saúde
EM DEBATE

ORIGINAL ARTICLE

Potential for health monitoring: Exploring SUS outpatient open data

Potencial no monitoramento em saúde: explorando os dados abertos ambulatoriais do SUS

Felipe Ferré[1,2], Amanda Oliveira Lyrio[2], Samara Helena de Carvalho[2], Laís Lessa Neiva Pantuzza[2], Jéssica Barreto Ribeiro dos Santos[2], Ana Carolina de Freitas Lopes[2], Luciene Fontes Schluckebier Bonan[2]

DOI: 10.1590/2358-28982025E110004I

ABSTRACT The Brazilian Unified Health System (SUS) has disseminated billions of administrative records corresponding to three decades of existence. Unlike hospital and notification data, which are fragmented as they are service-oriented rather than user-oriented, outpatient data can be linked via a pseudonymized identifier, enabling the tracking of therapeutic pathways. This paper presents an automated microdata processing tool from the Open Health Intelligence Platform (SABEIS) using open-source software to handle microdata provided through the TabWin/TabNet strategy from the file transfer directory, without relying on sophisticated Big Data. It describes the open data from the SUS Ambulatory Information System from 2008 to 2023, using modest hardware resources by public health and health informatics specialists. A total of 8,106,361,265 records were processed, corresponding to 3,135 procedures, 16,407 diagnoses, and 51,875,308 users, according to the files of APAC-SIA High-cost ambulatory procedures and high-cost medicines. There was a noticeable improvement in the quality of the pseudonymized identifier, especially from 2022 onwards, with 0.8% of users recorded as having more than one sex, more than one state of residence, over eight procedures, or five diagnoses, demonstrating the potential for public policy monitoring and knowledge generation. This approach demonstrated the potential for monitoring public policies using outpatient data with a pseudonymized SUS user identifier.

KEYWORDS Data science. Database management systems. Unified Health System. Documentation. Technology assessment, health.

RESUMO *O Sistema Único de Saúde (SUS) disseminou bilhões de registros administrativos correspondentes a três décadas de existência. Na contramão de dados hospitalares e de notificação fragmentados por serem orientados ao serviço, e não ao usuário do SUS, os dados ambulatoriais apresentam dados vinculáveis ao identificador pseudonimizado, viabilizando acompanhar o itinerário terapêutico. Este trabalho apresenta, uma ferramenta automatizada da Sala Aberta de Inteligência em Saúde (Sabeis) de processamento, com software livre, de microdados fornecidos via estratégia TabWin/TabNet a partir do diretório de transferência de arquivos, sem aparato sofisticado de Big Data; e descreve os dados abertos do Sistema de Informação Ambulatorial, de 2008 a 2023, empregando recursos modestos de hardware por especialistas em saúde pública e informática em saúde. Foram processados 8.106.361.265 registros, correspondentes a 3.135 procedimentos, 16.407 diagnósticos e 51.875.308 usuários, segundo o arquivo correspondente à Autorização de Procedimentos Ambulatoriais e Alta Complexidade/Custo. Verificou-se a crescente qualidade do identificador pseudonimizado, sobretudo a partir de 2022, com 0,8% dos usuários com mais de um sexo, mais de um estado de residência, acima de oito procedimentos ou cinco diagnósticos. A presente abordagem demonstrou a potencialidade para o acompanhamento de políticas públicas, utilizando os dados ambulatoriais com identificador pseudonimizado do usuário do SUS.*

PALAVRAS-CHAVE *Ciência de dados. Sistemas de gerenciamento de base de dados. Sistema Único de Saúde. Documentação. Avaliação de tecnologias em saúde.*

[1] Conselho Nacional de Secretários de Saúde (Conass) – Brasília (DF), Brasil.
labxss@gmail.com

[2] Ministério da Saúde (MS), Secretaria de Ciência, Tecnologia, Inovação e Complexo da Saúde (Sectics), Departamento de Gestão e Incorporação de Tecnologias em Saúde (DGITS) – Brasília (DF), Brasil.

Introduction

Although Brazil has an abundance of public records and open data systems, public administrators, civil society, and the academic community have spent the past three decades contending with fragmented and incomplete information. This persistent issue arises from the fact that not all public policies within the Unified Health System (SUS) have yet made consolidated microdata from different sources openly available. Historically, there have been significant gaps in open microdata from Primary Care and Pharmaceutical Assistance, both in their basic and strategic components. However, initiatives such as the Interagency Health Information Network (RIPSA) have been able to provide, over at least 20 years, analytical capacity with indicators using aggregated data that have shown the rapid transformation of Health Care Networks (RAS), especially outpatient services[1].

The strategic decision to make microdata publicly available, comprising detailed records of healthcare encounters and administrative activities without aggregation, has played a crucial role in enhancing the visibility and value of managers and professionals within the Unified Health System (SUS). When integrated, the historical data series that trace back to the earliest measures introduced under Unified Health System (SUS), such as the Basic Operational Norms (NOB), national policies, and other regulatory instruments, make it possible to conduct longitudinal assessments of patients' therapeutic journeys. Furthermore, the adoption of open data practices not only facilitates the monitoring of services provided by SUS but also fosters continuous improvement in the reliability of the data itself[2,3].

Although active transparency in public administration still faces obstacles and the dissemination of microdata remains partial[4], the SUS succeeded in developing, through TabNet and TabWin, an effective technology for tabulating and use of non-aggregated data. The dissemination tools implemented throughout the 1990s and 2000s have proven to be robust and relevant to this day, supported by a critical mass of public health professionals well-versed in their use. These tools have been essential for producing a significant share of the indicators and quantitative data used in health plans and management reports[5-7].

The official tools and datasets, however, do not present the available SUS microdata on Public Health Actions and Services (ASPS) in an integrated manner, which limits certain types of analysis. For instance, data from the Outpatient Information System (SIA) includes a pseudonymized identifier for SUS users, allowing for the tracking of patients' journeys within the system. Nevertheless, analyzing such data requires sophisticated methods and techniques capable of handling billions of records, expertise that is often not accessible to health council members, public health professionals, epidemiologists, or specialists engaged in Health Technology Assessment (HTA). This analytical capacity is crucial for evaluating and deciding whether new drugs, products, and procedures should be incorporated into the SUS[8].

The lack of integrated data and the absence of official curation methods compel each research or management unit to develop its own tools for extracting, transforming, and loading (ETL) open data produced through the SIA. As a result, the consolidation of these data may vary depending on the specific methods applied[5,9-16].

This study brings together the fields of health informatics and public health, a convergence that is strategically important for the country[17,18], focusing on the use of real-world data to inform public health decision-making. To date, there have been no reports of fully automated processing of SIA data using free and open-source software; existing experiences have been limited to data consolidation, often without an emphasis on integrating outpatient data or relying on proprietary tools[10].

No studies were found using the complete open SIA dataset that address the

pseudonymized identifier for SUS users, which is based on the National Health Card (CNS) and widely used for decision-making within the SUS, although there is a tool that applies deterministic linkage of diagnostic and treatment information from SIA specifically for oncology[19]. However, there is extensive literature on the use of deterministic-probabilistic record linkage with restricted-access data containing personally identifiable information, such as full name, mother's name, identification documents, and full address, among others[13,14,20–23], whose microdata with pseudonymized identifiers derived from linking tools are not available for comparison.

Although the volume of information available within the SUS is growing exponentially, gaps remain that require attention to support public policy and ensure comprehensive care for system users. Therefore, this study aims to explore the open data of the SIA and investigate potential opportunities for longitudinal monitoring, using the pseudonymized SUS user identifier to track patients. Such monitoring can provide valuable support for decision-making, particularly regarding the incorporation, removal, or modification of health technologies within the SUS.

Material and methods

This study is characterized as exploratory and descriptive research, based on the analysis of secondary data from SUS information systems, with a focus on the construction and structuring of an integrated database compiled from multiple subsystems.

Source and nature of data

The data used in this study are publicly accessible and were obtained from the repository of the Brazilian Ministry of Health. The files are disseminated and fragmented by subsystem, across the 27 Federative Units (UF), and by reference date formatted as year and month in two digits each (YYMM), resulting in thousands of files that must be processed.

Computational architecture and environment

The ETL process was carried out using GNU Bash 5.2.15 for routine automation; SQL (PostgreSQL 15.6, Ubuntu 23.10.1) for data structuring, transformation, and loading; Wine 8.0.1 to emulate the official dbf2dbc.exe decompression tool; and dbview for extracting DBF data into CSV format. The computational environment consisted of a Lenovo ThinkCentre M75q Gen 2 mini PC with 32 GB of RAM, an AMD Ryzen™ 5 PRO 5650GE ×12 processor, and a 4TB Kingston SNV2S4000G SSD. The source code and data repository are available under the General Public License (GPL 3.0) and the Open Database License (ODbL).

Data processing and integration (ETL)

Data consolidation followed the methodology of Sala Aberta de Inteligência em Saúde (Open Health Intelligence Room, SABEIS). The process consisted of the following steps[24]: (1) automated collection, including a file size check (in bytes) to ensure download integrity; (2) decompression and conversion of files from DBC to DBF format, followed by extraction to CSV; (3) loading into the Database Management System (DBMS) as a staging layer of raw or semi-processed data; (4) load validation, verifying that the number of records in the original file matched those loaded into PostgreSQL, ensuring consistency; 5) comparison of the approved number of procedures with the official data provided by the TabNet25 tabulator; 6) standardization of attributes according to the SUS data dictionary, maintaining a unified structure among the SAI subsystems; 7) loading of object-oriented SQL tables into the DBMS, using the Inherits feature to structure 'parent

tables' by subsystem; 8) generation of data marts[26,27], in accordance with business rules for analysis and reports; 9) addition of the National Health Card (CNS) to the Outpatient Care (PA) file from other subsystems, using the procedure authorization code, corresponding to the state manager and the reference month28. The user identifier is the same across the different subsystems.

Resolution No. 510 of April 7, 2016, issued by the National Health Council[29], Resolution No. 510 of April 7, 2016, issued by the National Health Council, stipulates in Article 1, item V, that research using 'publicly accessible information' does not require evaluation by a Research Ethics Committee (REC).

Results

SIA data from 2008 to 2023 (*table 1*) were processed, and the pseudonymized user identifier was qualified (*tables 2 to 4*), allowing assessment of users' trajectories within the SUS (*tables 2 and 5*). The extraction resulted in approximately 1 terabyte of disk storage. A total of 44,637 data files were processed, containing 8,106,361,265 records, of which 5,024,137,874 records from the Outpatient Care Administrative Records (SIA PA) file needed to be enriched with other reports to obtain the pseudonymized identifier.

According to the municipality of residence, all municipalities and the Federal District recorded at least one entry in the SIA. However, it is important to note that not all municipalities or administrative regions of the Federal District reported procedures during the evaluated period for every policy (*table 1*). The policy with the lowest coverage in terms of municipalities was the Multiprofessional Follow-up (AMP), with 1,050 municipalities included.

Table 1. Summary of administrative data generated by the Outpatient Information System from 2008 to 2023

Subsystem	Period	Original files	FTP Mbytes	Mbytes database**	Records	List of Procedures***	List of Diagnoses ****	Municipalities of residence
AB	01/2008-04/2017	544	438	897	211,252	1	-	1,425
ACF	08/2014-12/2023	2,945	14	29	315,631	1	-	5,378
AD	01/2008-12/2023	5,182	2,453	6,225	49,021,697	211	2,518	5,597
AM	01/2008-12/2023	5,131	13,115	35,958	287,923,699	608	660	5,591
AMP	03/2016-12/2023	541	3,821	6,570	64,900	2	-	1,050
AN	01/2008-10/2014	2,145	305	508	6,534,272	8	376	5,413
AQ	01/2008-12/2023	5,143	103,923	252,140	49,523,199	157	647	5,596
AR	01/2008-12/2023	4,683	137,080	563,504	3,615,964	39	719	5,595
ATD	08/2014-12/2023	3,047	51	322	12,391,833	6	-	5,551

Table 1. Summary of administrative data generated by the Outpatient Information System from 2008 to 2023

Subsystem	Period	Original files	FTP Mbytes	Mbytes database**	Records	List of Procedures***	List of Diagnoses ****	Municipalities of residence
BI	01/2008-12/2023	5,184	2,242	12,156	2,546,881,910	1,893	16,199	5,599
PA	01/2008-12/2023	5,447	24	62	5,024,137,874	3,135	16,184	5,599
PS	11/2012-12/2023	3,557	762	1,380	122,214,893	21	3,858	5,064
SAD	04/2012-10/2018	1,088	4	13	3,524,141	69	4,552	530
Total	**01/2008-12/2023**	**44,637**	**264,232**	**879,764**	**8,106,361,265**	**3,135**	**16,407**	**5,599***

Source: Own elaboration.

* To evaluate access, we chose to use the Administrative Regions (ARs) that comprise the Federal District, thus considering 5,569 municipalities and 30 ARs registered with the SUS (although there are 35 regions since Decree No. 38,094/2017). ** Data load with selected attributes. *** The RENASES (National List of Health Actions and Services) list is maintained by the SUS Management System for the Table of Procedures, Medications, Orthotics, Prosthetics, and Special Materials (SIGTAP), with distinct codes being counted. **** List of distinct codes according to the International Classification of Diseases, tenth revision (ICD-10).

AB = Bariatric Surgery Follow-up; ACF = Arteriovenous Fistula Construction; AD = Miscellaneous Reports; AM = Medications; AMP = Multidisciplinary Follow-up; AN = Nephrology; AQ = Chemotherapy; AR = Radiotherapy; ATD = Dialysis Treatment; BI = Individual Report; PA = Outpatient Production; PS = Psychosocial; SAD = Home Care. SIGTAP – SUS Table Management System for Procedures, Medications, and Orthotics, Prosthetics, and Special Materials (OPM). Primary diagnosis according to ICD-10. File Transfer Protocol (FTP).

Regarding the municipality of residence, all municipalities and the Federal District recorded at least one entry for Miscellaneous Reports (AD), Medications (AM), Chemotherapy (AQ), Individual Bulletin (BI), and, consequently, Outpatient Care Records (PA). However, no Radiotherapy (AR) records were found in the SIA for residents of Santa Isabel do Rio Negro-AM and Uiramutã-RR. The municipalities with the lowest relative coverage for SIA reports were Amaturá-AM, Nova Roma-GO, Serra Nova Dourada-MT, Nova Roma do Sul-RS, Putinga-RS, and Nova Castilho-SP, which never had records for Bariatric Surgery Follow-up (AB), Arteriovenous Fistula Creation (ACF), Multiprofessional Follow-up (AMP), Nephrology (AN), Dialytic Treatment (ATD), Psychosocial (PS), or Home Care (SAD).

Using the pseudonymized identifier, it is possible to follow the same user's access to different health policies, as well as their staging in oncological treatment recorded via the SIA[30]. *Table 2* presents absolute numbers and rates per 100,000 inhabitants for Brazil and its five regions. For example, it allows analysis of access to oncological treatment policies (AQ or AR) provided to the same AM user, even if it did not occur during the same period. It is important to note that most medications recorded in the SIA belong to the Specialized Component of Pharmaceutical Assistance (CEAF). Between 2008 and 2023, 266,500 users out of 2,666,836 (10.0%) AQ users accessed AM, while 170,633 users out of 1,666,261 (10.2%) AR users accessed AM at some point. These overall data can help highlight inequalities in Public Health Actions and Services (ASPS). Assuming the number of records as a proxy for access, it was observed that access for AQ was three times higher in the South region compared to the North, and 1.7 times higher when considering total PA records.

Table 2. Users of the Outpatient Information System by subsystem, 2008–2023. Absolute numbers are shown in the lower triangle (italics), with users per 100,000 inhabitants by Brazilian region (2a), and use of multiple outpatient service types (2b)

	AB	ACF	AD	AM	AMP	AN	AQ	AR	ATD	BI	PA	PS	SAD
Table 2a													
North	1.1	72.6	4,582.8	1,532.1	1.8	85.0	626.9	386.6	164.1	56,927.2	17,560.9	2,139.7	23.5
Northeast	2.1	108.8	6,219.6	3,105.7	4.2	123.6	1,090.9	673.8	232.5	38,917.8	28,031.8	2,512.8	44.1
Southeast	14.7	104.2	6,592.4	6,366.5	10.5	157.7	1,447.4	938.4	252.4	64,636.1	26,230.9	2,250.0	83.3
South	59.8	108.7	7,521.2	4,953.5	13.1	159.5	1,924.5	1,162.6	276.8	92,123.0	29,955.7	3,062.5	40.3
Central-West	0.1	98.8	5,546.3	3,407.7	3.9	136.2	1,029.7	542.8	262.3	45,555.3	18,145.2	2,173.0	32.6
	AB	**ACF**	**AD**	**AM**	**AMP**	**AN**	**AQ**	**AR**	**ATD**	**BI**	**PA**	**PS**	**SAD**
Table 2b													
AB	***31,757***	0.1	11.7	12.0	0.0	0.1	1.3	1.0	0.1	156.0	183.0	4.6	0.1
ACF	16	***208,654***	708.2	908.5	16.0	119.2	27.1	15.0	1,126.5	804.1	1,202.3	15.0	1.6
AD	2,036	122,909	***12,904,559***	7,966.8	32.8	947.2	1,846.9	1,156.6	1,523.0	42,692.8	74,326.7	2,738.2	75.6
AM	2,075	157,673	1,382,627	***9,362,825***	30.5	1,068.5	1,535.6	983.2	1,793.7	27,003.1	53,946.4	2,349.8	87.3
AMP	0	2,771	5,701	5,298	***16,034***	0.3	3.4	2.1	21.6	68.6	92.4	1.3	0.1
AN	11	20,688	164,393	185,436	55	***280,276***	32.5	21.0	522.7	722.6	1,614.9	4.6	3.6
AQ	225	4,696	320,530	266,500	588	5,641	***2,666,836***	6,588.9	82.0	10,055.9	15,365.0	144.8	27.9
AR	168	2,605	200,720	170,633	368	3,651	1,143,500	***1,666,261***	47.0	6,471.8	9,600.0	100.8	19.8
ATD	26	195,497	264,319	311,303	3,757	90,709	14,239	8,150	***485,605***	1,715.8	2,798.1	32.4	5.4
BI	27,078	139,554	7,409,285	4,686,353	11,897	125,403	1,745,184	1,123,166	297,768	***119,162,740***	175,307.6	13,843.0	350.7
PA	31,757	208,654	12,899,319	9,362,329	16,034	280,258	2,666,584	1,666,074	485,600	30,424,434	***51,875,308***	7,809.5	259.3
PS	807	2,609	475,205	407,802	223	797	25,137	17,501	5,615	2,402,434	1,355,333	***4,903,392***	9.6
SAD	23	285	13,126	15,154	16	628	4,846	3,433	944	60,872	45,003	1,673	***116,282***

Source: Own elaboration.

AB = Bariatric Surgery Follow-up; ACF = Arteriovenous Fistula Construction; AD = Miscellaneous Reports; AM = Medications; AMP = Multidisciplinary Follow-up; AN = Nephrology; AQ = Chemotherapy; AR = Radiotherapy; ATD = Dialysis Treatment; BI = Individual Report; PA = Outpatient Production; PS = Psychosocial; SAD = Home Care. Population according to the 2022 Census.

Tables 3 and 4 help characterize the quality of the pseudonymized identifier. The distribution of sex, race/color, and age shows consistency with what is typically observed within the SUS. Likewise, it is possible to clean certain data strata by disregarding pseudonymized identifiers with an unusually high number of records, which may indicate that the same CNS was used for different individuals. In addition, one can examine the number of states of residence, which may suggest the same number being used by different managers; an excessively high number of procedures and diagnoses that are statistically implausible for a single individual; and discrepancies in treatment duration, among other indicators. When assessing users of both sexes in administrative records, more than two states of residence, fifteen procedures, or ten diagnoses, only 3.05% fall into this category and may even be excluded from the analysis based on this exclusion criterion.

Table 3. Quality of the pseudonymized identifier based on the National Health Card in relation to the total number of users with Outpatient Care Administrative Records ('*Produção Ambulatorial*', SIA PA) in the Outpatient Information System, from 2008 to 2023, and the percentage of records requiring further investigation by period and by region*

Variable	Users	North	Northeast	Southeast	South	Central-West
Total	**51,875,308**	**3,000,390**	**15,093,470**	**21,985,990**	**8,810,592**	**2,903,022**
Sex**						
Female only	56.6%	56.3%	59.4%	55.0%	56.6%	53.8%
Male only	40.1%	38.6%	37.6%	41.4%	41.1%	41.6%
Both	3.4%	5.1%	3.0%	3.6%	2.3%	4.6%
Race/color***						
Unidentified	35.5%	30.6%	41.7%	33.2%	33.1%	32.9%
White	30.8%	5.1%	8.0%	40.2%	57.7%	23.4%
Black/brown	30.0%	58.7%	42.4%	25.2%	8.1%	38.8%
Yellow	3.7%	5.5%	7.9%	1.5%	1.1%	4.5%
Indigenous	0.1%	0.1%	0.0%	0.0%	0.0%	0.4%
Age***						
00-17	11.7%	14.5%	12.8%	10.2%	12.2%	12.0%
18-40	25.0%	27.4%	27.5%	22.6%	25.6%	25.1%
41-60	31.3%	31.6%	31.5%	31.5%	30.4%	31.9%
60+	31.8%	26.1%	27.8%	35.5%	31.5%	30.8%
Unknown	0.3%	0.3%	0.3%	0.3%	0.3%	0.3%
States of residence						
1	98.9%	98.5%	99.0%	99.0%	98.9%	98.1%
2-14	1.1%	1.5%	1.0%	1.0%	1.1%	1.9%
Procedures						
1	47.7%	48.0%	49.6%	46.8%	45.8%	49.7%
2	19.3%	16.8%	19.8%	19.0%	19.8%	20.9%
3	9.3%	7.6%	9.4%	9.3%	9.9%	8.7%
4	5.9%	4.5%	5.7%	6.2%	6.0%	5.5%
5	3.9%	3.0%	3.8%	3.9%	4.1%	3.7%
6	2.9%	2.3%	2.9%	3.0%	2.9%	2.6%
7	2.2%	1.6%	2.2%	2.2%	2.1%	2.0%
8	1.7%	1.5%	1.6%	1.9%	1.7%	1.6%
9	1.4%	1.3%	1.1%	1.6%	1.4%	1.1%
10 to 220	5.8%	13.3%	3.9%	6.1%	6.3%	4.1%
Diagnoses						
0	23.9%	26.9%	31.1%	17.4%	28.0%	19.8%
1	50.7%	48.6%	47.5%	53.3%	48.6%	56.1%
2	14.4%	13.6%	12.6%	16.0%	13.9%	14.8%
3	5.5%	5.2%	4.6%	6.4%	5.1%	5.1%
4	2.5%	2.3%	2.0%	3.0%	2.1%	2.0%
5	1.2%	1.3%	1.0%	1.5%	1.0%	0.9%
6	0.7%	0.7%	0.5%	0.8%	0.5%	0.4%
7 to 2,098	1.2%	1.4%	0.8%	1.6%	0.8%	0.8%

Table 3. Quality of the pseudonymized identifier based on the National Health Card in relation to the total number of users with Outpatient Care Administrative Records ('*Produção Ambulatorial*', SIA PA) in the Outpatient Information System, from 2008 to 2023, and the percentage of records requiring further investigation by period and by region*

Variable	Users	North	Northeast	Southeast	South	Central-West
Total	**51,875,308**	**3,000,390**	**15,093,470**	**21,985,990**	**8,810,592**	**2,903,022**
Investigate (users of both genders, from more than 2 states, 15 procedures, or 10 diagnoses						
2008	8.5%	11.2%	7.3%	8.7%	8.4%	9.6%
2009 to 2015	5.8%	8.5%	5.3%	6.1%	5.2%	5.9%
2016 to 2021	6.4%	16.1%	4.5%	6.8%	5.4%	7.4%
2022	4.5%	15.3%	2.7%	4.5%	3.6%	4.6%
2023	3.1%	11.4%	1.8%	2.9%	2.8%	3.3%

Source: Own elaboration.

*According to the municipality of residence at the first appointment. **Distinct sexes recorded at each appointment. ***At the first appointment.

In *table 4*, when applying stricter criteria for data cleaning in oncological treatment, considering users recorded as both sexes, with more than one state of residence, more than eight procedures, or more than six diagnoses, the proportion of non-unique users drops to 0.46%. These records should be cleaned using additional attributes (such as age and treatment profile) or excluded from the analysis, depending on the study's objective. The quality of the pseudonymized identifier has improved over time with the evolution of CadSUS, through the reconciliation of local and national registries, the integration of the Master Patient Index (MPI) with the Federal Revenue Service database starting in 2016, and the designation of the Individual Taxpayer Registry (CPF) as the single national identifier beginning in 2021. is worth noting that data cleaning requirements for SUS user records remain considerably higher in the North region compared to other regions, reaching 11.4% in 2023, while the lowest rate was observed in the Northeast, at 1.8%, according to the criteria applied. When analyzing specific policies, such as oncology, the proportion of records requiring cleaning in the North was 1.5% of cases, compared to 0.3% in the Southeast in 2023.

Table 4. Quality of the pseudonymized identifier based on the National Health Card for the total number of oncology users (chemotherapy – AQ and radiotherapy – AR) with Outpatient Care Administrative Records in the Outpatient Information System, from 2008 to 2023, and percentage of records requiring investigation by period and by region*

Variable	User	North	Northeast	Southeast	South	Central-West
Total	**3,168,693**	**135,069**	**701,393**	**1,455,544**	**672,032**	**204,655**
Sex**						
Female only	53.7%	57.5%	56.1%	52.7%	52.8%	53.5%
Male only	45.7%	41.2%	43.3%	46.8%	46.7%	45.8%
Both	0.6%	1.4%	0.6%	0.5%	0.5%	0.6%
Race/color***						
White	46.7%	8.7%	17.6%	52.6%	77.9%	27.0%
Black/brown	33.8%	54.5%	56.0%	31.8%	8.8%	40.1%

Table 4. Quality of the pseudonymized identifier based on the National Health Card for the total number of oncology users (chemotherapy – AQ and radiotherapy – AR) with Outpatient Care Administrative Records in the Outpatient Information System, from 2008 to 2023, and percentage of records requiring investigation by period and by region*

Variable	User	North	Northeast	Southeast	South	Central-West
Total	**3,168,693**	**135,069**	**701,393**	**1,455,544**	**672,032**	**204,655**
Unidentified	17.5%	35.8%	20.7%	14.5%	12.6%	31.9%
Yellow	2.0%	0.8%	5.7%	1.1%	0.7%	0.8%
Indigenous	0.0%	0.2%	0.0%	0.0%	0.0%	0.1%
Age* **						
00-17	2.6%	5.0%	3.3%	2.2%	1.9%	3.3%
18-40	8.9%	13.8%	10.2%	7.9%	8.3%	10.5%
41-60	35.5%	38.6%	35.4%	34.9%	35.8%	37.6%
60+	52.8%	42.5%	50.9%	54.9%	53.8%	48.4%
Unknown	0.1%	0.2%	0.1%	0.1%	0.1%	0.2%
States of residency						
1	99.2%	96.0%	99.1%	99.6%	99.4%	98.5%
2-4	0.8%	4.0%	0.9%	0.4%	0.6%	1.5%
Procedures						
1	38.2%	37.9%	37.0%	38.2%	39.1%	39.7%
2	17.8%	17.2%	18.1%	17.6%	17.9%	18.6%
3	9.3%	11.7%	10.1%	9.1%	8.2%	9.4%
4	8.5%	9.8%	8.0%	9.2%	7.1%	8.5%
5	10.8%	10.1%	10.8%	11.1%	10.5%	9.5%
6	8.3%	7.4%	8.2%	8.1%	9.2%	7.8%
7	4.1%	3.5%	4.7%	3.8%	4.4%	3.8%
8	1.7%	1.3%	1.9%	1.6%	2.0%	1.6%
9-18	1.3%	0.9%	1.3%	1.2%	1.5%	1.2%
Diagnoses						
1	70.5%	68.1%	69.8%	71.2%	70.3%	69.8%
2	24.0%	26.4%	25.2%	23.2%	23.6%	25.1%
3	4.5%	4.5%	4.2%	4.5%	4.9%	4.3%
4	0.8%	0.8%	0.6%	0.9%	1.0%	0.7%
5	0.2%	0.1%	0.1%	0.2%	0.2%	0.1%
6	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%
7+	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Investigate (users of both sexes, more than 1 state, 8 procedures, or 6 diagnoses)						
2008	3.8%	12.7%	3.6%	3.2%	3.5%	4.4%
2009 a 2015	3.2%	5.4%	3.6%	2.7%	3.2%	3.9%
2016 a 2021	2.4%	6.4%	2.4%	1.8%	2.5%	3.1%
2022	0.8%	3.8%	0.8%	0.5%	0.6%	1.1%
2023	0.5%	1.5%	0.4%	0.3%	0.4%	0.6%

Source: Own elaboration.

* According to the municipality of residence at the first appointment. ** Different sexes recorded at each appointment. *** At the first appointment.

The transparency provided by open data and their organization in repositories or data lakes allows for the exploration of how certain policies are implemented and enables both exploratory and data-driven analyses. For example, as shown in *table 5*, hemodialysis was provided for 466,918 users, whose average cost of BRL 81,945.25, according to official figures without deflation, corresponding to a long-term treatment, with a median duration of 818 days. Additionally, neuropsychomotor development rehabilitation stands out, with a median age of 6 years and an average cost per user of BRL 6,820.61, accounting for 426 procedures per user based on a simple average.

Table 5. Characterization of procedures performed more than three times per user among the SUS users with the highest numbers, from 2008 to 2023

Procedure	Users	Records per user	Quantity per use	Value per user (BRL)	Treatment days (median)	Age (median)
Hemodialysis (maximum 3 sessions per week)	466,918	36	445	81,945.25	818	56
Urea measurement	495,295	33	63	116.26	849	56
Potassium measurement	482,057	34	34	63.05	848	56
Care/monitoring of patients undergoing neuropsychomotor development rehabilitation	103,736	155	426	6,820.61	605	6
Glutamic-pyruvic transaminase (GPT) dosage	477,795	33	33	67.33	849	56
Phosphorus measurement	448,613	36	36	65.99	820	56
Calcium measurement	451,827	35	35	65.54	847	56
Formoterol 12 mcg + Budesonide 400 mcg inhalation powder (per 60-dose bottle)	457,793	33	35		1,034	59
Physiotherapy for motor disorders	512,731	27	69	324.80	147	50
Atorvastatin 20 mg (per tablet)	461,091	28	1,221		879	62
Alfaepoetin 4,000 IU injectable (per vial)	491,314	25	279		848	59
Creatinine dosage	495,165	19	19	35.97	819	56
Olanzapine 10 mg (per tablet)	235,140	38	1,554	1,423.08	1,153	39
Hematocrit	339,472	26	26	40.33	849	56
Hemoglobin measurement	336,777	26	26	40.21	849	56
Complete blood count	480,839	17	17	70.27	969	55
Specialized medical consultation	789,478	10	11	115.88	666	53
Formoterol 12 mcg + Budesonide 400 mcg (per inhalation capsule)	295,265	27	1,311		941	61
Glucose measurement	352,682	21	21	38.41	939	57
Azathioprine 50 mg (per tablet)	191,635	35	2,622		1,095	39
Hormone therapy for advanced prostate adenocarcinoma – 1st line	270,434	25	25	7,491.04	697	72
Atorvastatin 40 mg (per tablet)	321,568	21	751		606	64

Table 5. Characterization of procedures performed more than three times per user among the SUS users with the highest numbers, from 2008 to 2023

Procedure	Users	Records per user	Quantity per use	Value per user (BRL)	Treatment days (median)	Age (median)
Sodium measurement	247,959	27	27	49.72	757	57
Ferric hydroxide saccharate 100 mg injectable (per 5 ml vial)	345,532	19	82	355.58	817	57
Hormone therapy for stage II breast carcinoma	177,361	36	36	2,883.72	1,154	57
Leflunomide 20 mg (per tablet)	191,217	33	994	1,241.15	971	55
Tacrolimus 1 mg (per capsule)	115,763	51	6,692	229.86	1,549	46
Timolol 5 mg/ml ophthalmic solution (per 5 ml bottle)	219,855	25	25		909	65
Risperidone 2 mg (per tablet)	193,391	28	1,668	132.41	789	38
Alkaline phosphatase measurement	365,816	15	15	29.39	1,031	55

Source: Own elaboration.

Note: Values are not deflated. Medication prices are not maintained in the SUS Management System for Procedures, Medications, Orthotics, Prosthetics, and Special Materials (SIGTRAP). and, therefore, are not recorded in the Outpatient Information System (SIA).

Discussion

The study highlighted the ability to manage and improve open data without relying on professionals specialized in server administration, but rather with public health specialists who possess substantial knowledge of information technology, taking on a role similar to that of health data analysts or data engineers[27].

The SUS recorded 3,135 outpatient procedures in the SIA, according to the Management System for the Table of Procedures, Medicines, Orthoses, Prostheses, and Special Materials (SIGTRAP), performed under 16,407 primary diagnoses. The open data are not provided in a user-oriented format but are fragmented by the corresponding Public Health Actions and Services (ASPS). Consequently, when an administrative decision is made to discontinue a given ASPS in the SIA, the user may no longer be monitored in an integrated manner across other outpatient and specialized policies, as illustrated in *table 1* with data from AN reports, which only covers the period from 2008 to 2014.

When assessing the absence of procedures performed according to the municipality of residence, the data may reveal gaps in healthcare provision. However, they may also indicate a practice of registering patients outside their domicile to gain access, revealing weaknesses in the Integrated Regional Planning (PRI) and intergovernmental conflicts, and consequently, challenges in coordination between municipalities. Health services should ideally be resolved at the regional level rather than entirely within a single municipality, and access should not be restricted for individuals who do not reside in the municipality where care is provided[31,32].

Although open data are widely used in academic research and by SUS management, the use of the national health card has been questioned in various forums due to possible inconsistencies. These include duplicate records resulting from local information systems in municipalities and states that issue their own identification numbers for citizens, without synchronizing them with the national database maintained by the Ministry of Health[33].

Assessing the quality of pseudonymized identifiers is essential, as illustrated in *tables 3 and 4*. The analysis enables the identification and exclusion of records whose identifiers present inconsistencies, such as the possible use of the same identification number by different managers for different individuals. Several indicators can be used to refine the dataset, including the number of states of residence, discrepancies in sex or age (calculated from the first record and the date of care), divergent treatment durations, and a statistically implausible number of procedures or diagnoses for a single individual. For example, only 3.05% of records showed inconsistencies, such as both sexes recorded, more than two states of residence, fifteen or more procedures, or ten or more diagnoses, supporting the use of these exclusion criteria to ensure the integrity of the analysis.

Thus, this study demonstrated that the national data consolidation and the provision of the pseudonymized identifier can be leveraged to define population strata according to health policies, and even to establish cohorts based on treatment or diagnosis.

It is important to highlight the practical expertise of health informatics professionals in employing simple coding approaches and household-level hardware resources. One of the main challenges in processing outpatient data lies in handling the original DBC files, which contain millions of records. The official decompression tool provided by the Ministry of Health, dbf2dbc.exe, was the only one capable of processing all data files using domestic computing resources without errors. Consistent results could not be obtained when alternative tools were used for states with larger datasets, such as São Paulo and Minas Gerais, particularly after 2019, when employing alternative statistical or interpreted programming languages such as R and Python on household computers[9,34].

Bash is also an interpreted language, that is, it functions as a command interpreter, executing instructions provided by the operating system. However, Bash is primarily designed for interactive use, which means it is optimized for the immediate execution of system commands without requiring compilation or complex code structuring. The advantage of using Bash for ETL processing lies in its focus on automating tasks in Unix and Linux systems, including executing operating system commands, manipulating files, processing text, and creating terminal (shell) scripts.

The processing was carried out without using specialized big data tools, avoiding visual ETL platforms with 'drag-and-drop' functionalities. Instead, Bash ingestion techniques and SQL operations were employed, following best practices for development with open-source software. As a result, this approach is optimized for operations common to the command-line environment rather than ETL solutions from ecosystems such as Pentaho or Informatica PowerCenter®, including direct and efficient access to Unix/Linux operating system resources. This means that many operations can be performed without the overhead associated with initializing a virtual machine (as in Python) or processing complex data (as in R). Additionally, the approach is characterized by low memory and processing overhead due to its direct nature and focus on simple system tasks[35-38].

Deficiencies in financial transfers can create inequities in access to Dialytic Treatment (ATD) across different regions of Brazil and hinder evaluation efforts when data are missing from the information system, whether related to procedures or medications. Consequently, gaps in the dissemination of official data linked to the original information systems can lead, for example, to discrepancies in reported expenditures across different sources, with estimates varying according to the methods of the pharmacoeconomic study.

Conclusions

This study underscored the importance of standardizing open microdata within an integrated dissemination framework to facilitate longitudinal analyses of administrative data, serving as a shared and accessible resource for dialogue between the State and civil society. Unfortunately, hospital data, disease notifications, live births, deaths, and immunization records, although available in the TabWin/TabNet system, are not provided in an integrated, open format. While these records are individualized by care contact or user, their current structure prevents longitudinal analysis of patient pathways within the SUS. Nevertheless, when systematically compiled and organized into data lakes, such microdata offers considerable potential for ecological studies at the municipal level, supporting indicator systems and situational analysis platforms[40-43].

The study demonstrated the importance of assessing ETL quality at each stage and transparently reporting the resulting data in studies involving large volumes of disseminated information. This work advocates for the practice of open science with the provision of source code, not only to ensure reproducibility but, above all, to address the challenges posed by deep learning neural networks and the inevitable use of generative Artificial Intelligence (AI). Careful attention to data quality can serve as a safeguard against AI hallucinations and spurious assertions, preventing the creation of additional vectors of misinformation.

Digital transformation in health is expected to provide integrated data lakes, enabling researchers and managers to leverage health informatics knowledge using a comprehensive data foundation. This would prevent fragmentation of data by management service, care, or surveillance conducted before the establishment of the National Health Information and Informatics Policy (PNIIS), while promoting user-centered SUS data in line with the principles of tripartite agreements[44-46].

Collaborators

Ferré F (0000-0001-9879-4782)* contributed to conceptualization, investigation, data curation, formal analysis, methodology, and manuscript writing. Lyrio AO (0000-0001-7740-2524)* contributed to conceptualization, investigation, data curation, formal analysis, investigation, methodology, and manuscript writing. Carvalho SH (0009-0005-7716-8164)* contributed to conceptualization, investigation, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources provision, software provision, supervision, validation, visualization, writing, review, and manuscript editing. Pantuzza LLN (0000-0001-9831-8807)* contributed to conceptualization, data curation, writing, review, and manuscript editing. Santos JBR (0000-0002-5528-0658)* contributed to conceptualization, data curation, writing, review, and manuscript editing. Lopes ACF (0000-0001-5806-5155)* contributed to conceptualization, project administration, resource provision, supervision, validation, and manuscript writing, review, and editing. Bonan LFS (0000-0003-2857-2927)* contributed to project administration, resource provision, and manuscript supervision. ■

*Orcid (Open Researcher and Contributor ID).

References

1. Pinto LF, Freitas MPS, Figueiredo AWS. Sistemas Nacionais de Informação e levantamentos populacionais: algumas contribuições do Ministério da Saúde e do IBGE para a análise das capitais brasileiras nos últimos 30 anos. Ciênc saúde coletiva. 2018;23(6):1859-70. DOI: https://doi.org/10.1590/1413-81232018236.05072018

2. Fernandes FECV, Leal IS, Andrade JDA. Percepção dos profissionais da atenção primária à saúde sobre o sistema de informação ambulatorial. Rev Enferm Atenção Saúde. 2017;6(2):77-92. DOI: https://doi.org/10.18554/reas.v6i2.1673

3. Silva AR, Oliveira TM, Lima CF, et al. Sistemas de informação como instrumento para tomada de decisão em saúde: revisão integrativa. Rev enferm UFPE on line. 10(9):3455-62. DOI: https://doi.org/10.5205/1981-8963-v10i9a11428p3455-3462-2016.

4. Maduro-Abreu A, Litre G, Santos L, et al. Transparência da informação pública no Brasil: uma análise da acessibilidade de Big Data para o estudo das interfaces entre mudanças climáticas, mudanças produtivas e saúde. RECIIS. 2020;14(1):111-25. DOI: https://doi.org/10.29397/reciis.v14i1.1690

5. Silva NP. A utilização dos programas TABWIN e TABNET como ferramentas de apoio à disseminação das informações em saúde [dissertação]. Rio de Janeiro: Escola Nacional de Saúde Pública Sergio Arouca, Fundação Oswaldo Cruz; 2009.

6. Leandro BBS, Rezende FAV, Pinto JMC. Informações e registros em saúde e seus usos no SUS. Rio de Janeiro: Editora FIOCRUZ; 2020.

7. Garcia PT, Reis RS. Gestão pública em saúde: o plano de saúde como ferramenta de gestão. São Luís: EDUFMA; 2016.

8. Ferré F. Infoestrutura para apoio à decisão estratégica no SUS. In: Santos AO, Lopes LT, organizadoras. Reflexões e futuro. Brasília, DF: Conass; 2021. p. 114-27. (Coleção Covid-19; v. 6).

9. Saldanha RF, Bastos RR, Barcellos C. Microdatasus: pacote para download e pré-processamento de microdados do Departamento de Informática do SUS (DATASUS). Cad Saúde Pública. 2019;35(9):e00032419. DOI: https://doi.org/10.1590/0102-311X00032419

10. Santos RS, Santos RS, Gutierrez MA. MINERSUS Ambiente computacional para extração de informações para a gestão da saúde pública por meio da mineração dos dados do SUS. Rev Bras Eng Biomed. 2008;24(2):77-90. DOI: https://doi.org/10.4322/rbeb.2012.050

11. Franceschini PM, Porto JB, Kunst R. Ferramenta de Visualização de Dados Públicos da Saúde Disponibilizados pelo DATASUS. In: International Conference on Information Resources Management; 2021; [local desconhecido]: AIS Electronic Library; 2021.

12. Barbosa MN. Possibilidades e limitações de uso das bases de dados do DATASUS no controle externo de políticas públicas de saúde no Brasil [trabalho de conclusão de curso]. Brasília, DF: Instituto Cerzedello Corrêa, Escola Superior do Tribunal de Contas da União; 2019.

13. Cherchiglia ML, Guerra Júnior AA, Andrade EIG, et al. A construção da base de dados nacional em terapia renal substitutiva (TRS) centrada no indivíduo: aplicação do método de linkage determinístico-probabilístico. Rev Bras Estud Popul. 2007;24(1):163-67. DOI: https://doi.org/10.1590/S0102-30982007000100010

14. Barreto ML, Ichihara MY, Pescarini JM, et al. Cohort Profile: The 100 Million Brazilian Cohort. Int J Epidemiol. 2022;51(2):e27-e38. DOI: https://doi.org/10.1093/ije/dyab213

15. Moura L de, Prestes IV, Duncan BB, et al. Construção de base de dados nacional de pacientes em tratamento dialítico no Sistema Único de Saúde, 2000-2012. Epidemiol Serv Saude. 2014;23(2):227-38. DOI: https://doi.org/10.5123/S1679-49742014000200004

16. Moya J, Risi Junior JB, Martinello A. Salas de situação em saúde: compartilhando as experiências do Brasil. Brasília, DF: Organização Pan-Americana da Saúde; Ministério da Saúde; 2010.

17. Fornazin M, Joia LA. Articulando perspectivas teóricas para analisar a informática em saúde no Brasil. Saude Soc. 2015;24:46-60. DOI: https://doi.org/10.1590/S0104-12902015000100004

18. Giannotti EM, Fonseca F, Panitz LM. Sistemas de Informação da Atenção à Saúde: contextos históricos, avanços e perspectivas no SUS. Brasília, DF: Cidade Gráfica e Editora Ltda; 2015.

19. Atty ATM, Jardim BC, Dias MBK, et al. PAINEL-Oncologia: uma Ferramenta de Gestão. Rev Bras Cancerol. 2020;66(2); DOI: https://doi.org/10.32635/2176-9745.rbc.2020v66n2.827

20. Camargo Jr. KR, Coeli CM. Reclink: aplicativo para o relacionamento de bases de dados, implementando o método probabilistic record linkage. Cad Saúde Pública. 2000;16(2):439-47. DOI: https://doi.org/10.1590/S0102-311X2000000200014

21. Ali MS, Ichihara MY, Lopes LC, et al. Administrative Data Linkage in Brazil: Potentials for Health Technology Assessment. Front Pharmacol. 2019;10:984. DOI: https://doi.org/10.3389/fphar.2019.00984

22. Guerra Junior AA, Pereira RG, Gurgel EI, et al. Building the National Database of Health Centred on the Individual: Administrative and Epidemiological Record Linkage-Brazil, 2000-2015. Int J Popul Data Sci. 2018;3(1):446. DOI: https://doi.org/10.23889/ijpds.v3i1.446

23. Tomazelli JG, Girianelli VR, Silva GA. Estratégias usadas no relacionamento entre Sistemas de Informações em Saúde para seguimento das mulheres com mamografias suspeitas no Sistema Único de Saúde. Rev Bras Epidemiol. 2018;21:e180015. DOI: https://doi.org/10.1590/1980-549720180015

24. Ferré F, Oliveira G, Queiroz M, et al. Sala de Situação aberta com dados administrativos para gestão de Protocolos Clínicos e Diretrizes Terapêuticas de tecnologias providas pelo SUS. In: 20º Simpósio Brasileiro de Computação Aplicada à Saúde SBC; 2020 set 15-18; Porto Alegre. Porto Alegre: Sociedade Brasileira de Computação; 2020. p. 392-403.

25. TabNet [Internet]. Brasília, DF: DATASUS. c2008 [acesso em 2022 out 28]. Informações de saúde. Produção Ambulatorial do SUS por local de atendimento. Disponível em: http://tabnet.datasus.gov.br/cgi/deftohtm.exe?sia/cnv/qauf.def

26. Elmasri R, Navathe SB. Sistemas de banco de dados. 6ª ed. São Paulo: Pearson Addison Wesley; 2011.

27. Khan B, Jan S, Khan W, et al. An overview of ETL techniques, tools, processes and evaluations in data warehousing. J Big Data. 2024;6. DOI: http://dx.doi.org/10.32604/jbd.2023.046223

28. Franco TB. Trabalho, cuidado e transição tecnológica na saúde: um olhar a partir do sistema cartão nacional de saúde. Porto Alegre: Editora Rede Unida; 2021 (Série Micropolítica do Trabalho e o Cuidado em Saúde).

29. Conselho Nacional de Saúde (BR). Resolução nº 510, de 7 de abril de 2016. Dispõe sobre as normas aplicáveis a pesquisas em Ciências Humanas e Sociais. Diário Oficial da União, Brasília, DF. 2016 maio 24; Edição 98; Seção I:44-46.

30. Atty ATM, Tomazelli JG, Dias MBK. Análise Exploratória das Informações sobre Estadiamento nas Autorizações de Procedimentos de Alta Complexidade no Brasil e Regiões no Período 2010-2014. Rev Bras Cancerol. 2019;63(4):257-64. DOI: https://doi.org/10.32635/2176-9745.RBC.2017v63n4.126

31. Medeiros CRG, Saldanha OMFL, Grave MTQ, et al. Planejamento regional integrado: a governança em região de pequenos municípios. Saude Soc. 2017;26(1):129-40. DOI: https://doi.org/10.1590/S0104-12902017162817

32. Lima LD, Viana ALD, Machado CV, et al. Regionalização e acesso à saúde nos estados brasileiros: condi-

cionantes históricos e político-institucionais. Ciênc saúde coletiva. 2012;17(11):2881-92. DOI: https://doi.org/10.1590/S1413-81232012001100005

33. Cunha RE. Cartão Nacional de Saúde: os desafios da concepção e implantação de um sistema nacional de captura de informações de atendimento em saúde. Ciênc saúde coletiva. 2002;7(4):869-78. DOI: https://doi.org/10.1590/S1413-81232002000400018

34. Petruzalek D. read.dbc: Read Data Stored in DBC (Compressed DBF) Files [Internet]. New York: Datacamp; 2016 [acesso em 2022 out 28]. Disponível em: https://cran.r-project.org/web/packages/read.dbc/index.html

35. Park H, Lee S, Gim G, et al. Dataverse: Open-Source ETL (Extract, Transform, Load) Pipeline for Large Language Models. arXiv [csCL]. 2024. DOI: https://doi.org/10.48550/arXiv.2403.19340

36. Ernest A, Mensah E, Gilbert A. Qualitative assessment of compiled, interpreted and hybrid programming languages. Comm App Electronics. 2017;7:8-13. DOI: https://doi.org/10.5120/cae201765268

37. Sanner MF. Python: a programming language for software integration and development. J Mol Graph Model. 1999;17(1):57-61.

38. Prechelt L. An empirical comparison of seven programming languages. Computer 2000;33(10):23-29. DOI: https://doi.org/10.1109/2.876288

39. Souza Júnior EV, Santos GDS, Jesus ALO, et al. Tratamento hemodialítico e seus impactos financeiros no nordeste do Brasil. Rev Enferm UFPE On Line. 2019;13. DOI: http://dx.doi.org/10.5205/1981-8963.2019.239674

40. Aly CMC, Reis AT, Carneiro SAM, et al. O Sistema Único de Saúde em série histórica de indicadores: uma perspectiva nacional para ação. Saúde debate. 2017;41(113):500-12. DOI: https://doi.org/10.1590/0103-1104201711312

41. Brilhante OMA. Caldas LQ. Gestão e avaliação de risco em saúde ambiental. Rio de Janeiro: Editora FIOCRUZ; 1999.

42. Moya J, Risi Junior JB, Martinello A, et al. Sala de Situação em Saúde: compartilhando as experiências do Brasil. Brasília, DF: Organização Pan-Americana da Saúde; Ministério da Saúde; 2010.

43. Rosa L, Mrejen M, Franceschini MC, et al. Instituto de Estudos Para Políticas de Saúde [Internet]. [local desconhecido]: IEPS; 2022 [acesso em 2024 abr 5]. Disponível em: https://iepsdata.org.br/

44. Ministério da Saúde (BR); Conselho Nacional de Saúde. Resolução nº 659, de 26 de julho de 2021. Dispõe sobre a Política Nacional de Informação e Informática em Saúde (PNIIS). Diário Oficial da União, Brasília, DF. 2022 jun 15; Edição 113; Seção I:104.

45. Ministério da Saúde (BR). Estratégia de Saúde Digital para o Brasil 2020-2028. Brasília, DF: Ministério da Saúde; 2020.

46. Ministério da Saúde (BR), Gabinete do Ministro. Portaria nº 1.434, de 28 de maio de 2020. Institui o Programa Conecte SUS e altera a Portaria de Consolidação nº 1/GM/MS, de 28 de setembro de 2017, para instituir a Rede Nacional de Dados em Saúde e dispor sobre a adoção de padrões de interoperabilidade em saúde. Diário Oficial da União, Brasília, DF. 2020 maio 29; Edição 102; Seção I:231.

Received on 11/14/2024
Approved on 06/19/2025
Conflict of interest: Non-existent
Data availability: Research data are contained in the manuscript itself
Financial support: Non-existent

Editor in charge: Alessandro Jatobá