

REFERENCES

Anderson P 2000. CSDA linkage key trial, in *Disability support services provided under the Commonwealth/State Disability Agreement: national data*. Canberra:AIHW.

Armstrong B & Krickler A 1999. Record linkage — a vision renewed. *Australian and New Zealand Journal of Public Health*, 23, 5, 451–452.

Australian Bureau of Statistics (ABS) 2001. *Community services catalogue 1999–2000*, cat. no. 8696.0. Canberra:ABS.

Australian Institute of Health and Welfare (AIHW) 1999a. *SAAP National Data Collection annual report 1997–98 Australia*, SAAP NDCA Report Series 3, AIHW cat. no. HOU 24 Canberra:AIHW.

____ 1999b. *National Community Services Information Development Plan*, Standing Committee of Community Services and Income Security Administrators, AIHW cat. no. AUS 14. Canberra:AIHW.

____ 2000a. *SAAP National Data Collection annual report 1998–99 Australia*, SAAP NDCA Report Series 4, AIHW cat. no. HOU 38. Canberra:AIHW.

____ 2000b. *The use of linkage key for statistical work in community services: background paper for the Statistical Linkage Project of the National Community Services Information Management Group*. Unpublished paper, AIHW.

Brameld K, Holman CDJ, Thomas M & Bass J 1999. Use of state data bank to measure incidence and prevalence of a chronic disease: end stage renal failure. *American Journal of Kidney Diseases*, 34, 6, 1033–1039.

Department of Health and Aging (DoHA) 1999. *Health information and data protection issues*. Unpublished paper, DoHA.

Holman CDJ, Bass J, Rouse I & Hobbs M 1999. Population-based linkage of health records in Western Australia: development of a health services research linked database. *Australian and New Zealand Journal of Public Health*, 23, 5, 453–459.

Holman CDJ, Wisniewski Z, Semmens J, Rouse I & Bass J 1999. Mortality and prostate cancer risk in 19,598 men after surgery for benign prostatic hyperplasia. *British Journal of Urology*, 84, 37–42.

Holman CDJ, Wisniewski Z, Semmens J, Rouse I & Bass J 2000. Population-based outcomes after 28,246 in-hospital vasectomies and 1,902 vasovasostomies in Western Australia. *British Journal of Urology*, 86, 1043–1049.

Karmel R 2000. *Duplicates in the SAAP linkage key*. Unpublished report, AIHW.

Kelman CW, Bass AJ & Holman CDJ 2001. A ‘best practice’ for creation of linked health data drawn from two or more organisations. Under submission to the *Australian and New Zealand Journal of Public Health* (2001).

Marshall R 2001. *Business rules for the use of unique patient identifiers in statistical collections*. Draft paper prepared for the National Health Information Management Group (NHIMG).

National Health Information Management Advisory Council (NHIMAC) 2000. *A national health information standards plan for Australia: setting the standards*. Canberra:AGPS/NHIMAC.

O’Connor K 1999. *Protection of personal information by law in Australia: 11 years after the passage of the federal Privacy Act*. Paper to the Records Management Association Conference, Darwin.

Office of the Federal Privacy Commissioner (OFPC) 1994. *Plain English guidelines to information privacy principles 1–3: advice to agencies about collecting personal information*. Canberra:OFPC.

____ 1998. Guidelines for the use of data matching in Commonwealth administration. Canberra:OFPC.

____ 2001a. Guidelines approved under section 95A of the Privacy Act 1998: December 2001. Canberra, OFPC.

____ 2001b. Guidelines on privacy in the Private Health Sector: November 2001. Canberra, OFPC.

____ 2001c. Guidelines to the National Privacy Principles : September 2001. Canberra, OFPC.

Ryan T, Holmes B & Gibson D 1999. *A national minimum data set for Home and Community Care*. Canberra:AIHW.

United States General Accounting Office 2001. *Record linkage and privacy: issues in creating new federal research and statistical information*, cat. no. GAO-01-126SP.

APPENDIXES

APPENDIX A

Statistical linkage key Working Group membership

The following individuals were members of the Statistical Linkage Key Working Group (SLKWG):

Mr Andrew Stuart (Chair)	Commonwealth Department of Health and Ageing
Dr Ching Choi	Australian Institute of Health and Welfare
Mr James Jordan	Commonwealth Department of Family and Community Services
Mr Leon Pietsch	Australian Bureau of Statistics
Mr Paul Basso	Department of Human Services (South Australia)
Mr Allan Dernee	Department of Ageing, Disability, and Home Care (New South Wales)

The following individuals also contributed to the work of the SLKWG:

Mr D'Arcy Jackson	Commonwealth Department of Health and Ageing
Mr Mark Thomann	Commonwealth Department of Health and Ageing
Ms Trish Ryan	Australian Institute of Health and Welfare
Ms Margaret Fisher	Australian Institute of Health and Welfare
Mr Geoff Neideck	Commonwealth Department of Family and Community Services
Mr James Kemp	Commonwealth Department of Family and Community Services
Mr John Fulop	Commonwealth Department of Family and Community Services
Mr Proshanta Dey	Department of Ageing, Disability, and Home Care (New South Wales)

In addition, the Commonwealth Department of Health and Ageing (on behalf of the SLKWG) contracted Dr John Bass to provide expert advice and analysis on data linkage issues, especially in relation to the measure of the effectiveness of statistical linkage keys.

APPENDIX B

Measure of the effectiveness of statistical linkage keys

Previous measures of the effectiveness of SLKs in use in the community services sector have tended to focus on two measures of completeness and accuracy (for example, AIHW 2000b).

The first of these measures concerns the availability of data for the construction of an SLK. Client refusals to allow details to be used for linkage purposes, as well as incomplete/missing data items attached to a client's record, reduce the number of links that can be made. The missing data may be biased compared to the overall client population. Some demographic groups may have an increased aversion to allowing the use of their data, and the quality of data may also vary according to socioeconomic or demographic factors. The proportion of clients for which data are unavailable and the extent of selection bias amongst those clients are measures of the representativeness of an SLK. Most measures of effectiveness have examined the proportions of clients for which data are unavailable, with little information on whether these clients are representative of the whole population.

The second measure relates to the proportion of incorrect linkage keys being generated from the source data. These errors fall into two main types:

1. *errors in the source information leading to the generation of multiple keys for one individual, such as, when a surname is misspelt ('Smith' / 'Smythe') or when there is a name change (as often occurs at marriage or divorce); and*
2. *multiple clients sharing similar identifying information leading to the construction of a single linkage key.*

Errors of the second type will be more prevalent in linkage keys containing less information (that is, they are more likely with the SAAP key than the HACC key). As a measure of the effectiveness of linkage keys, these two errors are often added together as an overall 'mismatch' or 'duplication' rate.

Existing effectiveness measures of HACC, SAAP and CSDA SLKs

The quality of the HACC linkage key has been tested in terms of duplication rates using three sets of data: the Commonwealth Aged Care database, Silver Chain (a large HACC service provider in Western Australia) and the National Death Index. The testing found a key duplicate rate of between 0.6% and 1% against these collections, which was considered to be acceptable for statistical research purposes (Ryan, Holmes & Gibson 1999).

Two SAAP collections made in 1998–1999 and 1999–2000 reported 25% and 21% client refusals with a further 3.5% and 2.5% missing due to insufficient data. Estimates of duplication rates ranged from 3.3% to 5% (AIHW 2000). These estimates were within a level of accuracy acceptable to the SAAP Data and Research Advisory Committee.

A further test of the SAAP mismatch rate has been conducted by the AIHW (Karmel 2000). This involved testing the SAAP linkage key against a model based on synthetic populations of unique individuals that approximate the year of birth distribution of the SAAP population. These synthetic populations were constructed using data from the National Death Index. The mismatch (duplicate) rate was estimated to be about 3.3% over all year of birth groups. The mismatch rate also increased with the number of people within a particular year of birth, and was higher among younger SAAP clients than older clients. The test also shows that the mismatch rate is expected to be higher if data for more than one year are linked.

The CSDA-linked records from 1999 showed that about 3% of records were of insufficient quality to construct a linkage key. This was an improvement from the levels of invalid data in the 1998 test, which ranged from between 3.7% to 6% (Anderson 2000).

For the Reconnect program (using the HACC linkage key) a consent rate of 80% was achieved as at November 2000. Approximately 3% of the client group provided insufficient information for the construction of a linkage key. There has not been a test on mismatches using the linkage key, so information is not currently available on this aspect of the quality of the linked data. It is hoped that this linkage key will be used in the future to give an indication of multiple use and repeated use of services within the Reconnect program and perhaps to link to the SAAP data collection. Consultation with the community services sector will be undertaken if such linkage is to occur. Plans for this work have not yet been developed.

The results of these broad measurements of the completeness and accuracy of SLK methodology have generally been taken to indicate that these keys are adequate for statistical research purposes.

Current measures of the effectiveness of SLKs

As outlined above, existing measures of SLKs have usually focused on how well the linkage key represents the source population and on the extent of duplication that is, multiple keys for one individual as well as multiple individuals sharing the same key. It is a far more difficult task to ascertain whether the analysis of data linked by deterministic matching of SLKs leads to significantly different conclusions than would be obtained through analysis of 'real' linked data.

Dr John Bass is currently investigating this problem in collaboration with Professor D'Arcy Holman and members of the Data Linkage Unit in Perth (a collaborative project between the Health Department of Western Australia and the Department of Public Health at the University of Western Australia). Some preliminary results from the study have been made available for this paper.

A data set has been constructed containing seven years of hospital and death records (1993–1999) of individuals older than 19 years from Western Australia (2,844,030 hospital unit records). HACC and SAAP SLKs were created for all of these records, and deterministic linkages based on these keys were performed to link records within the hospital data as well as to a copy of the death register to which the HACC and SAAP SLKs had been added. The data also contain the project identifier (WA PID) created by the Data Linkage Unit, based on probabilistic linkage of full demographic data (all names, sex, date of birth, address, country of birth and Indigenous status). This WA PID has been improved by linkage to other data sets such as the State electoral roll that provides historical information on name and address changes. Significant effort has

also been put into validation of the links (Holman et al. 1999).

While not perfect, the WA PID and the associated demographic data are an excellent standard for assessing the comparative effect of the SLKs. Apart from the extensive resources that have gone into linking the Western Australia information, the data sets involved include the typical problems found in administrative data. The demographic information for an individual is often inconsistent, with varied dates of birth, names, addresses, race and (surprisingly) sex.

The files being used for analysis outside the Data Linkage Unit have had all identifying variables (including the SLKs) encrypted to ensure full protection of privacy. The files were obtained by a standard application to the Data Linkage Unit for de-identified linked data, a process which includes obtaining the signatures of the custodians of all data sets involved as well as that of the General Manager of the Health Information Centre at the Health Department of Western Australia.

The primary aim of the study is to compare the results of typical analyses of linked data from the same set of hospital and death records linked by means of the HACC and SAAP SLKs as well as the WA PID. The effects of increasing the time period over which data are collected, Indigenous status (a group where linkage is usually difficult and liable to an increased error rate) and sample size are all being examined.

Duplication rates for HACC and SAAP SLKs

Duplication rates of the HACC and SAAP keys in the Western Australian study are summarised in Table 2. For each key, the first row shows the percentage frequency of multiple HACC keys for one individual (that is, one WA PID) while the second row shows the percentage frequency of more than one individual sharing one HACC key. The third row shows the ratio of these two percentages while the fourth row shows their sum.

Table 2: Duplication rates of HACC and SAAP keys compared to WA PID

Duplication rate (%)	1 year 1993	2 years 1993–1994	3 years 1993–1995	5 years 1993–1997	7 years 1993–1999
HACC keys/ WA PID	2.1	3.3	4.3	5.7	6.7
WA PIDs/ HACC key	0.02	0.04	0.06	0.10	0.17
Ratio	105	83	72	57	39
Total	2.1	3.3	4.4	5.8	6.9
SAAP keys/ WA PID	1.4	2.2	3.0	4.1	4.9
WA PIDs/ SAAP key	4.6	7.6	9.8	13.0	15.4
Ratio	0.3	0.3	0.3	0.3	0.3
Total	6.0	9.8	12.8	17.1	20.3
Approximate number of WA PIDs	205,000	350,000	470,000	650,000	785,000

Table 2 shows that the rate of multiple HACC keys per individual PID increases steadily from 2.1 to 6.7% over periods of one to seven years. The rate of multiple WA PIDs per HACC key is very low, ranging from 0.02 to 0.17%. The ratio of the duplication types provides a measure of the prevalence of type 1 errors (multiple keys per individual) to type 2 errors (multiple individuals per key). For the HACC key this ratio ranges from 105 over one year to 39 over seven years.

The SAAP key displays a markedly different picture, with the ratio of the duplication types constant at 0.3. The rate of multiple SAAP keys per individual ranges from 1.4 to 5% (slightly lower than that for the HACC key), while the rate of multiple individual WA PIDs per SAAP key is much higher, ranging from 5 to 15%. This is to be expected because the SAAP key contains less information than the HACC key, increasing the chances of more than one individual having the same key.

These results show that the HACC and SAAP keys both produce inaccurate linkages compared to that resulting from the WA PID. The pattern and extent of these biases is different in the HACC and the SAAP keys, and the question arises as to whether analyses of different data sets linked by these two keys might produce different results.

Comparisons of analyses based on data linked on HACC and SAAP keys

Initial expectations of the group undertaking the Western Australian study were that analyses of data linked by SLKs would not vary greatly in terms of accuracy, but that they would be less precise (that is, have greater variance). If this turned out to be true, then data linked by SLKs would be expected to produce valid results with the finer details sometimes obscured by broader confidence limits. In statistical terms, it was expected that average values would not differ significantly but that there would be a significantly larger variance.

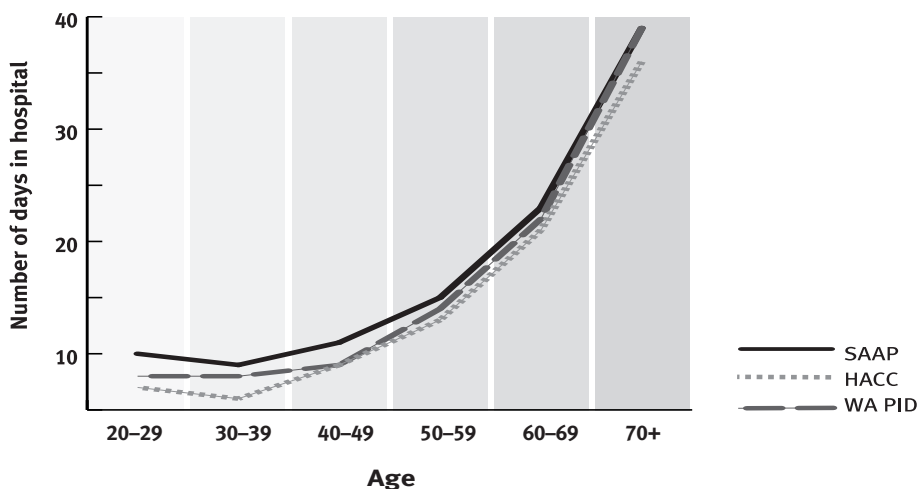
Results from the two analyses completed at the current time are presented here. The first, making use only of hospital data, looks at the total number of days in hospital per patient, a statistic commonly used in economic analyses of health and community services data. The second analysis, making use of death data as well as hospital data, looks at relative risk of death within the cohort of hospital patients.

Number of days in hospital

Figure 3 is a graph showing the number of days in hospital per patient by age group according to data linked by the HACC and SAAP keys and the WA PID.

It is quite clear that data linked with the HACC key under-estimate the number of days in hospital relative to the WA PID data. Data linked with the SAAP key consistently over-estimate the number of days in hospital, except for the oldest age group where the SAAP and WA PID data are virtually identical. These differences are significant at the 95% confidence level (in most cases, at the 99% confidence level) except for the SAAP/WA PID data in the oldest age group. In the age groups under 60 years of age the HACC results are closer to the WA PID data than are the SAAP results, but this is reversed in people of 60 years and older.

Figure 3: Number of days in hospital by age group according to data linked by HACC and SAAP keys and the WA PID



The number of unique HACC keys in these hospital data is higher than the number of unique SAAP keys. It follows that the average number of days in hospital per ‘individual’ will be lower in data linked by the HACC key than in data linked by the SAAP key.

These differences may be significant, but are they large enough to make an impact in practical applications? Table 3 shows the average number of days in hospital by age group (together with the 95% confidence limits) for the WA PID, HACC and SAAP linkages.

Table 3: Average number of days in hospital by age group (with 95% confidence limits)

Age group	WA PID		HACC		SAAP	
20-29	8.8	(8.7 – 8.9)	8.2	(8.1 – 8.2)	9.9	(9.8 – 10.0)
30-39	8.3	(8.2 – 8.4)	7.9	(7.8 – 8.0)	9.5	(9.3 – 9.6)
40-49	9.3	(9.2 – 9.5)	8.9	(8.8 – 9.0)	10.5	(10.3 – 10.6)
50-59	13.2	(13.0 – 13.4)	12.5	(12.3 – 12.7)	14.2	(14.0 – 14.5)
60-69	21.7	(21.4 – 22.0)	20.2	(19.9 – 20.5)	23.0	(22.6 – 23.3)
70+	39.0	(38.5 – 39.4)	35.8	(35.4 – 36.2)	39.8	(39.4 – 40.2)
All ages	14.6	(14.5 – 14.6)	13.7	(13.6 – 13.8)	16.0	(15.9 – 16.1)

Table 3 shows that, in the example of the 70+ years age group, the average number of days in hospital per patient according to the WA PID is 39.0 compared with 35.8 days for data linked by the HACC key and 39.8 days for data linked by the SAAP key. The 95% confidence limits for the average number of days according to the WA PID range from 38.5 to 39.4. This means that we can be 95% certain that the true value of the average (estimated at 39.0) occurs in this range.

Table 4 displays the percentage difference between the average number of days according to the WA PID and HACC keys, the WA PID and SAAP keys, and the HACC and SAAP keys, also indicating which comparisons are significantly different at the 95% confidence level.

Table 4: Percentage difference of days in hospital by age group

Age group	WA PID > HACC	WA PID > SAAP	HACC > SAAP
20–29	-6.9 *	13.5 *	21.8 *
30–39	-5.3 *	13.7 *	20.1 *
40–49	-4.7 *	11.9 *	17.4 *
50–59	-5.6 *	7.9 *	14.3 *
60–69	-6.8 *	5.8 *	13.6 *
70+	-8.2 *	2.1	11.2 *
All ages	-6.1 *	10.1 *	17.3 *

(* = 95% significant)

The HACC data average 6% less than the WA PID data with no consistent pattern except for a small rise in the oldest age group. The SAAP data are on average 10% greater than the WA PID data, with a clear pattern of larger differences in the younger age groups (over 13%) falling to 2% in the oldest age group. The only comparison not significant at the 95% confidence level was that between the WA PID and the SAAP data in the oldest age group.

Initial expectations that the different linkage keys would not produce significantly different results in terms of accuracy were clearly wrong.

What about the expectation that precision would be decreased in data linked by the SLKs? Table 5 shows the standard errors of the average values in Table 3.

Table 5: Standard errors of average values in Table 3

Age group	WA PID	HACC	SAAP
20–29	0.05	0.05	0.06
30–39	0.06	0.06	0.07
40–49	0.08	0.07	0.09
50–59	0.11	0.10	0.12
60–69	0.15	0.14	0.16
70+	0.22	0.20	0.22
All ages	0.04	0.04	0.05

The standard errors of the average values do not vary greatly or in a consistent pattern. The HACC averages are generally slightly more precise than the WA PID averages, with the SAAP linkage showing a slightly larger variance. If the data in Table 5 are normalised to remove the effect of differences in the average values, then the WA PID and HACC standard errors are virtually identical with the SAAP data displaying a consistent small (and not significant) increase.

The initial expectations were therefore wrong on both counts — this analysis shows significant differences between the three different linkages in the average values (that is, variation in accuracy) with virtually constant standard errors (that is, consistent precision) in these values. Analyses of three de-identified linked data sets based on the HACC or SAAP keys or the WA PID led to significantly different results in each case.

Indigenous status

Linkage of data from persons of Indigenous Australian descent is often more difficult compared to linkage of other cultural groups, with frequent name changes and relatively poor recording of dates of birth and other demographic details. Tables 6 through 8 show the results of an analysis of the number of days in hospital per patient by Indigenous status rather than by age group.

Table 6: Average number of days in hospital by Indigenous status (with 95% confidence limits)

Indigenous status	WA PID		HACC		SAAP	
Not Indigenous	14.2	(14.1 – 14.3)	13.4	(13.3 – 13.5)	15.7	(15.6 – 15.8)
Indigenous	27.7	(26.7 – 28.7)	22.1	(21.3 – 22.8)	26.3	(25.5 – 27.0)
Total	14.6	(14.5 – 14.6)	13.7	(13.6 – 13.8)	16.0	(15.9 – 16.1)

The results in Table 6 show that the estimates of number of days in hospital per Indigenous patient covered a wide range from just under 28 (WA PID) through about 26 (SAAP) to just over 22 (HACC). The significance and extent of these differences are summarised in Table 7.

Table 7: Percentage difference of days in hospital by Indigenous status

Indigenous status	WA PID > HACC	WA PID > SAAP	HACC > SAAP
Non-Indigenous	-5.7 *	10.6 *	17.2 *
Indigenous	-20.4 *	-5.3	18.9 *
Total	-6.1 *	10.1 *	17.3 *

(* = 95% significant)

Table 8: Standard errors of average values in Table 5

Indigenous status	WA PID	HACC	SAAP
Non-Indigenous	0.05	0.05	0.06
Indigenous	0.06	0.06	0.07
Total	0.08	0.07	0.09

Tables 7 and 8 show a similar pattern in the analysis for Indigenous status as that shown by the analysis for age groups, with significant differences between the average values (except for the WA PID/SAAP figures for Indigenous patients) and virtually constant precision.

The extent of the differences in average values is sufficient to raise serious concerns about the validity of some of these linkages. For instance, the estimate of the number of days in hospital for Indigenous patients is 20% lower for the HACC linkage than for the WA PID linkage.

Relative risk of death

The quality of the death data linkages was investigated by performing a Cox regression for the WA PID, HACC and SAAP linked data sets to show the relative risk of death by age group, sex and Indigenous status. Details of this analysis are provided in Figure 4 and Tables 9, 10 and 11.

As far as age groups are concerned, the HACC and SAAP keys display consistently lower estimates of the relative risk of death compared to the WA PID linkage. Differences between the HACC and WA PID linkages are less than 5% except for the 70+ age group where the HACC linkage has a difference of just over 9%. The variances of the relative risk estimates for the different age groups are relatively high and the differences are not significant except for the SAAP and WA PID linkages in the two oldest age groups (60–69 and 70+ years).

Estimates of the relative risk of death for males are remarkably similar for all three linkages, and there are certainly no significant differences.

Figure 4: Relative risk of death by age group for data linked by HACC and SAAP keys and the WA PID

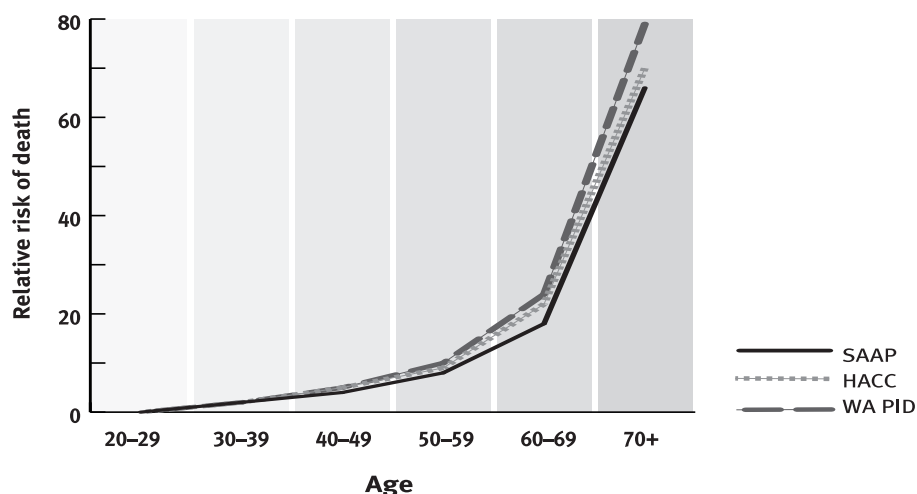


Table 9: Relative risk of death by age group compared to 20–29 year olds; males compared to females; and Indigenous patients compared to non-Indigenous patients (with 95% confidence limits)

Age group	WA PID		HACC		SAAP	
30–39	1.7	(1.6 – 1.8)	1.7	(1.6 – 1.8)	1.6	(1.5 – 1.7)
40–49	3.8	(3.5 – 4.1)	3.8	(3.5 – 4.1)	3.4	(3.2 – 3.7)
50–59	9.2	(8.6 – 9.8)	8.9	(8.3 – 9.5)	8.1	(7.5 – 8.7)
60–69	23.2	(21.8 – 24.7)	22.2	(20.8 – 23.7)	20.1	(18.8 – 21.5)
70+	75.7	(71.2 – 80.4)	68.8	(64.5 – 73.3)	65.5	(61.4 – 69.9)
Sex						
Male	1.5	(1.5 – 1.5)	1.5	(1.5 – 1.5)	1.5	(1.5 – 1.5)
Indigenous status						
Indigenous	2.30	(2.2 – 2.4)	1.2	(1.1 – 1.3)	1.5	(1.4 – 1.5)

Table 10: Percentage difference of relative risk of death

Age group	WA PID > HACC	WA PID > SAAP	HACC > SAAP
30–39	0.0	-7.1	-7.1
40–49	-0.8	-10.5	-8.8
50–59	-3.1	-12.0	-9.2
60–69	-4.4	-13.2 *	-9.2
70+	-9.1	-13.4 *	-4.7
Sex			
Male	0.7	0.7	0.0
Indigenous status			
Indigenous	-47.4 *	-36.5 *	20.7 *

(* = 95% significant)

For patients of Indigenous descent the figures are markedly different, ranging from 2.3 for the WA PID linkage through 1.5 for the SAAP key to 1.2 for the HACC key. These relative risk estimates are all significantly different from each other. This is emphasised when one considers that, according to the HACC key, Indigenous patients are 20% more likely to die than non-Indigenous patients but, according to the WA PID, this figure is increased to 130%.

Table 11 shows a marked increase in standard errors with increase in age group. This reflects the sharp increase in risk of death among older patients. For the WA PID linkage, the relative risk of death increases by a factor of 45 from the 30–39 age group to the 70+ age group, while the standard error increases by a factor of 35. Taking the increase in risk into account, there is therefore only a small increase in variance among the values for the older patients.

Table 11: Standard errors of average values in Table 8

Age group	WA PID	HACC	SAAP
30–39	0.1	0.1	0.1
40–49	0.1	0.1	0.1
50–59	0.3	0.3	0.3
60–69	0.7	0.7	0.7
70+	2.3	2.3	2.3
Sex			
Males	0.01	0.01	0.01
Indigenous status			
Indigenous	0.06	0.04	0.04

Conclusions

These results illustrate the need to consider the effects of using different linkage methods before undertaking any planning or research projects dependent on de-identified linked data. While the measures of effectiveness relating to duplication rates could easily lead to the conclusion that the HACC key provides a better linkage variable than the SAAP key, an analysis of bed use in elderly patients might well be more accurate using data linked with the SAAP key.

Variation in data quality between different demographic groups may result in marked differences after linkage by different methods. The estimation of the relative risk of death in Indigenous compared to non-Indigenous patients is 20% greater in data linked by the HACC key, compared to 50% greater for the SAAP key and 130% greater for the WA PID data.

Comparisons of analyses on data linked by different SLKs may be particularly doubtful if the two SLKs are affecting the analyses in opposite directions. For instance, Table 3 shows that, for all patients, the HACC key produces an estimate of average days in hospital that is 6% less than that produced by the WA PID. By contrast, the SAAP key produces an estimate that is 10% greater than that produced by the WA PID. If the corresponding estimates produced by the HACC and SAAP keys are compared, that of the SAAP data is 17% greater compared to the HACC data. Comparisons between two linked data sets based on different SLKs should be regarded with extra caution.

Decisions as to whether a particular linkage method is sufficiently accurate and precise need to be made separately for every distinct analysis. It is clear that some linkage/analysis combinations lead to results that are, at the very least, of dubious quality.

The causes of these marked differences are still being investigated. What these results do show is that the use of different linkage methods can lead to significantly varied (and unexpected) results. If SLKs are to be used for linkage, then the quality of that linkage in respect of any analysis should be routinely and thoroughly investigated. Ideally, linkage should be performed using probabilistic methods with as much demographic data as possible.

APPENDIX C

Western Australian Diabetes Linkage Project protocol

Model for cross-jurisdictional data linkage

Proposed approach

The process involves two separate stages. The first stage is a memorandum of understanding between participating agencies to share data for an agreed purpose and to prepare a linkage key file (using probabilistic methods). The second stage involves the production of linked, de-identified data files for an undefined number of separate (approved) research projects. Each project will be covered by its own agreement, the data for the project being supplied directly to the researchers by the various data custodians.

For each research project, a unique set of project IDs will be generated by the custodian of the linkage keys and will provide the only way of combining the data files into a single linked de-identified file.

This two-stage process will ensure that data custodians have full control over the distribution and usage of their data, as each project will need to be well defined and then individually approved before proceeding. No research will be undertaken without the written approval of every data custodian supplying data to the project.

Linked data files will be provided only to the individually identified researchers doing the analysis for each project, and will be destroyed when the analyses are complete. A different ID will be used in each project, thus making it extremely difficult to merge the linked data for two projects (such a process is, in any case, specifically prohibited).

Linkage key file

The linkage key file will be produced by a small technical team specialising in data matching, including personnel from several of the participating institutions. All people involved in the actual linkage and therefore requiring access to the data used in the linkage process will 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 will be allowed access to the files used in this process, as they will contain private and confidential information. The work will be done on an isolated computer, and all personal demographic data will be destroyed as soon as the linkage is complete. Transfer of these data files will be done only via tape, diskette or CD-ROM personally carried by those personnel taking part in the data matching. The linkage personnel will not be 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 will contain no actual data but will provide coded keys to the data sets involved. Every custodian will supply the approved records from their databases, together with a project ID number, directly to the nominated researchers for that project. These researchers will also sign confidentiality agreements. They will link the data together using the project ID, and will be the only people granted access to the de-identified linked information. They will be specifically forbidden to disseminate copies of the data files, and will be required to destroy these files on completion of the analysis.

Ethics approvals

Ethics approvals from the researcher's institution as well as the confidentiality or ethics committees of each of the participating institutions are mandatory.

APPENDIX D

Related legislation on health and privacy

Commonwealth health-related legislation

Australian Institute of Health and Welfare Act 1987

National Health Act 1953

Medicare Levy Act 1986

National Health and Medical Research Council Act 1992

Commonwealth and State privacy legislation

Commonwealth

Privacy Act 1988

Privacy Amendment (Private Sector) Act 2000

Data Matching Program Assistance and Tax Act 1990

State and Territory

Health Records (Privacy and Access) Act 1997 (ACT)

Health Rights Commission Act 1991 (ACT)

Freedom of Information Act 1992 (Qld)

Privacy and Personal Information Protection Act 1998 (NSW)

Health Administration Regulation 2000 (NSW)

Freedom of Information Act 1999 (Vic)

Data Protection Bill 1999 (Vic)

Information Privacy Act 2000 (Vic)

Health Records Act 2000 (Vic)

Health Commission Act 1976 (SA)

Housing Trust Act 1995 (SA)

Community Housing Authority Act 1991 (SA)

Agency-specific legislation

Social Security Administration Act 1999

— FaCS

Child Care Act 1972

— FaCS

Disability Services Act 1986

— FaCS

Home and Community Care Act 1985

— DoHA

Aged Care Act 1997

— DoHA

Public Sector Management Act 1995

— DHS

Family and Community Services Act 1997

— DHS

Public and Environmental Health Act 1987

— DHS

Transplant and Anatomy Act 1983

— DHS

Mental Health Act 1993

— DHS

Draft linkage documentation

Version 1: Data repository

MEMORANDUM OF UNDERSTANDING

(reference number XXXXX)

between

AAAAA

and

BBBBB

concerning

one-line description of overall project

THIS MEMORANDUM OF UNDERSTANDING

(reference number XXXXX) is made between:

AAAAA representing XXXXX (in this memorandum of understanding called AAAAA);

AND

BBBBB representing XXXXX (in this memorandum of understanding called BBBBB).

The parties have determined that they wish to cooperate to enable the completion of a project to link specific person – level data from data sets

list the data sets and the purpose of the project.

OVERVIEW

A brief overview of the memorandum, with emphasis on the protocol to be observed and stressing the attention paid to the protection of privacy and confidentiality. A sample follows:

This memorandum covers the linkage and extraction of AAAAA data relating to clients living in [State]. The data will be supplied to nominated analysts as de-identified linked files for use in planning and research on [Program — for example, aged and community care services]. The period of interest covers [time period 1] through [time period 2] and information will be included on individuals living in the State and registered as [program] clients.

The fundamental protocol aims to:

- maximise the conservation of individual privacy;
- minimise access to identified data;
- allow data custodians full control over the dissemination and use of de-identified data files;
- provide linked data files only to named analysts involved in specific approved projects;
- provide analysts with no more than the minimal data required for their analyses; and
- ensure that all copies of named data and all linked data files are destroyed immediately after use.

The process involves two separate stages. The first stage is a memorandum of understanding to share data for an agreed purpose and to prepare a linkage key file (using probabilistic methods) and a master copy of a de-identified linked data file to be stored in a safe repository.

The second stage includes a defined number of separate research projects.

Each separate research project will be covered by its own agreement, the data for the project being supplied by the data repository to analysts to conduct research on behalf of the steering committee. An agreement pro forma, to be completed by the researchers and sent to the steering committee by way of application for data, is in Attachment 2. For each research project, a unique set of project identification numbers (PID) will be generated by the repository custodian and attached to the copy of the master de-identified linked data file provided to the analysts.

This two-stage process will ensure that data custodians have full control over the distribution and usage of their data, as each project will need to be well defined and then individually approved before proceeding. No research will be undertaken without the written approval of every data custodian supplying data to the project. Linked data files will be provided only to the individually identified analysts from the data repository doing the analysis for each project, and will be destroyed when the analyses are complete. A different PID will be used in each project, thus making it extremely difficult to merge the linked data for two projects (such a process is, in any case, specifically prohibited).

The linkage key file will be produced by a small technical team specialising in data linkage. All people involved in the actual linkage and therefore requiring access to the data used in the linkage process will 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 will be allowed access to the files used in this process, as they will contain private and confidential information. The work will be done on an isolated computer, and all personal demographic data (with the exception of those items required to derive variables for further analysis) will be destroyed as soon as the linkage is complete. Transfer of these data files will be done only via tape, diskette or CD-ROM personally carried by those personnel taking part in the data linkage. The linkage personnel will not be permitted to take any part in the analysis of the linked data, or to have any communication about these data with the analysts.

The repository will supply linked data directly to the nominated analysts (from the data repository, or to a contracted third party) for each project. These analysts will also sign confidentiality agreements as required by all parties. They will be the only people granted access to the de-identified linked information. They will be specifically forbidden to disseminate copies of the data files, and will be required to destroy these files on completion of the analysis.

The arbitrary reference number (XXXXX) enables agreements for individual research projects to refer to this over-arching first-stage document.

INTERPRETATION

In this memorandum of understanding unless the contrary intention appears:

- ‘MOU’ means this memorandum of understanding signed by the parties and includes any schedules or attachments hereto.
- A reference to this MOU or another instrument includes any variation or replacement of them.
- The singular includes the plural and vice versa.
- The masculine includes the feminine and neuter; the feminine includes the masculine and neuter; the neuter includes the masculine and feminine.
- The word ‘person’ includes a firm, an unincorporated association or any authority.
- A reference to a person includes a reference to the person’s executors, administrators, successors, substitutes (including, without limitation, a person taking by novation) and assigns.
- An agreement, representation or warranty on the part of or in favour of two or more persons binds, or is for the benefit of them, jointly and severally.
- A reference to any thing (including, without limitation, any amount) is a reference to the whole of any part of it and a reference to a group of persons is a reference to any one or more of them.
- A reference to all clauses, exhibits, annexures or schedules shall, unless otherwise provided, be the clauses, exhibits, annexures or schedules of or to this MOU.
- Headings have been inserted for ease of reference only and shall not be regarded as forming any part of the context of this MOU.
- A reference to a statute shall include all statutes amending, replacing or consolidating the statute referred to.
- ‘Intellectual property’ includes all copyright, all rights in relation to inventions (including patents), registered and unregistered trademarks (including service marks), registered designs, confidential information and circuit layouts, and all other rights resulting from intellectual activity associated with the design, development, delivery or findings of this project, including applications or rights to apply for registration of any of these rights.

- ‘Personal information’ means information or an opinion (including information or an opinion forming part of a database), whether true or not, and whether recorded in a material form or not, about an individual whose identity is apparent, or can reasonably be ascertained from the information or opinion (refer to s.6(1) of the *Privacy Act 1988*).
- ‘Management committee’ means the committee established under clause 3.1 of this MOU.
- ‘Steering committee’ means the committee established under clause 11.1 of this MOU.
- *Add any further definitions that may be necessary.*

OPERATION

This MOU and its operations shall be managed by a steering committee, consisting of:

- The Secretary of AAAAA, or nominee.
- The Chief Executive Officer of BBBBB, or nominee.
- *Add further members as agreed.*

The Secretary of XXXXX, or nominee, will chair the steering committee which will meet if and as required, by teleconference or by face-to-face meeting. Parties are to meet their own costs.

The establishment of the steering committee is described at clause 11.

The work protocol to be followed is described in Attachment 1.

Data will be provided in accordance with the requirements of the agreed work protocol.

The provision of the data shall be subject to the requirements of *list here the appropriate acts, guidelines and ethics committees.*

All results of research or analysis undertaken through approved projects will be placed in the public domain as soon as is feasible according to specification in the Agreement for each project, with each party acknowledged as the source of the data.

OWNERSHIP AND USE OF MATERIAL

Analyses and resulting manuscripts based on each de-identified linked data set will be made available to the steering committee for comment within 30 working days. Any comments will be forwarded to the analysts for their consideration. Analysts should make every effort to take all comments into account and respond to any comments offered but not accepted.

The title to and intellectual property rights in all materials generated in relation to this MOU shall vest jointly upon its creation in each party to the MOU.

Steering committee members are required to disclose all intended publications and reports of results to the steering committee. The parties may decide to further disseminate material received, with full acknowledgment of the source, including through their own publications and other output.

CONFIDENTIALITY

Each party acknowledges that any data provided under this MOU are subject to the confidentiality provisions of the Act under which they were collected.

Officers handling identified data for the purposes of this MOU will be required to sign confidentiality agreements and/or unilateral deed polls as required by each of the Parties.

The parties recognise that in accepting identified demographic data from AGENCY A, AGENCY B, AGENCY C etc., each party becomes subject to the provisions of:

- *List here the appropriate acts and other documents that are appropriate and relevant. The following are some typical examples:*
- *Section 95 of the Privacy Act 1988;*
- *The Information Privacy Principles in section 14 of the Privacy Act 1988;*
- *Section 135A(4) of the National Health Act 1953;*
- *The Privacy Commissioner notes, May 1997;*
- *The National Statement on Ethical conduct in Research involving Humans — National Health and Medical Research Council, 1999.*

Each party will carefully observe these provisions and the conditions of the undertakings signed prior to receipt of data.

The steering committee will maintain appropriate administrative records of all data supplied to the linkage group and of all de-identified linked data supplied to researchers, as well as the undertakings signed by those people authorised to have access to the data.

On completion of the record linkage described in the work protocol, the de-identified linkage key file will be created and all files used by matching staff during the record linkage will be permanently and irretrievably destroyed. The master copy of the de-identified linked data file will then be created.

Agency C will act as custodian of the data repository, providing secure long-term storage for the linkage key file and the de-identified linked data file.

When a specific approved research program is complete, the analysts are to inform the steering committee in writing, attesting to the permanent and irretrievable destruction of the data set used in that project.

Each party will, at all reasonable times, give to the other parties, or to any person authorised in writing by the parties, permission to inspect the arrangements for storage and security of any data relating to the project. Researchers will be required to allow similar inspections of arrangements for the storage and security of their de-identified linked data.

DISPUTES

Where there is a conflict between the parties over any matter related to issues covered by this MOU, parties will seek to resolve the issue through the steering committee.

Should the parties fail to resolve a conflict, the matter shall be referred for resolution to the [relevant arbitrator].

ENTIRE MEMORANDUM OF UNDERSTANDING AND VARIATION

This MOU along with any attachments is to be extended by the addition of specific agreements for each research project proposed. This overarching MOU constitutes the entire agreement between the parties and supersedes all communications, negotiations, arrangements and agreements, either written or oral, between the Parties with respect to the matter hereof, except where otherwise required in law.

Each research project will require the signature of all parties on a detailed specification in writing as described in the project schedule (see Attachment 2).

No variation or extension to this MOU shall be binding upon any party unless in writing and signed by all parties.

TERM

This MOU shall commence when signed by all parties and shall remain in force and effect until otherwise agreed by all parties, or unless terminated in accordance with the terms hereof. At termination of the MOU all parties will destroy any de-identified linked data sets in their custodianship, and the Australian Institute of Health and Welfare will archive a copy of the linkage key file before destroying all other copies.

This MOU may be reviewed upon written request by any party.

TERMINATION

A party shall have the right to request termination of this MOU by written notice to the steering committee if any other party or parties fails to comply with any of the terms and conditions of this MOU.

Any party may terminate this MOU by giving the other parties three months notice in writing to terminate.

Upon termination pursuant to clauses 9.1 or 9.2 all materials and data relating to the project shall be destroyed by all parties, with the steering committee determining the timing and manner whereby any approved projects are to be terminated.

Any termination under clauses 9.1 or 9.2 by any party shall result in the termination of the MOU for all other parties.

ASSIGNMENT

Each party may not assign or otherwise deal with their rights under this MOU without the prior written consent of the other parties, which consent may be given on such terms or conditions as the other parties think fit.

STEERING COMMITTEE

The project shall be guided by a steering committee consisting of:

- (a) a nominee from AAAAA; and
- (b) a nominee from BBBBB.

Add further members to this committee as agreed by the parties to the memorandum.

A nominee of XXXX will chair the steering committee which will meet if and as required, by teleconference or by face-to-face meeting. Parties are to meet their own costs.

Each party must not represent itself, and must ensure that its employees do not represent themselves, as being employees or agents of the other parties.

Each party shall not by virtue of this MOU be, or for any purpose be deemed to be, an employee or agent of another party.

RESPONSIBILITIES OF AAAAA

List responsibilities for each agency as agreed by the parties. Examples are shown below:

AAAAA will identify records of the XXXX client population for the period from XXXX through XXXX inclusive, and will prepare files of demographic data for these individuals to be used in linkage to data supplied by the other parties.

AAAAA will encrypt the full demographic data from XXXX data collection to be used for linkage.

AAAAA will extract and supply de-identified, aggregated service data to the data repository (linkage team) to be merged into a single de-identified linked data file.

AAAAA will take the lead in collaborating with the other parties to select a suitable agency as custodian of the repository. This agency will link together the demographic files and provide secure storage and access to the de-identified linked data set for the nominated analysts.

RESPONSIBILITIES OF BBBB

BBBBB will identify records of the XXXX client population for the period from XXXX through XXXX inclusive, and will prepare files of demographic data for these individuals to be used in linkage to data supplied by the other parties.

BBBBB will encrypt the full demographic data from XXXX data collection to be used for linkage.

BBBBB will extract and supply de-identified, aggregated service data to the data repository (linkage team) to be merged into a single de-identified linked data file.

Executed as a memorandum of understanding:

SIGNED for and on behalf of Agency A

_____ Date _____

Xxxx X XXXXXXXX, Chief Executive Officer, Agency A

SIGNED for and on behalf of Agency B

_____ Date _____

Xxxx X XXXXXXXX, Chief Executive Officer, Agency B

Attachment 1 – Work protocol

Background

A brief description of the background to the project as a whole.

Objectives

A brief description of the objectives of the project.

Scope of study

List the populations and data sets on which the study will be based.

Work plan

Production of linkage key file and master de-identified linked data file

Demographic data from XXXXX to be supplied by AAAAA

List fields to be included in this file, e.g.:

Identification number (encrypted)

Surname First given name Second given name

Gender Date of birth

Country of birth Indigenous status

Address Suburb/town Postcode

First date of contact Last date of contact

Demographic data from XXXXX to be supplied by BBBB

List fields to be included in this file e.g.:

Identification number (encrypted)

Surname First given name Second given name

Gender Date of birth

Country of birth Indigenous status

Address Suburb/town Postcode

First date of contact Last date of contact

Date of death (if known to be deceased)

Linkage process

AAAAA and BBBB will deliver their demographic files to Agency C, where the primary linkage will be done. The linkage will result in the creation of a demographic file containing [add characteristics of linked file here—for example, linked pairs of hospital patients and aged and community care clients].

Delete all the demographic files used to create the linked file.

Merging of service data

The linkage team will merge together the service data, making use of the information contained in the linkage file, to create the master copy of the de-identified linked data file.

Extraction of de-identified linked data files

The linkage team will prepare extracts from the master de-identified linked data file to be sent to the nominated analysts (that is, either from the data repository or from a contracted third party). A unique project identification number (PID) will be generated and attached to an extract from the master file containing only the fields approved for the particular research project. The data file will be sent directly to the analysts named in the project.

Analysis

The analysts will be the only people to have access to the combined linked files, and will be under an obligation not to disseminate copies of the files, or to allow any other personnel to have access to the files. When the analyses are complete the data files will be destroyed.

Summary of entire process

Implementation	Data custodians	Agree to work together on project.
		Select agency to perform linkage and act as data repository.
		Draw up memorandum of understanding.
		Obtain ethical approvals where necessary.
		Sign memorandum of understanding.
		Create steering committee.
Linkage	Data custodians	Prepare demographic files containing at least one record for each individual. Each record will contain encrypted identifying demographic variable(s) and de-identified service experience data (may also be encrypted).
		Supply files to linkage/repository agency.
	Linkage agency	Supply list of linkage personnel to steering committee.
		Linkage personnel sign confidentiality forms and send them to steering committee.
		Link files by means of the identifying variables.
		Delete all identifying variables to create 'linkage key file'.
Master copy of linked data file	Data management team	Supply list of personnel who will have access to the data to steering committee.
		Personnel with access to data sign confidentiality forms and send them to steering committee.
		Use master copy of de-identified linked data file for specified analyses.
Supply of analyses	Steering committee researchers	Receive from steering committee completed analyses.
		Provide comments/requests for further analyses through steering committee within 30 days.
	Steering committee	Request data repository to provide subsequent analyses/incorporate comments.
	Linkage team	Add arbitrary PID to copy of linked data file.
		Extract copy of master de-identified linked data file containing only additional variables approved for the further analyses.
		Supply extracted data file to data management team.
	Data management team	Perform analysis and prepare reports for steering committee.
		Send copies of all output to steering committee.
		Delete data as soon as practicable, notifying steering committee.

Attachment 2 – Pro forma for agreement on research project

Agreement for research project

(Project reference number XXXXX/xx)

using de-identified linked data available as a result of the memorandum of understanding
put name of project here

(reference number XXXXX)

between

Agency A

and

Agency B

to approve a research project led by

enter chief investigator's name and (in parentheses) institution

INTERPRETATION

In this agreement, unless the contrary intention appears, all interpretation is as defined in the over-arching memorandum of understanding (reference number XXXXX).

OVERVIEW OF RESEARCH

enter a brief overview of the research project (no more than two pages)

PROPOSED OUTPUT / PUBLICATIONS

for example, internal report, planning paper, academic journal (name potential journals)

DATA TO BE USED IN ANALYSES

list all required fields

CONDITIONS FOR DATA ACCESS

Personnel listed under paragraph 6 of this agreement are the only people to be allowed access to these data files. These individuals will be required to sign confidentiality agreements as supplied by the steering committee. Any changes to this list must be notified in writing to the chairperson of the steering committee.

By accepting de-identified unit record data from **Agency A** and **Agency B**, these individuals and their host institution become subject to the provisions of:

list all appropriate acts, guidelines, etc.

The data files must be kept on secure computer systems requiring encrypted password entry.

No part of these data will be copied or made available (in any format) to any other individuals or institutions.

These data will not be linked or merged with any other data sets, including data sets generated under a separate agreement covered by the same memorandum of understanding as that for this research project.

These data will only be used for the analyses and output defined in sections 1 and 2.

No attempt will be made to identify any individual whose data are in these files.

Publications and other output will not contain tables or other information that might allow readers to identify any individual whose data have been used in this project.

All analyses and resulting manuscripts from this project will be made available (via the chairperson of the steering committee) to all the parties to the memorandum of understanding giving them opportunity to comment (within 30 working days). The researchers have the right to publish results that suitably address any qualifications or comments by the parties.

All reports and publications resulting from this project must acknowledge **Agency A** and **Agency B** as providers of the data.

If the analysis is not concluded within twelve months from provision of the data, a progress report and request for continued access must be provided to the steering committee.

At the conclusion of the analyses all copies of the data files will be destroyed with written notification to this effect to the chairperson of the steering committee. If necessary, new copies of the data files can be supplied by the data custodians subject to the approval of the steering committee.

The steering committee may request a progress report on this project at any time.

RESEARCH PROJECT PERSONNEL

By signing this agreement, the following personnel agree to observe the terms and conditions listed in this agreement and in the over-arching memorandum of understanding. These are the only individuals permitted to have access to the linked de-identified data files extracted for the purposes of this research project. Any changes to this list must be notified in writing to the chairperson of the steering committee. This notification must include full details (as under) as well as the dated signature(s) of any personnel added to this list. Signed copies of all confidentiality undertakings and/or unilateral deed polls must be supplied for these extra personnel who will have to comply with the terms of the agreement in the same manner as the personnel listed below.

List chief investigator first, then all others in the following format:

Chief investigator:

Title and name: _____

Address: _____

Telephone number: _____

Fax number: _____

Email: _____

Signed: _____ Date: _____

Other investigators:

Title and name: _____

Address: _____

Telephone number: _____

Fax number: _____

Email: _____

Signed: _____ Date: _____

Executed as a agreement:

SIGNED for and on behalf of Agency A

_____ Date _____

Xxxx X XXXXXXXX, Chief Executive Officer, Agency A

SIGNED for and on behalf of Agency B

_____ Date _____

Xxxx X XXXXXXXX, Chief Executive Officer, Agency B

Attachment 3 – Notifications of approval from ethics committees

Include copies of the approval documents from each ethics committee.

Abbreviations

ABS	Australian Bureau of Statistics
ACT	Australian Capital Territory
AIHW	Australian Institute of Health and Welfare
ASCII	American Standard Code for Information Interchange (a standard computer code)
CSDA	Commonwealth/State Disability Agreement
CSDA MDS	Commonwealth/State Disability Agreement Minimum Data Set
CSMAC	Community Services Ministers' Advisory Council
DoHA	Department of Health and Ageing (Commonwealth)
FaCS	Department of Family and Community Services (Commonwealth)
HACC	Home and Community Care
HACC MDS	Home and Community Care Minimum Data Set
HDWA	Health Department of Western Australia
HIC	Health Insurance Commission
IPP	Information Privacy Principles
IVF	in-vitro fertilisation
LLFF	(Canadian) Longitudinal Labour Force File
MCHRDB	Maternal and Child Health Research Data Base (Western Australia)
MOU	memorandum of understanding
NCSIMG	National Community Services Information Management Group
NPP	National Principles (for the fair handling of Personal Information)
NSW	New South Wales
OFPC	Office of the Federal Privacy Commissioner
PID	Project Identification Number
Qld	Queensland
SA	South Australia
SAAP	Supported Accommodation Assistance Program
SIDS	sudden infant death syndrome
SLK	statistical linkage key
SLKWG	Statistical Linkage Key Working Group
Vic	Victoria
WA PID	Western Australian personal identifier number

