

5 Collection procedures and data quality

Collation of the national data set

The collation of a national data set involves five distinct stages (see Figure 2).

1. The first stage is the collection of the agreed data elements by service providers for each client who is eligible for inclusion in the collection. Service providers then forward their collected information to the designated health authority for collation. This process will differ across jurisdictions, as service providers in some States/Territories are required to forward their data to an area or region coordinator, whereas in other States the data is forwarded directly to the central authority.
2. The second stage involves the designated health authority collating the data that was forwarded by the service providers. At this stage the data should also undergo a rigorous validation process to ensure the quality of the information. Health authorities are required to allocate establishment-level data elements. The collated unit record data is then forwarded to the AIHW.
3. At stage three the AIHW receives the collated State/Territory data for validation. When finished validating the data, the AIHW sends a summary report to each State/Territory health authority (which includes all queries and identified problems with their data) for resolution and clarification. The report includes an initial set of frequency and cross-tabulation tables.
4. At stage four, State and Territory health authorities send their responses to the validation report to the AIHW. The required changes are made to the data and revised frequency tables sent back to the State and Territories for final approval of the database.
5. At the final stage, the approved national database is stored by the AIHW ready for analysis and reporting.

Note that no data are to be directly submitted by service providers to the AIHW. Note also that the information transferred from service providers to health authorities and then to the AIHW does not include client names, only a person identifier code that is generated by the service provider.

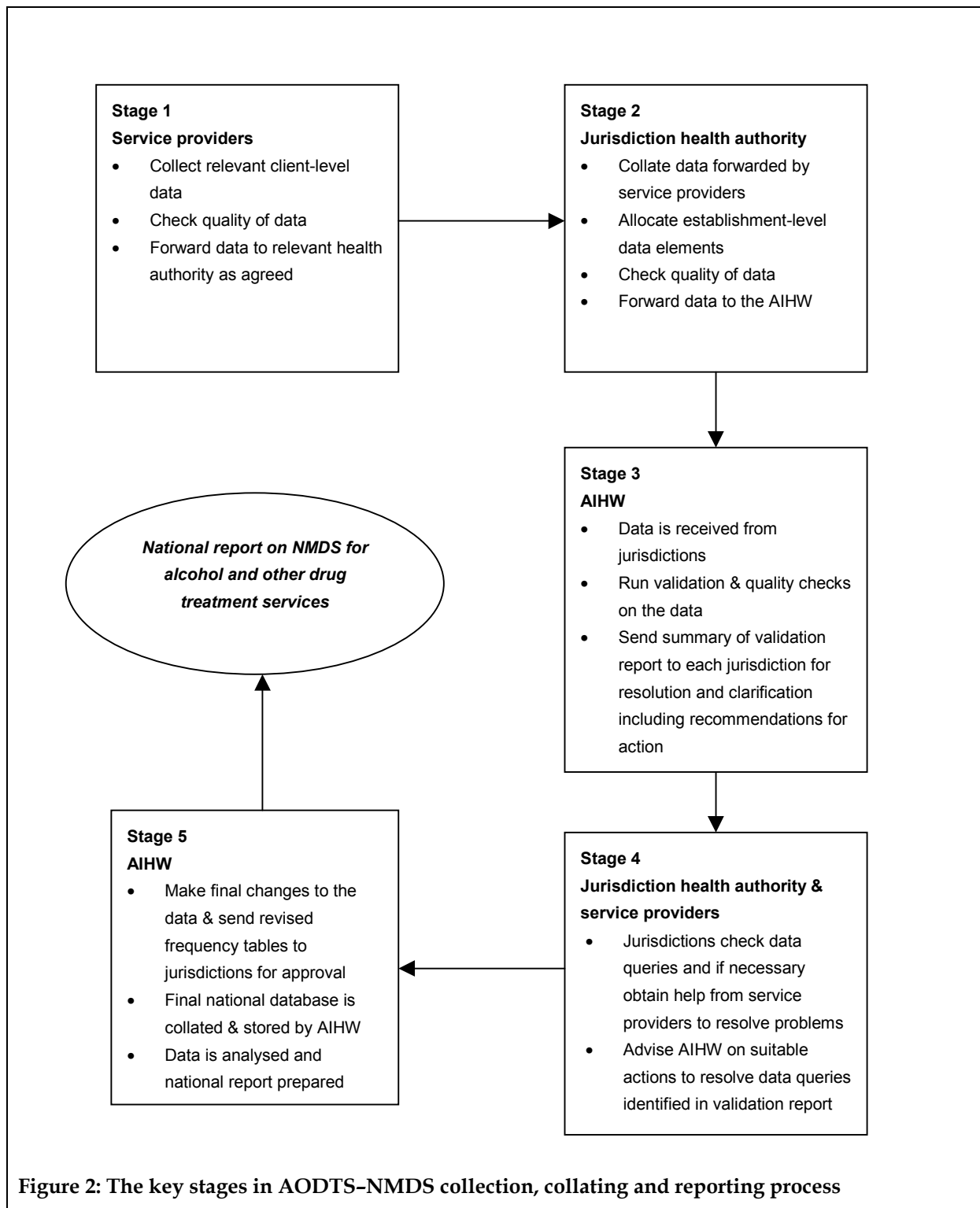


Figure 2: The key stages in AODTS-NMDS collection, collating and reporting process

Data transfer

Service providers to health authorities

Protocols for the transfer of data from alcohol and other drug treatment agencies to their jurisdictional health authority vary between States and Territories. Each health authority responsible for the AODTS-NMDS collection will contact service providers within scope for the collection to inform them of the required format and timing of the data transfer.

Health authorities to AIHW

The NMDS data will need to be forwarded to the AIHW annually by each jurisdiction. The data requested will be for a financial year reference period (1 July to 30 June). Data for the period 1 July 2002 to 30 June 2003 will be requested by the AIHW early in the 2003-04 financial year (September 2003). It is expected that State and Territory health authorities will supply this data to the AIHW by no later than December 2003. The results of the analysis of this data, at both the national and State/Territory levels, will be reported during 2004.

Transfer method

The preferred transfer method is by e-mail attachment or floppy disk. Note that floppy disks can only hold 1.4 Mb of data, and the AIHW can only accept files by e-mail that are less than 4 Mb. Files sent by e-mail or floppy disk should be compressed, preferably with WinZip, and password protected. Jurisdictions interested in sending the file via e-mail are requested to contact the AIHW, before sending the file, to advise the AIHW of the password.

If a file does not fit on a floppy disk or is too large for an e-mail attachment, the next preferred option is a CD-ROM.

File format

When jurisdictions are satisfied that their data are clean, and that all practical follow-up has been completed, unformatted data should be forwarded to the AIHW contact in the following form:

Comma Separated Variable (CSV) ASCII text records.

For example, a single client unit record will look like the following:

12A00101, PID99, 1, 05061977, 1101, 4, 19, 1, 01, 02092001, 03122001, 07, 02, 1, 3201, 0003, , , 2, 4, 2, 8, , , ,

The following file types can be accepted by AIHW:

- Microsoft Excel file
- Microsoft Access file
- Unformatted SAS file in transport mode indicating the appropriate platform (e.g. Unix, NT, MVS) and the SAS version used.

File content

There should be two files for each jurisdiction:

- establishment-level file (statistical unit = alcohol and other drug treatment agency/organisation)
- client-level file (statistical unit = closed treatment episode).

Accompanying information

When transferring data to the AIHW each jurisdiction should include the following documentation:

- a file with summary frequency tables and cross-tabulation tables, which can be used by the AIHW to verify information when compiling the national data set (the required frequency and cross-tabulation tables are listed in chapter 6);
- a description of the file including the total number of records it contains;
- identification of variables that do not conform to the standard definitions and any necessary translation or manipulation of the data to achieve national standards; and
- if non-standard names are used for any variables, please include a mapping of these names to the standard names.

File specification

The following tables specify the order in which the data should be provided to the AIHW and the range of valid responses.

Table 2: Specifications for data transfer to AIHW of establishment-level data

Item no.	Item name	Description	Valid response
1	Establishment identifier	Alphanumeric item. Consists of a unique identifier for each State/Territory (<i>State identifier</i>), an establishment sector code (<i>Establishment sector</i>), an alphanumeric, alpha or numeric code for the location of the service (<i>Region code</i>), and a unique service number within the State/Territory (<i>Establishment number</i>).	Maximum 9 characters NNANNNNNN where N = <i>State identifier</i> (valid range 1-9), N = <i>Establishment sector</i> (valid range 1-2), AA = <i>Region code</i> (jurisdiction specific), NNNNN = <i>Establishment number</i> (jurisdiction specific). Note that a 'not stated', 'null', 'missing' or 'coding error' response is not permitted. Where no Establishment identifier has been assigned, the establishment record should be excluded.

(continued)

Table 2 (continued): Specifications for data transfer to AIHW of establishment-level data

Item no.	Item name	Description	Valid response
2	Establishment type	Alphanumeric item. The type of service provider.	Valid code from data domain (e.g. R4.1, R4.2, N8.1.1, N8.1.2).
3	Geographical location of establishment	Numeric code to indicate the statistical local area (SLA) of the agency within the reporting State or Territory.	Five-digit valid code as defined in the <i>Australian Standard Geographical Classification (ASGC)</i> , ABS Cat. No. 1216.0 or 99999 (not stated/inadequately described). Use code 99999 as the default value for invalid codes and missing data.

Table 3: Specifications for data transfer to AIHW of client-level data

Item no.	Item name	Description	Valid response
1	Establishment identifier	As before.	As before.
2	Person identifier	Alpha, numeric or alphanumeric item. A client's unique identifier within an agency.	Alpha, numeric or alphanumeric code (agency determined). Note that a 'not stated', 'null', 'missing' or 'coding error' response is not permitted. Where no person identifier has been assigned, the client record should be excluded.
3	Sex	Numeric code to indicate the sex of a client.	1 (male), 2 (female) or 9 (not stated/inadequately described). Use code 9 as the default value for invalid codes and missing values.
4	Date of birth	Numeric item stating a client's date of birth.	DDMMYYYY format, right justified, zero-filled (e.g. 3 March 1965 would be 03031965). When an estimate is required use 0101 with a valid year. Use 01011900 as default date for invalid and missing dates.

(continued)

Table 3 (continued): Specifications for data transfer to AIHW of client-level data

Item no.	Item name	Description	Valid response
5	Country of birth	A numeric code to indicate the country in which a client was born.	Four-digit valid code as defined in the <i>Standard Australian Classification of Countries</i> (SACC), ABS Cat. No. 1269.0 (1998) and 0000 (inadequately described), 0003 (not stated), 0004 (unknown). Use 0000 as default value for invalid codes and 0003 for missing data.
6	Indigenous status	A numeric code indicating whether or not a client is of Aboriginal and/or Torres Strait Islander origin.	1-4 and 9 (not stated). Code 9 should only be used when the question was refused or not able to be asked.
7	Preferred language	A numeric code indicating the language most preferred by a client for communication.	Two-digit valid code 00-86 and 95 (Other languages nfd), 96 (Inadequately described), 97 (Non-verbal, so described including sign languages) and 98 (Not stated). Right justified. Use 96 as default value for invalid codes and 98 for missing values.
8	Client type— alcohol and other drug treatment services	A numeric code indicating the status of a person in terms of whether contact with the agency concerns their own alcohol and/or other drug use or that of another person or both.	1-3. Note that a 'not stated', 'null', 'missing', or 'not known' response is not permitted.
9	Source of referral to alcohol and other drug treatment service	A numeric code indicating the source from which a client was transferred or referred to an agency.	1-18 and 99 (not stated/ inadequately described), right justified. Use code 99 as the default value for invalid codes and missing values.
10	Date of commencement of treatment episode for alcohol and other drugs	Numeric item stating the date on which a client's treatment episode commenced.	DDMMYYYY format, right justified, zero-filled (e.g. 3 March 2001 would be 03032001). Note that a 'not stated', 'null', 'missing', 'not known' or an invalid date response is not permitted. Where an accurate date of commencement is not available, the client record should be excluded.

(continued)

Table 3 (continued): Specifications for data transfer to AIHW of client-level data

Item no.	Item name	Description	Valid response
11	Date of cessation of treatment episode for alcohol and other drugs	Numeric item stating the date on which a client's treatment episode ceased.	DDMMYYYY format, right justified, zero-filled (e.g. 3 March 2001 would be 03032001). Note that a 'not stated', 'null', 'missing', 'not known' or an invalid date response is not permitted. Where an accurate date of cessation is not available, the client record should be excluded.
12	Reason for cessation of treatment episode for alcohol and other drugs	Numeric item stating the reason a client's treatment episode ceased.	1-13, 98 (Other) and 99 (Not stated/inadequately described).
13	Number of service contacts within a treatment episode for alcohol and other drugs	Numeric item stating the number of service contacts within a treatment episode.	Valid digit (range 0-999).
14	Treatment delivery setting for alcohol and other drugs	Numeric item stating the type of treatment delivery setting.	1-4 and 8. Note that a 'not stated', 'null', 'missing' or 'not known' response is not permitted.
15	Method of use for principal drug of concern	A numeric code indicating a client's usual method of administering the principal drug of concern, as stated by the client.	1-6 and 9 (not stated/inadequately described). Use code 9 as the default value for invalid codes and missing values.
16	Injecting drug use	A numeric code indicating a client's use of injection as a method of administering drugs.	1-4 and 9 (not stated/inadequately described). Use code 9 as the default value for invalid codes and missing values.
17	Principal drug of concern	A numeric code indicating the drug that has led a client to seek treatment or advice from an agency. As stated by the client.	Four-digit valid code as defined in the <i>Australian Standard Classification of Drugs of Concern</i> , ABS Cat. No. 1248.0 (2000), 0000 (inadequately described) or 0001 (not stated). Use code 0001 as the default value for invalid codes and missing values.

(continued)

Table 3 (continued): Specifications for data transfer to AIHW of client-level data

Item no.	Item name	Description	Valid response
18a	Other drugs of concern (1st)	A numeric code indicating any drugs, in addition to the principal drug of concern, which a client perceives as a health concern.	Four-digit valid code from <i>Australian Standard Classification of Drugs of Concern</i> , ABS Cat. No. 1248.0 (2000) or 0001 (not stated) and 0003 (none/no other drugs of concern). Use code 0001 as the default value for invalid codes. – for up to 5 ‘other drugs of concern’ see items 14a-14e. To ensure correct data loading, 5 commas must be inserted even if blanks are present. (E.g. if there are no other drugs of concern it would be 0003,,,,, if 2 other drugs of concern it would be NNNN, NNNN,,,,). Note that code 0003 only needs to be used for the 1st field.
18b	Other drugs of concern (2nd)	A numeric code indicating the second other drug of concern.	Four-digit valid code as above. If blank response then insert a comma (,) without brackets.
18c	Other drugs of concern (3rd)	A numeric code indicating the third other drug of concern.	Four-digit valid code as above. If blank response then insert a comma (,) without brackets.
18d	Other drugs of concern (4th)	A numeric code indicating the fourth other drug of concern.	Four-digit valid code as above. If blank response then insert a comma (,) without brackets.
18e	Other drugs of concern (5th)	A numeric code indicating the fifth other drug of concern.	Four-digit valid code as above. If blank response then insert a comma (,) without brackets.
19	Main treatment type for alcohol and other drugs	A numeric item stating the main treatment type.	1-8. Each treatment episode will only have one main treatment type.

(continued)

Table 3 (continued): Specifications for data transfer to AIHW of client-level data

Item no.	Item name	Description	Valid response
20a	Other treatment type for alcohol and other drugs (1st)	A numeric code indicating any treatment that a client is receiving in addition to the Main treatment type.	1-5 and 8 (none/no other treatment type) or 9 (not stated/ inadequately described). Up to 4 other treatment types can be reported, see Item nos 20a-20d. To ensure correct data loading, 4 commas must be inserted, even if blank responses are present (e.g. if no other treatment types it would be 8,,,,. If 2 other treatment types it would be N, N,,,,). Note that code 8 only needs to be used for the 1st field.
20b	Other treatment type for alcohol and other drugs (2nd)	A numeric code indicating the second Other treatment type.	1-5 and 8 or 9. If a blank response then insert a comma (,) without brackets
20c	Other treatment type for alcohol and other drugs (3rd)	A numeric code indicating the third Other treatment type.	1-5 and 8 or 9. If a blank response then insert a comma (,) without brackets
20d	Other treatment type for alcohol and other drugs (4th)	A numeric code indicating the fourth Other treatment type.	1-5 and 8 or 9. If a blank response then insert a comma (,) without brackets

AIHW contacts for further information on file transfer

Dr Bradley Grant National Data Development Unit
 Phone: (02) 6244 1152 Email: bradley.grant@aihw.gov.au

Ms Gail Weaving National Data Development Unit
 Phone: (02) 6244 1050 Email: gail.weaving@aihw.gov.au

Mr Rod Hall Data and Information Technology Unit
 Phone: (02) 6244 1188 Email: rod.hall@aihw.gov.au

Data quality

Data collections require ongoing attention to quality. There is a need to attend to how questions are asked and information obtained, data entry, the handling of missing and erroneous information, edit checking and there is a need to follow-up with data providers to ensure the highest quality data possible.

To ensure that the AIHW is supplied with a useable national data set, it is essential that jurisdictions clean (edit) the data they receive from service providers before they transfer it to the AIHW. The quality of the NMDS data will also be enhanced if service providers check the quality of their data before sending it to their jurisdictional health authority. In collating the data into a national database, the AIHW also follows a formal validation process to maximise data quality (see Chapter 6).

Some general checks that should be conducted

Service providers and jurisdictions should perform the following quality checks before the data are sent to the AIHW.

- **Missing agencies:** Jurisdictions should ensure that all agencies within scope of the collection have sent data for the entire collection period.
- **Missing data:** Jurisdictions should investigate missing data to ensure that agencies are reporting all AODTS–NMDS data items. Where possible, a reasonable attempt should be made to resolve missing data issues, at both an agency level and at the unit record level.
- **Incorrect codes:** Where possible, jurisdictions should ensure that agencies use the correct codes for all data items. Coding errors that cannot be corrected should be coded to the appropriate default value (e.g. inadequately described).
- **Duplicate records:** Jurisdictions should check for duplicate unit records. When records are identified as possible duplicates, the agency should be consulted to ensure that unit records have not been mistakenly submitted on more than one occasion.
- **Reporting period:** The cessation dates of treatment episodes should be checked to ensure that only treatment episodes that closed within the valid reporting period (1 July 2002 to 30 June 2003) are included in the 2002–03 collection.
- **Data inclusion:** Jurisdictions should ensure that data not within scope of the AODTS–NMDS are excluded from the collated data set sent to the AIHW (e.g. methadone treatment).
- **Establishment identifiers:** Jurisdictions should ensure that establishment identifiers used on the establishment data file are the same as those used on the client data file and that there are the same number of establishments on each file.
- **Geographical location of establishment:** Jurisdictions should ensure that all geographic location codes begin with a valid State or Territory identification number.
- **Client type:** Jurisdictions should ensure that for clients who attend treatment because of another person's drug use (client type = 2), the following data elements are coded to Not stated: *Method of use for principal drug of concern* (code 9), *Injecting drug use* (code 9), *Principal drug of concern* (code 0001), and *Other drugs of concern* (code 0001).