Appendix 1: Technical notes

This appendix provides details on aspects of the Indigenous identification audit conducted in all states and in the Northern Territory.

A1.1 Sampling strategy

Sample size formula

The AIHW sought advice on determining the appropriate sample size from the AIHW's statistical consultant and the Australian Bureau of Statistics (ABS). The method used by the AIHW to calculate the sample size for this study was consistent with the methodology used by the ABS when the target sample size is relatively small compared to the total population.

The sample size formula used was:

$$Z \geq (1-s)/(s \cdot y^2 \cdot p)$$, where:

- Z is the required sample size
- s is the proportion of Indigenous patients correctly identified as Indigenous
- p is the proportion of total patients who were Indigenous
- y is the required relative standard error in estimating s.

Using this formula, the sample size was inversely related both to the proportion of patients who were Indigenous, and to the proportion of patients correctly identified as Indigenous. In other words, areas with lower proportions of Indigenous patients correctly identified, or with lower proportions of total patients who were Indigenous, required a larger sample size (Table A1.1).

For example, in an area where Indigenous persons accounted for 2.5% of all patients and 65% of these were assumed to be correctly identified (as Indigenous), the required sample size was Z = 2,154. However, if 95% of Indigenous patients were correctly identified then the required sample size was Z = 211.

Comparison of recommended and achieved sample sizes

There was some variation between the number of interviews recommended for the audit and the number completed by both jurisdiction and remoteness area (Table A1.2). Adequate sample sizes were obtained in the audit for all jurisdictions and for remoteness areas (nationally). However, the sample size was insufficient to allow assessment of the quality of Indigenous identification by remoteness areas within jurisdictions.

Selection of hospitals

The AIHW recommended that the audit include hospitals from each remoteness area within each jurisdiction, and provided a list of suitable 'candidate' hospitals, based on the anticipated level of hospital activity during the audit period, and the likelihood of including Indigenous admitted patients. Table A1.2 presents the numbers of candidate hospitals and

participating hospitals by state and territory, and remoteness area. In most jurisdictions, the majority of candidate hospitals participated in the audit.

Table A1.1: Sample size calculation

Sample size formula (Z >= (1-s)/[s.(y^2).p])	Proportion of Indigenous people correctly recorded (s) %	Proportion of total sample population who are Indigenous (p) %	Relative standard error (y)	Sample size (Z)
Vary proportion correctly recorded (s)				
Low level of identification	65	2.5	0.1	2,154
	70	2.5	0.1	1,714
	75	2.5	0.1	1,333
	80	2.5	0.1	1,000
	85	2.5	0.1	706
	90	2.5	0.1	444
High level of identification	95	2.5	0.1	211
Vary proportion correctly recorded (s) and proportion in sample (p)				
Low level of identification	65	2.5	0.1	2,154
	65	5.0	0.1	1,077
	65	10.0	0.1	538
High level of identification	95	2.5	0.1	211
	95	5.0	0.1	105
	95	10.0	0.1	53
Vary proportion correctly recorded (s) and relative standard error (y)				
Low level of identification	65	2.5	0.05	8,615
	65	2.5	0.1	2,154
High level of identification	95	2.5	0.15	94
	95	2.5	0.2	53

A1.2 Estimation

Weighting

Indigenous identification characteristics vary by jurisdiction, hospital and remoteness area. As the Indigenous identification quality project was based on a small sample of patients, the proportion of surveyed Indigenous patients in a hospital (or remoteness area) compared to the total for the jurisdiction or remoteness area may not be representative of the state or remoteness area overall. As a result, Indigenous patients may be over- or under-represented in the audit, potentially leading to biased estimates of correctness.

In order to account for this bias, the AIHW applied weightings to the audit results for each hospital and remoteness area within each jurisdiction. These were based on the observed number of Indigenous separations included in the audit, compared to the expected number

of Indigenous separations. These weightings were applied to the raw estimates of completeness, to produce the final estimates of completeness.

Table A1.2: Sample size distribution, by state and territory[a] and remoteness area

Jurisdiction and remoteness area	Candidate hospitals	Participating hospitals	Allocated sample size	Achieved sample size	Required sample size
New South Wales	**23**	**20**	**2,869**	**2,870**	**957[b]**
Major cities	5	5	1,630	1,646	
Inner regional	6	7	728	825[c]	
Outer regional	7	3	388	280[c]	
Remote and Very remote	5	5	123	119	
Victoria	**17**	**7**	**1,100**	**1,085**	**800[b]**
Major cities	6	3	675	678	
Inner regional	6	2	309	297	
Outer regional	5	2	116	110	
Queensland	**15**	**14**	**2,850**	**2,740**	**268[b]**
Major cities	4	4	1,103	1,108	
Inner regional	4	4	627	646	
Outer regional	3	3	730	722	
Remote and Very remote	4	3	390	264	
Western Australia	**15**	**12**	**1,401**	**966**	**63[b]**
Major cities	6	4	718	508	
Inner regional	3	3	143	126	
Outer regional	2	1	172	76	
Remote and Very remote	4	4	368	256	
South Australia	**26**	**11**	**601**	**610**	**135[b]**
Major cities	5	5	355	361	
Inner regional	8	4	64	67	
Outer regional	9	1	103	103	
Remote and Very remote	4	1	79	79	
Tasmania	**3**	**2**	**581**	**581**	**496[b]**
Inner regional	2	1	344	344	
Outer regional	1	1	237	237	
Northern Territory	**5**	**5**	**800**	**788**	**8[b]**
Outer regional	1	1	291	301	
Remote and Very remote	4	4	509	487	
Total	**104**	**71**	**10,202**	**9,640**	**5,392[d]**
Major cities	**26**	**21**	**4,481**	**4,301**	**3,315[d]**
Inner regional	**29**	**21**	**2,215**	**2,305**	**1,641[d]**
Outer regional	**28**	**12**	**2,037**	**1,829**	**423[d]**
Remote and Very remote	**21**	**17**	**1,469**	**1,442**	**13[d]**

Notes

(a) The estimation of Indigenous identification levels in the Australian Capital Territory was based on a separate linkage project.

(b) Required sample size as calculated in Table 3.2.

(c) There was a re-classification of the ASGC remoteness areas for some hospitals in the AIHW National Public Hospital Establishment Database. The reclassification changed the remoteness category of some hospitals from that assumed during the sample design stage.

(d) Required sample size as calculated in Table 3.1.

Variations in the weighting methods used

Victoria

The audit results from one hospital in Victoria were markedly different to the results from a similar survey in the same hospital, conducted in 1998. While this audit for the hospital indicated a within-hospital completeness of 33.3%, the 1998 survey had 100% within-hospital completeness. As the number of interviews conducted in the 1998 survey was approximately four times larger than this audit, the results of the 1998 survey were considered to be more reliable. Therefore, the level of completeness for the hospital was adjusted, based on the average of the levels identified in the two audits. The adjusted data for this hospital were then used in the weighted estimation process (as detailed below) to produce the weighted estimates for the relevant region, remoteness area and Victoria.

Queensland

The identities of the participating hospitals in Queensland were masked for privacy and confidentiality reasons. Therefore, within-hospital correction factors could not be calculated as the expected Indigenous proportions for the participating hospitals were unknown. However, Queensland Health provided information on the remoteness area of the participating hospitals, allowing the calculation of within-remoteness area correction factors for Queensland. It should be noted that the estimated results for Queensland are not directly comparable to the estimates calculated for other jurisdictions.

South Australia and Western Australia

For some hospitals in South Australia and Western Australia, the audit did not result in any interviews with Indigenous persons.

For these hospitals the completeness of Indigenous identification was assumed to be similar to the level of completeness for other participating hospitals in the same remoteness area and state.

Northern Territory

The Northern Territory excluded dialysis patients from the audit. It was suggested that separations for dialysis patients should also be excluded from the total separation numbers in the calculation of the weighted correctness factors, as this was more representative of the real distribution of patient numbers across the remoteness areas.

Therefore, a different methodology for estimating under-identification levels was employed for the Northern Territory, and these results may therefore not be comparable to results for other states and territories.

Completeness and correction factors

In this study, estimates of completeness (C) and correction factor (CF) were undertaken at four levels:

- Within-hospital C and CF
- C and CF for remoteness area (within a state or territory)
- C and CF by state or territory
- C and CF by remoteness area (within Australia).

The first level of estimates was an intermediate step to reach the second level of estimates. Like building blocks, the second level was then applied to the weighting system to form the third and fourth levels of estimates.

Within-hospital completeness and correction factor

Within-hospital *C* and *CF* were first estimated for each audited hospital with Indigenous patients identified in the interview, using the following formulas:

$C = A/(A+B)$ and $CF = (A+B)/(A+D)$, where:

- *A* was the number of patients identified as Indigenous in both interview and hospital records
- *B* was the number of patients identified as Indigenous in the interview but non-Indigenous in hospital records
- *D* was the number of patients identified as non-Indigenous in the interview but Indigenous in hospital records.

Completeness and correction factor by remoteness area (within a state or territory)

The *C* and *CF* for each remoteness area within a jurisdiction was estimated based on the estimated within-hospital *Cs* and *CFs* in the area, using either *Wi* or *AWi* as the weight.

- *Wi*, the weight for hospital *i* in the estimation of the remoteness area *CF*, was the proportion of separations for Indigenous persons in hospital *i*, out of the sum of separations for Indigenous persons from participating hospitals in the remoteness area. For this purpose, the number of separations for Indigenous persons was based on separations reported during the period February–April 2005 as recorded in the AIHW's NHMD.
- *AWi*, the weight for hospital *i* in the estimation of remoteness area *C*, was the proportion of adjusted separations for Indigenous persons in hospital *i,* out of the sum of adjusted separations for Indigenous persons from participating hospitals in the remoteness area (adjusted by the within-hospital *CF*).
- For each participating hospital with Indigenous patients identified in the interview, the adjusted number of separations for Indigenous persons was equal to the number of separations for Indigenous persons in the NHMD multiplied by the within hospital *CF*.

The *CF* for each remoteness area was calculated as a weighted average of the relevant within-hospital *CF*s, based on weight *Wi*.

The completeness for each remoteness area was calculated as a weighted average of within-hospital completeness, based on weight *AWi*.

Completeness and correction factor by state or territory

The *C* and *CF* for each jurisdiction was estimated, based on the remoteness area *Cs* and *CFs* in the state, using either *Wr* or *AWr* as the weight.

- *Wr*, the weight of remoteness area *r* in the estimation of jurisdiction level *CF*, was the proportion of separations for Indigenous persons in remoteness area *r*, out of all separations for Indigenous persons in the jurisdiction. For this purpose, the number of

separations for Indigenous persons was based on separations reported during the period February–April 2005, as recorded in the AIHW's NHMD.

- *AWr*, the weight of remoteness area *r* in the estimation of jurisdiction-level completeness, was the proportion of adjusted separations for Indigenous persons in remoteness area *r*, out of the sum of adjusted separations for Indigenous persons in all remoteness areas in the jurisdiction (adjusted by the regional *CF*).
- For each remoteness area, the adjusted number of separations for Indigenous persons was equal to the number of separations for Indigenous persons in the NHMD multiplied by the remoteness area *CF*.

The *CF* for the jurisdiction was calculated as a weighted average of *CF*s for all relevant remoteness areas in the jurisdiction, based on weight *Wr*.

The completeness for the jurisdiction was calculated as a weighted average of completeness factors for all relevant remoteness areas in the jurisdiction, based on weight *AWr*.

Completeness and correction factor by remoteness area (within Australia)

The *C* and *CF* for each remoteness area (within Australia) was estimated based on the remoteness area (within jurisdiction) *Cs* and *CFs*, using either *Wj* or *AWj* as the weight.

- *Wj*, the weight of remoteness area *j* in the estimation of remoteness area level *CF*, was the proportion of separations for Indigenous persons in remoteness area *j*, out of all separations for Indigenous persons in the same remoteness area category.
- *AWj*, the weight of region *j* in the estimation of remoteness area completeness, was the proportion of adjusted separations for Indigenous persons in remoteness area *j*, out of the sum of adjusted separations for Indigenous persons in all remoteness areas from the same remoteness area category (adjusted by the remoteness area *CF*, as defined above).

The *CF* for the remoteness area (within Australia) was calculated as a weighted average of *CF*s for all regions belonging to the remoteness area, based on weight *Wj*.

The completeness for the remoteness area was calculated as a weighted average of remoteness area completeness factors, for all areas belonging to the remoteness area, based on weight *AWj*.

Confidence intervals

The weighted completeness proportions are reported with 95% confidence intervals, calculated using the Normal approximation method for remoteness areas, and for New South Wales, Queensland, Western Australia and the Northern Territory.

The formulas used were:

$$\text{Lower bound} = p - Z_{a/2} * \sqrt{(p*(1-p)/n)}$$

$$\text{Upper bound} = p + Z_{a/2} * \sqrt{(p*(1-p)/n)}$$

For Victoria, South Australia and Tasmania, 95% confidence intervals were calculated using Wilson's score interval to accommodate the small numbers of Indigenous patients identified at interview in those states and territories.

The formulas used were:

$$\text{Lower bound} = p + (1/2n)*(Z_{a/2}^2) - Z_{a/2}*\sqrt{[(p*(1- p)/n) + (Z_{a/2}^2/4n^2)]}$$

$$\text{Upper bound} = p + (1/2n)*(Z_{a/2}^2) + Z_{a/2}*\sqrt{[(p*(1- p)/n) + (Z_{a/2}^2/4n^2)]}$$

Where:

- p is the weighted correctness proportion
- n is the number of Indigenous persons at interview and
- $Za/2 = 1.96$

A1.3 Possible sources of error or bias

Random and systematic error

Random errors occur due to chance variations in the sample. They are not a source of bias, as there is an expectation that the number of hospitals with Indigenous identification levels less than the true value would be balanced by a number of hospitals for which the Indigenous identification levels were greater than the true value.

Systematic errors are introduced when, as a result of the sampling method, the sample consistently underestimates or overestimates the true value. For example, if the participating hospitals in a jurisdiction systematically excluded patients from taking part in the survey on the basis of age or sex, the resulting Indigenous identification levels may be biased.

Assumptions

The project method was underpinned by the following assumptions that:

1. The patient's Indigenous status reported during the interview was correct.

The accuracy of the answer to the Indigenous status question at interview could vary due to factors including:

- the patient's reaction to the interviewer when asked about his or her Indigenous status
- interview conditions
- carer's knowledge of the Indigenous status of the patient.

Any violations of this assumption could introduce non-systematic (random) sampling errors, necessitating larger confidence intervals for the estimation results.

2. There was no change in admission practices during the audit period.

A systematic change in admission practices at a participating hospital could introduce bias into the estimates of Indigenous identification derived from the audit.

It was assumed that admission practices were consistent throughout the audit period, and that these were indicative of the usual admission practices at the hospital.

3. There was no change in the conduct of admission interviews by staff during the audit period.

It was assumed that staff did not alter the way in which they asked patients about their Indigenous status, or their explanation of the question.

A violation of this assumption would lead to biased estimates.

Assumptions 2 and 3 above allow the assumption that the information obtained from the audit was relatively consistent with the usual level of accuracy of Indigenous identification in the hospital.

Sampling

In random sampling, all public hospitals and all patients within the hospitals would have an equal chance of participating in the audit. However, the sampling strategy used for the audit (for participating hospitals) gave preference to larger hospitals with sufficient admitted patient activity (both Indigenous and non-Indigenous), to allow the audit to be conducted in a timely manner.

Potential sources of bias introduced as a result of the sampling strategy were:

- over-representation of hospitals with high proportions of separations for Indigenous persons to ensure that sufficient Indigenous persons would be included in the audit. The results obtained from the study were expected to include a higher proportion of Indigenous persons than in the NHMD overall
- exclusion of hospitals with small admitted patient populations from the sample. Therefore the completeness estimates obtained from the audit may not be typical of the level of Indigenous identification in smaller hospitals.

Estimation

The over-representation of hospitals with high proportions of separations for Indigenous persons could potentially lead to a bias in the estimate of the correction factor.

If there was a systematic relationship between the Indigenous proportion of the hospital's admitted patient population and the within-hospital correction factor, then the resulting estimated correction factor could be biased.

Figure A1.1 shows the within-hospital correction factor plotted against the proportion of Indigenous separations (for February–April 2005). While the figure shows a clear non-linear relationship between the proportion of Indigenous separations and the estimated correction factor, this reflects the quality of identification within the hospital, rather than a systematic error introduced by the sampling strategy. That is, hospitals with very low proportions of separations for Indigenous persons had higher correction factors than hospitals with higher proportions.

The figure also shows that there is larger variation in the within-hospital correction factors for hospitals with low Indigenous proportions, and little variation in correction factors for hospitals with higher proportions.

The results of this audit are consistent with past studies, finding that hospitals located in catchment areas with a low proportion of Indigenous patients often have higher levels of Indigenous under-identification.

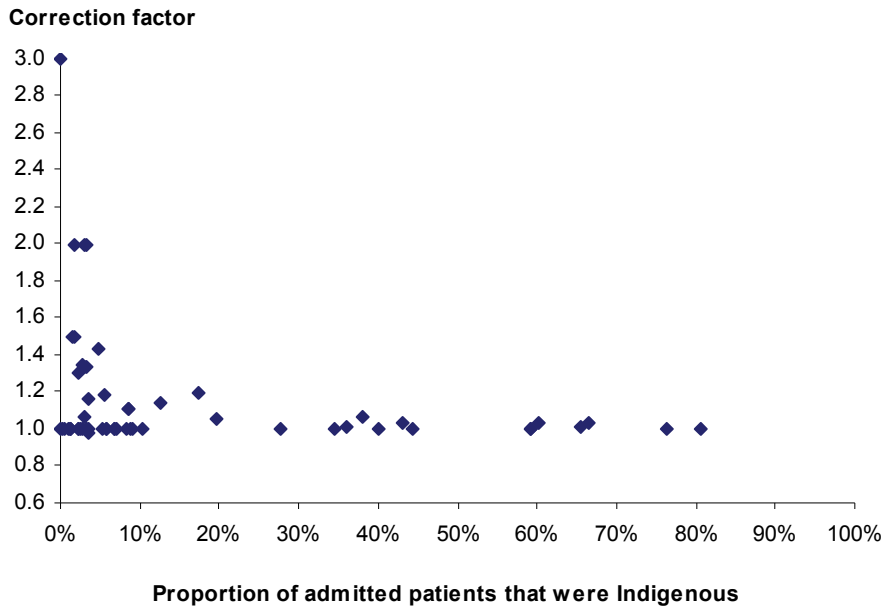

Figure A1.1: Correction factor vs. proportion of separations for Indigenous persons in hospital admissions record for February to April 2004–05

During the estimation process, the AIHW applied weightings to the audit results for each hospital and remoteness area within each jurisdiction, and for each jurisdiction within the total sample. The weightings were based on the number of Indigenous separations observed during the audit compared to the expected number. These weightings were applied to the raw estimates of completeness, to produce the final weighted estimates.

Conduct of the audit

Timing of the interviews

Due to administrative arrangements, the audit was performed during different months of the year for different states and territories. The sampling strategy was not adjusted for seasonality or variation in admission practices over time.

Workforce

There was some variation in the approaches used by the jurisdictions in assigning staff to conduct the interviews. Some jurisdictions used existing hospital staff members to complete the interviews, and some recruited interviewers specifically for the study. For most jurisdictions, a project team was assigned to oversee the conduct of the interviews across hospitals. As the jurisdictions were supplied with identical training materials, the effect of these differences was assumed to be minimal.

Variation in the conduct of the audit by states and territories

Some states and territories reported other issues that may have affected the results of the audit, including:

- small patient turnover in some hospitals

- low participation rates in some hospitals (one reported a refusal rate of around 50%)
- communication difficulties with patients (due to language barriers).

A1.4 Hospital separations for care involving dialysis

Hospital separations for care involving dialysis comprise a large proportion of same-day separations, and result in multiple admissions for the same patient during any given period.

Using the patient sampling strategy as outlined earlier, a patient was only interviewed once during the audit, and therefore the inclusion of dialysis patients may have skewed the results. As Indigenous persons are 12 times more likely to be admitted for dialysis than other Australians (AIHW 2009), the inclusion of dialysis patients may have resulted in an under-representation, relatively, of Indigenous patients.

The exclusion of admitted patient cases for dialysis from both the sample and the total population resulted in a lower proportion of separations for Indigenous persons in the admitted patient data (3.8% compared to 5.0% including dialysis), and consequently a larger sample size being required in every jurisdiction (Table 5.1). The number of interviews completed by each of the participating jurisdictions exceeded the total sample size required by state or territory, after excluding admissions for dialysis.

For most jurisdictions, dialysis patients were included in the audit. The Northern Territory excluded dialysis patients from the audit.

Table A1.3: Sample size calculation by state and territory, based on separations for admitted patients (excluding separations for dialysis), 2004–05

	Separation-based Indigenous proportion			
State/territory	Estimated separations correctly recorded for Indigenous persons[(a)] (%)	Proportion of separations that were for Indigenous persons (%)	Relative standard error	Sample size
New South Wales	77	2.99	0.1	999
Victoria	80	0.66	0.2	951
Queensland	83	6.05	0.1	339
Western Australia	94	7.93	0.1	81
South Australia	95	2.97	0.1	177
Tasmania	70	2.34	0.2	459
Australian Capital Territory	70	1.47	0.2	728
Northern Territory	95	49.07	0.1	11
Total		**3.84**		**3,743**
Australia	**82**	**3.84**	**0.1**	**571**

Note: (a) See Table 4.6 for notes about these estimated proportions.