

7 Data quality

Three main factors affect the quality of data reported in this chapter:

- service type outlet response rates
- service user response rates
- ‘not stated’ and ‘not known’ rates for individual questions on the forms.

The first two of these affect the accuracy of counts of service users – nationally and by jurisdiction and service type – and all three affect the accuracy of results for individual data items.

7.1 Service type outlet response rates

Jurisdiction-reported response rates are based on the number of service type outlets that responded out of the total number of outlets in the jurisdiction. These response rates are shown in Table 7.1.

The national response rate for service type outlets has increased from 82% in 2002–03 to 94% in 2006–07. The majority of jurisdictions now achieve a 100% service type response rate. The response rate in New South Wales has increased from 70% in 2002–03 to 89% in 2006–07; over the same period, the Queensland response rate has increased from 93% to 100%. In Victoria the service type outlet response rate has decreased from 94% in 2003–04 to 90% in 2006–07. Table 7.1 shows the effect that response rates in the larger jurisdictions have on the overall response rate.

Table 7.1: Response rates for service type outlets by jurisdictions, 2002–03 to 2006–07 (per cent)

	NSW	Vic	Qld	WA	SA	Tas	ACT	NT	Aus Gov	Total
2002–03	70	79	93	100	100	100	98	97	100	82
2003–04	80	94	97	100	100	100	93	95	100	93
2004–05	85	92	99	100	100	96	98	70	100	94
2005–06	89	90	99	100	100	100	100	100	100	94
2006–07	89	90	100	100	100	100	100	100	100	94

Notes

1. Response rates are based on figures provided by jurisdictions.
2. The total response rate is based on the number of outlets in the data set divided by the number of total outlets that would have been in the data set if all jurisdictions had a 100% response rate.
3. The response rate for the Australian Capital Territory in 2003–04 is based on agency response rates rather than service type outlets.
4. During 2003–04, Queensland reported 38 and the Australian Capital Territory 35 service users as not providing consent for their data to be transmitted.
5. During 2004–05, Queensland reported 133 and the Australian Capital Territory 36 service users as not providing consent for their data to be transmitted.
6. The 2006–07 NSW response rate is an estimation. Because of the way NSW organises its funding allocations, particularly in relation to individualised funding arrangements and the limitations of the existing funding management system, it is impossible to accurately measure the rate of the return.
7. Although the response rate for the Northern Territory was 100% in 2006–07, this jurisdiction did not submit Internal Department of Health and Community Services data for that year, meaning numbers of users and outlets are undercounted.

7.2 Service user response rate

Service user information may be missing from the data set for a number of reasons. There are outlets that, through administrative or other error, do not report on all of their service users. It is not possible to estimate the number of service users who may be missing from the data set for this reason. Note also that particular service types with a high volume of users and minimal contact (for example, information/referral services) are not required to report service user information under the CSTDA NMDS. The data item 'number of service users' on the service type outlet form was designed to provide this information for all service types, but it is apparent, both from examination of the data for this item and reports from jurisdictions, that the quality of analysis is somewhat hindered by the questionable reliability of the data.

7.3 'Not stated' and 'not known' rates

Service user data items

'Not stated' rates for service user data in 2006–07 showed notable changes from rates in 2005–06 (Table 7.2; see also AIHW 2007: Table 7.2). In particular, there were improvements in the quality of data for the following items:

- 'Indigenous status' had 5.0% not stated in 2006–07; importantly, this rate has now improved beyond the previous low of 8.0% in 2003–04.
- 'Country of birth' had 5.5% not stated (compared with 12% in 2005–06).
- 'Need for interpreter services' and 'method of communication' items had 'not stated' rates of 9.5% and 11.7% – down from 14% and 16% respectively in 2005–06.
- 'Living arrangements' had 11% not stated, down from 15% in 2005–06.
- 'Primary disability' fell from 20% in 2005–06 to 9.2% in 2006–07. The biggest improvements were in Victoria where the 'not stated' rate fell from 42% to 24% and in the Australian Capital Territory where it fell from 21% to 9.5%.

In 2006–07, data collection relating to carers generally improved in terms of reduced 'not stated' responses, with the exception of 'carer – primary status' and 'receipt of Carer Allowance (Child)'. The 'not stated' rate for 'receipt of Carer Allowance (Child)' increased to 46% and is by far the highest of figures reported against items of interest, mainly because of a high proportion of 'not stated' in Victoria (84%). 'Carer – primary status' 'not stated' responses increased to 10% from 4.7% in 2005–06. This was largely due to an increase in the figures from Victoria and Western Australia, 24% and 9.2% respectively (up from 2.7% and 0.7% in 2005–06). Western Australia also reported a significant increase in the 'not stated' rate for 'labour force status', from 17% in 2005–06 to 75% in 2006–07.

In the 'not known' category, Western Australia had a high rate for 'receipt of Carer Allowance (Child)' (93%) and the Australian Capital Territory improved its 'not known' rate from 72% in 2005–06 to 60% in 2006–07. Victoria had a very low 'not known' rate for 'receipt of Carer Allowance (Child)' (0.7%); however, as discussed above, the 'not stated' rate was very high. South Australia had a high recorded 'not known' rate (19%) for 'main income source' compared with other jurisdictions. The aggregate rate across all jurisdictions was 5.3%.

Table 7.2: 'Not stated' and 'not known' response rates for service user data items, 2006–07

Data item	NSW	Vic	Qld	WA	SA	Tas	ACT	NT	Aus Gov	Australia
Not stated										
Age	—	0.0	—	0.4	—	—	—	—	—	0.0
Date of birth	—	0.4	—	0.4	—	—	—	—	—	0.2
Sex	0.5	0.5	0.2	—	—	0.3	—	—	—	0.2
Indigenous status	4.5	12.5	3.0	1.6	2.5	0.3	0.1	2.4	0.1	5.0
Country of birth	5.8	11.9	1.7	4.1	3.0	0.2	3.1	1.6	1.3	5.5
Need for interpreter services	9.4	24.9	1.5	5.4	2.6	1.0	2.1	6.7	0.0	9.5
Method of communication	14.0	21.6	1.6	4.9	3.6	2.0	5.2	5.6	6.8	11.7
Living arrangements	10.2	13.8	1.2	2.9	3.0	1.4	0.6	1.4	13.3	10.8
Postcode of usual residence	0.1	2.6	0.0	2.9	1.2	0.4	0.0	—	—	1.1
Residential setting	9.2	21.1	1.1	3.5	3.4	1.2	2.0	9.8	7.8	11.0
Primary disability group	9.1	23.8	0.9	1.0	1.4	4.5	9.5	26.6	0.7	9.2
Frequency of support or assistance needed										
Self-care	21.8	35.9	2.5	4.8	5.4	0.5	17.2	27.2	8.2	18.0
Mobility	21.4	35.0	2.4	4.8	5.3	0.4	16.6	27.0	7.6	17.4
Communication	20.3	34.7	2.7	4.7	5.4	0.5	11.7	26.9	7.1	17.0
Interpersonal interactions and relationships	29.3	36.9	3.2	5.1	5.6	0.7	16.3	26.9	8.2	19.4
Learning, applying knowledge and general tasks and demands										
Education	18.5	34.9	4.2	6.0	6.0	3.5	15.8	27.1	9.8	18.0
Community (civic) and economic life	19.9	35.8	5.6	6.6	6.2	4.8	16.0	27.2	8.6	18.3
Domestic life	22.6	35.3	5.3	8.6	6.0	3.2	38.9	27.2	11.4	19.9
Working	30.3	43.0	4.4	6.6	5.8	2.5	2.3	22.2	11.2	22.4
Working	36.4	44.6	7.9	10.0	6.4	4.9	2.9	22.5	7.0	23.0
Carer—existence of										
Carer—primary status	0.1	23.3	1.9	3.7	0.2	2.3	13.0	27.2	1.5	8.2
Carer—residency status	1.2	24.3	5.5	9.2	1.5	4.3	4.6	16.2	12.3	10.3
Carer—relationship to service user	16.0	23.6	4.9	4.9	9.3	2.6	1.6	17.8	19.4	13.6
Carer—age group	4.9	24.0	5.0	4.6	2.1	0.9	1.9	2.1	7.7	9.8
Main income source (adult)										
Main income source (adult)	27.4	28.6	9.3	9.3	9.8	3.8	21.2	18.7	25.6	19.6
Receipt of Carer Allowance (Child)										
Receipt of Carer Allowance (Child)	13.0	30.5	3.6	5.0	4.3	4.2	3.7	10.1	1.3	11.5
Labour force status										
Labour force status	52.2	84.4	7.7	3.0	2.9	4.1	15.9	17.8	6.4	46.0
Individual funding status										
Individual funding status	13.4	19.0	4.4	74.5	4.4	7.9	7.4	16.1	—	11.9
Individual funding status	11.4	15.3	—	4.3	—	6.1	18.2	11.7	—	6.9
Not known										
Main income source (adult)	4.5	0.1	3.0	0.5	19.2	2.2	4.8	5.1	6.9	5.3
Receipt of Carer Allowance (child)	21.3	0.7	28.3	93.0	7.3	55.4	59.8	54.5	0.5	26.7
Individual funding status	14.5	0.0	12.6	56.5	—	10.1	36.7	13.3	—	8.1

Notes

1. Figures are the percentage of total data item responses for each data source.
2. Service users accessing service type 3.02 were only required to report on data items relating to age and sex. Service users who accessed only this service type over the 12-month period are therefore excluded from calculations of 'not stated' rates for all other data items.
3. Service types 6.01–6.05 and 7.01–7.04 did not collect service user data and are therefore excluded from this table.
4. Service types 5.01 and 5.02 were not required to collect data on *carer—primary status*, *carer—residency status*, and *carer—age group*. 'Not stated' rate calculations therefore exclude 5.01 and 5.02 service types for these data items.
5. 'Not stated' rates for *carer—primary status*, *carer—residency status*, *carer—relationship to service user*, and *carer—age group* are based only on those service users who answered 'yes' to the item *carer—existence of*.
6. The higher levels of not stated/not known on some data items for the Australian Capital Territory reflect in part the more limited data collection for some services within agencies.

Service use data items

For three of the six service use data items the 'not stated' rates decreased between 2005–06 and 2006–07 – 'service start date' decreased from 0.7% to 0.1%; 'snapshot date flag' decreased from 13% to 7%; and 'main reason for cessation of services' decreased from 28% to 7% (Table 7.3; see also AIHW 2007: Table 7.3). For 'date service last received', the 'not stated' rate remained relatively steady at around 14%. The 'not stated' rates for 'hours received in the reference week' increased from 23% to 30%, and for 'hours received in a typical week', it increased from 2.2% to 8.8%.

Table 7.3: 'Not stated' response rates for service use data items, 2006–07 (for applicable service types)

Data item	NSW	Vic	Qld	WA	SA	Tas	ACT	NT	Aus Gov	Australia
Service start date	0.1	0.2	—	—	—	2.8	—	3.9	—	0.1
Date service last received	—	34.6	—	40.7	—	2.8	—	3.9	—	13.7
Snapshot date flag	3.0	8.1	—	46.6	0.1	3.4	2.0	—	0.1	7.0
Main reason for cessation of services	11.1	12.3	—	—	—	7.8	29.2	9.5	3.1	6.9
Hours received in the reference week	—	46.2	0.2	52.5	35.7	0.4	9.7	0.1	n.a.	30.2
Hours received in a typical week	13.4	n.a.	—	27.3	0.2	0.4	5.1	—	n.a.	8.8

Notes

1. Figures are the percentage of total data item responses for each data source.
2. Service users accessing service types 3.02, 6.01–6.05 and 7.01–7.04 were not required to collect service use data, and are therefore excluded from this table.
3. Some service types were not required to report on hours received (reference week) and hours received (typical week). These service types are therefore excluded from calculations of 'not stated' rates for these data items (namely service types 1.01–1.04, 1.08, 2.01–2.05, 2.07, 5.01 and 5.02).
4. Victoria did not collect data on hours received in a typical week.
5. 'Not stated' rates for *main reason for cessation of services* are based only on service users who entered a service exit date. Exit dates left blank were treated as indicating that the service user had remained with the service; therefore the response rate for this item was 100%.
6. A response of '0' was considered as a 'not stated' response, except for *snapshot date flag* and *hours received* (both reference week and typical week).

Service type outlet data items

Increases from 2005–06 were seen in the 'not stated' rate for four of the service type outlet items in 2006–07 – 'full financial year operation' (up from 1.8% to 4.3%), 'staff hours in a typical week' (6.4% to 10.5%), 'operating weeks per year' (7.3% to 7.9%) and 'number of service users' over the year (3.5% to 11.3%) (Table 7.4; see also AIHW 2007: Table 7.4). 'Not stated' rates improved between 2005–06 and 2006–07 for 'agency sector' (2.5% to 0.1%), 'staff hours in the reference week' (10.4% to 8.3%), 'operating days per week' (7.9% to 7.7%) and 'operating hours per day' (7.0% to 5.4%). As in 2003–04, 2004–05 and 2005–06, there were no missing service types in 2006–07.

Table 7.4: 'Not stated' response rates for service type outlet data items, 2006–07

Data item	NSW	Vic	Qld	WA	SA	Tas	ACT	NT	Aus Gov	Australia
Agency sector	0.3	—	—	—	—	0.5	—	—	—	0.1
Service type	—	—	—	—	—	—	—	—	—	—
Full financial year operation	1.1	12.2	—	7.0	—	6.8	1.8	—	—	4.3
Staff hours in the reference week	5.0	24.5	0.4	2.6	1.8	—	1.8	0.7	0.3	8.3
Staff hours in a typical week	10.5	3.1	11.4	2.6	1.0	—	—	17.7	8.6	10.5
Operating weeks per year	7.5	10.7	18.3	2.3	—	—	—	—	0.3	7.9
Operating days per week	10.1	9.8	14.0	2.3	—	—	—	—	0.3	7.7
Operating hours per day	0.4	10.8	14.1	2.3	—	—	—	—	0.3	5.4
Number of service users over the year	23.1	1.2	1.9	8.6	—	6.5	0.6	—	33.0	11.3

Notes

1. Figures are the percentage of total data item responses for each data source.
2. Service types 7.01–7.04 were not required to report on the data item *number of service users over the year*, hence these outlets are excluded from the 'not stated' calculations for this data item.
3. Victoria did not collect data on *staff hours in a typical week*.
4. A response of '0' was considered as a 'not stated' response, except for *staff hours* (for both reference week and typical week)—if only one staff hours (paid or unpaid) variable was missing, it was assumed to be validly recorded as zero.

Statistical linkage key validity

The proportion of invalid linkage keys in 2006–07 was 0.8%, ranging from 0% to 2.4% of the total number of service user records across jurisdictions (Table 7.5). This was a similar result to the 2005–06 data set (0.9% were invalid overall, ranging from 0% to 2.6%). A small number of invalid linkage keys were recovered by comparison with records from current or previous year's data.

Table 7.5: Validity of the statistical linkage key in the CSTDA NMDS, 2006–07

	NSW	Vic ^(a)	Qld	WA	SA	Tas	ACT	NT	Aus Gov	Total
Number of service user records (unlinked)	43,400	90,552	26,281	24,803	22,815	5,788	4,175	1,231	82,770	301,815
Number with invalid linkage keys	9	2,161	235	107	18	0	0	0	7	2,537
% invalid linkage keys^(b)	0.0	2.4	0.9	0.4	0.1	0.0	0.0	0.0	0.0	0.8

(a) For Victoria, 'pseudo' linkage keys are included as valid—see Appendix 4 for details.

(b) Statistical linkage keys missing sex only are counted as valid.