# Costs and Timeframes

Known factors that impact project approval timeframes

DATA CONNECT AUSTRALIA

| Data quality | Shorter | Medium | Longer |
|---|---|---|---|
| Quality of person identifiers in data collections being linked | Good quality data with no missing or misspelt details – little or no manual clerical review required | Some errors or missing data – a moderate number of links need manual review<br><br>Some intervention required to import data into data linkage unit systems | Missing identifiers and/or many errors in identifiers - large amount of manual clerical review required |
| Format of data provided for linkage | All data items able to be easily read into the linkage unit data systems with no issues with format or misspelt delimiters | | Data not readily able to be imported into data linkage unit systems with manual input required to correct formats and ensure all data items are read in correctly |
| **Size of request and amount of linkage required** | | | |
| Number of records that require linkage for a projec | No linkage required outside of core data collections (all part of MLK) | Some custom linkage required, and datasets are moderately sized (hundreds of records)<br><br>Longer, but recent time period (eg last 5 years) | Thousands or millions of records required to be linked |
| Time period requested | Short, recent time period (eg 1-2 years) | | Long time period, especially if multiple versions of ICD codes are included and including years of data going back in time that may be stored in different systems. Very recent data that has not yet been submitted by those collecting the information or validated |
| **Complexity of request** | | | |
| Number of jurisdictions involved | Single state or territory | Multi-jurisdictional linkage of state data collections | Cross jurisdiction linkage or inclusion of Commonwealth data (e.g. MBS, PBS) |
| Number of data collections involved | Two datasets or single custodian responsible for approving and extracting all data collections requested | | Multiple data collections and custodians need to approve data access and extract data for linkage. A delay with the provision of a specific data set for linkage may lead to a delay for the whole project. |
| Types of datasets requested | Linkage of core data collections | Linkage of core data collections with frequently linked data collections | Project specific data collections need to be linked for the projectSome data collections such as Commonwealth data collections may have particular governance and security requirements Some data collections require external data custodian approval |
| Study design | Simple cohort specifications; no control group | | Control group or complex cohort specifications, for example involving characteristics from across multiple data collections or requiring large amounts of coding to describe inclusion and exclusion criteria (eg using a list of hundreds of ICD codes for selection of hospital records; select suitable comparison records; or to generate data items in the required format) |
| **Capacity issues** | | | |
| Data linkage team capacity and competing priorities | Few other requests and experienced staff available | Some experience but limited capacity | Inexperienced team, staff on leave, many other requests to process, urgent departmental requests |
| Data custodian capacity and experience with providing data extracts | Custodian has staff experienced in extracting and formatting data | Some experience but limited capacity | Data collected for operational purposes and limited resources allocated for extracting and formatting data |

**Clarity of request**

| | | | |
|---|---|---|---|
| Descriptions of requirements provided | Variable names clearly specified using the terminology used by the data custodian, research questions and proposed analysis and use of the data clearly described and/or required format of data items clearly described | | List of data items described only in generic terms and inadequate justification provided for the data requested or detail for how it will be used |
| ICD and ACHI (diagnosis and procedure codes) provided | All provided, including accounting for changes in edition and version for all time periods requested | Some ICD/ACHI codes provided but no consideration of potential changes over time | No ICD/ACHI codes provided - just broad descriptions of conditions/procedures of interest |