Data Starter

**A project of the Advisory Committee on Australian
and International Disability Data**

Collecting disability data: A guide for service providers

Inside Issue 2

Data Starter 2 is intended to provide a resource for service providers by outlining the steps that are involved in planning and developing a disability data collection. The Starter outlines the types of information that may be gathered, gives practical advice on how to get started, and provides guidance on data management and analysis and on privacy implications of collecting data. This document is a companion to Data Starter 1, which describes existing disability data and how they can be obtained. To access Data Starter 1, please go to <www.aihw.gov.au/publications/index.cfm/title/6361>.

1   Why collect data?

The most important thing to consider when developing a data collection is what information you need and what you are going to use the information for.

Determining the main purposes is essential when designing any data collection. There may be a wide range of reasons for collecting data, such as planning, program evaluation, and monitoring objectives and agreed priorities. More specific reasons could include:

- a need for more complete information about clients and/or services;

- a need for information about the population being served;

- a need for information about a new population a service is moving into;

- to find out about needs or unmet needs of clients;

- meeting the reporting requirements of government departments, boards or other regulatory bodies;

- comparing information at different levels (e.g. state, national, international);

- updating a previous data collection or database;

- agency and service planning; and

- developing a Management Information System for agencies or services.

2   Do you really need to collect data, and what data do you need?

When designing your data collection it is important to first ask 'do I really need to collect the data?' In order to answer this, consider whether the information already exists. Data Starter 1 provides an overview

of national and other large disability-related data collections in Australia, as well as some national special-interest collections. Check to see if any of these existing collections answer your information needs.

If you decide that you do need to collect data, answering the following questions may help you to determine exactly what data you need to collect:

- How will the data be used? — are the data going to be used to inform policy, track trends in service usage, monitor the number of service users, or some other purpose?

- What do you need to know specifically? — talk to potential users of the data and encourage them to be specific about the information they need. For example, 'I need to know if the number of people with an intellectual disability accessing this service is increasing' or 'how many service users have a primary carer?'

- Who will use the data? — are the data going to be used by service users, agency staff, a government body, or others?

- How will the data be analysed? — the analysis you choose to perform may influence the data items you include and/or the coding categories you use to capture information.

Setting up and running a data collection can be costly. To make the best use of resources it is important to clearly define your objectives and your data and information needs from the outset.

3   What should you collect data about?

It is important to define the scope of the collection—that is, about whom or what you will collect information. This is also known as determining the 'population' of

the collection. Your population will depend largely on what you need to know from the data. If you need to track service usage, then your population will be people with a disability who use your services. If you need to know how many staff hours go into caring for people with disabilities at your service, your population will be the staff that work at your agency. If you need to know the geographic distribution of agencies, your population will be the agencies in a defined group or organisation.

To put it simply, the population is the group of people, agencies or other 'entities' for which you need questions answered.

4   Privacy implications of collecting data

Privacy and confidentiality must be considered whenever data are collected or disseminated about individuals or service provider organisations. The *Privacy Amendment (Private Sector) Act 2000* regulates the way that private sector organisations can collect, use, keep secure, and disclose personal information. For the first time it gives individuals the right to know what information an organisation holds about them and a right to correct that information if it is wrong.

A set of 11 Information Privacy Principles (IPPs) and 10 National Privacy Principles (NPPs) set out rules relating to the collection, use and disclosure, storage, handling, quality and security of personal information held by Commonwealth agencies and private sector organisations, respectively.

The IPPs and NPPs can be summarised by three basic principles:

1 Agencies and organisations must tell people what information they are collecting and what they will do with it (i.e. the purpose and uses of that collection).

2 Agencies and organisations should obtain an individual's consent to collect the data, for instance by giving them the opportunity to 'opt out' before collecting, using or disclosing information about them. There are exclusions 'authorised by law' which agencies can investigate if thought necessary.

3 Agencies and organisations should give people confidence that they respect their personal information and will treat it accordingly (for instance, in terms of storage access and disclosure).

Some states and territories also have legislation governing the handling of different types of personal, health, or other sensitive information. Additionally, some data collections already have privacy principles developed and agreed (AIHW 2004).

Further information about the IPPs, NPPs and other privacy practices can be found at www.privacy.gov.au.

5 Designing a data collection

Before delving into the technical side of developing a data collection, the following steps may provide a useful guide as a starting point:

1 Look again at what data are *already available*. Has a similar collection been undertaken, and if so can you use that experience in developing this collection?

2 Focus only on the *relevant information* required — revisit the purpose and objectives of your collection, and ensure that your plans for data collection address these. That is, if the purpose is to monitor the age distribution of your service users, then make sure you ask for their date of birth. Avoid collecting any information that you don't really need and aim to collect the minimum amount of information required.

3 Plan your *analysis* — this will help to ensure your results address the objectives of the collection. Think about which other data collections you may want to be able to compare your data with. How will you achieve this? Will the use of national data standards help?

4 Consider the *simplest methods* of collection first. For example, it may be more appropriate to contact each service user personally and ask them two questions than devise an elaborate 3 page survey that is posted out for completion.

5 Also consider the *context* of the collection — factors such as who asks the questions, how they are asked, who answers them, and what audience the results are to be conveyed to. All these factors will affect the design of the collection. For example, if asking about the Indigenous status of clients, think about the appropriate person to ask the question and the way to phrase the question, do not make assumptions based on appearances.

6 Think about what *questions* you will ask. Careful attention needs to be given to the survey item development as slight differences in the wording of a question can result in quite different responses. For example, asking what 'plans' carers have for their disabled family member in order to estimate future service demand,

would elicit different responses from asking what carers 'expect' to happen.

7 Identify what *resources* (time, money and people) are necessary and available to run the collection. Recognise the particular skills that will be required for designing, running and administering the data collection, as well as how much it will cost and who will be doing the work. If resources are very limited, it may be best to consider a simpler option for satisfying at least some information requirements, rather than abandoning the collection altogether.

8 Decide how *often* the data will be collected. Is it a one off data capture to provide information on a particular issue at a point in time, or will the data be collected every week, month, or year to monitor trends over time? The data may be collected on a continuous basis as a by-product of administrative processes.

9 Determine and define who the *target population* for the data collection is going to be and work out how best to access this population.

6 How should you collect data?

This section provides an overview of the different methods available to collect data. When determining how best to collect data, it is important to consider the potential impact of the collection on everyone involved, for example, staff time taken up and service users being asked to provide personal information.

There are a variety of data collection methods available; some of the main ones are:

- *Questionnaires/surveys* — are made up of a list of standard questions that are answered by your target population. Questionnaires can be used to reach a large population (especially if sent through the post or via email), but often involve a time lag before you receive most of the responses and questionnaires that are sent out often have a low response rate which may affect your analysis. Questionnaires can also be used in a service setting, e.g. clients may be asked to complete a questionnaire when they visit an agency to access a service.

- *Administrative collections* — these are based on information that is collected as part of your service administration procedures. For example, information collected on standard client intake forms can be combined with data about staffing or finances to provide a profile of your client group and/or information about the average number of staff hours per client. This method requires all your data needs to be captured on your administrative system; different components of your administrative system can be combined by, for example, using the same client identifier on each system.

- *Interviews* — require a set of standard questions to be personally asked of your target population either face-to-face or over the phone. This method takes a personal approach and can be effective in some settings; it can also be expensive and time consuming.

- *Focus groups* — involve gathering representatives from your target population to a central location and discussing, as a group, a list of issues you would like explored. This method may involve a lot of preparation to organise a group of people to meet at the same time; but is an effective way of obtaining complex qualitative information. When running focus groups, it is important to obtain representatives from your target population that have a range of views. It is often desirable to hold multiple focus

groups, both for practical reasons (e.g. because of the size or geographical distribution of the target population) and also to allow a wider range of views to be captured. Ideally, focus groups should be conducted until no new opinions or issues are raised.

7 Developing questions and data items

Once you have determined the purpose of the data collection, what data will be collected, and how you are going to collect the data, the following points should also be considered in order to develop questions and data items:

- Develop data collection materials — you will need data collection materials for your collection, whether it is a list of questions, a survey form, an interview checklist, or (if you are running an administrative data collection) service intake forms, staff hours forms, etc. When developing collection materials remember:

  ‣ Make sure every question is in some way linked to the objectives of the collection.

  ‣ Consider the language style, tone and level of difficulty of every question. Try to keep your questions as short and as simple as possible, using language that will be easily understood by the target population. Also avoid leading questions that give respondents the impression there is a correct response, and questions in which the wording is such that effectively two questions are being asked at once.

  ‣ Test your questions to ensure the data items are appropriate and can be

collected accurately. Testing techniques commonly used include:

  ✧ pre-testing — informal testing of the questionnaire with colleagues or experts

  ✧ focus groups — informal discussions with people who represent the target population

  ✧ cognitive interviewing — in depth one-on-one interviewing in which participants are asked questions about the survey questions being tested

  ✧ observational studies — completion of the questionnaire in the presence of an observer

  ✧ pilot testing — formal testing of the questionnaire with a small representative sample of the target population.

- Data definitions — the *National Community Services Data Dictionary*, *National Housing Assistance Data Dictionary*, and the *National Health Data Dictionary* provide nationally consistent and endorsed definitions for a wide range of data items relevant to the community services, housing and health sectors in Australia. A comprehensive list of possible responses is given for each data item (to access the dictionaries see Section 11). It is strongly recommended that you consult these dictionaries when developing your data collection, so that the terminology you use is consistent with national standards. Using standard definitions and response categories increases the chance that you will meet the information requirements of multiple funding bodies. It also means you may be able to relate your data to data from other collections. The AIHW metadata on-line registry (METeOR) contains all Australian data standards in the community services, housing and health

fields. The Commonwealth State/Territory Disability Agreement National Minimum Data Set (CSTDA NMDS) Data Guide (AIHW 2004) and redevelopment report (AIHW 2003b), and the Australian Bureau of Statistics (ABS) Survey of Disability, Ageing and Carers (ABS 2004), may also be useful.

8 Data management considerations

It is important to consider how the data will be managed once collected. The following are some issues that require consideration when collecting and storing information.

- **Electronic collection** — depending on how much data you are collecting it may be worth using an electronic tool (software) to capture and/or store the data. An electronic tool can range from a basic electronic spreadsheet to a purpose built database tool. No matter what tool you use or how you use it (data capture or data storage) it is always advisable to test the tool to ensure it is fully operational, complies with your data needs and is easy to use.

- **Storing data** — under the *Privacy Amendment (Private Sector) Act 2000* agencies collecting data have responsibility for securely storing identifiable information about people— for example, information on age, sex or address that could be related back to an individual. For paper-based data holdings, identifiable information should be kept securely locked away when not in use. The minimum requirements are that information should be accessible only to those who are authorised, and that outside normal working hours, the

information must be stored in locked draws or cabinets. For data holdings that are stored on computers, security practices include, but are not limited to, providing individual user accounts with password protection, and automatic screen shutdown or automatic log-off.

- **Support and/or contact person** — it is always advisable when collecting data that people are fully aware of the purposes of the collection and their roles and responsibilities under the *Privacy Act*. It is important to have a key contact person who can provide further details on why the data are being collected.

- **Data guide** — good documentation of your data collection can help ensure that the data collected are of good quality. A data guide may contain any or all of the following:

  ▸ background information, including an explanation of the overall purpose, content and structure of the collection;

  ▸ information on the scope and counting rules of the collection;

  ▸ definition of key concepts;

  ▸ data definitions, coding classifications, and coding guidance, to ensure that data are recorded in a consistent way.

The data guide should be user-friendly and easy to understand. Ideally, data providers should have input into the development of the data guide, and it should be field tested and reviewed to ensure that it 'works'. The existence of a data guide will ensure that data can still be used and collections administered even if staff change (see AIHW 2004).

- **Training** — training sessions for those involved in collecting the data can help to ensure that information is coded and recorded consistently.

9 Analysing data and feeding back to data providers

A most important thing to consider when collecting data is: what are you going to do with the data once they have been collected?

The data you collect need to be organised in a meaningful way. The best way of doing this is placing the data into a central location—for instance in an electronic spreadsheet. When entering your data into a central repository it is important to assign codes or numbers to all responses. This will enable the data to be readily analysed using standardised tools—analysing data stored as text is generally more difficult. The national data dictionaries include coding responses for most data items, which could be used when you are entering data.

Ideally you should plan your data analysis after you have determined your collection objectives and *before* you design the collection, or begin collecting data. The analysis you want to do may influence the items you decide to include and/or the coding categories you use to capture information.

For example, one of your objectives may be to determine the age distribution of clients. If you plan to do detailed analysis of client age you may decide to record date of birth. If you just need more general information on client age, it may be sufficient to use broad age groups (e.g. '0–14','15–25', etc.)—this may be a less invasive approach to collecting personal information, but would not support very detailed analysis; further, your data would not be useful in later years as clients' ages would become increasingly difficult to estimate.

Reliable data analysis is dependent upon the amount and quality of the data. Some data collections may be very small and can support only basic counts of frequencies, such as numbers, percentages or averages. Most data analyses will require comparisons among groups, for example the average age of people accessing accommodation services compared to the average age of people accessing respite services.

As detailed in Data Starter 1, when interpreting data it is important not to:

- generalise across groups or from small samples to large samples;
- jump to conclusions about the results you have found; or
- assume a cause and effect relationship where there is no evidence that one exists; in particular correlation between two items does not prove causality.

When reporting data it is essential that you address the objectives of the collection, that is, answer the questions you set out to answer!

It is also recommended that you feed back the findings of your analysis to the people who provided you with data. Doing so will show that you have used the information as you intended, as well as providing them with useful information. Data providers are more likely to be supportive of a collection if they can see that it is providing useful information.

10 Data collection checklist

Points to remember when collecting data:

- ✓ Identify the objectives of the data collection, your information needs and the question you want answered
- ✓ Plan your data analysis
- ✓ Design the collection
- ✓ Prepare the collection materials
- ✓ Test the materials
- ✓ Run the collection
- ✓ Compile the data
- ✓ Analyse the data
- ✓ Interpret and report the results
- ✓ Feed back to data providers
- ✓ ANSWER YOUR QUESTIONS!

11 References

ABS (Australian Bureau of Statistics) 2004. Disability, ageing and carers: summary of findings. ABS cat. no. 4430.0. Canberra: ABS.

AIHW (Australian Institute of Health and Welfare) 2000. Data Starter: A project of the Disability Data Reference and Advisory Group. Issue 1, December 2000. <www.aihw.gov.au/publications/index.cfm/title/6361>

AIHW 2003a. National Housing Assistance Data Dictionary version 2. AIHW cat. no. HOU 89. Canberra: AIHW. <www.aihw.gov.au/publications/hou/nhadd02/nhadd02.pdf>

AIHW 2003b. Australia's national disability services data collection: redeveloping the Commonwealth State/Territory Disability Agreement National Data Set (CSTDA NMDS). AIHW cat. no. DIS 30. Canberra: AIHW. <www.aihw.gov.au/publications/dis/cstdanmds/cstdanmdsc00.pdf>

AIHW 2004. Data Guide: data items and definitions 2004–05. Commonwealth State/Territory Disability Agreement National Data Set collection. Canberra: AIHW. <www.aihw.gov.au/disability/csda_public/0405_data_guide.doc>

AIHW 2005. METeOR: metadata on-line registry. <meteor.aihw.gov.au/content/index.phtml/itemId/181162>

National Community Services Data Committee 2004. National Community Services Data Dictionary. Version 3. Canberra: AIHW. <www.aihw.gov.au/publications/hwi/ncsddv3/index.html>

National Health Data Committee 2003. National Health Data Dictionary. Version 12. AIHW cat. no. HWI 43. Canberra: AIHW. <www.aihw.gov.au/publications/index.cfm/title/9812>

Acknowledgments