

# THE COMBINED PROTOCOL AND APPLICATION FORM

This document describes what to look out for in each section

## 1. INVESTIGATORS AND PARTICIPATING INSTITUTIONS

### Checks

- All the investigators on my project are listed
- All the sites where data is held are detailed
- I have indicated which researchers have access to CHeReL linked data and how they will access the data

### Tips

- ☞ The linked dataset is usually stored at a research institution on a secure server or the SURE / ERICA
- ☞ If an investigator is accessing the SURE or another researcher institution through VPN, please indicate this in this section.
  - e.g. “ Accessing the SURE from UNSW”, or “Accessing UNSW secure servers via VPN”

## 2. METHODS

### Tips

- ☞ The linked dataset is usually stored at a research institution on a secure server or the SURE / ERICA.
- ☞ If an investigator is accessing the SURE or another researcher institution through VPN, please indicate this in this section.
  - e.g. “ Accessing the SURE from UNSW”, or “Accessing UNSW secure servers via VPN”

## DATA GOVERNANCE

- ☞ Ensure that all linkage and integration processes performed by the CHeReL are described correctly. Please seek advice of CHeReL clients services if you are unsure.
- ☞ Specify all the sites where data will be collected, stored and analysed.
- ☞ If you have a researcher owned external cohort and you are receiving linked data from the CHeReL, you need to clearly describe the separation principal (and name the data custodian), and how you will prevent re-identification of participants.
- ☞ Please detail the type of data storage e.g. secure servers or SURE / ERICA.
- ☞ Ensure you have a clear data flow, showing where different data is travelling. If your project has an AIHW component please also specify the PPN flow. Often a flow chart is helpful to attach to your application. There are some example flow charts on the CHeReL website.

## 3. PROJECT FUNDING/SUPPORT

- ☞ Fill out this section, even if funding has not been confirmed. The data custodians like to be sure that if they approve data release the research has funding to be completed.

## 4. DATA REQUEST SECTION

- ☞ Only fill out section 13 if you are requesting UNLINKED unit record data. If you are requesting linked data please complete section 14.

### (SECTION B) EXTRACT FROM CHEREL MLK COLLECTIONS

- ☞ Please ensure the date ranges you are requesting are available.
- ☞ If you would like the most recent data available at the time we do your linkage you can write “latest available” in the ‘To’ column.
- ☞ Ensure that all datasets you are requesting have been mentioned in the study design and analysis, and helps to meet the objectives of your research.
- ☞ For the APDC data, you need to request a time period based on either admission date or separation date. If you select admission date, you will receive people who have been admitted to hospital between the two dates (but could have been separated after the dates requested). You will not receive information on people who have been admitted but not separated. If you select separation date, you will receive patients who have been separated between the two dates, but they may have been admitted at a time prior.

### COHORT AND RESTRICTIONS

#### Checks

- My cohort is clearly defined – this includes which data sets make up the cohort, the year range of the data, the approximate amount of people in the cohort, and ICD code list if you cohort is based on ICD codes

#### Tips

- ☞ If you are unsure of the exact number of people in your cohort, an estimate is fine. You might be able to search available literature, or extrapolate from a smaller/large study.
- ☞ If you only want records for certain condition returned you will need to provide a code list. E.g. if you had a cohort of people who had a heart attack you might only want hospital records with related ICD codes, or you might all their hospital records, regardless of what they are admitted for.

### OTHER COLLECTIONS (NOT HELD IN THE CHEREL MLK)

#### Checks

- I have included all data sets that I am linking to if they are not in the CHeReL MLK
- I have described the identifiers available for linkage
- I have listed the custodian details

#### Tips

- ☞ If the CHeReL has not linked the dataset you are requesting before it is important to fill in all the data custodian details accurately.
- ☞ The identifiers we generally use for linkage are Full name, DOB, Address (+ date associated with the address) and Sex. We can also use other identifiers like MRN (in conjunction with

hospital name/code) if they are available. The more identifiers that are available for linkage the better you can expect the linkage rate and quality to be.

## Supporting Documents

### 5. VARIABLE LISTS

#### Checks

- I have completed a variable list for every dataset getting linked including external datasets
- My variable list only contains variables I want to receive back for analysis (not identifiers used for linkage)
- I have consulted the data dictionary for information on the definitions and limitations of the variables.
- All requested variables have a justification

#### Tips

-  The CHeReL has a blank variable list on the website, for any researcher owned data sets or other data collections that are not in the MLK.
-  Postcode and DOB variables often need a strong justification. Indigenous status can be requested for use as a confounder but if you want to report the results by indigenous status it will require approval from AHMRC Ethics Committee.

### 6. HREA (NOW GENERATED VIA REGIS)

#### Checks

- My HREA aligns with my protocol across all sections (i.e. sites, personnel, datasets, aims and objectives).
- The details I give about the CHeReL linkage process is accurate and consistent with the protocol.