FACTSHEET 5: *An example of Record Linkage*

THE MIDDLE CHILDHOOD SURVEY (UNSW)

Responses (research data) from the Middle Childhood Survey are separated from the 'Minimal Identifier Set' of personal identifiers that is sent to the *Centre for Health Record Linkage (CHeReL)* for any requested linkages of the survey data

A research team intends to conduct a LINKAGE PROJECT

They obtain ETHICAL APPROVAL and contact each DATA CUSTODIAN (holder) of records for permission to use them in the linkage project

THE PERINATAL (MIDWIVES) DATA SET (NSW Health)

Responses (research data) from the Perinatal Data Collection are separated from the 'Minimal Identifier Set' of personal identifiers that is sent to the *Centre for Health Record Linkage (CHeReL)* for any requested linkages of the perinatal data

MCS Identifiers

[MCS ID]
Name;
Date-of-Birth;
Demographics

Only the personal identifiers go to the record linkage provider (not the survey data)

Only the personal identifiers go to the record linkage provider (not the perinatal data)

Perinatal Identifiers

[Health ID]
Name;
Date-of-Birth;
Demographics

THIRD PARTY RECORD LINKAGE PROVIDER (CHeReL)

The Linkage Provider detects any of the same 'minimal identifier sets' among the files using a computerised matching process.

A Project-Specific *Linkage ID* is assigned to each matched record.

THEN,
ONLY the Project-Specific *Linkage IDs* and the matching (linked) Data Collection *Record IDs* (e.g., Health ID, MCS ID) are sent back to each Data Custodian

MCS Links

[MCS ID]
[Linkage ID]

Perinatal Links

[Health ID]
[Linkage ID]

MCS Data

[Linkage ID]
Social Skills;
Emotion & Behaviour;
Sleep Patterns

Perinatal Data

[Linkage ID]
Birth weight;
Delivery type;
Birth complications

The research team receive only the *Linkage IDs* and associated research data for each set of records, directly from the Data Custodians. The research team then match the records from these multiple datasets using the anonymous *Linkage ID*, without ever seeing the personal identifiers. The Linkage Provider saw only the personal identifying information, but not the research data.