GUIDE TO HEALTH RECORD LINKAGE SERVICES

Contents

1. Overview 2

2. Aims and functions of the CHeReL 2

3. The Master Linkage Key 3

4. Services provided by the CHeReL 3

5. Contact details for custodians of selected health data collections 4

6. Human Research Ethics Committee approvals 5

7. Applications to Human Research Ethics Committees 5

8. Release of Unit Record Data owned by the NSW Department Of Health 6

9. Charges 6

10. Further Information 6

Appendices: Information to support development of project proposals and ethics applications

Appendix 1: Background on the CHeReL 7

Appendix 2: How the CheReL works 8

Appendix 3: Examples of Project Plans 11

Appendix 4: Security at the CHeReL 13

Appendix 5: Ethical issues relating to record linkage using population health databases 14
1. OVERVIEW

In order to obtain record linkage services from the Centre for Health Record Linkage (CHeReL), investigators are required to have:

1. a project that is consistent with the aims of the CHeReL
2. approval of the relevant data custodians; and
3. approval of the relevant Human Research Ethics Committee (HREC)

The CHeReL will charge a fee for record linkage services.

2. AIMS AND FUNCTIONS OF THE CHeReL

The aims of the CHeReL are to:

(a) facilitate research which may contribute to the promotion, protection and maintenance of the health of the public;
(b) facilitate the planning, evaluation and delivery of health services;
(c) contribute to knowledge regarding research methodologies relating to health data collection, linkage of health-related data and compilation and use of health-related statistics generally;
(d) to allow the outcomes of activities related to the above objects to be available for the benefit of the public and to contribute to knowledge on an open and equitable basis.

The CHeReL has the following functions:

(a) The CHeReL is a repository for personal information (such as name, address, sex and date of birth) compiled from health-related databases. These records do not include information on the health of an individual, other than the fact that the record is derived from a specified database. Records held by the CHeReL have an identification number that points to the original record on the source database.

(b) Subject to ethics committee approval and approval of data custodians, the CHeReL collects and links these records and maintains a database of linked and unlinked records as a resource.

(c) For specific projects, and subject to relevant ethics committee and data custodian approval, the CHeReL supplies lists of specified records to data custodians then provide approved data to the project investigators.

(d) Subject to relevant ethics committee and data custodian approval, linkage can be carried out with other health-related data collections that are not included in the repository.
3. **The MASTER LINKAGE KEY**

The NSW Population and Health Service Research Ethics Committee has given approval for the CHeReL to establish a *Master Linkage Key* of records from the following data collections:

1. NSW Admitted Patient Data Collection
2. NSW Emergency Department Data Collection
3. NSW Midwives Data Collection
4. NSW Perinatal Death Review Database
5. NSW Central Cancer Registry
6. NSW Pap Test Register
7. NSW Registry of Births, Deaths and Marriages birth registration data
8. NSW Registry of Births, Death and Marriages death registration data
9. ABS Mortality Data
10. ABS Perinatal Mortality Data
11. The 45 and Up Study

Personal information from these databases will be linked to create the *Master Linkage Key*, which contains the following information:

- a person number, which is unique to the CHeReL and does not leave the CHeReL
- a code indicating the source database
- the encrypted record number from the source database.

On receipt of ethical approval for a project, the CHeReL will provide each relevant data custodian with a list of encrypted record numbers related to the project, with each record number having a Person Number specific to that project. The data custodians will, subject to any further approvals required within their own organisations, decrypt the record number, match the Project Person Number to the database using the record number, extract the approved information for the project, and forward the project database to the researcher. The researcher will then match the project databases received from each custodian using the Project Person Numbers.

4. **SERVICES PROVIDED BY THE CHeReL**

The CHeReL offers three types of record linkage services:

1. **The Master Linkage Key**
2. Linkage of other datasets with the *Master Linkage Key*
3. Linkage of datasets not included in the *Master Linkage Key*

All research projects that use CHeReL services require the approval of a HREC, even if the data that are proposed to be provided to the researcher are de-identified.
5. CONTACT DETAILS FOR CUSTODIANS OF SELECTED NSW DATA COLLECTIONS

The contact details below are for custodians of the main population health data collections.

1. NSW Admitted Patient Data Collection
2. NSW Emergency Department Data Collection
   Mr Tony Dunn
   Director Demand and Performance Evaluation Branch
   NSW Department of Health
   Locked Bag 961, North Sydney NSW 2059
   Phone: 02 9391 9590
   Email: tdunn@doh.health.nsw.gov.au

3. The NSW Central Cancer Registry
4. The NSW Pap Test Register
   Dr Paul Jelfs
   Director, Population Health and Registries
   Cancer Institute NSW
   PO Box 41, Alexandria NSW 1435
   Phone: 02 8374 5607
   Email: paul.jelfs@cancerinstitute.org.au

5. NSW Midwives Data Collection
6. NSW Perinatal Death Review Database
7. NSW Birth Defects Register
   Dr Lee Taylor
   Manager, Surveillance Methods
   Centre for Epidemiology and Research
   NSW Department of Health
   Locked Bag 961, North Sydney NSW 2059
   Phone: 02 9391 9223
   Email: ltayl@doh.health.nsw.gov.au

8. Registry of Birth, Deaths and Marriages birth registration data
9. Registry of Birth, Deaths and Marriages death registration data
10. ABS mortality data relating to NSW registrations
11. ABS Perinatal Death Data relating to NSW registrations

For information concerning RBDM and ABS data please contact:

Dr Lee Taylor
Manager, Surveillance Methods
Centre for Epidemiology and Research
 NSW Department of Health
 Locked Bag 961, North Sydney NSW 2059
 Phone: 02 9391 9223
 Email: ltayl@doh.health.nsw.gov.au
12. Notifiable Disease Database

Dr Jeremy McAnulty
Director, Communicable Diseases
Centre for Health Protection
NSW Department of Health
Locked Bag 961, North Sydney NSW 2059

For information concerning the NDD contact:

Mark Bartlett
Manager Surveillance
Centre for Health Protection
NSW Department of Health
Locked Bag 961, North Sydney NSW 2059
Phone: 02 9391 9675
Email: mbart@doh.health.nsw.gov.au

13. The 45 and Up Study

Dr Sonia Wutzke
Associate Director (Policy and Operations)
The 45 and Up Study
The Sax Institute
PO Box 123
Broadway NSW 2007
Phone: 02 9514 5943
Email: Sonia.Wutzke@saxinstitute.org.au

6. HUMAN RESEARCH ETHICS COMMITTEE APPROVALS

Ethical approval from the Population and Health Research Ethics Committee should be sought for research projects that use the Master Linkage Key or projects that involve use of data collections owned or managed by the NSW Department of Health or the Cancer Institute NSW.

For other projects, investigators should contact the Manager of the CHeReL for advice on requirements for ethical approval.

The Population and Health Research Ethics Committee is a joint committee of the NSW Department of Health and the Cancer Institute NSW. Further information and application forms can be obtained from: www.health.nsw.gov.au/cancer_inst/research/ethics.html.

7. APPLICATIONS TO HUMAN RESEARCH ETHICS COMMITTEES

This section includes information that may be included in HREC applications.

7.1 Project plan

Background information on the CHeReL is shown at Appendix 1.

Project investigators will need to demonstrate how the record linkage will be carried out in order to complete the project plan. An overall description of how
the CHeReL works in shown at Appendix 2. Examples of approaches to wording for a project plan are shown at Appendix 3.

There are many variations on how record linkage projects may work. A HREC is likely to prefer a project plan that separates the processes of record linkage (which requires access to personal identifiers) from data analysis (which does not require access to personal identifiers). Investigators are encouraged to discuss their project with CHeReL staff prior to submission of any application to a HREC.

7.2 Security
Information on security at the CHeReL is shown in Appendix 4.

7.3 Ethical issues
Advice on ethical issues in relation to population-based record linkage studies is shown at Appendix 5.

8. RELEASE OF UNIT RECORD DATA OWNED BY THE NSW DEPARTMENT OF HEALTH

Approval of an application by a HREC does not by itself constitute authority for disclosure of unit record data. In the case of linkage of data owned by the NSW Department of Health and data owned by another organisation, approval of the project by a HREC is a prerequisite for an authorisation for disclosure to occur.


Release of Department of Health data always has conditions, typically:
- the data are to be used only for [project];
- the data are to be kept in a secure physical and electronic environment that is accessible only by persons directly involved in the project;
- a confidentiality undertaking will be completed prior to the information being released;
- the NSW Department of Health is to be acknowledged in any publication or report;
- the data will not be matched with information on individuals from another source; and
- no information will be released with which it may be possible to identify an individual person.

9. CHARGES

The CHeReL will charge a fee for record linkage services and will provide a quotation on request. CHeReL Member organisations receive a quota of free or discount services. For each Member organisation, these services are coordinated by a nominee of the Member organisation. The CHeReL can provide contact details on request.
10. FURTHER INFORMATION

Ms Katie Irvine
Manager
Centre for Health Record Linkage
Cancer Institute NSW
PO Box 41
ALEXANDRIA NSW 1435
Phone: 02 8374 3515
Fax: 02 8374 3500
e-mail: katie.irvine@cancerinstitute.org.au
APPENDIX 1: BACKGROUND ON THE CHeReL

Record linkage will be carried out by the Centre for Health Record Linkage (CHeReL), which is hosted by the Cancer Institute NSW. The CHeReL is a collaborative venture funded by eight member organisations (ACT Health, the NSW Clinical Excellence Commission, the Cancer Institute NSW, NSW Department of Health, the Sax Institute, the University of Newcastle, the University of NSW and the University of Sydney).

The CHeReL provides a mechanism for access to linked health data for research. The CHeReL uses probabilistic record linkage techniques to link personal information from a defined set of health-related datasets (including hospital inpatient and emergency department data, cancer registry data, birth and death data) in order to create a Master Linkage key, consisting of ‘pointers’ to records for persons on the source databases. For specific projects, and subject to relevant ethics committee and data custodian approval, the CHeReL supplies lists of specified records to data custodians, who then provide approved data to the project investigators. Subject to relevant ethics committee and data custodian approval, the CHeReL also carries out ad hoc linkages of other health-related datasets.

Creation of the Master Linkage Key by the CHeReL has been approved by the NSW Population and Health Research Ethics Committee (a joint committee of NSW Health and the Cancer Institute NSW). The work of the CHeReL is informed by a Community Advisory Committee.
APPENDIX 2: HOW THE CHeReL WORKS

1. The data custodian provides an encrypted source record number and personal information to the CHeReL.

Data custodian → CHeReL

- Cancer register: a1 + name + address + DOB + sex + ...
- c3 + name + address + DOB + sex + ...
- d4 + name + address + DOB + sex + ...
- e5 + name + address + DOB + sex + ...
- Hospital data
- Death registrations: f6 + name + address + DOB + sex + ...

2. The CHeReL links the records and assigns a CHeReL person number to groups of records for the same person. In the example, CHeReL person 35 has one Cancer Register record, two hospital records and a death record, and CHeReL person 78 has a one Cancer Register record and one hospital record. The CHeReL number never leaves the CHeReL.

Data custodian → CHeReL

- Cancer register: a1 35
c3 35
d4 35
- Hospital data: f6 35
- Death registrations: b2 78
e5 78

CHeReL person no.
3. When the data custodians and a human research ethics committee approves a project, the CHeReL assigns a person number for each person in the project. In the example, CHeReL person number 35 becomes project person number 1, and CHeReL person number 78 becomes project person number 2. The project person numbers are different for each project.

4. The CHeReL returns to each data custodian a list of encrypted source record numbers for their database, with each record having a project person number for the project.
5. The data custodian decrypts the source record number, and merges the project person number with the health information that has been approved for use in the project.

6. The data custodian removes the source record number and provides the researcher with the project person number and the health information. The researcher then groups the records together for the same person using the project person number.
APPENDIX 3: EXAMPLES OF PROJECT PLANS

Example 1: Linkage of datasets not already held in the CHeReL Master Linkage Key

We seek access to linked data of [dataset 1] and [dataset 2] for the [financial]/[calendar] years [year] to [year] as follows:

- [dataset 1] eg all records for people aged 65 years and over
- [dataset 2] eg all records that link to [dataset 1]

The variables that will be used for the project are listed in Section [X].

Note: All variables, including personal identifiers should be included in the list, even if the data that will be provided to the researchers is de-identified.

The following fields will be provided to the CHeReL for the project:

- [dataset 1]: encrypted record number from the source dataset, [variable 1] ...[variable X]
- [dataset 2]: encrypted record number from the source dataset, [variable 1] ...[variable X]
- [dataset 3]: encrypted record number from the source dataset, [variable 1] ...[variable X]

Note: Only personal identifiers are to be provided to the CHeReL.

The CHeReL will not hold information on individuals other than the information needed for the linkage.

The databases deposited with the CHeReL will be placed into a Microsoft SQL database. Records will be linked using the probabilistic record linkage software Choicemaker. From the linkage a linkage key for the project will be created, which contains the following information:

- a Project Person Number for each person identified through the linkage
- a code indicating the source database
- the encrypted record number from the source database.

Two project keys will be created from the linkage key:

1. a [Dataset 1] Key containing the encrypted record number from [Dataset 1] and the Project Person Number for that record. The [Dataset 1] Key will be returned to [dataset 1] custodian.
2. a [Dataset 2] Key containing the encrypted record number from [Dataset 2] and the Project Person Number for that record. The [Dataset 2] Key will be returned to [dataset 2] custodian.

The custodians of [dataset 1] and [dataset 2] will decrypt the encrypted source record numbers from the datasets and attach the relevant information from the dataset that was previously approved to be released to the study investigators.

The data custodians will forward the two datasets to the study investigators, who will match the datasets using the Project Person Numbers.
Example 2: Linkage of datasets held in the CHeReL Master Linkage Key

We seek access to linked data of [dataset 1] and [dataset 2] for the [financial]/[calendar] years [year] to [year] as follows:

- [dataset 1] eg all records for people aged 65 years and over
- [dataset 2] eg all records that link to [dataset 1]

These datasets form part of the Master Linkage Key held by the CHeReL.

The variables that are requested for the project are listed in Section [X].

*Note:* List only those variables to be provided to the study investigators.

Two project keys will be created from the Master Linkage Key:

1. a [Dataset 1] Key containing the encrypted record number from [Dataset 1] and the Project Person Number for that record. The [Dataset 1] Key will be returned to [dataset 1] custodian.

2. a [Dataset 2] Key containing the encrypted record number from [Dataset 2] and the Project Person Number for that record. The [Dataset 2] Key will be returned to [dataset 2] custodian.

The custodians of [dataset 1] and [dataset 2] will decrypt the encrypted source record numbers from the datesets and attach the relevant information from the dataset that was previously approved to be released to the study investigators.

The data custodians will forward the two datasets to the study investigators, who will match the datasets using the Project Person Numbers.
APPENDIX 4: SECURITY AT THE CHeReL

CHeReL databases are held on a data linkage computer server located in Australian Centre for Advanced Computing and Communications (ac3) at the Australian Technology Park (ATP), Eveleigh in Sydney.

The CHeReL is physically located in a suite in Bay 9 at the ATP. There is swipe card access to the suite. Access to desktop computers requires username and password. No data are stored on the desktop PC as all data are stored on the computer servers in the ac3 data centre.

For additional information on security please contact us.
APPENDIX 5: ETHICAL ISSUES RELATING TO RECORD LINKAGE USING POPULATION HEALTH DATABASES

The ethical issues of population-based record linkage projects relate to:

1. the use of personally identifying information, such as name and address, on individuals without their consent;
2. whether the collection, use and disclosure of personal information as described in the project application is lawful; and
3. whether the potential benefits to the community outweigh the risks to privacy.

These are discussed below.

1. the use of personally identifying information, such as name and address, on individuals without their consent

Personally identifying information is required to carry out the data linkage. Obtaining individual consent is often difficult to obtain because of the size of the population involved in the research, but also because some individuals will have changed address or died since the health information was originally collected, particularly for the more historic years of data.

2. Whether the collection, use and disclosure of personal information as described in the project application is lawful

Personal information held by public sector agencies is subject to the NSW Privacy And Personal Information Protection (PPIP) Act 1998, and personal health information is subject to the NSW Health Records and Information Privacy (HRIP) Act 2002.

Note: The PPIP Act may be found at: www.austlii.edu.au/au/legis/nsw/consol_act/papipa1998464/ and the HRIP Act may be found at: www.austlii.edu.au/au/legis/nsw/consol_act/hraipa2002370/

In relation to the PPIP Act, under the Privacy Commissioner’s Direction On Disclosures Of Information By Public Sector Agencies For Research Purposes (www.lawlink.nsw.gov.au/lawlink/privacynsw/l_l_pnsw.nsf/pages/PNSW_03_s41rp2), agencies may disclose personal information for research purposes under certain circumstances, such as where ‘the proposed research has been approved by a committee established for the purpose of giving ethical approval to research projects after such a committee has considered the privacy implications of the collection and subsequent use of such information by the researcher in the absence of express consent.’

In relation to the HRIP Act, projects may seek to meet the criteria for the ‘research exemption’ from Health Privacy Principles 10 (Limits on use of health information), and 11 (Limits on disclosure of health information) as set out in the NSW Privacy Commissioner’s Statutory Guidelines on Research:

- the use or disclosure is reasonably necessary for research, or the compilation or analysis of statistics, in the public interest
- the purpose of the research cannot be served by using or disclosing de-identified information
• it is impracticable to seek the consent of the person to the use or disclosure (it is impracticable to seek the consent of the several million people whose admissions are recorded on the Admitted Patient Data Collection and persons whose information is held on the Registry of Births Deaths and Marriages death registration data are deceased)

• reasonable steps will be taken to de-identify the information. Only de-identified linked datasets will be made available for analysis

• all use or disclosure will be in accordance with the Statutory Guidelines on Research. In essence this requires the project to be approved by a Human Research Ethics Committee that operates according to NHMRC guidelines.

Similarly, projects may also seek to meet the criteria for the ‘management of health services exemption’ from Health Privacy Principles 10 and 11 as set out in the NSW Privacy Commissioner’s Statutory Guidelines on Management of Health Services.

t#4b

The NSW Public Health Act 1991 permits the disclosure of data held by the NSW Central Cancer Registry, the NSW Pap Test Register, the NSW Midwives Data Collection and the NSW Birth Defects Register for epidemiological purposes with the approval of the NSW Department of Health Chief Health Officer.

The NSW Health Administration Act 1982 permits the disclosure of personally identifying health information held in Department of Health data collections under certain circumstances, including where ‘the information is required for the purpose of medical research and the Director-General is satisfied that the research is being conducted in accordance with any guidelines of the National Health and Medical Research Council that the Director-General considers relevant, in particular any guidelines relating to the circumstances where the consent of an individual the subject of research need not be obtained and the protection of individual privacy.’

The birth and death registration data are collected under the Births, Deaths and Marriages Registration Act and may be used for research purposes under the PPIP Act and the HRIP Act as described above.

3. Whether the potential benefits to the community outweigh the risks to privacy.

The benefits of projects using linked population health data include the ability to: describe the total health burden of disease in the population and to assess risk factors and causal pathways of outcomes; assess infrequent outcomes or exposures; determine small effects of common exposures and outcomes; overcome some of the disadvantages of a single cross-sectional data set such as under-ascertainment or misclassification of exposures and outcomes; and capture important longitudinal outcomes including mortality and major morbidities.

The risks associated with record linkage projects relate to the possible breach of privacy through the disclosure of personally identifying information. We have minimised the risk to personal privacy by: ensuring that the smallest number of people have access to identifiable data; ensuring that data are securely stored;
and that access to identified data by officers in the CHeReL is password
protected and subject to signing of a confidentiality agreement. Separation of
the record linkage process (where access to personally identifying information
is necessary), from the analysis of health data, (where such access is not
necessary), helps to preserve personal privacy by limiting the number of people
with access to personally identifying information.