Chapter 4.4
Collection and management of good quality data

Fernando Gouvea-Reis

Marcelo Farah Dell'Aringa

Virginia Murray

Learning objectives

To understand key aspects of data collection for research in health emergency and disaster risk management (Health EDRM), including:

- Different sources and methods for data collection, along with their advantages and limitations.
- Challenges involved in collecting data in disaster settings, and how these might be overcome.
- The importance of data quality, data storage and data sharing.

Introduction

- Collection and management of data are essential steps in research.
- Inappropriate collection or management of data can lead to inadequate analysis and compromise the results of the research.
- Major challenges arise when analyzing and managing data in emergency and disaster contexts.

Initial steps

- Formulate a research question (see chapter 3.5).
- Define the data collection strategy.
- Develop a written protocol stating objectives and methods.
- Define how data will be collected and managed.

Types of data

- **Primary Data:** data collected for the first time and for the purpose of a specific study.
- **Secondary Data:** data already collected or produced by others (e.g., government reports, scientific articles, books, general media or routine data).

Primary data

Primary data can be **qualitative** or **quantitative**.

Advantages of using primary data:

- Gives the researchers more control on the information to be gathered and on the quality of the data
- Can be gathered having specific research questions in mind

Disadvantages of using primary data:

- Time and cost
- Limited access to the affected region, population, and sources of data
- Ethics

Secondary data

Some examples of secondary data are demographic, health event data, circumstantial data and national reference data.

Advantages of using secondary data

- Useful and cost-effective if time or resources are constrained
- Routine data can be used for comparisons before and after an event
- Can be used to develop statistic models

Disadvantages of using secondary data

- Researchers do not have full control on the quality of data
- Multiple data sources may be needed

Challenges for data collection in disasters

- Safety and emergence of new hazards, such as disease outbreaks, water and sanitation problems, violence and aftershocks of earthquakes.
- Lack of data or complete loss of data.
- Damage to transportation network.
- Language barriers.
- Political barriers.

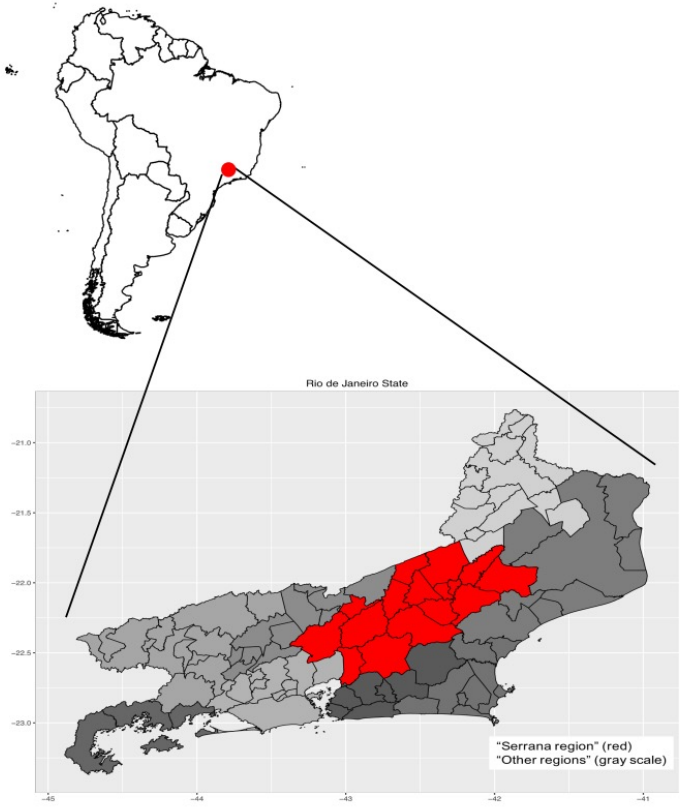
Data storage

- A reliable curation system is needed that allows for standardization and adaptation to the context.
- There should be clear rules for data entry.
- A variety of types of data variable should be allowed.
- Data should be as disaggregated as possible.
- Data collection and storage should safeguard individual privacy.

Data sharing

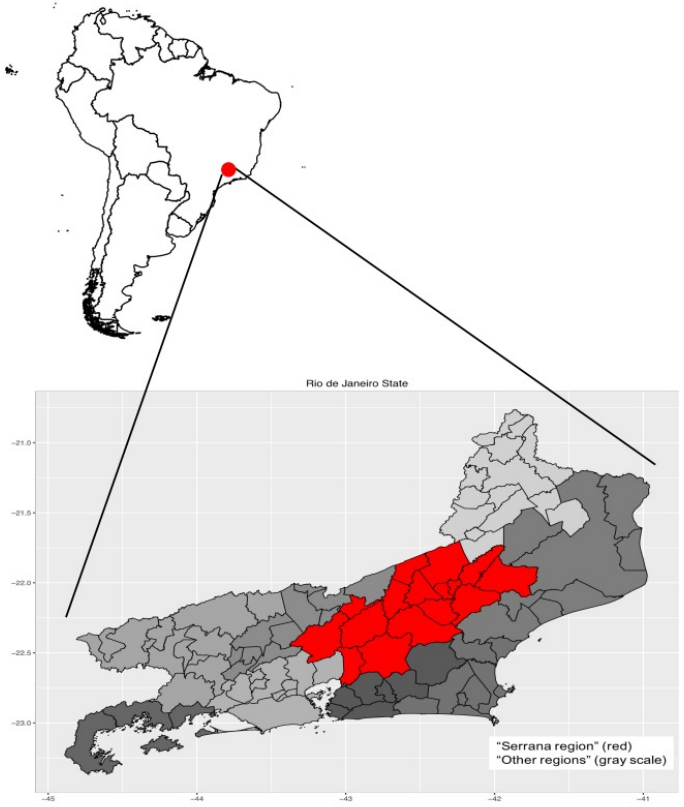
- Data sharing with other researchers allows replication and confirmation of study results.
- Science in general is moving towards open and transparent data.
- FAIR principles: findability, accessibility, interoperability and reusability.
- Timely access to data is still a major constraint.
- Need for clear standards and definitions, such as the hazards investigated in Health EDRM research.

Case study: *Evaluation the impact of the 2011 Rio de Janeiro landslides in the utilization of public mental health services (1)*



- South and Southeastern regions of Brazil are frequently hit by events linked to hydrometeorological hazards.
- The 2011 landslides in the mountainous region of Rio de Janeiro State were the largest disaster by immediate death count in recent Brazilian history: 845 immediate deaths and around 30 000 people left homeless.
- 11 municipalities were affected, causing damage to agricultural and industrial activities.

Case study: *Evaluation the impact of the 2011 Rio de Janeiro landslides in the utilization of public mental health services (2)*



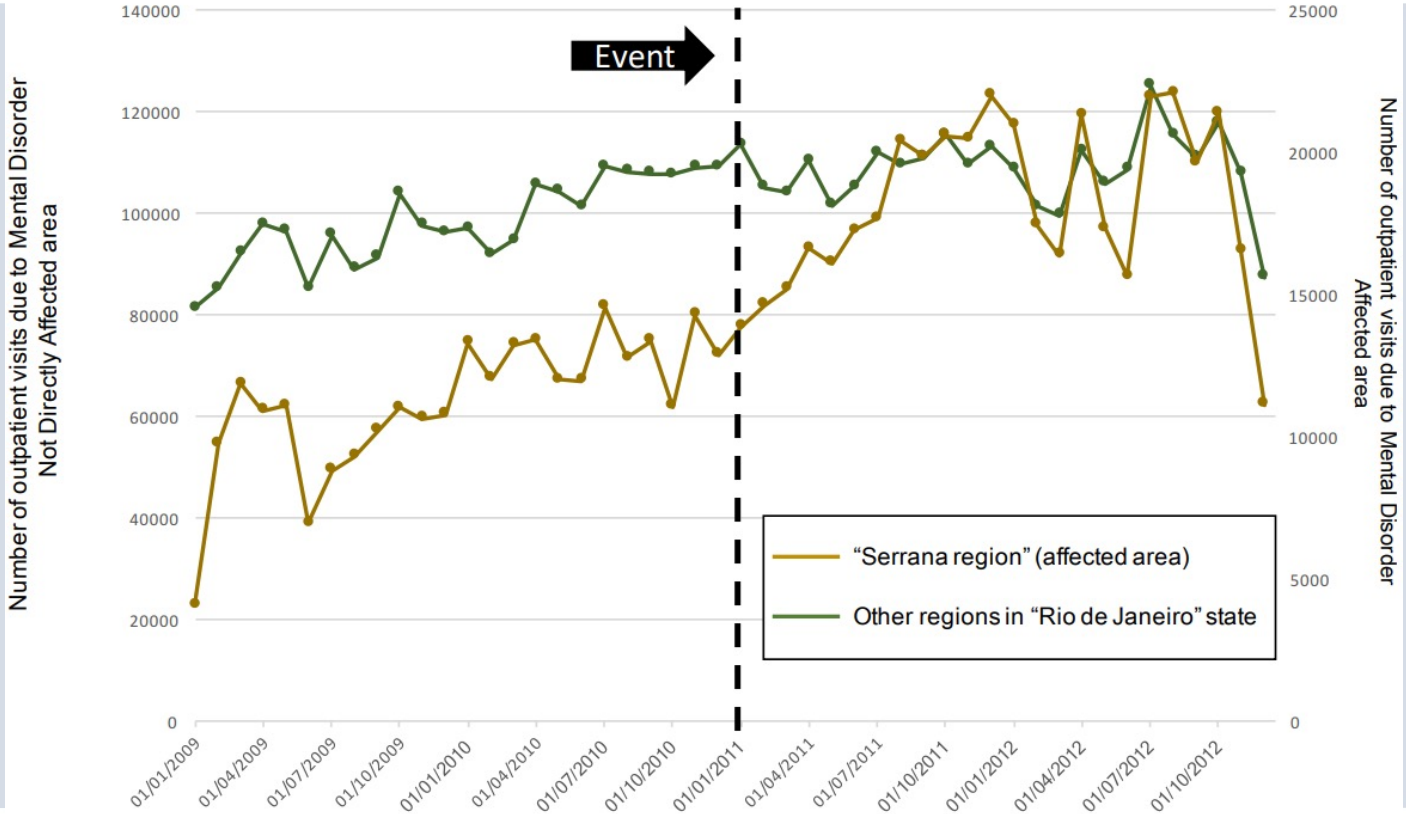
Main data source: DATASUS (Brazilian public health information system), providing routine open-access secondary data, for 2 years before and after the event.

Study design: Interrupted time series using a comparison group.

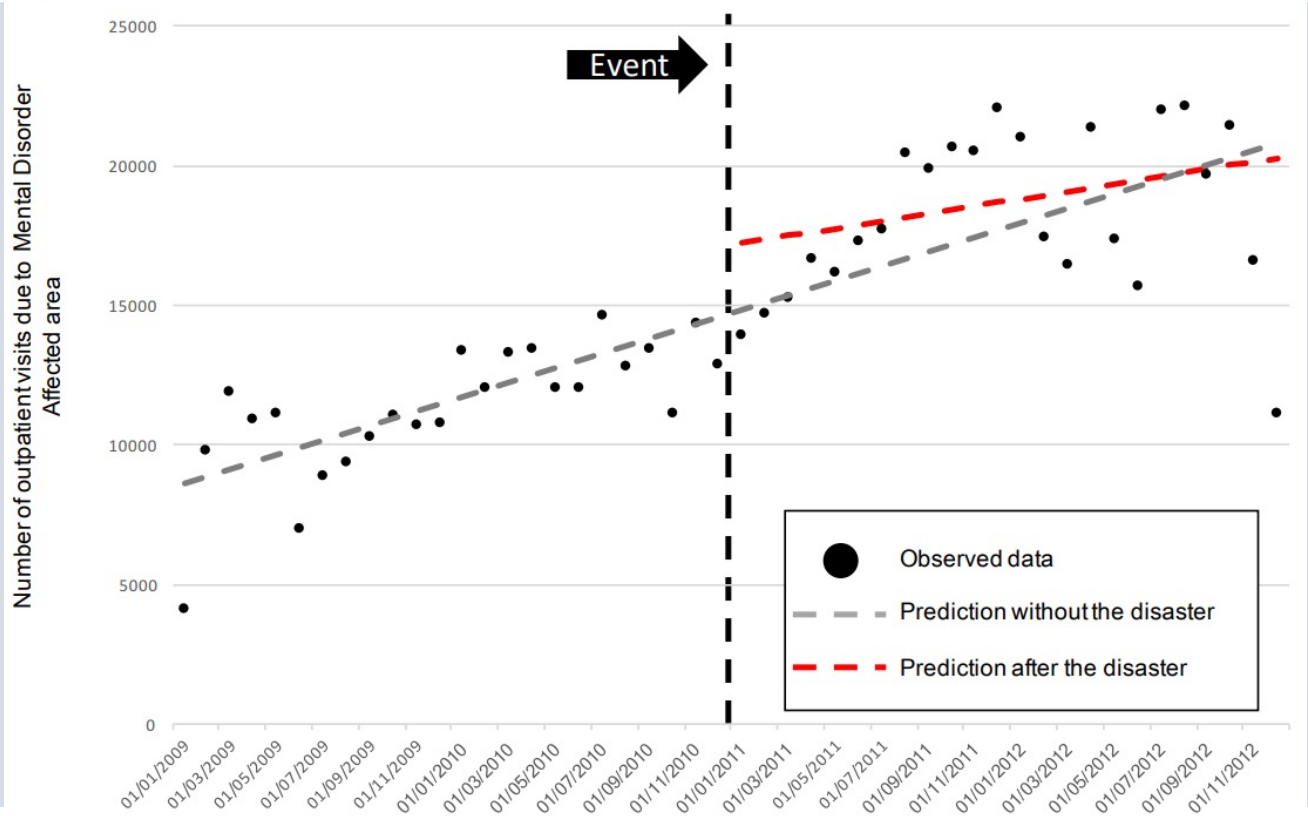
Population: Population of Serrana region in Rio de Janeiro State (affected region).

Comparison Group: Population of other health regions of the State.

Case study: *Evaluation the impact of the 2011 Rio de Janeiro landslides in the utilization of public mental health services (3)*



Case study: *Evaluation the impact of the 2011 Rio de Janeiro landslides in the utilization of public mental health services (4)*



Case study: *Evaluation the impact of the 2011 Rio de Janeiro landslides in the utilization of public mental health services (5)*

Result: There was a sustained increase in the use of public mental health services in the affected region after the landslides, which was not seen in the unaffected region.

Key messages

- High quality research needs to have a specific research question and a data collection strategy that will provide adequate and sufficient information to answer this question with the available resources.
- Even with good preparation, challenges may occur and anticipating how to deal with these can help researchers to overcome future barriers.
- A careful plan on how the collected data will be stored and shared in the long term will ensure that others benefit from the study.

Further readings

Fakhruddin B, Murray V, Gouvea-Reis F. Disaster loss data in monitoring the implementation of the Sendai Framework [Policy brief]; 2019. https:// council.science/publications/disaster-loss-data-in-monitoring-theimplementation-of-the-sendai-framework

This policy brief focuses on the use of disaster data archives and loss data collection in monitoring the implementation of the Sendai Framework for Disaster Risk Reduction 2015-2030.

References

This chapter: Gouvea-Reis F, Dell'Aringa MF, Murray V. Chapter 4.4: Collection and management of good quality data.
Rio de Janeiro Landslides Study: Dell'Aringa M, Ranzani O, Bierens J, Murray V. Rio's Mountainous Region ("Região Serrana") 2011 Landslides: Impact on Public Mental Health System. PLoS Currents: Disasters; 2018: 10.
FAIR Principles: https://www.go-fair.org/fair-principles/

Contact information

Marcelo Farah Dell'Aringa
EMS Physician Piedmont, Italy
PhD Student CRIMEDIM, University of Eastern Piedmont
Email – mardellaringa@gmail.com

Health EDRM Research Network Secretariat
WHO Centre for Health Development (WHO Kobe Centre)
Email: wkc_tprn@who.int