**Intro:**

The project aims at demonstrating how different NCRIS facilities (i.e., PHRN and AURIN) can work together and obtain SA1 level linked health data that would otherwise be difficult to obtain. With this data it is possible to generate health indicators at a finer spatial granularity, bringing insights that are beneficial to research communities and decision-making scenarios. From this perspective, this project is a pilot project for helping researchers in the future to follow the procedure we are going through (which will be documented and shared by the end of this project) and obtain data in their own research.

**Risks:**

The data used in this study are de-identified, and thus the risk of re-identification is low. In order to assess and further minimize any disclosure risk, we apply and echo here the Five Safes framework in this project:

1. Safe people: We have a highly experienced team of researchers covering various disciplines who are aware of issues of confidentiality, privacy, and the conditions of data usage through our numerous similar studies that have addressed these issues. Investigators either have experience in working with linked health data (e.g., David, Melanie, Derrick), or will take course for linked health data in order to gain experience (e.g., Flavia, Hao).
2. Safe projects: the data is used in this project to derive fine spatial granularity indicators and identify urban and regional settlements that are vulnerable to extreme heat events. The benefit of this project will be explained more in the next section. There will be no contact with members of the study cohort. Names and addresses are not included in our datasets. Potentially re-identifiable fields full admission and separation dates are required for identifying hospital transfers and for accurate calculation of time to event for the outcomes analyses.
3. Safe settings: Data will be stored on the secure computer network at SURE facility. Access to the datasets is restricted to research team members who are named in the ethics application and have signed a confidentiality document. Users also need to install a personal digital certificate on each computer from which they access SURE.
4. Safe data: Any data transaction procedure will be vetted through SURE’s secured Curated Gateware by a nominated SURE data custodian, for compliance and the SURE system records and archives all transactions for future reference. The health data itself will not leave SURE during the project and will be archived in SURE and destroyed within 7 years after publication of the final output.
5. Safe outputs: the output health data indicators will be aggregated and smoothed in order to reduce the risk of disclosure to a minimum. A risk assessment procedure will be conducted through the project advisory committee, which is independent from the project steering committee. All publications and presentations arising from this project will not contain any identifying information, and no individual, medical practice or hospital will be identified or identifiable in such material.

Specifically, since we are using SA1 level linked health data in this project, the reporting of the health indicator data consider two key issues in order to minimize identification risk: (i) data privacy and (ii) statistical stability. Data privacy relates to the responsibility to protect the identity of individuals in their data, and ensure that this is not compromised by the release of that data for reporting purposes. Statistical stability relates to the inherent random fluctuation of statistics based on small numbers of cases; the smaller the numbers, the more they fluctuate, potentially leading to incorrect interpretation. These issues are particularly relevant when considering geographical data. To address both these issues for geographical data, we will use a specific statistical method known as “spatial smoothing”. While standard methods typically only adjust for age and sex in each area, spatial smoothing recognises the geographical structure of the data and includes data from the neighbouring geographical areas when calculating the spatial estimates. This additional data provides greater stability to the estimates. In addition, because the spatial estimates are modelled, rather than observed, spatial smoothing reduces any risk of identifiability for specific individuals. Smoothed estimates are designed to reflect the real differences in the underlying rate or risk between areas. For this study, the spatial smoothing will be adjusted for age, sex and comorbidities (determined from the principal and secondary discharge diagnosis fields).

**Benefits:**

This project will develop new indicator data assets to improve our understanding of the health of urban populations and identify incidence patterns and key risk factors across the population. We will integrate health, socio-economic, environmental, climate and built environment datasets to provide a holistic spatially-explicit understanding of urban population health. These indicators will allow health, urban and social infrastructure planners and policy makers to develop targeted policies and actions, and the outcomes will be shared with the research community.

The added value of this project contributes to better understanding of effects caused by heatwaves, which are Australia's deadliest natural hazard and the occurrence and severity of is predicted to increase and become more frequent through climate change, including effects such as extreme heat in densely built urban areas with low vegetation through the urban heat island effect in Australian cities and towns. It aligns with the federal government’s goal to reduce green-house emissions by 43% by 2030. It’s also important to understand the importance of heatwaves regarding environmental population health considering all aspects that might play a role, so they can address better with their policies.

The findings are expected to lead to improved understanding of urban planning and design options for positive public health outcomes, leading to improved urban planning and decision making. The research will also contribute assets to the research community that will help other researchers investigate similar phenomena in the built environment at every stage, from data integration and analysis methodologies. This project will provide us with a precision medicine approach to identify local areas of vulnerability to heat which could be used to mitigation interventions (e.g. local government planting more trees, or erecting shade cloths in areas with high heat vulnerability)”.