How to Improve Access to Data from Epidemic Surveillance and Electronic Health Records in order to Advance Research in Portugal

Como Melhorar o Acesso a Dados de Vigilância Epidemiológica e Registos de Saúde Eletrónicos para Avançar a Investigação em Portugal

Vasco RICOCA PEIXOTO✉[1,2], Nina de SOUSA SANTOS[3], Tiago L. SOUTO[4], Margarida GIL CONDE[5,6], Bruno HELENO[7], João V. CORDEIRO[1,2,8], Pedro AGUIAR[1,2], Alexandre ABRANTES[1,2]
Acta Med Port 2022 Nov;35(11):786-788 ▪ https://doi.org/10.20344/amp.18381

Keywords: Data Collection; Delivery of Health Care; Epidemiologic Research Design; Health Services Research; Information Storage and Retrieval; Medical Records Systems, Computerized; Portugal; Public Health Systems Research; Routinely Collected Health Data
Palavras-chave: Armazenamento e Recuperação da Informação; Dados de Saúde Recolhidos Rotineiramente; Investigação em Serviços de Saúde; Investigação em Sistemas de Saúde Pública; Portugal; Prestação de Cuidados de Saúde; Projectos de Investigação Epidemiológica; Recolha de Dados; Sistemas Computorizados de Registos Médicos

INTRODUCTION

There is a large number of healthcare data collected in Portugal through various services, which can be more accessible and better used for research by health services and academia in order to inform public policies, improve the quality of services or advance scientific knowledge.[1,2]

To answer research questions relevant to the health and well-being of populations, it is necessary to use secondary health data that exists in different information systems.[3] Its treatment enables more consensual health policies, better care, maximization of benefits and minimization of risks and costs of interventions.[4,5] The use of aggregated, non personal data for surveillance and continuous improvement of health services quality is an important practice but is beyond the scope of this article.

Social utility of data processing

Obstacles to the generation of evidence can entail a high burden of preventable disease, social and economic damage, as well as less appropriate policies and practices and worse quality in services. The SARS-CoV-2 pandemic has reinforced the need for knowledge. Outside the pandemic there is a comprehensive field of health research: real-world studies, cohort and case-control studies, assessment of the effectiveness and safety of several treatments and interventions, and healthcare quality research.

According to the European Centre for Disease Prevention and Control (ECDC), generating knowledge to support decision-making, strategy definition and continuous improvement are the main uses of surveillance systems[6] and the potential for using health records for surveillance purposes is being mapped.[7] This usefulness of surveillance systems is enhanced by the existence of teams dedicated to research projects in research centers and health services, contributing to the governance and advocacy aimed at improving the quality of policies and care. The Terms of Reference of contractualization for the Portuguese National Health System also include scientific research,[8] which is one of the World Health Organization (WHO) Essential Public Health Operations.

Routine health records are considered by the WHO a global public good[9] and are treated as such in several countries. After being anonymized, they are accessible through specific procedures.[5] There is an ethical imperative for maximizing the benefit of data for the good of societies, reducing and mitigating risks. This balance is referred to in the General Data Protection Regulation (GDPR).

In Portugal, the scope and centrality of many health records is recognized: The National Epidemiological Surveillance System (SINAVE), primary health care (SClínico), Medicines Prescribing (PEM), vaccination (VACINAS), hospitals (BDMH, SONHO, emergencies), screening programs (SiiMA), cause of death (SICO), referral from primary care

1. Centro de Investigação em Saúde Pública. Escola Nacional de Saúde Pública. Universidade NOVA de Lisboa. Lisboa. Portugal.
2. Comprehensive Health Research Centre. Universidade NOVA de Lisboa. Lisboa. Portugal.
3. Instituto de Bioética. Universidade Católica Portuguesa. Porto. Portugal.
4. Unidade de Saúde Pública. Agrupamento de Centros de Saúde de Sintra. Administração Regional de Saúde de Lisboa e Vale do Tejo. Sintra. Portugal.
5. Comissão de Ética para a Saúde. Secção de Investigação. Administração Regional de Saúde de Lisboa e Vale do Tejo. Lisboa. Portugal.
6. Unidade de Saúde Familiar Jardins da Encarnação. Agrupamento de Centros de Saúde de Lisboa Central. Administração Regional de Saúde de Lisboa e Vale do Tejo. Lisboa. Portugal.
7. Clínica Universitária de Medicina Geral e Familiar. Faculdade de Medicina. Universidade de Lisboa. Lisboa. Portugal.
8. Centro Interdisciplinar de Ciências Sociais. Universidade NOVA de Lisboa. Lisboa. Portugal.
✉ **Autor correspondente:** Vasco Ricoca Peixoto. vrf.peixoto@ensp.unl.pt
Recebido/*Received*: 08/04/2022 - **Aceite/*Accepted*:** 13/07/2022 - **Publicado *Online*/*Published Online*:** 17/08/2022 - **Publicado/*Published*:** 02/11/2022

(AlertP1), among others.

In the context of the SARS-CoV-2 pandemic, the need to generate evidence enhanced access to data and cross-referencing from different information systems (using unique identifiers) in unique databases, with the patient as the observation unit, to answer research questions. An example of this is the research on the effectiveness of vaccines against COVID-19 in the real world based on the linkage of health records (SINAVE, VACINAS, SONHO, SICO). Data were also shared with university institutions to support the generation of evidence with some relevant outputs. Yet other data appear to be more underutilized in terms of producing scientific knowledge: the specific causes of mortality coded by doctors in death certificates and available in real-time with minor limitations that are probably stable through time and do not impede the analysis of trends of specific causes of mortality (SICO) data, which could help to understand the causes of excess non-COVID mortality; prescription data (PEM) for real-world evidence of drugs (efectiveness, safety, natural history of disease); primary health care data (SClinic) to study risk and protective factors for various health outcomes; vaccination data, to understand vaccination determinants.

After data linkage and anonymization or appropriate pseudonymization, observational epidemiological studies (cohort and case-control) of high importance are possible.

The processes for accessing this data should be easy to consult, simplified and facilitated, towards their widespread use by public health and other competent healthcare services, academia, and other entities promoting research in Portugal.

Lawfulness of the processing of personal and health data in the face of the General Data Protection Regulation (GDPR): privacy protection mechanisms and national legal framework

Data collected without identification of its holder may be used without the obligations of the GDPR. This Regulation is intended only for personal data, defined as relating to an identified or identifiable person.[10]

If the study does not need identification of data subjects, as is the case in most studies in the field of Public Health and in most observational research, irreversible anonymization may be carried out (without the creation of a key, or through its destruction).[10] That process and the possible linkage of databases must be carried out by dedicated professionals in the institutions which provide the data before sharing with the researchers.

After irreversible anonymization, personal data can no longer be considered as relating to an identified or identifiable person and is excluded from the scope of the GDPR, and may be used, for example, without the need for informed consent.

The detailed characterization of personal data through multiple variables should be avoided, alternatively, using the categorization of continuous variables such as age, or the non-inclusion of geographic variables from small areas or rare specific conditions.

The processing of data must comply with the principle of minimization. For this, the metadata of the various databases must be public so that researchers and research centers can request only the essential data, and being aware of their potential and limitations.

Implementation proposal in Portugal

In view of the foregoing, procedures for health services and academia to access health data for scientific research purposes should be easily available and include the following steps:

1. Submission of access request to the entity responsible for the data, selecting the variables of interest (from the metadata).
2. The application should be instructed with the following documentation:
   a) Summary of the research protocol.
   b) Approval of the ethics committee.
   c) Approval of the data protection officer, if the GDPR applies (if they are not irreversibly anonymized).
   d) Authorization for the conduct of the research project by the employer/institution of the researcher.
   e) Filling out the application form.
3. The request shall be answered within a period of not more than 10 working days in accordance with the Portuguese Code of Administrative Procedures.
4. In the event of rejection, the reasons shall be mentioned in writing and the manner in which they can be overcome shall be mentioned where possible.
5. The approval or conditional approval shall determine the immediate submission of the database to the applicant.

In the institutions that hold the data, it is necessary to ensure the existence of dedicated professionals, responsible for the following activities:

1. Maintenance and updating of a public register of available metadata.
2. Anonymization of data and the performance of identification tests.
3. Irreversible anonymization or pseudonymization of data (with or without key custody).
4. Application of privacy protection techniques including generalization, concealment and permutation.
5. Public availability of forms required for access requests.
6. Integration/cross-reference of data from multiple sources. Protocols for secure linkage of data between institutions and assigned responsibilities must be ensured.

CONCLUSION

The public entities responsible for the various types of data of health information systems in Portugal should have clear and simple procedures for requesting data and

guarantee the existence of responsible personnel for the processing and sharing of secondary data. Available health data should be publicly listed with corresponding responsible institutions and link to metadata and data access request. Processes must be agile and guided by easily available public documents, produced by the responsible entities, in accordance with the law. Metadata must be listed in those documents.

Harmonized access to health data for research by services and academia and the linkage of data from different systems before sharing with researchers are urgent necessities for the advancement of health knowledge. Respect for the ethical principles of research and national and international data protection standards should be ensured, while promoting the social usefulness of registries.[6] National law and the GDPR protect the use of this data for research, thus ensuring the applicable assumptions. The GDPR does not apply to irreversibly anonymized data even though its general principles of risk reduction should always be considered.

ACKNOWLEDGEMENTS

Carlos Carvalho, Alexis Sentís, Raquel Vareda, Vera Pereira Machado.

AUTHORS CONTRIBUTION

VRP: First draft.

NSS, TLS, MGC, BH, JVC, PA, AA: Conception and critical review of the work, literature review.

PROTECTION OF HUMANS AND ANIMALS

The authors declare that the procedures were followed according to the regulations established by the Clinical Research and Ethics Committee and to the Helsinki Declaration of the World Medical Association updated in 2013.

DATA CONFIDENTIALITY

The authors declare having followed the protocols in use at their working center regarding patients' data publication.

COMPETING INTERESTS

The authors have declared that no competing interests exist.

FUNDING SOURCES

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

1. Guimarães R, Dinis-Oliveira RJ, Pereira A, Rodrigues P, Santos A. Reutilization of clinical data for research: the footprint scientific model of the hospital center of São João. Acta Med Port. 2017;30:159-62.
2. Meneses-Oliveira C. Use of electronic medical records for research: new ethical challenges and possible solutions. Acta Med Port. 2019;32:332-4.
3. Hoffmann W, Bobrowski C, Fendrich K. Sekundärdatenanalyse in der Versorgungsepidemiologie: Potenzial und Limitationen. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2008;51:1193-201.
4. Organisation for Economic Co-operation and Development. Enhancing access to research data during crises: lessons learned from the COVID-19 pandemic. OECD;2021. [cited 2021 Jun 18]. Available from: https://www.oecd.org/sti/inno/enhance-access-research-data-during-crises.htm.
5. Obel C, Olsen J, Jensen UJ. Use of existing health data in epidemiologic research–issues of informed consent under normal circumstances and at a time of health crises. Clin Oncol Res. 2020;2020:1-4.
6. European Centre for Disease Prevention and Control. Data quality monitoring and surveillance system evaluation – A handbook of methods and applications. Stockholm: ECDC; 2014. [cited 2021 Jun 18]. Available from: https://www.ecdc.europa.eu/en/publications-data/data-quality-monitoring-and-surveillance-system-evaluation-handbook-methods-and.
7. European Centre for Disease Prevention and Control. Assessment of electronic health records for infectious disease surveillance: final mapping exercise report. Stockholm: ECDC; 2021. [cited 2021 Jul 4]. Available from: https://www.ecdc.europa.eu/en/publications-data/assessment-electronic-health-records-infectious-disease-surveillance.
8. Administração Central do Sistema de Saúde. Contratualização Categoria. [cited 2022 Mar 29]. Available from: https://www.acss.min-saude.pt/category/prestadores/normas-e-metodos/?lang=pt.
9. World Health Organization. Health Data as a global public good – a call for Health Data Governance 30 September. [cited 2022 Apr 6]. Available from: https://www.who.int/news-room/articles-detail/health-data-as-a-global-public-good-a-call-for-health-data-governance-30-september.
10. European Parliament. Regulamento (UE) 2016/679 do Parlamento Europeu e do Conselho, de 27 de abril de 2016, relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados e que revoga a Diretiva 95/46/CE (Regulamento Geral sobre a Proteção de Dados) (Texto relevante para efeitos do EEE). 2016. [cited 2019 Oct 16]. Available from: https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=CELEX:32016R0679.