

A DATA MANAGEMENT PLAN OF HYPERTENSION HEALTH LITERACY IN AUSTRALIA

1. PLAN OVERVIEW



Title: Hypertension health literacy in Australia

Creator: Zhenyu Zhang

Affiliation: UNSW Sydney (unsw.edu.au)

Data Manager: Zhenyu Zhang

Project Administrator: Dr. Sanja Lujic

Funder: UNSW Sydney (unsw.edu.au) Centre for Big Data Research in Health (CBDRH)

Grant: UNSW seeding grant (ID 1234/56).

Project abstract:

Hypertension, also known as high blood pressure, poses a significant risk to individuals' cardiovascular health. It can lead to heart disease, stroke, and other related conditions. According to the National Health Survey, 22.8% or 4.3 million Australians adults had high blood pressure in the 2017-2018 financial year [1]. This prevalence rate has remained relatively unchanged since the previous survey conducted in 2014-2015, which reported a rate of 23.0%. It concerns to note that the prevalence of hypertension has not significantly decreased in recent years, despite efforts to address it. Raising awareness and improving prevention strategies are necessary for hypertension management. Patient knowledge of hypertension is vital for effective management [2]. Individuals can make informed decisions about lifestyle, treatment, and prevention by understanding hypertension. However, despite the significance of hypertension health literacy, limited research has been conducted to investigate this area specifically in Australia.

Start date: 02-Jan-2023

End date: 22-Dec-2024

Last modified: 25-Jun-2023

2. DATA COLLECTION

1. Self-Reported Questionnaires:Data

Type: Primary data

Format: Structured questionnaire responses

Volume: The study population included patients aged 30 and over attending select GP practices in Australia. The aim is to recruit 3,000 patients from 150 GP practices across Australia between May and September 2023. A sample of the Australian Medical Publishing Company (AMPCo) GP members will be randomly selected from the AMPCo mailing list. This sampling approach is similar to the methods employed in a previous research study [3]. However, the specific quantity will be determined by the number of participants. Modifying the length and complexity of the questionnaire and increase electronic survey platforms are considerable if cannot meet the population target on time.

Data Entry Validation: Implement data validation process to detect and correct errors during data entry. This process includes range checks, consistency checks, and data verification procedures. Data verification procedures include randomly comparing the entered data against the original data and cross-checking values.

Sharing and Long-Term Access: It is important to follow protocol and consider restrictions to ensure data privacy, confidentiality, and compliance with ethical and legal requirements. The collected questionnaire data will be transcribed into CSV (Comma-Separated Values) format and submit/update to the UNSW secure network drive for long-term access.

Reuse: This data is collected specifically for your study and may not be reused directly for other purposes unless proper consent and ethical considerations are in place.

Versioning: Include version numbers and timestamps in file names to manage different iterations of files.

2. Australian Bureau of Statistics (ABS) Mid-Year Estimated Residential Population:

Data Type: Secondary data

Format: Tabular or database format, provided by ABS

Volume: The ABS population data covers the entire Australian population and is accessible at various geographic levels (e.g., national, state, regional).

Sharing and Long-Term Access: The ABS data is publicly available, ensuring broad sharing and long-term access. It is important to comply with any licensing or usage restrictions imposed by ABS.

Reuse: The ABS population data can be reused for other studies or analyses, as it provides population estimates for different geographic areas and time periods.

Versioning: Include version numbers and timestamps in file names to manage different iterations of files.

3. DOCUMENTATION AND METADATA

Metadata will accompany the data all the time (include backup). About documentation details, please see **6. Selection and Preservation**.

Please see **Metadata_of_HHLA_V1.0_25_Jul_2023** for Metadata details.

4. ETHICS AND LEGAL COMPLIANCE

The use of Hypertension survey data in this research project is subject to rigorous adherence to ethical and legal standards to ensure the protection of privacy, confidentiality, and the rights of research participants.

Ethical Review: The established research procedures involving the use of Hypertension survey data has undergone ethical review by the UNSW Human Research Ethics Committee (HREC) in accordance with the National Health and Medical Research Council (NHMRC) guidelines. The ethics approval ensures that the project meets the principles of ethical research design and conduct, as outlined in the [National Statement on Ethical Conduct in Human Research \(2007, updated 2018\)](#).

Compliance with Legal Requirements: This research project will comply with all relevant laws, regulations, and guidelines related to the use of personal information in research. The project team will stay updated on any changes to applicable legislation and ensure ongoing compliance throughout the duration of the project.

Informed Consent: All participants in the Hypertension survey provided informed consent for the use of their data in research. The consent process followed the guidelines set forth by the [National Statement on Ethical Conduct in Human Research \(2007, updated 2018\)](#) and relevant institutional policies. Consent forms clearly outlined the purpose of the survey, the potential risks and benefits, and the voluntary nature of participation. After clear explaining these, the investigator will obtain written informed consent from study participants (both the patient and GP). The original signed consent will be retained in the UNSW premises.



Privacy and Confidentiality: [The Privacy Act 1988](#) and the [Australian Privacy Principles \(APPs\)](#) govern the handling of personal information in this research project. Personal information will be de-identified and securely stored in compliance with applicable data protection laws and institutional policies. Access to the data will be restricted to authorised project members only.

Data Security: Please see **5. Storage and Backup** for details.

Data Sharing and Transfer: Please see **7. Data Sharing and Transfer** for details.

5. STORAGE AND BACKUP

Adequate data security measures will be employed to protect the confidentiality and integrity of the Hypertension survey data. Only authorised staff can get access to the data. Staff are required to comply with applicable privacy regulations and codes of ethics, and ensure relevant training is completed before being authorised to access data.

- Completed paper questionnaires and consent forms will be stored in locked filing cabinets at Dr. Lujic's office, accessible via a swipe card.
- Scanned electronic copies of the questionnaire and consent forms as well as entered and analysed data will be kept on the UNSW secure network drive. Access controls, encryption, and secure storage protocols will be implemented to prevent unauthorised access, data breaches, and data loss.  

6. DATA DE-IDENTIFY AND DATA CLEAN

Personally identifiable information of patient is not included in questionnaire. Instead, a unique Participant ID was assigned to each participant for identification purposes. Anatomical Therapeutic Chemical (ATC) Classification System was used to coding hypertension medications in data clean. This will not allow easy identification of the actual medications and simplify subsequent analysis. Numeric codes (integer) will replace the categorical data, also the specific codes for missing or implausible values. These will help to simplify subsequent analysis. Limitations of Numeric data will be applied during data clean. If a data value falls outside of the predefined limitations but is still considered reasonable:

- Contact the GP: If you have access to the GP who collected the data, it can help to confirm whether the value was accurately recorded. Or the GP may provide additional insights to reach and verify the value.
- Use judgment: Depending on analyser expertise and knowledge of the subject matter, analyser can exercise professional judgment to assess the reasonableness of the value. If the value appears to be within a clinically acceptable range despite exceeding the predefined limitations, you may choose to retain it. PLEASE NOTE, document the reasons consider the value to be reasonable and any other supporting information. This documentation will help ensure transparency and allow others to understand the decision-making process.
- Sensitivity analysis: If appropriate for your analysis, you can conduct sensitivity analyses by evaluating the impact of including or excluding values that fall outside the limitations. This can help assess the robustness of your results and provide insights into the potential influence of these data points.

7. SELECTION AND PRESERVATION

Following long-term value dataset should be retained and/or preserved.

- Primary Data: The primary data collected from the self-reported questionnaires filled by patients and GPs are valuable for the study and should be retained for long-term use. These data include patient demographics, health literacy measures, disease implications, and other relevant variables. Keeping the primary data allows for data validation, re-evaluation, future analysis, and comparison with new research findings.
- Metadata and Documentation: Metadata and documentation are related closely to the research. Metadata, along with data dictionary, study protocol, survey instrument, and data collection procedure, are crucial for understanding and interpreting the data. These materials should be preserved and shared along with the data to ensure the reproducibility and transparency of the study in the future.
- Derived Data: Derived data, such as aggregated statistics, calculated variables, and transformed variables, that are generated during data analysis can be valuable for long-term reference. These derived data reveal trends, patterns, and relationships during the study. Keeping derived data allows further analysis and comparison of future studies.
- Anonymized Data: Considering the importance of confidentiality, privacy, and security in managing health data, the anonymized versions of data can be shared and used for secondary analysis while protecting the privacy of the participants.
- External Data Sources: The Australian Bureau of Statistics (ABS) mid-year estimated residential population, were used in this study to produce adjusted prevalence estimates. It is valuable to retain and document the used data. As these dynamic data will be changed after years, keep the original data allows other researchers to replicate and validate the study's findings.


The long-term preservation plan for the dataset

To ensure the long-term value and accessibility of these data, it is recommended to store them in a secure and standardised format, such as a structured database or file formats like CSV (comma-separated values) or XML (extensible markup language). Maintain good file naming habits will make organise and categorise data easier and help to enhance data traceability. In addition, regular physical backups of data can be applied and be stored in locked filing cabinets at Dr. Lujic's office with paper questionnaires and consent forms.

8. DATA SHARING AND TRANSFER

Data sharing and transfer can be facilitated through allow the access of data repository - UNSW SECURE NETWORK DRIVE. When sharing or transfer data, it is important to consider restrictions to ensure data privacy, confidentiality, and compliance with ethical and legal requirements.

Restrictions on data sharing

- Access Control: Data can be shared with restricted access; thus, only authorised individuals or groups have permission to access and download the data.
- Anonymization and De-identification: Personal identifiers can be removed or anonymized from the shared data to protect individual privacy. 
- Following informed consent to maintain participant privacy.

When a data breach or suspects that a data breach has occurred, following the UNSW [Data Breach Management Procedure \(unsw.edu.au\)](https://www.unsw.edu.au/data-breach-management-procedure).

9. MODIFICATION AND REVISIONS

A modification request is to be submitted to the approving HREC or HREAP if revisions are made. Please refer to [Modifying or Amending an approved Human Research Project | UNSW Research](#) to clarify whether the revisions would require a modification submission.

10. DATA ATTRIBUTION

Centre for Big Data Research in Health (CBDRH), UNSW

REFERENCES:

1. Australian Bureau of Statistics. *National Health Survey: First results* [Internet]. Canberra: ABS; 2017-18 [cited 2023 June 25]. Available from: <https://www.abs.gov.au/statistics/health/health-conditions-and-risks/national-health-survey-first-results/latest-release>.
2. Hocking A, Laurence C, Lorimer M. Patients' knowledge of their chronic disease-the influence of socio-demographic characteristics. *Australian Family Physician*. 2013;42:411-6.
3. Schäfer WL, Boerma WG, Kringos DS, De Maeseneer J, Gress S, Heinemann S, Rotar- Pavlic D, Seghieri C, Svab I, Van den Berg MJ, Vainieri M, Westert GP, Willems S, Groenewegen PP. QUALICOPC, a multi-country study evaluating quality, costs and equity in primary care. *BMC Fam Pract*. 2011;12:115.