Research & Innovation Advancement

How to de-identify data

In many cases, an IWK Health Researcher is required to work with de-identified data. De-identifying data requires Researchers to assign a unique identification (ID) number for each participant. This can be done manually or via data software that can create unique IDs (e.g., excel).

Information included in a data set will depend on sample size and characteristics of the sample. In addition to a de-identified data set, there must be a data file linking the unique ID to the participant's identifiable information. This data file must be kept in a separate location (physically or virtually) from the study file and must be password protected.

Participants should be assigned a unique ID once they have consented to the study. The unique D should be used throughout the study rather than identifiable information (e.g., participant's name).

Below is an example of what de-identified data and its accompanying data file could look like.

Example of de-identified data:

ID	Sex	Province	QOL
1000	Male	NS	80.05

Example of accompanying data file:

ID	Name	DOB	Health Card
1000	Joe Smith	1/7/2003	0003650560

Note. If this was a study of a rare disease Sex and Province combined could be considered potentially identifiable characteristics.