazilian population. In 2019, the Ministry of Health published the results of the most recent edition of the Risk and Protective Factors for Chronic Diseases by Telephone Survey – VIGITEL 2018, highlighting the city of Rio de Janeiro as the federative unit with the highest prevalence of hypertension (31.2%) and diabetes (9.8%), and above the mean of the country’s capitals (24.7% and 7.7%, respectively). It is relevant to highlight that data in medical records are sources with greater power and representativeness of the population’s health (clinical morbidity) than population surveys such as VIGITEL, which consider the referred morbidity.

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, 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 J
neiro. These sheets contained administrative and sociodemographic data, essential to define the list of people responsible for each primary health care team—which had a reference limit of 3,000 people—allowing the construction of a baseline for the “management of duplicate user roster”. This management was facilitated when the City of Rio de Janeiro defined the CPF for people over 16 and the number of the Live Birth Certificates (DNV) for people below 16 years of age as the unique identification number of each person.

The Workshops were held after the printing and binding on A3 paper of the list of variables in Sheet A, by 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 one, 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 OTICS-RIO/SMS-RJ).

After the realization of these Workshops, the projection carried out brought evidence of approximately 340,000 records of homonymous people with duplicate records; that is, 13.7% of the people registered in the municipality (Table 1).

These records were revised and deleted over the following months, thus expanding the possibility of access and registration of new residents in the areas covered by the ESF. With this set of workshops developed in 2013, it was possible to create a baseline that determined the denominators to be used for the calculation of clinical and epidemiological indicators in the ensuing years.

In this study, we opted to analyze two of these population-based indicators, considering the resident population aged 18 years and over, which totaled 115,280 inhabitants and were monitored by the Family Health Teams in a planning area in the city of Rio de Janeiro.

Results and Discussion: the use of electronic medical records in Primary Health Care and the management of ‘duplicate’ records

Prevalence of hypertension and diabetes

The analysis of the health situation of a given area or territory is in constant construction/reconstruction. Thus, it can only be understood if analyzed as a result of a historical process in constant transformation. Ferreira defines the health situation as the knowledge and interpretation of the quality of life of the population of a given territory, historically produced and in a permanent process of transformation. Two of the leading chronic diseases monitored by teams in primary health care refer to hypertension and diabetes. Of the total number of people monitored by the Family Health Teams, 4.1% of Rio de Janeiro’s dwellers had both morbidities in the period studied. When observed in isolation, hypertension and diabetes rates stood at 16.7% and 9.6%, respectively (Table 2).

Our population-based findings report higher rates in females in most of the age groups considered, except for the last two age groups (75-79 years and 80 years and over), in which an under-registration of diagnosed cases is probable, with a positive association between prevalence rates and age group, in both genders. Older adults in these age groups have a harder time traveling to health facilities, and this evidence suggests that family health teams could review and intensify the needs for home visits for this population subgroup in the area in question.

The quality of the records that allowed the calculation of the prevalence of hypertension and diabetes for the population in the south region of the city of Rio de Janeiro was facilitated by the importance attributed by resident doctors working in most of the researched health units to the data records for the generation of information for decision-making.

The estimates found by gender and age range allow its use for future studies in the area covered by the Family Health Teams in the south region of the city of Rio de Janeiro. Both our findings and those of the VIGITEL-2018 study evidenced that women had higher prevalence rates. However, for hypertension, the situation presented by the Ministry of Health for the municipality of Rio de Janeiro is quite different: 35.3% vs. 16.7% for the area surveyed in the municipality of Rio de Janeiro, among PHC users. We believe that this occurs mainly because the middle and upper-class population is present in the Brazilian study and, proportionally, less frequent in the study with the Family Health Teams in the city of Rio de Janeiro, as well as the issue of measurement (self-informed vs. registered in medical records).

On the other hand, in the period studied for the municipality of Rio de Janeiro, the management of duplicate records allowed calculating prevalence more appropriately than the traditional way in local health systems in primary care that do not “overhaul” their databases and end
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.

| Planning Area (AP) | Population registered in electronic medical records | Duplicate population (for example, homonyms) |
|--------------------|------------------------------------------------------|-----------------------------------------------|
|                    | Total | Total with valid CPF/DNV | % valid CPF/DNV | Total records | Total with valid CPF/DNV recorded | Total with same CPF/DNV | % duplicate records (CPF/DNV) | Projected duplicate records among the registered population | Projected registered population without duplicate CPF/DNV |
| 1.0                | 123,456 | 33,323 | 27.0% | 10,071 | 2,607 | 151 | 5.8% | 7,151 | 116,305 |
| 2.1                | 157,819 | 55,119 | 34.9% | 30,124 | 9,950 | 1,007 | 10.1% | 15,972 | 141,847 |
| 2.2                | 57,736 | 15,234 | 26.4% | 10,383 | 2,739 | 910 | 33.2% | 19,182 | 38,554 |
| 3.1                | 375,544 | 79,777 | 21.2% | 60,293 | 13,870 | 2,029 | 14.6% | 54,937 | 320,607 |
| 3.2                | 225,945 | 113,023 | 50.0% | 29,243 | 13,599 | 1,873 | 13.8% | 31,120 | 194,825 |
| 3.3                | 393,579 | 123,069 | 31.3% | 69,162 | 20,183 | 2,495 | 12.4% | 48,654 | 344,925 |
| 4.0                | 125,821 | 43,930 | 34.9% | 18,357 | 5,738 | 873 | 15.2% | 19,143 | 106,678 |
| 5.1                | 332,569 | 87,496 | 26.3% | 57,789 | 15,401 | 1,478 | 9.6% | 31,916 | 300,653 |
| 5.2                | 355,289 | 93,156 | 26.2% | 76,515 | 20,361 | 3,598 | 17.7% | 62,783 | 292,506 |
| 5.3                | 329,618 | 104,697 | 31.8% | 62,855 | 21,567 | 2,825 | 13.1% | 43,176 | 286,442 |
| Total              | 2,477,376 | 748,824 | 30.2% | 424,792 | 126,015 | 17,239 | 13.7% | 338,908 | 2,143,343 |

Source: CEMAPS-RJ/SUBPAV/SMS-RJ, July 2013 (Rio de Janeiro, 201325).

Note: PAs correspond to the areas (health districts) for municipal health management in the Rio de Janeiro Municipality.

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 care27. 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)
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.

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

1. Canada Health Infoway. The emerging benefits of electronic medical record use in community-based care. Toronto: Health Canada Infoway; 2013.
2. McGinn CA, Grenier S, Duplantie J, Shaw N, Sicotte C, Mathieu L, Leduc Y, Légaré F, Gagnon MP. Comparison of user groups’ perspectives of barriers and facilitators to implementing electronic health records: a systematic review. BMC Medicine 2011; 9:46.
3. Pinto LFS. Estratégias de integração e utilização de bancos de dados nacionais para avaliação de políticas de saúde no Brasil [tese]. Rio de Janeiro: Fiocruz; 2006.
4. Hendy J, Reeves BC, Fulop N, Hutchings A, Masseria C. Challenges to implementing the National Programme for Information Technology (NPfIT): a qualitative study. BMJ 2010; 331(7512):331-336.
5. Rozenblum R, Jang Y, Zimlichman E, Salzberg C, Tamblyn R, Buckeridge D, Forster A, Bates DW, Tamblyn M. A qualitative study of Canada’s experience with the implementation of electronic health information technology. CMAJ 2011; 183(5):E281-E288.
6. Webster PC. National standards for electronic health records remain remote. CMAJ 2010; 182(9):888-889.
7. Webster PC. National electronic health records initiative remains muddled, auditors say. CMAJ 2010; 182(9):E383-E384.
8. Chang F, Gupta N. Progress in electronic medical record adoption in Canada. Can Fam Physician 2015; 61(12):1076-1084.
9. Freudmann A, Studer S, editors. Future-proofing Western Europe’s healthcare - a study of five countries: A report from the Economist Intelligence Unit. 2011. [accessed 2019 Nov 12]. Disponível em: http://www.reforminghealthcare.eu/uploads/Modules/Media-room/eiu-eucomed-future-proofing-healthcare-finalv2-web-5-10-11.pdf
10. Schoen C, Osborn R, Doty MM, Squires D, Peugh J, Applebaum S. A survey of primary care physicians in eleven countries, 2009: perspectives on care, costs, and experiences. Health Aff (Millwood) 2009; 28:w1171-1183.
11. Birtwhistle R, Williamson T. Primary care electronic medical records: a new data source for research in Canada. Can Med Assoc J 2015; 187(4):239-240.
12. Birtwhistle RV. Canadian Primary Care Sentinel Surveillance Network: a developing resource for family medicine and public health. Can Fam Physician 2011; 57(10):1219-1220.
13. World Health Organization (WHO). A Global Brief on Hypertension: Silent Killer, Global Public Health Crisis 2013 [Online]. [accessed 2019 Nov 19]. Disponível em: http://www.who.int/mediacentre/factsheets/fs183/en/
14. Ministério da Saúde (MS). Vigilância de Fatores de Risco e Proteção para Doenças Crônicas: Atividades Físicas. Vigilê 2016. [accessed 2019 Nov 5]. Disponível em: https://www.endocrino.org.br/ministerio-da-saude-divulga-dados-do-vgitel-2016/
15. Ministério da Saúde (MS). VIGITEL Brasil 2018: vigilância de fatores de risco e proteção para doenças crônicas por inquérito telefônico: estimativas sobre frequência e distribuição sociodemográfica de fatores de risco e proteção para doenças crônicas nas capitais dos 26 estados brasileiros e no Distrito Federal em 2018. Brasília: SVS/MS; 2019.
16. Stewart M, Thind A, Terry A, Chevendra V, Marshall I. Implementing and maintaining a researchable database from electronic medical records: a perspective from an academic family medicine department. Health Policy 2009; 85(2):26-39.
17. Biro S, Williamson T, Leggett JA, Barber D, Morkem R, Moore K, Belanger P, Mosley B, Janssen I. Utility of linking primary care electronic medical records with Canadian census data to study the determinants of chronic disease: an example based on socioeconomic status and obesity. BMC Medical Informatics and Decision Making 2016; 16:32.
18. Jutte DP, Roos LL, Brownell MD. Administrative record linkage as a tool for public health research. Annu Rev Public Health 2011; 32:91-108.
19. Justino ALA, Oliver LL, Melo TP. Implementation of the Residency Program in Family and Community Medicine of the Rio de Janeiro Municipal Health Department, Brazil. Cien Saude Colet 2016; 21(5):1471-1480.
20. Howe A, Padula AMI, Ribeiro JM, Pinto LF. 450th anniversary of the city of Rio de Janeiro: Primary Health Care Reform. Cien Saude Colet 2016; 21(5):1324-1326.
21. World Conference of Family Doctors (WONCA). 21st WONCA World Conference of Family Doctors, [accessed 2019 Nov 8]. Disponível em: http://www.wonca2016.com/ingles/index.php
22. Soranz D, Pinto LF, Oliveira GO. Themes and Reform of Primary Health Care (RCAPS) in the city of Rio de Janeiro, Brazil. Cien Saude Colet 2016; 21(5):1327-1338.
23. Brasil. Ministério da Saúde (MS). Secretaria de Atenção Primária à Saúde. Portal eGestor Atenção Básica. (eGESTOR AB). [accessed 2019 Nov 14]. Disponível em: https://egestorab.saude.gov.br/paginas/acessopublico/relatorios/relHistoricoCoberturaAB.xhtml
24. Pereira MG. Morbidade. In: Pereira MG. Epidemiologia: Teoria e Prática. Rio de Janeiro: Editora Guanabara Koogan; 1995. p. 320-330.
25. Ferreira SM. Curso de Aperfeicoamento em Gestão de Saúde. In: Programa de Educação à Distância. Sistema de Informação em Saúde. Rio de Janeiro: ENSP; 1998. p. 98.
26. Rio de Janeiro, Prefeitura da Cidade do Rio de Janeiro, Secretaria Municipal de Saúde. Cadernos de Estatísticas e Mapas da Atenção Primária em Saúde do Município do Rio de Janeiro (CEMAPS). Contribuições para a construção de uma linha de base para os cadastros das microredes da estratégia de Saúde da Família do município do Rio de Janeiro. Rede de Estações-Observatório das Tecnologias de Informação e Comunicação em Serviços de Saúde. Rio de Janeiro: Secretaria Municipal de Saúde; 2013.
27. Brasil. Portaria MS nº 2.979, de 12 de novembro de 2019. Institui o Programa Previne Brasil, que estabelece o modelo de financiamento de custeio da atenção primária à saúde no âmbito do SUS. Diário Oficial da União 2019; 13 nov.
28. Brasil. Decreto nº 9.795, de 17 de maio de 2019. Diário Oficial da União 2019; 20 maio.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License