National Network for Rare Diseases in Brazil: The Computational Infrastructure and Preliminary Results

Diego Bettiol Yamada[1], Filipe Andrade Bernardi[2], Márcio Eloi Colombo Filho[2], Mariane Barros Neiva[3], Vinícius Costa Lima[2], André Luiz Teixeira Vinci[1], Bibiana Mello de Oliveira[4,5], Têmis Maria Félix[5], Domingos Alves[1]

[1] Ribeirao Preto Medical School, University of Sao Paulo, Ribeirao Preto, Brazil
[2] Bioengineering Postgraduate Program, University of Sao Paulo, Sao Carlos, Brazil
[3] Institute of Mathematics and Computer Sciences, University of São Paulo, São Carlos, Brazil
[4] Postgraduate Program in Genetics and Molecular Biology, Federal University of Rio Grande do Sul, Brazil
[5] Medical Genetics Service, Porto Alegre Clinical Hospital, Brazil

Abstract. According to the World Health Organization, rare diseases currently represent a global public health priority. Although it has a low prevalence in the general population, this type of condition collectively affects up to 10% of the entire world population. Therefore, these pathologies are numerous and of a diverse nature, and some factors imply significant challenges for public health, such as the lack of structured and standardized knowledge about rare diseases in health units, the need for communication between multidisciplinary teams to understand phenomena and definition of accurate diagnoses, and the scarcity of experience on specific treatments. In addition, the often chronic and degenerative nature of these diseases generates a significant social and economic impact. This paper aims to present an initiative to develop a network of specialized reference centers for rare diseases in Brazil, covering all country regions. We propose collecting, mapping, analyzing data, and supporting effective communication between such centers to share clinical knowledge, evolution, and patient needs, through well-defined and standardized processes. We used validated structures to ensure data privacy and protection from participating health facilities to create this digital system. We also applied systems lifecycle methodologies, data modeling techniques, and quality management. Currently, the retrospective stage of the project is in its final phase, and some preliminary results can be verified. We developed an intuitive web portal for consulting the information collected, offering filters for personalized queries on rare diseases in Brazil to support evidence-based public decision-making.

Keywords: Rare Disease, Health Network, Digital Health, Public Health Observatory.

1 Introduction

Rare Diseases (RDs) are conditions that, although individually presenting a low prevalence, together affect an expressive part of the world population. It is estimated that the set of all known RDs affects approximately 10% of all individuals on the planet [1]. Such pathologies are numerous and diverse, demanding the relationship of different fields of knowledge. These factors imply significant challenges for public health, such as the lack of structured learning in health units about each of these medical conditions, the difficulty and multidisciplinary nature involved in the

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08757-8_4

accurate establishment of RD diagnoses, and the chronic and degenerative nature of these diseases, like malformation syndromes, morphological, and biological anomalies, which causes a significant impact and social burden [2].

According to the World Health Organization (WHO), to overcome the barriers, research and development of studies about RDs must involve the formation of multidisciplinary networks of collaboration between health professionals from different areas, reference centers, public health managers, and patient groups and associations [3;4]. So, initiatives are developed to provide informational support to these networks, and the best known is Orphanet [5].

In Brazil, the Unified Health System ('SUS'; Portuguese: 'Sistema Único de Saúde') is responsible for offering comprehensive public care to all citizens and at all levels of complexity. Nevertheless, only in 2014 the Brazilian Policy for Comprehensive Care for People with Rare Diseases was established and instituted within the scope of the SUS [6].

However, governmental and socio-political financial support is not sufficient to guarantee the population's rights to health in practice. Problems with human, technological, and infrastructure resources are observed in health units in all Brazilian regions [7]. Therefore, the Brazilian Network of Rare Diseases ('RARAS'; Portuguese: 'Rede Nacional de Doenças Raras') project presents itself to reverse this scenario. This project aims to carry out a national representative survey about the epidemiology, clinical overview, diagnostic and therapeutic resources used, and costs for RD of genetic and non-genetic origin in Brazil and create a national RD surveillance network [1].

This paper aims to present the plan's preliminary results that involve all the computational and procedural infrastructure used to create the RARAS, following the WHO guidelines for developing digital health observatories [4].

2 Literature Review

RDs are a highly complex problem for health organizations all over the planet. The intrinsic multidisciplinarity of this area requires that health services articulate in an interoperable, transparent, and agile manner. In addition, communication between the various health units involved in the care of these patients must be precise and unambiguous to generate savings in resources and time, and consequently, greater patient satisfaction. Many collaborative RD health network initiatives exist, and others are emerging. Such networks are of great importance and value for the development of science and technology in this area [8].

The International Rare Diseases Research Consortium (IRDiRC) is a collaborative network that seeks new developments to generate knowledge and improvements in therapeutic, diagnostic methods, and consequently, quality of life for patients with RD [9]. The European Reference Networks (ERNs) were established to expand their scope of attention to rare and low-prevalence conditions in Europe [10]. At the same time, other initiatives already existed and were functioning at the national level.

Italy was one of the first European countries to create specific regulations for RDs. Its successful experience evidenced a reduction in the costs of health services, lower mobility of patients among the different health units, and better planning of public

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08757-8_4

health policies [8]. In France, establishing the French National Plan for Rare Diseases contributed to structuring processes, integrating care models and epidemiological models for research [11].

Consequently, such projects promoted greater engagement in models of care for RDs. Similar contributions emerged in countries such as the United Kingdom, Germany, and the United States [12]. The Cooperation Rare Disease Action Plan was created based on the premises of the Asia-Pacific Economic Cooperation to improve the social and economic inclusion of people with RDs [13].

In Brazil, the Ministry of Health established the Brazilian Policy for Comprehensive Care for People with Rare Diseases. However, the non-standardization of processes, the heterogeneity of the health information systems, and the lack of terminologies make complex the task of guaranteeing the premises of the Brazilian plan [6;7]. Therefore, we develop and offer the necessary computational and procedural infrastructure for what specialized centers in RDs from all regions of Brazil can collect, analyze, generate and share knowledge about RDs.

3 Methods

3.1 Study Design

This paper is part of a larger project, RARAS, which is an observational cohort study. The development of computing infrastructure is characterized as applied basic research because it uses validated scientific knowledge to develop methods and technologies to improve the understanding of events and phenomena, increasing our scientific knowledge base [14]. The RARAS project covers all Brazilian demographic regions, with 39 participating health centers. The inclusion criteria for the retrospective phase refers to patients seen at RDs services between 2018 and 2019.

For the prospective phase, the inclusion criterion refers to patients seen in these services in 2022, between April and September. Participating health centers include health units in 16 state capitals and other municipalities, totaling more than 47 million people living in these areas [15]. To centralize information from all participants, we used standardized electronic capture of health data techniques.

3.2 Electronic Data Captures Systems and Data Collection

The use of Electronic Data Capture (EDC) systems for data collection eliminates risks associated with paper-based instruments and enhances the collection of high-quality data necessary for conducting health research [16]. So, the EDCs systems REDCap [17] and KoBoToolbox [18] were selected in this study because they are free, stable, widely used, and well-documented software, with Application Programming Interfaces (API) available for integration with other systems.

We designed the electronic Case Report Forms (eCRFs) in REDCap to collect data from the health centers, and we structured processes to ensure the monitoring of data quality indicators through a methodology of auditory Early Hearing Detection and Intervention (EHDI) [19]. The classical database architecture Structured Query

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08757-8_4

Language (SQL) was chosen to store the data, using the MYSQL database [20] once the REDCap supported it.

We define internationally validated standards for the interoperable structuring of the data collection stages. To identify a confirmed RD diagnosis, we use 3 selectable options: International Statistical Classification of Diseases and Related Health Problems-10 (ICD-10), Online Mendelian Inheritance in Man (OMIM), and ORPHA code. For the proper mapping of the signs and symptoms caused by pathologies, we use the Human Phenotype Ontology (HPO) [21].

Automatic and manual validations are performed to verify the consistency of the data entered in the data collection instruments. The automatic validation is done by a script that searches for outliers in critical fields on filled forms. For manual validation, we create a monitor process using the Business Process Model and Notation (BPMN), to standardize each step of the data collection process [22]. Six data managers are responsible for doing the manual validation.

3.3 Web Portal Development and Data Analysis

The processed information and analyzed data are made available in a web portal that is in development using open source technologies and languages, being PHP7 for the back-end side and HTML5, CSS3, and JavaScript for the front-end side, following the W3C guidelines [23]. The data is retrieved from an internal server, and Python scripts were developed to summarize the characteristics of the collected information.

We emphasize evaluating the distribution of human resources and laboratory procedures for each state and health center using statistical methodologies. Patients' data distributions such as age, initial symptoms, and diagnosis are also available. The preliminary web portal is available at the domain: https://raras.org.br/. However, as the study is still in progress, all tools have not yet been made available because we are following the project schedule.

4 Results and Discussion

The health center characterization form showed us that just over half of the participating centers (59%) have Electronic Health Records (EHRs) established in their work process. The rest of the centers use physical health records on paper. Less than half of the centers indicated that the data generated internally are shared with other organizations or systems (46.2%), showing the inadequacy of the SUS in guaranteeing adequate communication through the health network.

The retrospective data collection has 10,442 records entered from patients with RDs from 39 referral centers spread across all regions of Brazil, in a data collection time of 433 days. To verify the consistency of the data entered by the typists, we design a BPMN process. Figure 1 shows the clinical collection monitoring diagram.

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08757-8_4

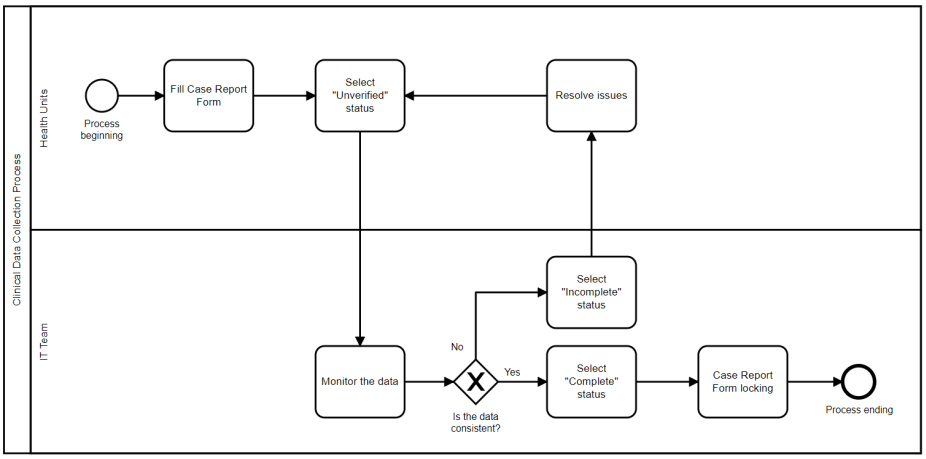

Fig. 1. Clinical data collection monitoring diagram

From the collected data, the distribution of the diagnostic status of all patients entered so far can be graphically visualized, as shown in Figure 2. We can see that, although more than half of the patients entered have a confirmed diagnosis of RDs, 35.5% of the patients do not have an actual confirmed diagnosis.

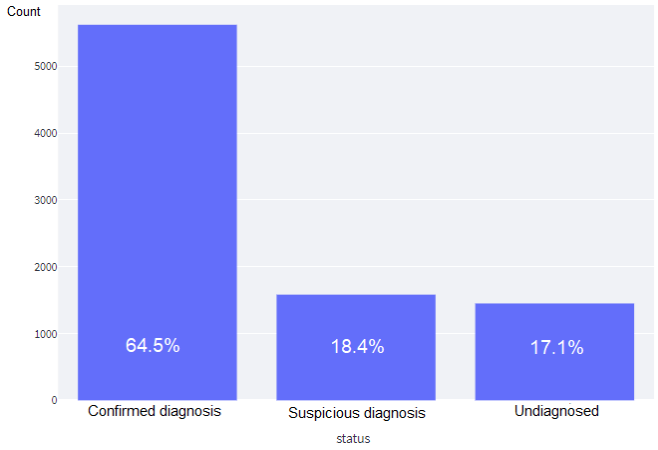

Fig. 2. Diagnostic status distribution

We also identify the moment of diagnosis of these patients, offering information of great use to direct public investments. This moment can be prenatal (1.01%), neonatal screening (8.79), postnatal (70.9), and 19.3% of the records did not have this information. Regarding the terminology used in these health units, we found that the ORPHA code was the most used (62.1%), followed by the ICD-10 (29.9%), and later, the OMIM (8%).

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08757-8_4

In the map presented in Figure 3, it is possible to select a RD through one of its variables (disease name, ICD-10, OMIM, ORPHA code) and verify how the selected disease is distributed in Brazil. In this example, the RD selected is Phenylketonuria (ORPHA code - 716.0).

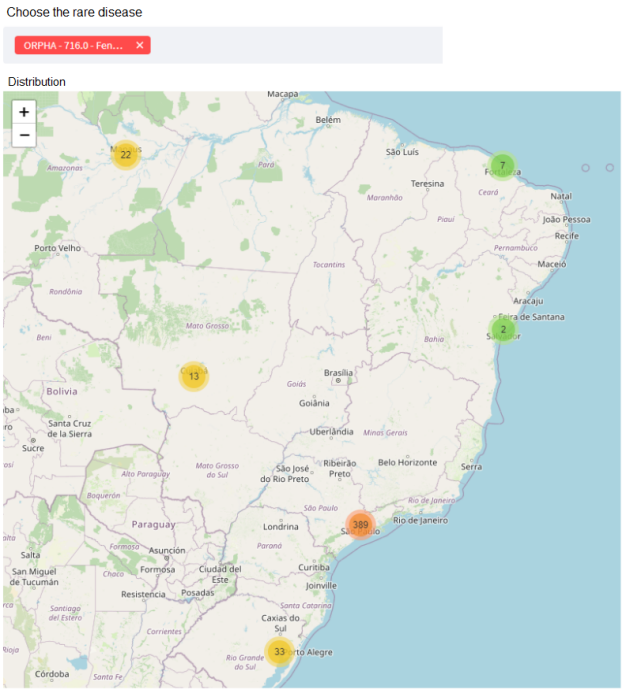

Fig. 3. Filter for diagnoses check geographically

5 Conclusion

In conclusion, the computational infrastructure for implementing this health network promotes the ability to map the actual scenario of RD in all regions of Brazil. From these data, it is possible to carry out analyses to generate intelligence in health to support public planning and assist clinical and managerial decisions based on evidence. As future work, we intend to complete the prospective data collection stage and make all information available in a dynamic and user-friendly way on our web portal. Thus, we believe that improving health management, quality of care, and allocation of technological and human resources is possible.

References

1. Félix, T. M., de Oliveira, B. M., Artifon, M., Carvalho, I., Bernardi, F. A., Schwartz, I. V., ... & Alves, D. (2022). Epidemiology of rare diseases in Brazil: protocol of the Brazilian Rare Diseases Network (RARAS-BRDN). Orphanet Journal of Rare Diseases, 17(1), 1-13.
2. Nguengang Wakap, S., Lambert, D. M., Olry, A., Rodwell, C., Gueydan, C., Lanneau, V., ... &

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08757-8_4

Rath, A. (2020). Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. *European Journal of Human Genetics, 28*(2), 165-173.
3. World Health Organization, Priority diseases and reasons for inclusion, https://www.who.int/medicines/areas/priority_medicines/Ch6_0Intro.pdf, last accessed 2022/02/16.
4. World Health Organization - Guide for the establishment of health observatories, https://apps.who.int/iris/handle/10665/246123, last accessed 2022/04/05.
5. Orphanet inventory of rare diseases, https://www.orpha.net/orphacom/cahiers/docs/GB/eproc_disease_inventory_PR_R1_Nom_04.pdf, last accessed 2022/02/16.
6. Portaria nº 199, de 30 de Janeiro de 2014, Ministério da Saúde, https://bvsms.saude.gov.br/bvs/saudelegis/gm/2014/prt0199_30_01_2014.html, last accessed 2022/02/16.
7. Iriart, J. A. B., Nucci, M. F., Muniz, T. P., Viana, G. B., Aureliano, W. D. A., & Gibbon, S. (2019). Da busca pelo diagnóstico às incertezas do tratamento: desafios do cuidado para as doenças genéticas raras no Brasil. *Ciência & Saúde Coletiva*, 24, 3637-3650.
8. Simone Baldovino, M. D., Domenica Taruscio, M. D., & Dario Roccatello, M. D. (2016). Rare diseases in Europe: from a wide to a local perspective. Isr Med Assoc J, 18, 359-63.
9. IRDiRC - Vision and Goals. https://irdirc.org/about-us/vision-goals/, last accessed 2022/02/16
10. Héon-Klin, V. (2017). European Reference networks for rare diseases: what is the conceptual framework?. *Orphanet journal of rare diseases, 12*(1), 1-9.
11. Choquet, R., & Landais, P. (2014). The French national registry for rare diseases: an integrated model from care to epidemiology and research. *Orphanet journal of rare diseases, 9*(1), 1-2.
12. Alves, D., Yamada, D. B., Bernardi, F. A., Carvalho, I., Colombo Filho, M. E., Neiva, M. B., ... & Félix, T. M. (2021). Mapping, Infrastructure, and Data Analysis for the Brazilian Network of Rare Diseases: Protocol for the RARASnet Observational Cohort Study. *JMIR Research Protocols, 10*(1), e24826.
13. APEC Rare Disease Network. Asia-Pacific Economic Cooperation, https://www.apec.org/rarediseases, last accessed 2021/01/08.
14. Roll-Hansen, N. (2009). *Why the distinction between basic (theoretical) and applied (practical) research is important in the politics of science*. London: London School of Economics and Political Science, Contingency and Dissent in Science Project.
15. IBGE - Cidades, https://cidades.ibge.gov.br/, last accessed 2022/02/16.
16. Pasalic, D., Reddy, J. P., Edwards, T., Pan, H. Y., & Smith, B. D. (2018). Implementing an electronic data capture system to improve clinical workflow in a large academic radiation oncology practice. *JCO Clinical Cancer Informatics, 2*, 1-12.
17. Wright, A. (2016). REDCap: A tool for the electronic capture of research data. *Journal of Electronic Resources in Medical Libraries, 13*(4), 197-201.
18. KoBoToolbox Website, https://www.kobotoolbox.org/, last accessed 2022/02/16.
19. Chen, H., Hailey, D., Wang, N., & Yu, P. (2014). A review of data quality assessment methods for public health information systems. *International journal of environmental research and public health, 11*(5), 5170-5207.
20. MySQL - Documentation, https://dev.mysql.com/doc/, last accessed 2022/02/16.
21. Köhler, S., Doelken, S. C., Mungall, C. J., Bauer, S., Firth, H. V., Bailleul-Forestier, I., ... & Robinson, P. N. (2014). The Human Phenotype Ontology project: linking molecular biology and disease through phenotype data. *Nucleic acids research, 42*(D1), D966-D974.
22. Zarour, K., Benmerzoug, D., Guermouche, N., & Drira, K. (2019). A systematic literature review on BPMN extensions. *Business Process Management Journal*.
23. W3C Standards, https://www.w3.org/standards/, last accessed 2022/04/04

Funding

This study was funded by the National Council for Scientific and Technological Development – CNPq and the Ministry of Health of Brazil – MoH.

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08757-8_4