

Appendix B

MODEL FOR CROSS-JURISDICTIONAL DATA LINKAGE FOR MEDICAL RESEARCH

The approach for cross-jurisdictional data linkage was developed in 2000 by staff at Department of Health and Ageing, the Health Insurance Commission, the Health Department of Western Australia, the Department of Public Health at the University of Western Australia and the Australian Institute of Health and Welfare in response to a request for data for an approved public health research project. It may be regarded as a model of 'best practice' in the utilisation of administrative data for the production of de-identified linked data files for specific approved purposes.

PROPOSED APPROACH

The process involves two separate stages. The first stage is a memorandum of understanding to share data for an agreed purpose. The second stage includes the production of linked, de-identified data files for approved projects. Each project is covered by its own agreement and the various data custodians supply the data for the project directly to the researchers. For each research project, a unique set of project identifiers is generated by the custodian of the linkage keys and provides the only way of combining the data files into a single linked de-identified file. These project identifiers, being unique to each project, cannot be later used to link additional data from subsequent projects.

This two-stage process ensures that data custodians have full control over the distribution and usage of their data, as each project must be well defined and then individually approved before proceeding. No research can be undertaken without the written approval of every data custodian supplying data to the project. Linked data files are provided only to the individually identified researchers doing the analysis for each project, and must be destroyed when the analyses are complete.

LINKAGE KEY FILE

The linkage key file is produced by a small technical team specialising in data matching, preferably including personnel from all participating institutions. All people involved in the actual linkage and therefore requiring access to the data used in the linkage process must sign confidentiality agreements and be named on a list provided to the steering committee. Any changes to this list will be reported in writing to this committee. No other personnel are allowed access to the files used in this process, as they contain private and confidential information. The work is done on an isolated computer, with all personal demographic data destroyed as soon as the linkage is complete. Transfer of these data files is only done via tape, diskette or CD-ROM personally carried by those personnel taking part in the data matching. The linkage personnel are not permitted to take any part in the analysis of the linked data, or to have any communication about these data with the researchers.

LINKED DE-IDENTIFIED DATA

The linkage key file contains no actual data but does provide coded keys to the data sets involved. Every custodian supplies the approved records from their databases, together with the project identifiers, directly to the nominated researchers for that project. These researchers are also required to sign confidentiality agreements. They can link the data together using the project identifiers, and are the only people granted access to the de-identified linked information. They are specifically forbidden to disseminate copies of the data files, and are required to destroy these files on completion of the analysis.

ETHICS APPROVAL

Ethics approvals from the researchers' institution as well as signed approval from the CEOs of each of the participating institutions are mandatory.