

What is linked health data?

What is data linkage?

Data linkage allows researchers and policy makers to study emerging health trends and problems in a more cohesive and comprehensive way. Linking information helps in the design and development of health services for the community.

The data linkage process brings together information collected routinely about individuals and the services they use after identifying details have been removed.

For example, use of health services such as hospitals can be linked with other information collected by state and territory health departments, births, deaths and marriages registries, cancer registries and research studies such as cohort studies and clinical trials. The results can be used to provide a fuller picture on disease causes and risk factors, the best approaches to prevention or early intervention, and the effectiveness and safety of health programs.

Who manages the information?

Each set of data is held by a data owner, or custodian, who is responsible for the management and security of the information. Data custodians work within organisations that hold data such as government departments or universities

Who links the data and how does linkage work?

Organisations called *Data Linkage Units* work with data custodians to link information together without identifying individual patients.

Data custodians hold three types of information about people: personal identifying information (for example, a person's name and demographic details), ID numbers for people's records, and the content of the records themselves, which could include tests ordered, medication taken, or hospital admission information.

Custodians provide the data linkage units with encrypted record ID numbers and personal identifying information – but they do not supply them with the record content or any clinical information.

Linkage Units then use the information from custodians to create a unique linkage ID or key – made up of a random set of numbers and letters – for each person. These keys replace the identifying information for each person and are stored securely in the data linkage unit.

How do researchers access the linked data?

Before custodians can provide researchers with any information, they must approve the research project that will use it. The project must also have approval from a Human Research Ethics Committee.

Custodians and linkage units then decide which records are needed for the project. The linkage units then use the unique ID keys they have created to devise another set of linkage IDs – called Project Linkage IDs – that are specific to the study in question.

These Project Linkage IDs and the encrypted record ID numbers needed for the project are then sent back to the Data Custodian, who decrypts the record IDs, links them – together with the project linkage keys – to the record content, and then removes the record IDs altogether.

Each data custodian then sends the researcher the Project Linkage IDs and their corresponding record content without any identifying information. This allows the researcher to link together data that relates to the same person from different sets of data without knowing who that person is.

What role does the Sax Institute play in data linkage?

The Institute has developed and runs SURE – the Secure Unified Research Environment – which allows multiple researchers to securely access linked data without receiving it directly from data custodians. Instead, custodians will send the information to SURE, where researchers can access it to carry out their analysis from any location. Custodians have control over what records go into and out of the SURE facility.

Using SURE to access and analyse the linked data will offer a consistent, high level of security for researchers, data custodians and the people to whom the records relate. Researchers who already have linked data sets can also choose to put them in SURE to take advantage of its benefits.

Privacy and security of information

Information entering SURE has already been through the data linkage process and is therefore anonymous. Data custodians and linkage units also have their own strict privacy and security measures.

The Population Health Research Network (PHRN), an Australian Government initiative working across states and territories to develop national infrastructure in health data linkage, adopts a range of privacy and security measures. These include an extensive governance and management system, and ongoing privacy impact assessments. The PHRN collaboration includes seven data linkage units.

Further resources

More information on Data Linkage Units, how data linkage is used and data security can be found at:

- The Population Health Research Network (PHRN) www.phrn.org.au
- SA/NT Data Link www.santdatalink.org.au has developed an animation that explains the data linkage process: <https://www.santdalalink.org.au/animation>
- The Centre for Health Record Linkage (CHeReL) www.cherel.org.au

For information about SURE visit <http://www.saxinstitute.org.au/our-work/sure/> or the SURE website www.sure.org.au