v. 4, n. 31, 2024

International Journal of Health Science

APPLICATION OF BUSINESS INTELLIGENCE FOR THE AUTOMATION IN THE EPIDEMIOLOGICAL SURVEILLANCE: XBI A SUCCESSFUL CASE STUDY FROM MUNICIPAL HEALTH DEPARTMENT OF RIBEIRÃO PRETO, BRAZIL (2020-2024)

Samuel Sullivan Carmo
Municipal Health Department, Ribeirão Preto, Brazil

Daniel Cardoso de Almeida e Araújo
Municipal Health Department, Ribeirão Preto, Brazil

Luzia Marcia Romanholi Passos
Municipal Health Department, Ribeirão Preto, Brazil

Abstract: The emergence of the Covid-19 pandemic prompted the Epidemiological Surveillance (ES) unit within the Municipal Health Department (SMS) of Ribeirão Preto to address the necessity for enhanced data management strategies. In August 2020, the XBI methodology was created to analyze the epidemiological landscape of Covid-19 with more speediness, subsequently extended to encompass all other ES priorities. XBI aims to engender SMS in automating the generation of paginated reports detailing epidemiological indicators and developing interactive dashboards improve the ES singularity and internal analyzes and broadcasting of epidemiological trends. Implementation of this methodology necessitates access to the Minimum Data Set (MDS) with the notifications specific to Ribeirão Preto within national surveillance systems (SINAN, SIVEP, SIM, SINASC, ESUS, etc.). Moreover, XBI effectively functions as an instrument capable of automating the generation of reports that incorporate epidemiological indicators, quality metrics, and the identification of inconsistencies within data sets. This involves establishing connections with the MDS, processing data with relational and non-relation databases, and pivot-table analyses. For a suitable reproducibility, can be used MS Excel, Power BI and ArcGIS or other software with support scripting languages (e.g., VBA, R, Python, or M) and facilitate the creation of dynamic dashboards, preferably alongside secure publishing services. The implementation of the XBI methodology has enabled the automated generation of epidemiological reports covering all ES priorities. Since 2021, these dashboards have been integrated with Geographic Information Systems (GIS). Furthermore, in 2024, the establishment of a Public Epidemiologic Observatory was accomplished, featuring interactive dashboards for monitoring various health indicators including live births, mortality rates, incidences of infectious diseases, and sentinel respiratory virus surveillance. In conclusion, our findings suggest the feasibility and reproducibility of this model, particularly for Brazilian municipalities seeking to transition from surveillance models reliant on pre-existing DATASUS tabulators to the XBI framework.
Keywords: Epidemiological Surveillance (ES); Business Intelligence (BI); Geographic Information System (GIS)

INTRODUCTION AND JUSTIFICATION

Ribeirão Preto is a city of 698,642 inhabitants located in the interior of the state of São Paulo, Brazil. With the advent of Covid-19, the methodological definitions for producing epidemiological reports had to be forcibly updated in the Epidemiological Surveillance Division (DVE) of the Municipal Health Department (SMS).[1] In March 2020, our COVID-19 bulletins had to be published daily.[2] After unsuccessful negotiations to contract a specific information technology (IT) service, SMS, strongly guided by the Health Surveillance (DEVISA), made the important decision to invest in the XBI project. It is a data science methodology through which the authors themselves, using available IT tools [3], such as MS Excel and Power BI, were able to apply the concepts of business intelligence (BI) to automate data generation processes. reports of epidemiological indicators through a notification systems database. The general objective is to automate the data science process in epidemiological surveillance. [4, 5] Specific objectives are: publish automated reports without using multiple applications; create connection with notification system databases without using 2x2 tabulators such as TABNET and TABWIN by DATASUS [6, 7]; Create an interface for technical decisions; Generate

inconsistency reports for specific recipients; Generate analytical reports with interest segmentations; Automate analyzes of specific groups, such as healthcare professionals and outbreaks in epidemiologically relevant institutions; Automate routine checks of private laboratories; Create keys for relationships between different databases.

In recent years, the integration of advanced data analysis tools has revolutionized epidemiological surveillance practices, increasing the effectiveness of disease surveillance [8]. MS Excel, with its versatile data manipulation capabilities, has long been used in epidemiological research for data management, analysis and visualization [9]. Similarly, the advent of Power BI has provided unprecedented opportunities for visualizing real-time data and creating interactive dashboards, facilitating informed decision-making in public health settings [10, 11]. Furthermore, the application of data science methodologies is an important part of today's epidemiology, enabling the extraction of insights from complex data sets and driving evidence-based policymaking [12, 13].

The XBI project, launched in August 2020 with Excel licenses available at SMS, began with the automation of the Covid19 Bulletin published on Sundays and holidays. Rapidly evolving into a robust operational model, XBI methodologies were gradually integrated into Epidemiological Surveillance practices.

In 2021, weekly indicator reports and vaccination coverage updates were easily transferred to the XBI framework. With the adoption of tools such as Power BI and ArcGIS [14, 15], the implementation of the XBI methodology, with these tools, it was possible to automate more complex processes, generating, in addition to paginated reports, dynamic dashboards.

Thus, in 2024 the municipality shows that it is possible to improve its approach to disease monitoring and public health management as an alternative to the traditional model based on 2x2 data tabulators (TABNET and TABWIN by DATASUS), hosting an Observatory with Interactive Epidemiological Dashboards in SMS Portal of Ribeirão Preto.

OBJECTIVES

The general goal is to provide the municipality with the capability to analyze its own epidemiological scenario without relying entirely on external analyses. Specific objectives include producing automated paginated reports of epidemiological indicators relevant to the conditions of interest to the municipality, generating automated reports for continuous verification of inconsistencies, and publishing well-structured epidemiological dashboards from its minimum data set.

METHODS

The techniques used involved, firstly, exporting data on the relevant condition of interest to the municipality into a file. This can be done through official epidemiological surveillance information systems, for example, a DBF file in the DBF Base directory, on Sinan-Net, or a ZIP/DBF from Sinan-Online, SIVEP, or even a CSV from CEVESP, TBWEB, ESUS. Next, it is necessary to establish a connection with this data in a spreadsheet application equipped with a pivot table and process the variables of interest, as well as define the meanings of their stratified values. Lastly, the data scientist must write automation code in a language compatible with the application used to reproduce a routine that updates the data and concludes by publishing a paginated PDF report or an interactive dashboard of epidemiological indicators.

The methods used can be summarized as follows: appropriating the data of interest by identifying and constructing non-native

variables that may be necessary, such as the epidemiological week of symptom onset, specific age groups, and the like; using the appropriate data dictionary to build correct indicators by understanding the variables and values involved; and creating the routine to be automated through the appropriate sequence of events, avoiding inaccuracies in the final result. Finally, work processes should be adapted for new inconsistency reports, and it is likely that the correction of input data may also require automation, as showed in Figure 1.

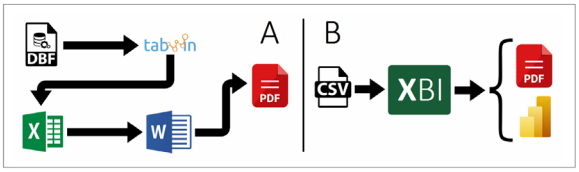

Figure 1: Situation Diagram; A) before and B) after - the figure illustrates a flowchart for each situation; A) by the previous methodology and B) by the XBI methodology.

Among the main achievements, we highlight firstly that the XBI experience of Ribeirão Preto has achieved its overall goal, so that today the municipality can carry out timely monitoring of its epidemiological scenario without needing this analysis to be done by the state and/or federal sphere. Ribeirão Preto now has an Interactive and updated Dashboard of Epidemiological Indicators, which applies to Covid-19, influenza-like illness (ILI), severe acute respiratory syndrome (SARS), Covid-19 vaccination coverage, mpox, measles, influenza, dengue fever, chikungunya, Zika, yellow fever, and more recently, even to violence notifications.

Through the application of the XBI concept, it was also possible to achieve the desired result of transforming the municipality's death TABNET into the Ribeirão Preto Mortality Dashboard. The first result was the internal XBI report of Covid-19 used daily, during the most severe period of the epidemic, by municipal authorities to assist decision-making regarding measures to restrict the movement of people.

Then, the most published result came from the automation of the Covid-19 bulletin with control of deaths under investigation, starting with the automated bulletin for Saturday, Sunday, and holiday shifts. Followed by the weekly, a summary report of Covid-19 indicators, published on the health page every week. In addition to Covid-19, there were results resulting from the automation of the internal report on arboviruses, with dengue, zika, chikungunya, and yellow fever indicators, and from it came the results for the automation of continuous verification reports of inconsistencies for correction and feedback of evidenced non-conformities, such as suspected birth dates (year-to-date), old symptom onset dates (year of birth), duplications (notifications related to the same event) among others. There was also, through the application of XBI, results from the automation of reports on monitoring contacts of index cases (Covid-19 cases with severe acute respiratory infections), outbreaks in long-term care facilities for older adults (LTCF), and schools.

RESULTS

Until July 2020, in order to produce a document with any epidemiological indicator, the data obtained in database file (DBF) by the originating system had to be tabulated in TABWIN, processed in Excel, and formatted in Word before being published in portable document format (PDF).

Due to the volume of notifications, much higher than that of dengue - previously the record holder in notifications - DATASUS did not insert COVID-19 notifications into the Notifiable Diseases Information System (SINAN), but rather developed e-SUS VE, and

in view of the constant instability of this system, Ribeirão Preto chose to use the Municipal System (J166) to manage the Influenza-like illness (ILI), which until then was only treated with pivot tables in Excel, and similarly, the data from severe acute respiratory infections (SARI) from the Influenza Epidemiological Surveillance Information System (SIVEP). In August 2020, the Emergency Bulletin was created, the first product that automated ILI and SARI data.

The development was based on the XBI model through requirements gathering, case studies of the intended scenario, connection to the data source, and consultation of the data dictionary - a documentation that describes the terminology of variables used in the database and identifies them in the data entry form, including data validation rules and value categories - of the listed notification forms.

Through the data dictionary, the available variables and the codes assigned to discrete values were identified, and with the final model, it was possible to know the variables of interest and the necessary filters for each indicator and thus create the functions that generate unavailable variables - e.g., the epidemiological week of symptom onset, age group, reinfection, and others - from existing variables in the source.

In a second stage, in 2021, the XBI concept was applied with more robust computational tools, and it was possible to create keys for the relationship of various databases, thus internal reports of biostatistics between molecular and antigenic tests were automated, genomic surveillance reports, correlations between vaccine coverage, hospitalizations, and deaths, and identification of deaths among the vaccinated. In the third moment, the XBI concept provided for the development of the Epidemiological Dashboard of Ribeirão Preto, an interactive Dashboard for Covid-19, influenza, measles, dengue fever, Zika, chikungunya, and yellow fever that replaced the publication of a monthly paginated bulletin.

Since 2022, we have also achieved the result of publishing a public and a restricted Dashboard for the monitoring of the monkeypox scenario for each new confirmed case.

Since 2023, we achieved the inclusion of geographic information systems (GIS) in restricted epidemiological panels for continuous monitoring of space and time for Covid-19 outbreaks in boarding institutions.

We have in a restricted area dashboards of cases with geographically referenced addresses plotted on a map with layers segmented by attributes of interest. Today, the results achieved by XBI in Ribeirão Preto range from the automation of repetitive work processes, such as generating a PDF certificate for each name and sending an individualized message with each certificate to its respective recipient from a attendance list, to the publication of an Epidemiological Dashboard with data segmentation interaction through a choropleth map, passing through periodic reports of notified diseases and of laboratory results of dengue NS1 antigen ELISA tests by region of residents.

It is worth mentioning that we currently also achieved the creation of the Ribeirão Preto Mortality Dashboard through data from the mortality system (SIM) and the Brazilian Institute of Geography and Statistics (IBGE) and we applied the automated generation of paginated reports of epidemiological indicators for various diseases of compulsory notification or municipal interest, namely viral hepatitis, acquired syphilis, in pregnant women and congenital, influenza, accidents with venomous animals, American cutaneous leishmaniasis, yellow fever, zika, chikungunya, dengue fever, measles, whooping cough,

meningitis, chickenpox, scarlet fever, mumps, leptospirosis, spotted fever, leprosy, mpox monitoring of outbreaks in schools, violence notifications and children exposed to HIV.

Finally, we continue with the automation of indicator reports for tuberculosis, leprosy, meningitis, and other diseases of interest to health surveillance. Summarily the results targeted the goals by methods as shown in Figure 2.

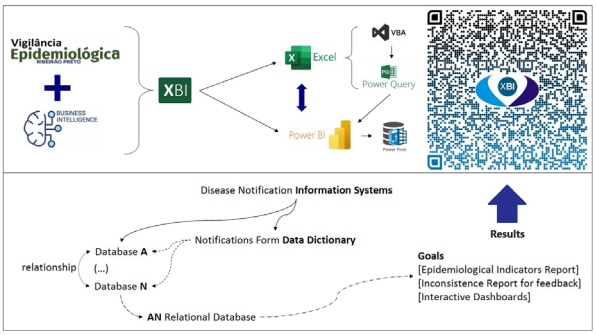

Figure 2: By accessing the disease notification information system, we obtain the clone databases of interest. Using a data dictionary, we identified each variable in the database according to the disease notification form. Using a key, we create the relationship between the databases of interest and thus reach the goals. The results can be views by QR-code access.

Therefore, the general goal of XBI is to ensure that the municipality can carry out its own analyzes for its epidemiological scenario, it is clear that Ribeirão Preto since 2020 has experienced the independence from ready-made scripts on tabulators defined by external agencies. And likewise, the specific objectives of producing automated paginated reports of epidemiological indicators plausible to the municipality's interests; automating continuous verification reports of inconsistencies; and publishing interactive panels on its minimum set of data were successfully achieved. Finally, in 2024, the Ribeirão Preto SMS by XBI methods, provides a Public Epidemiologic Observatory with interactive epidemiological dashboards and vital statistics panels, the latter with records of mortality and too live births.

CONCLUSIONS

We conclude, therefore, that currently, except for the major cities of the country, which have robust support from specialized data science teams, as well as state secretariats supported by entities such as PRODESP and the Ministry of Health with notable support from DATA-SUS, municipalities do not have epidemiologists, data scientists, or entities providing biostatistical, data science, and epidemiological information support within their technical staff to conduct epidemiological surveillance within their jurisdiction.

Generally, municipal epidemiological surveillance relies on public health physicians and public health nurses who act as health vigilantes, diligently ensuring the quality of data entry and thorough case investigation for correct notification and feeding of surveillance information systems. However, since there is no data scientist on the team, the Ministry of Health, through DATASUS, has developed tabulators such as TABNET (more limited, for general use) and TABWIN (more robust, for health professionals) and provides manuals with ready-made routines for producing epidemiological indicators of interest to the Union and states about municipalities.

Therefore, we strongly recommend that municipalities interested in producing their own personalized epidemiological analyses and advancing to the level of major cities with timely and responsive epidemiological surveillance invest in local data science and, through the application of the XBI concept, feel free to reproduce this methodology, centered on integrating a data science professional with the current team of public health physicians and nurse vigilantes, in order to automate repetitive processes and

appropriate municipal data in surveillance information systems to generate their own reports and epidemiological dashboards on Public Epidemiological Observatory.

With the COVID-19 pandemic, the Ribeirão Preto Public Health Service faced unprecedented challenges. Health Surveillance did not shy away from its role in this scenario and worked diligently on assertive solutions that added the necessary dynamism, without forgetting its commitment to intrinsic responsibility towards the population.

In Ribeirão Preto, the adoption of advanced data science tools represented a significant change in the field of epidemiological surveillance. Using the XBI methodology, processes for generating paginated reports and creating dynamic dashboards were automated. This municipality shows that it is possible to improve its approach to disease monitoring and public health management, with XBI as an alternative to the traditional model offered to Brazilian municipalities with 2x2 data tabulators (such as TABNET and TABWIN by DATASUS) to produce reports on epidemiological indicators.

The successful implementation of the XBI project highlights the transformative potential of BI in modernizing epidemiological surveillance systems. By automating processes and improving analytical capabilities, XBI allowed public health authorities in Ribeirão Preto to respond effectively to the challenges posed by pandemic times and beyond.

ACKNOWLEDGMENTS

The authors extend their gratitude to the Epidemiological Surveillance Division (DVE), especially for: Juliana Queiroz Rocha de Paiva, Camila Balsero Sales, Lilian Christiane Andrade Grimm, Silvio Fernando Schiavetto, Alexandre Mishima dos Santos, Elaine de Fatima Giordano, Helena Barbosa Lugão, Viviane da Mata Pasti Balbão, Erika Aparecida Catoia, Paula Marília Afonso Torres, Daniela Borges Bittar, Mayra Fernanda de Oliveira, Carmelita Aparecida de Oliveira Zainin Felix, Giuliana Bin, Andreia Arantes Guedes de Campos, Soraya L. F. Fiamengui, Danielle Cristina Dacanal Gentil, and Denise Bergamaschi Giomo.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

1. DONIDA, Bruna et al. **Making the COVID-19 pandemic a driver for digital health: Brazilian strategies.** JMIR Public Health and Surveillance, v. 7, n. 6, p. e28643, 2021

2. CRISAN, Anamaria. **The importance of data visualization in combating a pandemic.** American Journal of Public Health, v. 112, n. 6, p. 893-895, 2022.

3. ETAATI, Leila. **Descriptive Analysis in Power Query with R. In: Machine Learning with Microsoft Technologies.** Apress, Berkeley, CA, 2019. p.121-135

4. ARANGURÍ, María et al. **A Prospective Model of a Digital Epidemiological Surveillance 4.0 Based on the Peruvian Public Health System.** In: Communication and Smart Technologies: Proceedings of ICOMTA 2021. Springer Singapore, 2022. p. 40-50.

5. COSSIN, Sébastien et al. **Public Health and Epidemiology Informatics: Recent Research Trends Moving toward Public Health Data Science.** Yearbook of Medical Informatics, v. 29, n. 01, p. 231-234, 2020.

6. SILVA, Norberto Peçanha da et al. **A utilização dos programas TABWIN e TABNET como ferramentas de apoio à disseminação das informações em saúde.** 2009. Tese de Doutorado.

7. LIMA, Areta Cristina et al. **DATASUS: o uso dos Sistemas de Informação na Saúde Pública.** Refas-Revista Fatec Zona Sul, v. 1, n. 3, p.16-31, 2015

8. LAGUARDIA, Josué et al. **Sistema de informação de agravos de notificação em saúde (Sinan): desafios no desenvolvimento de um sistema de informação em saúde.** Epidemiologia e Serviços de Saúde, v. 13, n. 3, p. 135-146, 2004.

9. DE NONI, Agenor et al. **A two-wave epidemiological model of COVID-19 outbreaks using MS-Excel®.** medRxiv, p. 2020.05. 08.20095133, 2020.

10. COLLIE, Rob; SINGH, Avichal. **Power Pivot and Power BI: The Excel User's Guide to DAX, Power Query, Power BI & Power Pivot in Excel 2010-2016.** Tickling Keys, Inc., 2015

11. DUNLOP, Neil**. Power Query. In: Beginning Big Data with Power BI and Excel 2013.** Apress, Berkeley, CA, 2015. p.145-172

12. DA SILVA, Leonardo José Tomaz; BRILHANTE, Andreia Fernandes; MELCHIOR, Leonardo Augusto Kohara. **Business Intelligence in supporting strategic health management: an experience report.** Revista de Epidemiologia e Controle de Infecção, v. 12, n. 4, p. 180-183, 2022.

13. OLUOCH, Tom et al. **Implementation of an HIV Case Based Surveillance Using Standards-Based Health Information Exchange in Rwanda.** Studies in Health Technology and Informatics, v. 310, p. 875-880, 2024.

14. MEI, Zhang; ZIRONG, Yang. **Design of epidemic monitoring platform based on ArcGIS**. In: 2015 14th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES). IEEE, 2015. p. 380-383.

15. ZHAN, F. Benjamin et al. **Match rate and positional accuracy of two geocoding methods for epidemiologic research**. Annals of epidemiology, v. 16, n. 11, p. 842-849, 2006.