Responsibilities for the Secondary Sharing of Clinical Trial Data in Australia

Data Glossary

Term	Definition
Aboriginal Health Research Ethics Committee	A specialised HREC that reviews human research with Aboriginal or Torres Strait Islander participants or communities.
Administrative data	A data collection made up of information that is routinely collected during the delivery of a service. These data are collected via departments and agencies and used for policy, planning, management, monitoring, evaluation and research purposes. These data are not typically collected for research purposes but may be used for clinical trials or future research using clinical trial datasets.
Adverse event	Any unfavourable medical occurrence in a trial participant administered the investigational product. The adverse event does not necessarily have a causal relationship with the treatment.
Aggregate data	Data that has been grouped together so it no longer identifies specific individuals. 'Aggregate data' is sometimes used interchangeably with 'summary data' or 'tabulated data'.
Bundled Consent	A bundled consent is a single request from an individual, organisation or agency that includes several requests within it to collect, use and disclose someone's personal information. It does not let the person to whom the information relates, choose which requests they consent to and which they do not.
Capacity	The ability of a person to consent to a clinical trial, research or other forms of medical procedure. A person has decision making capacity if they can understand the information relevant to the decision, retain that information to decide something, weigh that information as part of a decision-making process and communicate that decision.

Cell size	The number of cells (rows or columns) in a dataset that all share a particular value. Where this cell size is small, de-identification should be applied to the data to reduce disclosure risks.
Clinical trial	Any research study that prospectively assigns human participants or groups of human participants to one or more health-related interventions to evaluate the effectiveness on health-related outcomes.
Clinical trial dataset	A dataset collected or generated during a clinical trial.
Clinical trial protocol	A plan that explains the purpose and procedure of the clinical trial protocol. This may include the rationale, design, methodology for the trial, inclusion and exclusion criteria, the length of the trial, procedures, methods of analysis, and safety/quality monitoring.
Clinical trial research agreement	A legally binding agreement that governs the relationship between a sponsor and the institution that is responsible for the conduct of the trial.
Cloud Computing	Computing infrastructure and applications that are available over a network. Cloud computing services are typically scalable to meet demand. This term may be used interchangeably with ‘cloud environment’.
Confidentiality, Duty of	A legally enforceable obligation to maintain the confidence of certain information. This obligation can apply to an individual, an organisation, an agency or the officer of an agency.
Consent	An expression of a person’s willingness to participate in a particular activity. Consent must be informed, voluntary, specific and given by a person with decision-making capacity or their guardian.
Consumer Group	A group or organisation that represents people who use health services or carers for people who use health services. Representatives of these groups may provide consumer perspectives, contribute consumer experiences and take part in decision making processes. This term is synonymous with ‘participant groups’.
Data Asset	A data set that is formally recognised and governed by a data custodian. The terms ‘data collection’ or ‘data holding’ are sometimes used interchangeably with ‘data asset’. A data asset may be collected or created internally by an organisation, gathered through a survey, obtained from or held by one or more government agencies, or merged from several data sets.

Data Linkage	The process of joining or linking together two or more data sets because they refer to the same entity (such as the same individual or institution). This term is sometimes used interchangeably with ‘record linkage’ or ‘data integration’.
Dataset/Data Set	A collection of structured data in a single file or repository.
Data Breach	Where a dataset is accessed, modified, used, disclosed or misused by an entity that does not have authority to access that data.
Data Custodian	Any entity, including a person, agency or organisation, responsible for managing the use, disclosure and protection of data in a dataset. A data custodian may also be responsible for ensuring the dataset adheres to quality and security standards. The term ‘data custodian’ may be used interchangeably with ‘data steward or ‘data access committee’ (depending on the size of the dataset).
Data Sharing Agreement	A legally binding arrangement between a data custodian and another agency, organisation or individual that specifies how a dataset may be shared and used.
Data Sharing Policy	A policy that addresses the types of data held by a data custodian, the principles and strategy around the sharing of that data and the governance and management procedures that should be adopted when this data is being shared.
Data management plan	A document that addresses how data and information will be generated, collected, accessed, used, analysed, disclosed, stored, retained, disposed of and shared, as well as the risks associated with these activities, and any strategies for minimising those risks.
Data repository	The final location where data will be published or stored at the conclusion of a research project. A data repository can either be a subject or domain specific repository, an institutional data repository or a generalist repository.
Data retention	The storage of data for archiving or auditing purposes at the end of a research project.
Data security	The use of strategies to protect data from disclosure risks, alteration and destruction.
Data sharing	The process of making data available to another person, agency or organisation.
Data storage	When an agency, organisation or person physically holds or otherwise controls data.

De-identification	The process of removing identifying details from a dataset so that it no longer contains identifiable information. This can involve multiple processes, including aggregating data or applying privacy enhancing technologies. De-identification should be understood more as a process rather than an output. The appropriateness of different de-identification techniques will depend on the setting in which the data is made available, as well as who is performing the de-identification process.
Disclosure risk	The likelihood that an individual may be re-identified from a dataset. This is sometimes synonymously referred to as an identifiability risk.
Direct personal identifier	A value that can be used to identify an individual without being combined with any other value. A direct personal identifier could include a name, an address, a postcode, a full date of birth or a unique personal identifier such as a Medicare number.
Dynamic consent	An informed consent framework where participants can adjust their consent choices for different research projects over time. This framework is given effect through a technological platform.
Electronic health record	An individual health record for a person stored in an electronic platform that is accessible by their health care team. This term is sometimes used interchangeably with ‘electronic medical record’ or ‘electronic patient record’.
Ethical approval	A determination by an Ethics Review Board that a research project is ethically acceptable and in accordance with relevant standards and guidelines.
Ethics review	A process conducted by an Ethics Review Board to ensure that research has scientific merit, the benefits outweigh any risks, the research team are appropriately qualified, and respect is shown for the participants.
Ethics Review Board	A body that is responsible for assessing the ethical merit associated with a particular research activity. This definition can include but is not limited to human research ethics committees and low risk ethics review panels.
Extended consent	Consent by a participant for the use of their information for future research projects that are closely related to the original research project or in the same area of research.
Free Prior and Informed Consent	A framework for consent when working with Indigenous communities that understands consent as an agreement concluded without coercion or undue pressure. Under this framework, an individual should have clear mechanisms to modify or withdraw their consent at any time.

Guardian	A person with legal authority to make decisions on behalf of another person.
Health information	Personal information about the health of an individual, what healthcare they have received or might want in the future, as well as information about organ donation or genetic information.
Health related interventions	Surgical treatments, pharmacological treatments, the use of medical devices or other therapeutic goods, preventative strategies or education campaigns.
Human Research Ethics Committee (HREC)	A body registered with the NHMRC to review whether a research project is ethically acceptable and in accordance with relevant standards and guidelines.
Indigenous Data Sovereignty	The right of Indigenous people to govern the collection, ownership and use of data about Indigenous communities, peoples, lands, cultural knowledge, and resources.
Indirect personal identifier	A value that may be combined with other values to identify individuals from a dataset. These values could include age, ethnicity, sex, marital status, occupation, height or weight. These values could also be included in images, such as MRI scans and X-rays.
Investigator	A researcher or other person who is responsible for the conduct of a clinical trial and managing participants. Where there are multiple investigators attached to a clinical trial, the lead investigator is referred to as the principal investigator.
Law	Legislation or a common law decision that is legally binding in Australia.
Legislation	A law made by the Commonwealth Parliament of Australia, or a state and territory Parliament.
Low risk research	Research in which there is no risk of harm, but in which there is a risk of discomfort and in which there may also be a foreseeable burden.
Metadata	Any data about a dataset. In the context of a clinical trial, metadata includes any data variable labels, as well as the clinical trial protocol , the statistical analysis plan or any code generated during the trial.

Minimal risk research	Research where there is no risk of harm or discomfort, but only a potential for minor burden or inconvenience for participants.
Non-Identifiable Data	Data that has been rendered non-identifiable so it is impossible without legal or technical means to re-identify the data.
Opt-out approach	A method used in the recruitment of participants into research where information is provided to the potential participant regarding the research and their involvement and where their participation is presumed unless they take action to decline to participate.
Participant	An individual who participates in a clinical trial.
Participant Information and Consent Form	A form which provides participants with information about a clinical trial, such as the trial's purpose, duration, interventions, risks and potential benefits. The form also allows participants to provide informed consent to participate in the trial.
Participant-level data	Any raw data collected or generated about participants. This data can include any tests conducted on participants, any care or diagnosis received by participants or any adverse events experienced. Direct personal identifiers are usually removed before sharing the data.
Personal information	Information or an opinion about an identified individual, or an individual who is reasonably identifiable.
Private sector organisation	A natural or legal person that is not a federal government agency, or a state or territory government agency. In the context of clinical trials, a private institution could include a private research organisation or hospital.
Privacy enhancing technologies	Technical measures that are designed to reduce the risk of individuals being re-identified from a dataset. The term 'privacy enhancing technologies' is sometimes used interchangeably with 'privacy-preserving techniques'.
Quality improvement	An activity involving a healthcare system that will lead to better patient outcomes, system performance or professional development. Quality improvement should be treated as a separate activity to research.
Quasi-identifiers	A value associated with a dataset that does not identify an individual on its own but when combined with other data, could be used to re-identify individuals from that dataset. The more quasi-identifiers for each record in a dataset, the more likely that they could be used to identify individuals from these records.
Rare disease	A disease that affects less than five in 10,000 people.

Regulation	Any attempt to achieve a publicly stated objective or set of objectives, including but not limited to guidelines issued by a government department or public agency.
Research	Any activity undertaken to gain knowledge and understanding.
Regulatory agency	A body that has the power to investigate a clinical trial or research project. Relevant regulatory agencies for clinical trials in Australia include but are not limited to the Therapeutic Goods Administration (TGA), the Office of the Australian Information Commissioner (OAIC), and any state or territory privacy regulators.
Re-identifiable data	Data where individual identifiers have been replaced with a code or pseudonym. Individuals in the dataset can be re-identified using that code, or if the dataset is combined with another dataset.
Reproducibility	The ability to analyse existing data from a previous study to verify the study's findings or uncover potential concerns.
Risk	The likelihood of a particular event (either positive or negative) occurring.
Secondary research	Research where existing data is used to answer new research questions, validate findings or inform policy and practice. This research must be separate from the original purpose for which the data was collected or generated.
Secondary use	Uses of data for a purpose other than that for which it was specifically collected. Secondary uses can include developing or improving health services or products, developing government policy, evaluating health programs, recruiting participants for clinical trials and conducting research.
Sensitive personal information	Any personal information about an individual's racial or ethnic origin, political opinions or membership of a political association, religious or philosophical beliefs, membership of a profession, trade association or trade union, sexual orientation, criminal record. Sensitive personal information also includes health information, genetic information or biometric information used for biometric verification or identification.
Specific Consent	Consent by a participant to participate in a specific research activity.
Sponsor	An organisation or consortium responsible for initiating, managing and financing a clinical trial. A sponsor can include a private organisation, a health service organisation or a university.

State government agency	A Minister or Department of a state or territory government of Australia, as well as a body established under state or territory legislation.
Study protocol	A document that describes the objective, design, methodology, statistical considerations and organisation of a clinical trial. The term ‘study protocol’ may be used interchangeably with ‘clinical trial protocol’ or ‘protocol’.
Trusted Research Environment	A computing environment that allows researchers to remotely access and analyse a dataset without the dataset leaving that environment or the researchers being able to link it with other datasets. This may be used synonymously with ‘secure research environment’ or ‘secure remote access’.
Unspecified consent	Consent to use a participant’s personal information for any future research project.
Value	An attribute associated with the records in a dataset.
Verification	The process of independently assessing any evidence collected or generated.
Waiver of the requirement for consent	Authorisation for a research team to undertake research without the consent of study participants. In Australia, a waiver of the requirement for consent may be authorised by an ERB. For personal information in medical research, or personal health information, a waiver must be authorised by a human research ethics committee.