*'FROM DATA TO INFORMATION'*

# DATA USE GUIDELINES

## FOR STANDARD MEASURES COLLECTED IN THE NEW ZEALAND MENTAL HEALTH SYSTEM

Te Pou o Te Whakaaro Nui
The NATIONAL CENTRE of MENTAL HEALTH RESEARCH, INFORMATION and WORKFORCE DEVELOPMENT
www.tepou.co.nz

INFORMATION
USING INFORMATION

# Table of Contents

# Table of Figures



# List of Tables

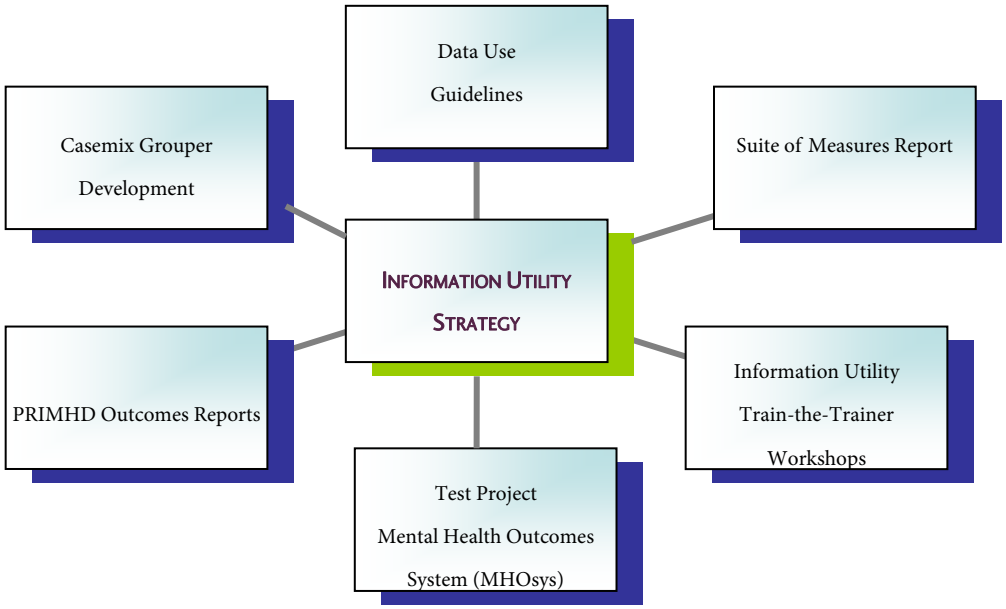
# THE AIM OF THIS DOCUMENT

This document aims to assist all people who work with and/or use mental health services (including service users, their significant others, clinicians and other mental health workers, managers, analysts, and policy-makers for mental health services at all levels) to become more adept at turning the data collected from Mental Health 'Standard Measures' of Assessment and Recovery (MH-SMART) into useful information.

Maximum benefit can be obtained from standard measures when the data is being collected, presented and used consistently across all mental health services. This document offers guidance to the types of questions data may help us to answer and aims to facilitate the development of consistent approaches to the collection, analysis, presentation, and use of mental health information throughout New Zealand.

These guidelines sit alongside Te Pou's Information Utility Strategy (2008), and a range of national clinical data utility training strategies (as shown below) to support all stakeholders to use information accurately and consistently.

**FIGURE 1: TE POU'S KEY INFORMATION UTILITY WORKSTREAMS**




In recent years there has been considerable effort in New Zealand to establish the collection of standard measures to give more information about mental health services, the people who use them, and the outcomes for these people.

The Programme for the Integration of Mental Health Data (PRIMHD) is the latest stage of this initiative and will incorporate, refine, and extend the data that has been collected nationally for several years through the Mental Health Information National Collection (MHINC) and, in District Health Boards (DHBs) since 2005, through the Mental Health Standard Measures of Assessment and Recovery (MH-SMART) initiative.

The MH-SMART initiative includes collection of the Health of the Nations Outcomes Scale (HoNOS) family of measures. HoNOS is the first of a proposed series of standard measures of assessment (also commonly known as 'outcome measures').

Further standard measures including a service user self-rated measure, a measure of functioning, an Alcohol and Other Drug Use measure, and the Hua Oranga Maori mental health measure are at varying stages of development as described in the Te Pou Outcome Measures Report to the Ministry of Health, 2009.

The collection and use of standard measures data in New Zealand mental health services is still early in its development. The specific approaches used with data are likely to evolve over time. This evolution will be reflected in future versions of documents similar to this one. However, many of the guidelines and principles outlined will remain consistent as further standard measures are introduced and new approaches are developed and refined.

# Acknowledgements

This document has been written by Malcolm W. Stewart PhD, DipClinPsych, with guidance from a steering group consisting of Professor Graham Mellsop (Psychiatry Department, University of Auckland), Allen Morris-Yates (Director PMHA-CHMS, South Australia), Jim Burdett (Mind and Body Consultants, Auckland), Helen Mitchell-Shand (Wairarapa DHB), Alison Bower (Contractor, Te Pou), Monique Faleafa (National Manager, Le Va, Te Pou), Hine Martin (National Manager Information, Te Pou), Laurie Hakiwai (Clinical Lead Implementation, Te Pou), Sandra Baxendine, Information Analyst, Te Pou) and Mark Smith (Clinical Lead Specialist, Te Pou). Guidelines for the use of Maori Data are based on material written by Leah Whiu. This document draws on the unpublished document *"The Use of MH-SMART Standard Measures Data: A Guide to the Improvement of Mental Health Services"* authored by Jennifer Chipps, Graham Mellsop, and Malcolm Stewart in 2003-5 and extensively reviewed by an expert group under the auspices of the Mental Health Research and Development Strategy for the MH-SMART initiative. Our thanks to many others who have provided kōrero, ideas, and inspiration for this document.

# Introduction

## Data and Standard Measures

### What is data?

For the purposes of this document, data is defined as information that is collected about a person or another entity (such as a mental health service) that is able to be represented as numbers. The numbers may represent a variable amount (e.g. the number of days a person is in an inpatient unit) or a category (e.g. all males are coded as "1" and all females are coded as "2"). This is not the only possible definition of data (for example, narrative data such as information about a person's life story is not always easily quantifiable in numbers) but applies well to the data discussed in these guidelines. Data, once analysed, provides the type of information that prompts further questions or enquiry and guides informed decision-making linked to service improvement.

### What are standard measures?

For the purposes of this document, standard measures are assessment tools that allow consistent quantification of some aspect of the service user's experience or status. They are often standardised rating scales and other types of measures (such as demographic information, biomedical measures, information about service activities on behalf of the service user, and satisfaction measures) that are completed using consistent rating criteria by or for all service users. Some standard measures are collected once only and others are completed by or for the service user at consistent times during the service user's contact with a service (e.g. on admission, at determined intervals, and on discharge).

To ensure they are measured consistently between different raters and over time, the definition of what is being measured, and the definition of what constitutes various scores on the measure, need to be well-defined. Standard measures are usually chosen only if they have shown evidence of **reliability** (they yield the same result in the same situation if completed by different people or at different times) and **validity** (they measure what they are purported to measure). Standard measures give us a way of speaking a common language about our services. To make this common language effective we need to collect standardised measures and data elements in accordance with an agreed protocol, and then analyse and report that data in accordance with agreed frameworks and protocols.

### What are outcomes?

The outcome for a service user is the change in their life. This change can be defined and measured in many ways. The outcome we can readily measure is the change brought about by the combination of the efforts of the service, the service user and many other factors that might influence the status of the service user (e.g. the efforts of other social support services and changes in the service user's situation). Mental health services may contribute to these outcomes through direct interventions, by supporting service users to access resources outside the mental health system that contribute to positive outcomes, and by encouraging or supporting the service user to bring about positive gains for themselves (Miller & Rollnick, 1991).

Although the service may contribute to these outcomes in various ways, we need to be aware that the outcome we measure is not just the *service outcome* (i.e. changes for a service user that result from contact with that service) but is the outcome of all contributing factors. The extent to which outcomes result specifically from the activities of the service often cannot be determined unless rigorous research methods are employed (Mirin & Namerow, 1991).

From a recovery perspective (and a client-centred perspective) it is the *outcome* rather than the *service outcome* that is the most important, because this reflects how the service user's life has changed. Also, the service outcome cannot be larger than the "overall" outcome. Therefore standard measures data does help us understand something of the outcomes of service users and the outcomes achieved by services.

### Why collect standard measures in mental health?

As providers, funders, and/or users of mental health services, our primary aim is that appropriate and effective services are provided to those who need such services. This is clearly a challenging task because not only must services be delivered effectively; they must also be delivered efficiently and equitably.

To be able to achieve this we need to be able to understand:

- the needs of the person with a mental health issue,
- the costs - in all senses - of the services provided, and
- the outcomes arising from services (Coombs, 2001).

We need reliable information to be able to understand these. Standard measures are an important way this range of information can be obtained.

Albert Einstein is quoted as having said

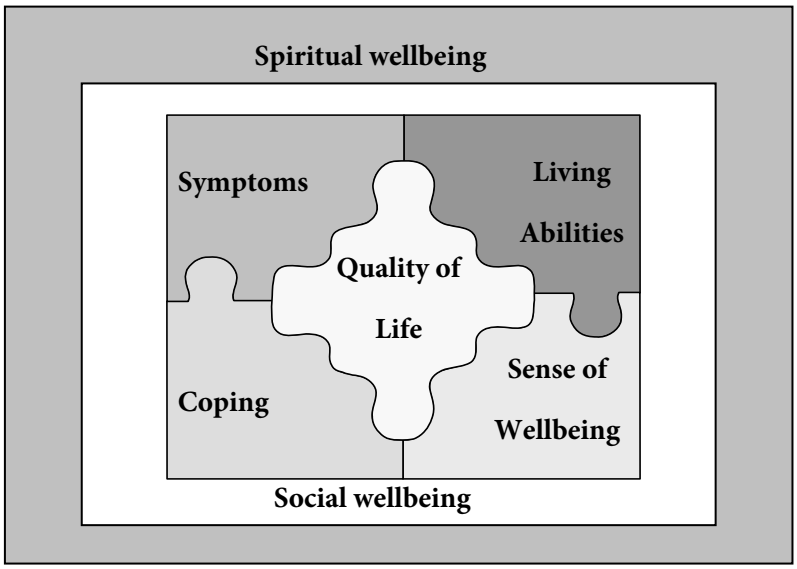
*"Not everything that can be counted counts, and not everything that counts can be counted."*

This is true in mental health. Some aspects of the experience of mental health or illness and the complex process of providing mental health services are difficult if not impossible to measure and count (Lakeman, 2004). However, if Einstein's statement was put another way it would also be true: *"Many things that can be measured are important, and many things that are important can be measured."*

For example, some subtleties in the interaction between a clinician and a service user or some of the "ingredients" of an effective mental health intervention are hard to measure. However, most would agree that this intervention should lead to outcomes for the service user that are significant enough to them to be measurable in some way. These measurable outcomes may include elements such as increased wellness, increased ability to live life well, reduced experience of symptoms that are distressing or disabling, and satisfaction with services. This is consistent with a Recovery Approach (Deegan, 1996). The "change" may in some instances be

an absence or slowing of further deterioration, but should usually include improvement on one or more aspects of the person's broadly-defined wellness. Some important aspects of wellness in mental health are shown in the following diagram.

**Figure 2: Some dimensions of mental health**




The *Blueprint for Mental Health Services* (Mental Health Commission, 1998) identified the development and use of routine outcome measurement (i.e. the use of standard measures) as a priority for mental health services in New Zealand. This was because of the key role that standard measures and outcome measurement can play in identifying and learning from areas of best practice and in guiding other quality improvements for mental health services. Since the *Blueprint,* considerable progress has been made toward developing standard outcome measures. However, further work is now needed to ensure that the data being collected is used in ways that justify the effort of collection. Data being collected should meet the needs of service users/tangata whaiora, their family/whanau/significant others, mental health services and funders.

The data we collect will not answer all our questions. Often its major role is to generate questions rather than answering them. However, it can be helpful in answering some of our questions and in making us aware of other questions we should be exploring. Therefore, the use of standard measures data can help us reflect and inquire more about our practice and wider issues. Examples of how this type of data can be presented and how the resulting information can be used to improve service delivery are provided later in this document.

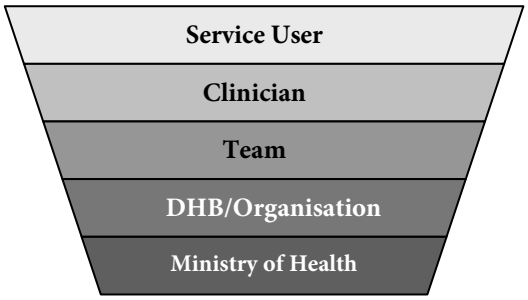
### Who can use and benefit from the use of standard measures?

The main reason for collecting standard measures is to help with better decision making. Data can help improve our decision making at several levels, from helping to inform the recovery plan made by clinicians and an individual service user, right through to assisting the Ministry of Health (MoH) and the government to better plan the future of health services in New Zealand. Stakeholders who may use and benefit from the use of standard measures data fall into five general levels as follows:

1. The service user (and their family/whanau)
2. The clinician/mental health worker
3. The team
4. The healthcare provider/organisation
5. The national level/Ministry of Health

### Levels of Data Use

**Figure 3: Levels of Data Use**




Users of data at each of these levels of analysis will have access to different amounts and types of information. For example, the service user will primarily have access to their own information, which can contribute to comprehensive clinical and needs assessments of that individual. In contrast, at the healthcare organisation and MoH level, data that is aggregated across groups of service users will hold more interest (Mellsop & O'Brien, 2000). At the clinician and team levels, individual data will help with recovery planning whilst aggregated data will inform service delivery. Users at each level can use standard measures data to answer both similar and different types of questions. Even when the same question is being asked at different levels the span and nature of the answer (for instance, the degree of detail and the extent to which it applies to particular service users or to whole services) will vary with the level of data use. Some applications of standard measures data are summarised below, organised by the level of user they may be particularly relevant to.

- **Individual service users** benefit because the collection of standard measures data as part of the ongoing assessment process can improve quality and comprehensiveness and strengthen the monitoring of changes in their mental health situation and experience over time. Thus planning for recovery can be strengthened.

  Standard measures data also gives service users a point of feedback and dialogue with mental health staff that can be very useful in improving shared understanding.

The PRIMHD data set will eventually include service user self-ratings of outcome and satisfaction with service. Service users will then have another way of feeding back on their experience and influencing the development of services.

Service user advocacy groups can use aggregated data to examine the value of service practices and to lobby for planning and funding decisions to be backed up by statistical evidence (Burdett, 2005).

Insights gained from research or analysis of standard measures data can lead to improvements in the delivery of care. (Cheung & Strachan, 2007; Trauer, Eager, & Mellsop, 2006).

- **Service users' significant others**, with due attention to confidentiality, can also benefit from the use of standard measures. The data gained from collection of these measures can assist significant others to understand and be constructively and appropriately involved in the service user's recovery process.

  Standard measures data may give family/whanau a point of feedback and dialogue to optimally support the service user's recovery and possibly influence the care received by the service user.

  For detailed guidance about confidentiality and information sharing with significant others, please see the *Health Information Privacy Code* (New Zealand Privacy Commissioner, 1994) and subsequent amendments.

- **Clinicians and mental health workers** also benefit because information gained from standard measures data can make it easier to do their jobs effectively and efficiently. Standard measures data adds value as a summary of a competent and comprehensive clinical assessment.

  Clinicians and mental health workers are interested in providing optimal treatment and improving outcomes for service users. Standard measures data contributes to problem/need identification, solution finding, better support, and recovery planning. Standard measures data helps clinicians develop and revise treatment decisions and monitor service user progress.

  Using standard measures data provides a method of evaluating the effectiveness of interventions. It allows the mental health worker to reflect on their own performance and to make decisions that improve performance and satisfaction. Being able to work in this way increases work satisfaction and decreases the risk of frustration, high levels of stress, and burnout (Felton, 1998).

  The measures can assist clinicians in identifying service users who might benefit from additional/alternative approaches.

  Using standard measures such as HoNOS can provide data that can be a useful point of dialogue with the service user as described above.

  While there is a cost in terms of time taken to complete the measures, when the data is well used the benefits of collecting standard measures is likely to well outweigh the "cost" in time.

- **Mental health teams'** benefit by better understanding what they are doing and what they are achieving. They can use the standard measures data to develop better ways of delivering safe, effective, and efficient service.

  Standard measures can also be used to improve discussion within the clinical team and to facilitate planning for the individual service user. The service user and, when appropriate their family/whanau, can also be involved in this team discussion and planning.

  As this data is typically available to all team members, it can be used collaboratively to develop service improvements. This promotes inclusive approaches to service development and clinical governance that can increase both satisfaction and satisfactory outcomes.

- **The health care organisation and the government** also benefit because information gained from aggregated data can be used to better understand the needs of service users at population level. An important objective of collecting standard measures is to enhance mental health service delivery. Standard measures data can contribute to decisions around planning, developing and funding services that best meet the real needs of the people they serve.

  Data from standard measures can be aggregated (combined for groups of service users) to better understand many aspects of service function and to understand how to improve services (e.g. Ganju, 2006).

  Data, particularly related to groups showing less improvement, can be studied to look for commonalties and possible solutions.

  Data on groups showing improvement can also be studied for commonalties in the hope of achieving further enhancement for this group or others.

  Data from standardised measures should be incorporated into quality improvement feedback loops at all levels. Often standard measures data will not provide us with answers, but will help us ascertain what we need to know and guide us towards possible answers.

### What needs to happen for standard measures to be useful?

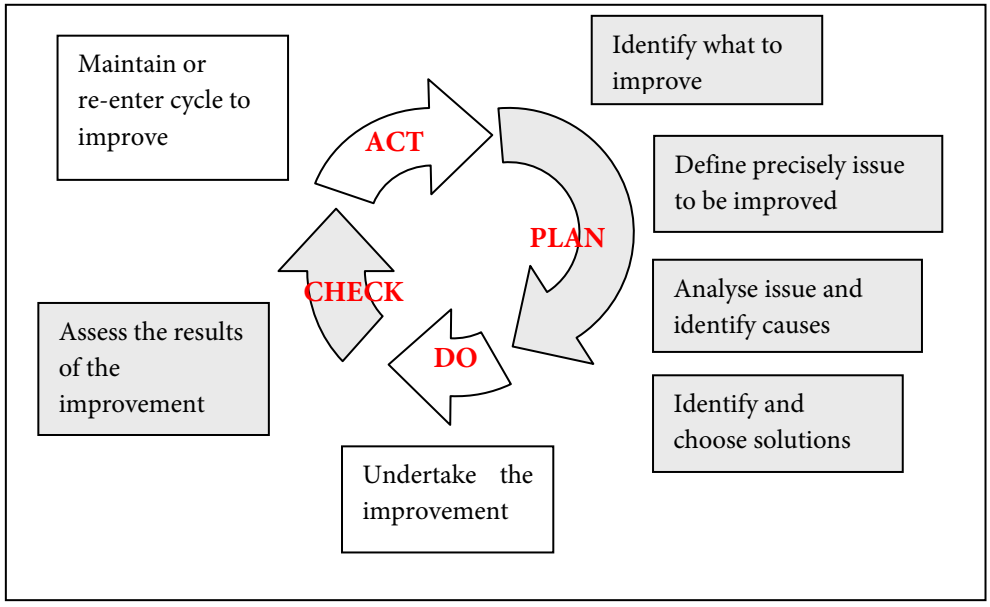
As discussed above there is huge potential for the data collected from standard measures to help us improve mental health service provision, making outcomes better and more satisfying for the stakeholders described above. However, to achieve this potential, data needs to be collected and used routinely, reliably, and intelligently. To achieve these three processes must occur:

- Training in, and use of, standard measures and documentation to **collect** data that is relevant to improving clinical and management decision making.
- Development of systems that allow relevant, appropriate and prompt **reporting** of the information collected.
- Clinicians, managers, and others then need to **use** the information collected to inform clinical practice and the management of resources. Training in information utility will be required for all stakeholders.

When these three processes occur we can get a clearer picture of the complex and dynamic activities that are involved in the mental health system. This kind of information is important for successful completion of the quality improvement cycle to facilitate improvement (Britz, Emerling, Hare, Hoerl, Janis, & Shade, 2000). The processes involved in the quality improvement cycle are shown below.

### The Quality Cycle

**Figure 4: The Quality Cycle – Continuous Quality Improvement – (CQI)**

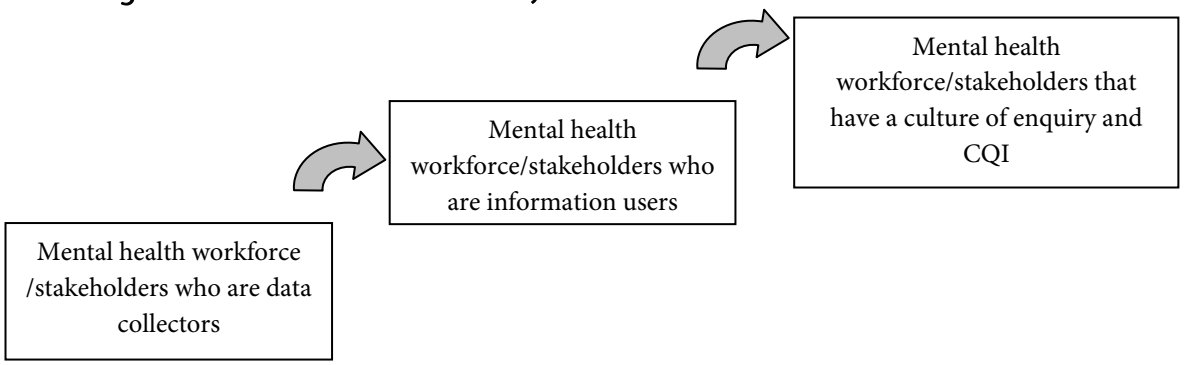



**Note:** Processes shaded in grey are most strongly dependent on information such as standard measures data.

## Moving from data collection to a culture of enquiry and CQI

Mental health and related services in New Zealand are required to routinely collect a range of data including standard measures. This data collection is important, but is only the first step towards really gaining from the effort involved in collecting the information. The journey we are on towards *really using this information* could be shown as follows:

### Figure 5: The Information Utility Journey




This document aims to help us move from being collectors of data, such as standard measures, to being users of this information to benefit service users/tangata whaiora, mental health staff, the mental health system, and the funders of that system (that is, eventually, all of us who pay taxes!) The next step on the journey is a system where questioning how things can be made better is a routine part of the culture of the mental health system. Good quality data and well-used information underpins the development of this kind of culture.

## Why have Guidelines for the Use of this Data?

As with almost any tool there is the potential for using or misusing the data collected from standard measures. The analysed data provides information that can help us better understand the mental health system, what it does and how to make it work better. It can also be used to raise questions that will require answering by other means (such as discussion, self-reflection and trying things out) and that may lead to improvements in the mental health system. However, there is also the risk that data is misused, leading to unjustified or incorrect conclusions. This document offers guidelines to using standard measures data (mental health information) in constructive, valid, and ethical ways.

The development of analysis and reporting tools to use this information is ongoing, and therefore there is limited opportunity to describe the exact tools that will become available. However, this document presents general approaches to the collection of data and the use of information arising from data.

# Guidelines for the collection and analysis of standard measures data

(Modified from Chipps, Mellsop, and Stewart [unpublished]),

1. Data should be collected as part of the clinical process and support a recovery approach. It should contribute to better outcomes for service users/tangata whaiora.
2. For standard measures to be useful, all services and individuals collecting them must consistently use standard data definitions, classifications and coding systems.
3. Data should be collected once, as close to the source as possible, and used as many times as required to meet different information requirements, in keeping with the purpose for which it was collected.
4. The data collected should be meaningful in a New Zealand context, and should be collected in a culturally sensitive manner.
5. Collection, analysis and use of data are an integral part of mental health service delivery.
6. Staff collecting data must be trained in the use of standard measures to ensure consistency across services and between healthcare providers. This training will be updated or repeated as required to maintain consistency.
7. The coding of the data must be accurate and there should be mechanisms for checking on data authenticity.
8. When codes for categorical data are developed, the categories should be unambiguous and non-overlapping and clear classification guidance should be supplied.
9. All data collection, analysis, and reporting systems should protect service user/tangata whaiora confidentiality and privacy.
10. Compliance with data collection protocols should be high (above 80%) otherwise the utility of data for decision making, service development and planning is compromised.
11. Completion of standard measures should be high (above 90% with all items in the individual measures completed). Partial collection of data also limits its usefulness for decision making and planning.
12. Data collection, verification, management and analysis processes should be developed and implemented for all healthcare providers.
13. All data collection and verification processes should ensure appropriate data standards are utilised and a high level of confidence, quality, and consistency of information collected.
14. All data collected under PRIMHD will be owned by the healthcare provider that collects it but must be stored in a Ministry approved database and format and be available for quality management purposes.

15. Data should be shared wherever possible to foster the development of networks and partnerships.
16. Complete meta-data documentation is required for every dataset. Meta-data is the information that is necessary to understand the data that is collected. It includes elements such as the description of the data fields, categories within a variable and information about the characteristics of the data storage method.

## Improving data collection quality

There is an old adage in computing that goes *"Garbage in, Garbage out"*. In this context it means that if standard measures data is not accurate (or complete) then the conclusions that are drawn from that data will also be unreliable or inaccurate. Therefore, to be of value to any of the stakeholder groups, standard measures that are collected need to be as accurate and complete as possible. Ways of ensuring that data collection is as accurate and complete as possible include:

**The service user**

- *Be as truthful and open with your clinicians and mental health workers as you are able.* There is much important information that they can only know if you tell them.
- *Be involved in the standard measures collection process by*
    1. making sure that information you give is accurate,
    2. letting people know if information is recorded inaccurately, and
    3. ask to discuss your standard measures (e.g. your HoNOS results) with the clinician who has completed it.

**The clinician/mental health worker**

- *Ensure that a comprehensive assessment is undertaken.* The quality of information and care that can be delivered depends on the quality of the assessment. This does not mean that every clinician should be asking every question, but it means that a comprehensive assessment should be undertaken and its results should be recorded in a way that is accessible by those who need the information to support recovery.
- *Record the information accurately and in a timely fashion.* Data that is recorded as soon as it is collected is more likely to be recorded accurately and is more likely to be of use to others.
- *Become an accomplished data collector.* Complete the standard measures use/collection training. New ways of collecting or recording data may seem difficult and unwieldy at first but typically, with practice, you will find yourself able to complete the information requirements quickly and more naturally as part of your assessment. Use website/resources to remain up-to-date.

- *Use the information and encourage others to do so also.* One of the best incentives for improving (or encouraging the improvement of) the quality of standard measures collection is knowing that information will be used by you, your colleagues and service users. Other sections of this document describe some ways this can be done. Getting in the habit of using this information and encouraging others too also will help improve the quality of the data.

**The team**

- *Make use of standard measures data as a regular part of your team process.* Use of standard measures can have very positive impacts on team process and outcome (Lambert, Harmon, Slade, Whipple, & Hawkins, 2005). Other sections of this report describe some ways this can be done.
- *Be prepared to sustain efforts to introduce routine use of standard measures.* Typically, just training staff is not enough. Embedding routine use of standard measures for long enough so staff see benefits takes sustained team leadership and championship by "early adopters".
- *Provide team members with feedback about their completion of measures.* Periodic feedback to team members about their completion of measures is helpful. In some settings this has involved celebrating or rewarding high levels of completion and problem-solving when levels are low.
- Run team inter-rater reliability sessions using case studies to discuss and ensure quality of interpretation of standard measure rating.

**The organisation / national level**

- *Ensure collection processes are as straightforward, "intuitive" and efficient as possible.* The more integrated with service user-related data collection requirements the standard measure collection is, the better. Also, the simpler and easily used the system is by people with a wide range of attitudes and aptitudes towards computers, the better.
- Ensure that data is provided back to organisations, teams and, as appropriate, individuals in a timely manner, and in a form that is comprehensible and useful to them. Perhaps the best way of ensuring improved data quality is to ensure that the data is of use to the people who are collecting it. Feeding back analysed data in useful ways, for instance to assist with service planning and development in a timely manner, becomes an incentive for higher quality data collection.
- *Build in data checking, but avoid it from becoming disruptive.* Point of collection data checking and feedback are powerful ways of improving data quality. If data errors or inconsistencies are detected at the time of data entry, ideally they can be fed back to the clinician immediately for review and correction. Detection of errors should not however, lead to the loss of data so it needs to be re-entered. This causes frustration and lowers compliance with data collection.

- *Develop clear, straightforward processes for later signalling and correction of data errors and inconsistencies.* Where correction or completion of data is needed later, clear and efficient processes should be developed for making the needed changes. The additional burden that these processes add to various parts of the mental health workforce needs to be carefully considered.
- *Ensure information users have adequate training and support to be able to use data effectively*. This has been recognised by clinicians as important for improving the acceptance of standard measures (Callaly, Hyland, Coombs, & Trauer, 2006).

## Data Analysis

(Modified from Chipps, et al. [unpublished])

1. When analysing data, care should be taken to ensure that appropriate clinically and epidemiologically informed statistical and presentational techniques are employed so that accurate conclusions can be drawn. If not available within the organisation, competent statistical and analytical advice should be sought when new analyses are undertaken.
2. When analysing aggregated data, statistical significance should be considered in evaluating the difference between groups and the difference of results from a particular level. If groups are not statistically different, they should not be interpreted as being different. Even when analyses do show statistical significance, the clinical significance should also be considered.
3. For samples of less than 30, many statistical techniques may become unstable or imprecise. These difficulties can sometimes be overcome by combining multiple years of data, collapsing data categories, or expanding the geographic area under consideration. All statistical techniques should be used appropriately.
4. Data should not be aggregated and reported when working with very small numbers due to confidentiality issues in small populations and the lack of data reliability associated with a small number of events.
5. Analysis and reporting on ethnic data should meet the criteria described in the "Guidelines for the Use of Maori Data" included later in this report. Maori input and advice should be included throughout the process of preparing, analysing and interpreting of data. Important issues include articulating the reasons for analysing data by ethnic groups, and clarification of how ethnicity was recorded.

## Asking questions – growing a culture of enquiry & CQI

Stakeholders most benefit from standard measures by using the data either to answer specific questions, or to stimulate new questions.

Some of the questions that are stimulated by looking at the data may be answered from standard measures data. Other questions may require other approaches, such as reflection and discussion, seeking expert opinions, literature reviews and trying out and evaluating change. It is very important to be clear about the question that you are posing in either of these situations. If you are not clear about the question you are asking, it will be much harder to answer.

Another reason for being very clear about the question you are posing is that collected data can be analysed in thousands of different ways, some of which may be helpful or not. It is easy to be overwhelmed by the number of possible analyses. One of the best ways to avoid being overwhelmed or producing unhelpful results is to *start with a clear question that will be useful to have answered.* This will make it much easier to decide what analysis may help you to answer the question.

One guideline for posing questions is that a relatively straightforward "simple" question is often better than a more complex question. In the words of one health statistics consultant,

> *"Ask a simple question, you will often get a complex answer; ask a complex question, and you will often get no answer at all"*

Some general questions that might be answered by using standard measures include:

- **Who do we serve?** Getting a better understanding of the people who use mental health services is potentially helpful for many of the levels discussed above. The PRIMHD data set contains a wide variety of demographic, clinical, and service information that can be used to help characterise the people served by different mental health services. It will often be of value to use this information.

- **What are the outcomes of our service?** As discussed earlier, changes in standard measures can be regarded as an outcome for the service user. Changes are not purely a result of the activities of the service, also being the result of the efforts of the individual and other factors. Despite this, repeated use of standard measures data can give valuable insights to service delivery. At an individual level, the results of standard measures data will inform needs assessment and planning for recovery. At an aggregate data level, for example, if outcomes appear small for particular groups or on particular measures, this may indicate the need for closer consideration of what is being done to increase future outcomes.

- **Who do we work more, and less, effectively with?** Standard measures data can be used to identify groups with whom a service works more or less effectively. These groups may be defined on age, diagnosis, ethnicity, chronicity of condition, level of social disadvantage or other characteristics. Groups with whom a service works more effectively could generally be expected to show greater (improvement) change. Using standard measures data to identify groups that show less (improvement) change sets the stage for developing more effective services for these groups.

- **How do the outcomes of a service compare with similar services?** After casemix adjustment (see page 32), the outcome of different services can be compared to identify where there is best or improving practice which can be shared with other services. In this situation casemix adjustment of the data is required to ensure that the outcomes observed are likely to be due to differences in services rather than differences in the types of people served. Benchmarking is a technique used to answer this type of question. Benchmarking often involves:

  1. comparing different services with each other on the same factors (e.g. performance on standard measures), or
  2. identifying best practice in a certain area and using this as a standard which other services can be compared with.

  Benchmarking aims to encourage the spread of best and improving practice, and to reduce unhelpful variation in practice across different services or programs.

- **How do the outcomes of services relate to service costs?** Standard measures can include data on cost of service provision, including both the fixed costs (costs of running a service that are not attributable to any one service user) and variable costs (costs that directly relate to providing service to one individual service user).

The following table lists more examples of questions that might be asked and answered by different stakeholders using information from standard measures. It also indicates the type of information that may be used to answer these questions.

**Table 1: Levels of Analysis for Different Stakeholder Groups and Types of Questions that may be addressed**

| Level of analysis | Question(s) addressed by report | Type of report |
|---|---|---|
| Service user level | • As below (Clinician level)<br>• Assessing and enhancing a shared view between service user, clinician and family regarding current status and concerns. | • Timely longitudinal data for the individual service user showing progress over time and current areas of concern |
| Clinician level | • How can the service user's recovery pathway best be supported at present (data to assist planning for recovery)?<br>• How has the status of this service user changed over time?<br>• What are present areas of concern with this service user? | • Timely longitudinal data for the individual service user highlighting change/progress and current areas of concern |
| | • What outcomes are achieved by service users, compared to expected outcome as defined by Focus of Care? | • Aggregated outcomes for all people on the clinician's caseload differentiated by Focus of Care |
| | • What groups of service users (as determined by chosen characteristics) achieve greater or lesser outcomes? | • Aggregated outcomes for all people on the clinician's caseload differentiated by characteristics of the service user or service. |
| Team level | • What outcomes are achieved by service users who access this team, compared to expected outcome as defined by Focus of Care? | • Aggregated outcome for all service users who access this team differentiated by Focus of Care |
| | • Do service users with different characteristics or who receive different interventions achieve different outcomes from other groups (could be Casemix adjusted). | • Aggregated outcomes for groups of service users, differentiated by clinically relevant characteristics (e.g. demographics, treatment type or clinical characteristic). |

| | | |
|---|---|---|
| | • What outcomes do people on this team's caseload achieve compared with people on others' caseloads after adjusting for any differences in cases the teams work with? | • Aggregated casemix adjusted outcomes for service users who access the team compared to average outcomes of similar teams (e.g. similar teams in the same healthcare provider, region or nationally) |
| | • What outcomes do people on different clinicians' caseloads achieve, compared to expected outcomes defined by Focus of Care, after adjusting for differences in the types of service users worked with? | • Casemix-adjusted comparison of outcomes for people on similar clinicians' caseloads differentiated by Focus of Care. |
| | • Uses acuity to better balance workload (case weighting) instead of basing it on number of service users. | • Number of clients in Full-Time-Equivalent (FTE) caseload scoring 3 or 4 on two or more items on the HoNOS at last assessment. |
| Healthcare Provider level | • How do the outcomes of service users who access the different teams compare after adjusting for the differences in the types of service users between the teams? | • Casemix-adjusted comparisons of outcomes for service users who access similar teams of the healthcare provider. |
| | • What outcomes do service users who access different teams achieve compared to expected outcomes defined by Focus of Care, after adjusting for differences in those service users? | • Casemix adjusted comparisons of the outcomes for service users who access similar teams of the healthcare provider differentiated by Focus of Care |
| | • What groups of service users of the healthcare provider (as determined by Casemix classification) achieve greater or lesser outcomes? | • Comparisons of the outcomes for service users of all services, differentiated by Casemix classifications. |
| | • Which teams appear to produce better outcomes with particular types of service users? | • Comparisons of the outcomes for service users who access different teams differentiated by Casemix classification. |

| | | |
|---|---|---|
| National/ MoH level | • How do the outcomes of service users of different healthcare providers compare after adjusting for differences amongst service users? | • Casemix-adjusted comparisons of the outcomes for service users of different healthcare providers. |
| | • Which healthcare providers appear to produce better outcomes with particular types of service users? | • Comparisons of the outcomes for service users of different healthcare providers differentiated by Casemix classification. |

**Notes:** In the table, outcome is defined as change in the standard measures over one or more episodes of care. The types of questions addressed in this table relate to questions that the HoNOS may assist with answering. Casemix adjustment involves adjusting the data to take account of different kinds of service users seen by different clinicians or services. This allows more meaningful comparison across different services. After Chipps, Mellsop, and Stewart (unpublished).

.

# Horses for courses: different data uses for different questions

## Types of data

Different types of data can be used to answer different questions and require different types of analysis. Some dimensions on which data can be differentiated and, in general terms, analysed follow.

### Dimension 1: Individual or aggregated data

**Individual data** is data from a single person. It is often collected at repeated points in time so that changes over time can be observed. Some individual data only needs to be collected at one point in time. Individual data will not usually require any particular statistical analysis. Where there is longitudinal data (the same information collected at different times) the data is often easier to understand if it is presented as a graph rather than as a series of numbers.

**Aggregated data** is data from more than one individual (e.g. more than one service user) which is grouped together so it can be analysed. For example, the admission HoNOS scores for all service users using a respite service might be aggregated so the mean admission HoNOS score could be calculated. Aggregated data often needs statistical analysis to make it interpretable. For example, if two services were found to have different mean admission HoNOS scores, statistical analysis would be needed to know if the difference in means reflected a real difference between the services, or if it was not really significant.

### Dimension 2: Categorical, ordinal or continuous data

**Categorical (nominal) variables:** The different levels of a categorical variable define different groups (categories) but these different levels can not be meaningfully ranked from lesser to greater. For example, apples, oranges and pears could be assigned values 1, 2 and 3 respectively, but the specific values tell you nothing about the relationship between the three groups. Presentation and statistical analysis of categorical data often relates to the number (frequency) in each category.

**Ordinal (Ranking) Variables:** The different levels of an ordinal variable tell you something about their relative value, but do not tell you the exact relationship between the values. For example, if a satisfaction rating scale had four levels, “not at all”, “a little”, “moderately” and “very”, and these were assigned values 1, 2, 3 and 4 respectively, we know that the higher numbers tell us that the person was more satisfied, but we cannot say that someone with a value of 4 (very satisfied) is twice as satisfied as someone with a value of 2 (a little satisfied). Special statistical techniques are used for analysing ordinal data.

**Continuous (Interval) Variables:** The different levels of a continuous variable do tell you about their relative size and also the interval between them. For example, a person's age in years is an interval scale, the numbers not only tell you which is older but also by how much. Parametric statistics such as the mean and standard deviation, t-tests, analysis of variance, regression analysis and similar statistics are most commonly used with continuous variables.

In some instances ordinal variables may be treated as if they are continuous and statistical techniques most appropriate for continuous variables may be used. Competent statistical advice should be sought around the choice of statistical analysis.

### Dimension 3: Matched or unmatched data

Outcomes data in PRIMHD data collection is based on an episode of care which is defined either in terms of event (e.g. admission or discharge) or by time (e.g. a review at the end of three months of care from the last data collection). **Matched data** refers to data that is available at the beginning and end of a particular episode. **Unmatched data** is data that is available for only one point in time. If we wish to assess the outcomes of a period of time we need to compare the beginning and end of that time, so in general we will only be able to include matched data. For many other questions (e.g. understanding the demographics of users of a service) unmatched data can be used.

## Statistical analysis

To be able to draw valid conclusions from data such as standard measures in mental health, we need to be sure that phenomena observed in the data, such as differences between groups and relationships between variables or trends over time, reflect real phenomena. For example, if in comparing two inpatient units it was noted that (after casemix adjustment) one of the services showed a four day longer average admission duration than the other, your first question would be "is the difference statistically significant?" If statistical analysis showed that the difference is statistically significant you would feel reassured that there was an actual difference between the admission duration in the two units. If statistical analysis showed no statistically significant difference then you would understand that the difference observed was most likely due to sampling factors and not reflective of a real difference. In this case you would report the groups as not being significantly different.

One of the best ways of evaluating statistical significance is through the use of **confidence intervals**. In this context the confidence interval is the range on either side of the mean (or other statistical parameter) in which the true value is likely to be found. A variety of factors, including measurement issues and the impact of incomplete and/or missing data means that the actual value of a parameter may not be exactly the same as the value that is found. The confidence interval shows the range in which we can expect that the true value will be found. The practical interpretation of confidence intervals for parameters (such as the averages of the length of stay for the two units) is that if they do not overlap, then you assume (are confident) that the things you are measuring are actually different. Confidence intervals can be chosen to reflect different levels of certainty. The 95% Confidence interval is most commonly used. This means that there is a 95% chance that the true value lies within the confidence interval that is set. The higher the confidence level, the wider the confidence interval.

If you found (in the example above) that the confidence intervals for the average length of stay in the two inpatient units did not overlap, you would assume that one unit does have a longer average length of stay than the other. On the other hand, if the confidence limits did overlap, you would assume there is no real difference between their average length of stay. Confidence limits can also be shown on graphs as lines above and below the data point (or top of the bar in a bar graph). In this case, if the confidence limits overlap (i.e. cover some of the same values on the y axis), then we assume there is no real difference between the means. If the confidence limits do not overlap we assume that there is a real difference between them. Further description and a further example of confidence intervals are given in the glossary at the end of this document.

## Casemix

Comparison of different services is important for many reasons, including recognising aspects of services that lead to improved outcome, assessing the impact of innovation, allowing benchmarking between services and to help gather other information that can inform the development of more effective and efficient services. However, variation between the outcomes of services is not only determined by what services do, but also by whom they serve.

Variation in the outcomes (or other factors of interest) achieved by different services can be due to either:

1. differences in how the service works or
2. differences in the types of people that the service assists.

To be able to validly compare the effectiveness of the contribution made by different services it is important to control for the variation caused by differences in the characteristics of service users they serve (i.e. their mix of cases).

**Casemix** is a term used to describe a system for grouping service users together in categories that have similar characteristics (Buckingham & Eager, 1999). In some fields of healthcare, diagnoses are used as variables for casemix classification. However, in mental health diagnosis on its own is not effective for case classification, and a combination of variables is needed (Hunter & McFarlane, 1994).

**Casemix adjustment** is the process of using a casemix system to statistically adjust for the difference between types of people using different services so that the impacts of differences in service delivery can be seen more clearly (Herz, Euseb, Breckenridge, & Rosen, 2006). If the outcomes of two services were compared without such adjustment we could not be sure better outcomes in one service were due to differences in what the team did or due to differences in who it served. If one team showed a better outcome that the other after the data had been casemix adjusted, we would have more confidence that the better outcome was due to what the team was doing. We might study the teams to identify what that team was doing that was working better. Examples showing the use of casemix adjustment are shown later in this document.

In 2002, the Classification and Outcome Study (CAOS) developed a set of Casemix groups based on a large amount of New Zealand data (Eager, Gaines, Burgess, Green, Bower, Buckingham, & Mellsop, 2004). Work is currently underway to adapt this work and develop a casemix system that can be used with the New Zealand standard measures now being collected.

# Guidelines for the Presentation of Standard Measures Data

(Modified from Chipps, et al. [unpublished])

1. Systems should be developed so that the information from standard measures is readily available when mental health information users need it.
2. Data should be presented in a way that is clear, well formatted and as understandable as possible by the wide range of people using this type of information.
3. Aggregated data should be presented in ways that allow statistically significant differences between groups (or differences from key levels - for instance differences from zero in change scores) to be identified. The preferred method is to present data (graphically or numerically) with 95% confidence limits shown.
4. In general when aggregated data is based on small numbers of cases this should be signalled to aid accurate interpretation.
5. When subscale data is presented in graphic form on computers (i.e. not on paper versions) it is valuable to be able to access the data for individual items in the subscale by clicking on the data bars on the graph. Irrespective of the method of presenting the data, it is important that data can be explored at greater levels of detail (e.g. seeing the data at item level rather than at subscale level) to further explore the meaning of the data.
6. Aggregated data (such as means) based on small numbers of cases must be handled carefully as misleading results can frequently be produced. In many areas of statistics 30 cases is regarded as a minimum group size for stable results. In mental health and similar fields it may often be necessary though to use groups with considerably fewer cases. Standard measures data may be used in a variety of techniques and applications. This makes establishing firm guidelines difficult but a general guide might be:
    a. If you have fewer than 30 cases be particularly careful with your interpretation, and
    b. If you have less than 10 cases it is not advisable to use aggregated data from that group.
    c. Following this guideline, groups with less than 10 cases should generally not be presented numerically or graphically.
7. A set of national report formats has been developed to cover key standard reports this can be utilised by all healthcare provider organisations. This will enhance the interpretation of reports as either service users or staff move around the country.
8. Representations of data in graph form should specify clearly which direction indicates improvement or deterioration.

9. Representations of change data should be standardised so that improvement is indicated by a longer upward line. This would help reduce the likelihood that errors of interpretation are made.
10. Confidentiality and encryption of data is maintained according to privacy guidelines.

## Presenting clear consistent and honest graphical data

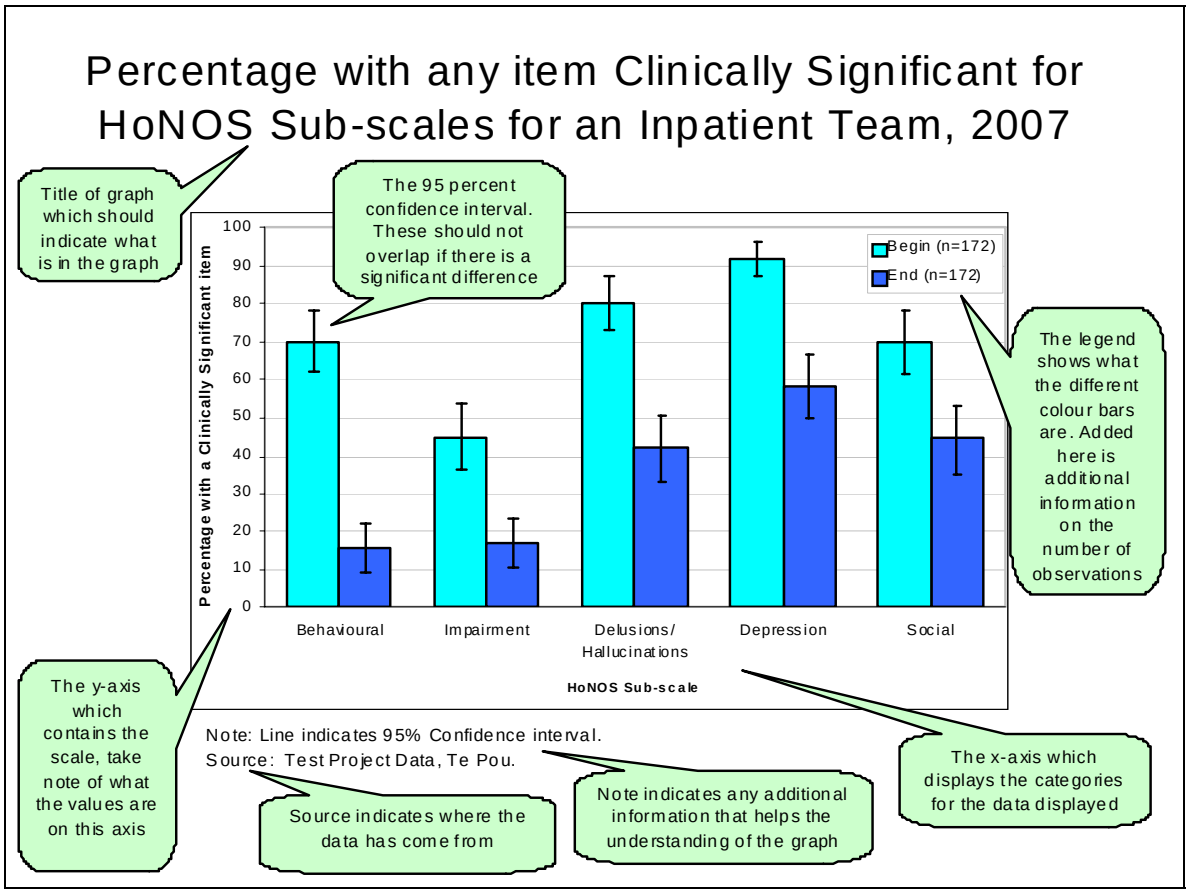
The main reason for presenting standard measures information is to inform good decision making. The interaction between different measures can be complex, so methods of presentation that allow accurate interpretation are needed. This will often involve graphical interpretations of data. Few (2006) provided very useful suggestions on how to present graphical information in practical contexts. Tufte (1983) provided the following guidelines for graphically presenting data:

- show the data
- induce the viewer to think about the substance of the data (what it means)
- avoid distorting what the data has to say
- present the data clearly in a compact space
- make large datasets coherent
- encourage the eye to compare different pieces of data
- reveal data at several levels of detail, from broad overview to fine structure
- serve a reasonably clear purpose - description, exploration, tabulation or decoration
- be closely integrated with statistical and verbal descriptions of the data set

### Making graphical presentation work well

The optimal form of graphical presentation for any set of data depends on the data and the purpose of the graph. The following example (supplied by Sandra Baxendine of Te Pou) shows a common form for the presentation of standard measures data and details several important characteristics.

FIGURE 6: PERCENTAGE WITH ANY ITEM CLINICALLY SIGNIFICANT FOR HONOS SUB-SCALES FOR AN INPATIENT TEAM, 2007




## AVOID MISLEADING GRAPHS

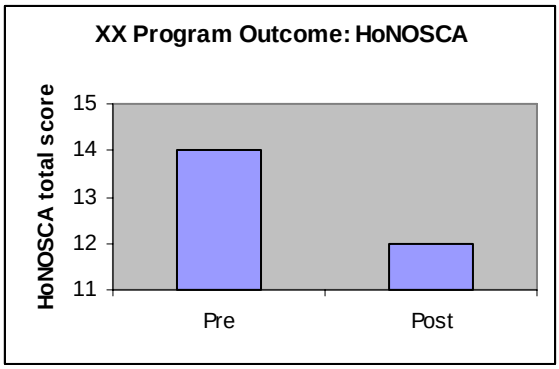
While graphical presentation of data can be extremely useful it is important to make sure data is not presented in ways that are misleading. The following are suggestions of things to look out for that may contribute to misinterpretation of graphs.

- **Number of cases:** Graphical or numeric data based on small numbers has a risk of being misleading. It is often valuable to include an indication of the number of cases making up a data point on a graph. If one category on the graph has very few cases, usually it should not be presented. In many graphs within this document the number of cases is shown as a box starting with "n=". The number of cases for that data point is adjacent to where it is shown on the graph.

- **The scale chosen for graphs:** By selecting where the scales on the side of the graph start differences on a graph can be made to look either larger or smaller than they really are. Take the following example:

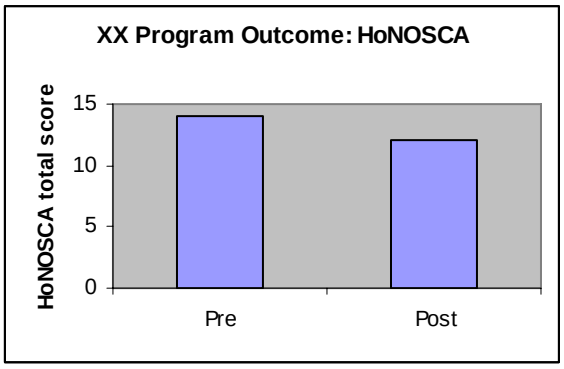
A child and adolescent service introduced a new four-week intensive treatment approach that had been evaluated overseas as making a large difference to the overall mental health status of non-suicidal adolescents. The service evaluated the outcome of this programme using several measures including the total score for the HoNOSCA. After completion of the initial trial period, the data was presented back to the team as follows:

### Figure 7: Change in HoNOSCA Total Score – Example 1




At an initial glance, this looked quite promising. However, the team noticed that the scale on the y axis of the graph had been set to start at 11, rather than at 0, which is the starting point for the HoNOS. The graph was then redrawn as follows.

### Figure 8: Change in HoNOSCA Total Score – Example 2




This gives a more accurate representation of the real impact of the programme. While the team did not have a comparison group to compare these results with their expectation was that the change in the HoNOSCA total score would be larger after four weeks of treatment so they were not convinced that the programme was performing well in their environment. This lead them to consider carefully how the programme was being implemented and to consider if there were any local factors that might have reduced potential outcomes. Feedback from service users and their families was sought as part of this review. As a result, several revisions were made to the programme to optimise it for local conditions. Six months later outcomes (as evaluated by the total HoNOSCA score) were much larger.

### Showing the significance of differences

Whether aggregated data is presented either numerically or graphically it is generally important to show which differences in the data are statistically significantly different from each other. The preferred means of doing this is by indicating the 95% confidence intervals for each data point. See above for a more detailed description of confidence intervals. The risk of not including confidence intervals is that information users may draw false conclusions if they assume that a difference that shows up in data is meaningful when in fact it is just due to random variation.

# Guidelines for the use and interpretation of standard measures data

(Modified from Chipps, et al. [unpublished])

1. Data should only be used for the purposes it was intended for consistent with relevant privacy law and guidelines.
2. Training in how to interpret and use data is important to ensure that information is used accurately and consistently by all parties. Gaining access to information alone may not lead to people becoming skilled users of that information.
3. Data should only be used in ways consistent with the measure's design and psychometric properties.
4. When establishing standard forms to present data or utilising the data to answer new questions, consultation should be sought from resource people with expertise in the analysis of mental health standard measures and the PRIMHD data set. This will help ensure that data is used appropriately and that valid information and interpretations are derived.
5. Interpretation of data should consider all relevant contextual factors such as socio-economic factors and data should be adjusted for these factors.
6. Analysts should avoid discussions of differences due to underlying traits or other explanations without clear evidence of such. This is important to prevent over-interpretation or misinterpretation of causal or other processes behind the standard measures data.

## Ethical collection and use of data

- It is important to note that standard measures will only ever be a part of what is required for a quality comprehensive assessment of a service user (Onyett, 1992). The collection of standard measures is not intended as a substitute for comprehensive assessment. Some standard measures such as the HoNOS can only be completed adequately if a reasonable assessment has been undertaken.
- **Informed consent:** At admission to services and on other occasions, as appropriate, service users should be informed that standard measures are routinely collected as part of clinical care. They should also be informed that these measures are used both for assisting with planning for their care and, in an anonymous form, for assisting with the planning and improvement of the mental health services (New Zealand Privacy Commissioner, 1994).
- Clinicians and other collectors of information have an obligation to provide valid and reliable information for standard measures and the clinical record.

- Service users, clinicians and mental health workers have the right to expect that managers and administrators will use information appropriately and wherever possible make explicit their rationale for decisions based on this information.

# MAORI DATA USE

Maori access the full range of mental health services in Aotearoa, New Zealand. How services can best use mental health information to support the health and well-being of these tangata whaiora and their whanau is a complex discussion. These guidelines (while not definitive) highlight key concepts and examples of sound practice to inform this discussion.

## GUIDELINES FOR THE COLLECTION, ANALYSIS AND USE OF MAORI STANDARD MEASURES DATA

- Whanau ora is the korowai of recovery for tangata whaiora. Wellness may not be conceptualised as relating to an individual alone, but to wellness within family/ whanau and social relationships and being able to contribute and meet whanau responsibilities.
- In the processes of data collection and use mental health services should work in collaboration with whanau to further the recovery of tangata whaiora wherever possible. For example, this could simply be ensuring that whanau know they are able to be part of the process.
- Ensure tangata whaiora and their whanau are supported to make sense of:
  - ❖ what is happening
  - ❖ the purpose and use of the information being gathered
  - ❖ the substance of the information that is being shared by mental health workers
  - ❖ the processes to address any issues that arise.
- Ensure that the use of the mental health information is understood within the particular cultural context of tangata whaiora and their whanau. E.g. access cultural supervision and seek the input of cultural workers.
- Ensure that Whanau experience and knowledge is considered and incorporated in the use of mental health information to further the recovery of tangata whaiora. E.g. listen to whanau interpretations of what is happening for tangata whaiora.
- Ensure that mental health information is used to strengthen whanau capacity in order to develop pathways to wellbeing for tangata whaiora. For example, the full resources of the multidisciplinary team are used to support whanau to be able to access resources/services.

- Good health for Maori involves balancing the dimensions of Te Whare Tapa Whā not just Te Taha Whanau.
  - ❖ Te Taha Wairua
    - ➢ E.g. offer tangata whaiora and their whanau opportunities for karakia and whakawatea.
    - ➢ E.g. recognise the potential impact of medication and other interventions upon Te taha wairua.
  - ❖ Te Taha Hinengaro
    - ➢ E.g. talk to whanau about how they see the behaviours, thoughts and feelings of tangata whaiora. Are these out of character? Are they distressing to tangata whaiora and whanau?
  - ❖ Te Taha Tinana
    - ➢ E.g. establish links between physical wellbeing and mental health issues.
    - ➢ E.g. establish links with GPs and other health care providers and explore any barriers to accessing services.
- Accurate identification of ethnicity is essential.
  - ❖ E.g. it is necessary that tangata whaiora confirm their ethnicity.
- Do not assume that cultural identity and ethnicity are the same. Some tangata whaiora may be Maori by ethnicity but may not identify with Maori cultural activities or readily be able to access Maori knowledge and skills.
- Ensure that cultural identification is determined by tangata whaiora and their whanau.
  - ❖ E.g. encourage tangata whaiora to talk about their whanau. This might include whether they are connected to marae or not.
- Ensure cultural identity of tangata whaiora and their whanau is acknowledged.
  - ❖ E.g. with whanau involvement opportunities for cultural expression through hui can support tangata whaiora in this.
- The tino rangatiratanga of tangata whaiora and their whanau is about having control over the direction and shape of their lives, their mental health and mental health information and processes of recovery.
  - ❖ E.g. discuss the results of outcome measurements with tangata whaiora. Incorporate their feedback and perspectives in your reflection on results.

# Key Concepts and Examples of Sound Practice for the Use of Maori Data

## Cultural Safety and Competence

Culturally safe practice ensures that mental health information is consistent with Maori values and experiences. This does not need to be a daunting task for clinicians; it may be as simple as pronouncing names correctly, using services' cultural resources (e.g. Kaumatua, cultural workers) and undertaking cultural supervision.

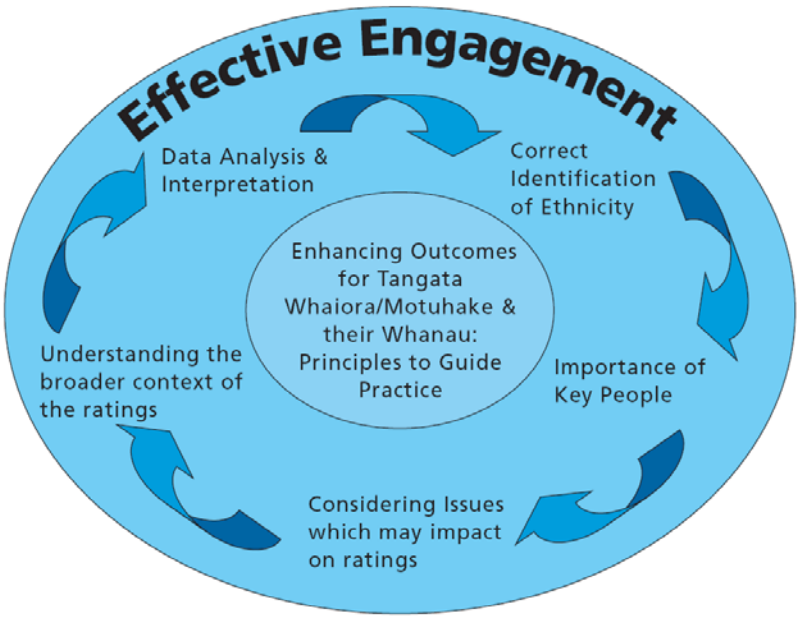
*Let's get real* is a foundation document for the mental health and addiction workforce. It envisions a workforce that embraces recovery, is person centred, is culturally capable, and assures and improves quality of services for people. The framework stipulates that 'every person working in a mental health and addiction treatment service contributes to whanau ora for Maori'.

Each of the *Let's get real* skills have a broad definition linked to performance indicators. Mental health workers should be familiar with the seven essential indicators in the "Working with Maori guide" and be working or training to them.

## Effective Engagement

Through culturally safe practices, rapport and good engagement can be developed. The information obtained through effective engagement will support the clinician to understand the relevant cultural context of tangata whaiora and their whanau. Where possible, whanau and tangata whaiora should be encouraged to take part in all stages of data collection, presentation and analysis to ensure engagement.

**Figure 9: Effective Engagement Model (Levy, Mar, van Altvorst, Sokratov, Eade, Meihana and Winiata, 2005)**



### Environment

For effective engagement to be sustained, the environment where services are being provided needs to be considered. Services need to canvas and listen to the experiences of tangata whaiora and their whanau. This may mean conducting surveys via various mediums and responding to these experiences, thereby making necessary changes.

Ensuring accessibility and acceptability of mental health services affects the ability of clinicians to gain accurate information that enhances engagement and recovery. It may mean considering such issues as access to parking and public transport; meeting spaces that can accommodate whanau, inclusion of Te Reo, including mediums such as posters and signage, and incorporating practices of karakia.

However, if services feel that they are unable to effectively meet cultural needs, other options that support cultural inclusion need to be considered. This could mean arranging alternative meeting places or meeting at someone's home or marae. This may not be possible for all clinical settings (e.g. acute inpatient settings). In these circumstances ensure that key people are involved at the assessment stage.

### Ethnicity

The identification of ethnicity is an important first step in understanding who is accessing mental health services. Failure to identify ethnicity accurately affects the quality of mental health services and the mental health information gathered from tangata whaiora.

Ethnicity data that is not accurately and consistently collected will not truly reflect service utilisation figures and the number of Maori accessing services (Levy, et al., 2005). This has consequences at all levels of information use.

Correctly gathering ethnicity data does not need to be a complicated task. It can be as simple as asking. Relying on physical appearance or surnames that sound Maori or Pākehā or not asking because of awkwardness is not acceptable. Tangata whaiora and their whanau will determine how they identify.

## Te Whare Tapa Wha

Maori view the world in a holistic way. Four dimensions, as recognised in the widely used Te Whare Tapa Whā model, (Durie, 1998) contribute to the total wellbeing of Maori:

- Te Taha Wairua refers to the spiritual wellbeing of a person. It determines one's identity. It provides a direct link with one's tūpuna (ancestors) and whanau group and strengthens the taonga (treasure, special possessions) and tikanga (customs, values) of one's cultural system.
- Te Taha Hinengaro refers to the mental and emotional well being of a person.
- Te Taha Tinana, the physical well-being of a person, cannot be dealt with separately from the other dimensions.
- Te Taha Whanau acknowledges the importance of whanau in providing sustenance, support and an environment that is conducive to good health.

The HoNOS family of measures are the only outcome measures currently mandated by the Ministry of Health. The MH-SMART initiative identified other types of outcome measures that require further review and development prior to implementation. One of these is a Maori measure called Hua Oranga. This tool will further support the information gained from best practice and HoNOS to advance whanau ora.

## Social Domains of HoNOS Family of Measures

The social domains of the HoNOS family of measures provide a significant opportunity to address key aspects of the wider context of tangata whaiora which is important to achieve positive health gains and outcomes that further recovery.

As alluded to previously, whanau ora encompasses the principle that individual well-being is affected and influenced by the whanau as well as the individual. It highlights the importance of working with people within their social context, not just working with their physical symptoms.

It is widely accepted that social and economic factors of employment, housing and poverty have a significant impact upon the wellbeing and recovery of people using mental health services. In the 2006 census key social and economic disparities between Maori and non-Maori are identified.

When working with Maori exploring the possible existence and extent of key risk factors such as: exposure to violence, sub-standard and/or over-crowded housing, poverty, addiction, long term unemployment and a lack of access to services may be useful.

The social domains of the HoNOS measures may go some way to identify the impact of social risk factors on tangata whaiora. At the assessment stage the clinician will need to engage the full resources of the multi-disciplinary team (MDT) when rating the social domains. For example comprehensive assessments from allied health members will be required to establish

levels of occupational, social and cultural functioning, identifying social stressors and barriers to recovery. Whilst services may not be able to remediate every social stressor linked to social risk, services have a responsibility to broker plans for change. Using HoNOS in this way will support and progress whanau ora.

## Levels of Information use

The accurate recording and rating of Māori information is important at multiple levels of service delivery. It enables an understanding of how Māori cultural dimensions contribute to the mix of outcomes/recovery. At an individual level this information will provide a holistic picture that will aid recovery and treatment planning for tangata whaiora. There may also be opportunity for practitioners to discuss the potential impact that culturally sound practices have on outcomes.

- At whanau level mental health information will help assess whanau capacity to support and care for tangata whaiora and identify the necessary resources needed to achieve the goal of whanau ora.
- Information at local, regional and national levels also has the potential to underpin strategies to advance whanau ora.
- At a service/team level this information will highlight staff skill mix and resource requirements for strategic planning and the potential to contribute to bench mark against comparable services.
- At a funding/provider level accurate information will better enable targeted resourcing that ensures value for money. Collaborative interagency funding arrangements to address social risk issues should impact positively on mental health outcomes.
- At a national level data can be interrogated to better understand whanau ora as an outcome, identifying what and where gains are being achieved.

Mental health information should incorporate cultural, clinical and social dimensions. The following guidelines should not be used in isolation. All the principles of the general data use guidelines apply to Māori; these extra items support the utility of mental health information for Māori

## Summary

Collecting standard measures is not about reinventing the wheel or giving fatigued clinicians more work. Due to the technical nature of rating standard measures there is potential for 'rating' to be perceived as mundane and driven by organisational need as opposed to measuring change for tangata whaiora. Where tangata whaiora have used services for some time with little observable change a clinician may be tempted to rate on past history in the absence of a current, clinical assessment. This is not professional practise and falls foul of information collection protocols for standard measures.

Robust outcome measurement underpins evidence based practice that is focused and leading edge. Māori mental health information can underpin the development of responsive services that measure and change according to the needs of tangata whaiora and whanau. Equitable health status can only be achieved when we understand what is working well, where gaps are and what strategies best address them.

When whanau ora is acknowledged as the korowai of recovery and is embedded in service delivery, tangata whaiora and their whanau can expect nothing less than excellent support from services. The utility of Maori mental health information collected by services rests on many factors.

- Sound cultural practices
- Effective engagement
- Accessible and excellent services
- Tangata whaiora and their whanau leading recovery and having active input at each stage of information processes.
- Accurate collection, analysis and use of data that is understood with in the cultural context of tangata whaiora and whanau.

Careful analysis of the social domains of the HoNOS family of tools and how they relate to cultural constructs and realities broadens clinicians' views about what factors influence well-being. Wellness is seen as a balance of many dimensions which encourages varied interventions to be offered rather than medication alone. Future tools will further add to a holistic understanding of the experiences of tangata whaiora and their whanau.

Everyone benefits when services best meet the needs of those who access them. Services that measure are services that can prove what they do well and can highlight areas to develop. At best outcome measurement is about celebration and transformation.

# Examples of the use of standard measures data

## Individual Data – Discussing an individual's data with the individual and their family/whanau[1]

The data collected from standard measures such as the HoNOS suite of measures can provide a point of feedback and dialogue between service users, clinicians and with the agreement of the service user, their family/whanau or significant others. The clinician-completed HoNOS standard measure should be discussed with the service user and used as an opportunity to identify areas of agreement and difference between service user and clinician perspectives. This can also guide a collaborative process of planning for recovery and give the service user another method to monitor their progress.

Even with good intentions and good practice, mental health staff and service users (and/or their family/whanau) often have different understandings of the current situation, how things have changed over time, what is driving that change, and the service user's own goals for change. This difference in understanding can have negative impacts on service user well-being and achievement of optimal outcomes.

Data collected from standard measures such as HoNOS, HoNOS65+ or HoNOSCA can be helpful as part of a review with a service user (and their family/whanau) to give a point of discussion about the service user's current situation, change over time, and factors that might be driving both positive and negative changes. It is not necessary that this discussion leads to complete agreement between the parties, (although it is likely to lead to better agreement). On rare occasions it may generate some heat but usually much more light is generated as the parties come to understand each other's perspectives better.

Some clinicians believe that service users they work with will not be interested in engaging in this kind of discussion. However, service user representatives and clinicians using this approach report that the majority of service users welcome this kind of discussion, and both parties gain from having it.

In general this discussion should be quite a bit broader than reporting on the HoNOS data, and would typically involve a thorough discussion of the current situation and the recovery goals for the service user (for example, the discussion may be more based around reviewing and updating the Wellness Recovery Action Plan (WRAP; Copeland, 2002). However, the HoNOS data can provide a picture of how things are going and generate useful discussion towards an extensive wellness planning exercise.

[1] Thanks to Shirl Corles-Davis, and Mike Wilson for their comments and input into this section.

Remember that the HoNOS standard measure data forms a framework for a conversation with the service user, and possibly their family/whanau, and the two major aims of that conversation are to:

- understand better each other's point of view and perhaps come to a shared perspective and
- use this as a springboard for developing, revising or perfecting shared plans for further recovery.

To achieve these aims it is important to really engage with the service user to determine how they see things - in the words of Stephen Covey (1989: 7 Habits of Highly Effective People) *"Seek first to understand, **then** to be understood"*. Be prepared also to bring into the discussion information (other than the HoNOS scores) that you think is important, and be receptive to other information supplied by the service user.

The following section contains suggestions on how to feed back HoNOS data /results to service users (and family/whanau, where appropriate to increase understanding between mental health staff, service users and whanau.

### Getting Ready

- Have a look at the data and think about what it means before starting discussion with the service user (and/or family/whanau).

- HoNOS data is probably easier for most people to understand if it is presented as a graph showing each individual item. If data from more than one time is available, the bars for each item over time should be shown next to each other. Print out this graph or be ready to show it to the service user on a computer screen.

### Setting the Scene

- Provide the service user and their whanau/family with information, written and verbal, about outcome measures

- Explain to the service user that the HoNOS or (HoNOSCA or HoNOS 65+) measure is collected for all service users at entry, every 3 months/ or at changes in wellness and at discharge. Explain that it shows a record, over time, of their journey to recovery and is useful for helping us monitor change/ levels of improvement and understand how successfully the service helps people achieve positive outcomes.

- Explain the following:
    1. The HoNOS data (scores) is your view (as a mental health clinician) of their current situation and the changes over the collection period
    2. They may or may not agree with every rating you have given. You are really interested in finding out about how they see things.
    3. You are not able to change the HoNOS scores that has been entered but, if they see things differently you would be happy to record their view in the notes.

- Explain that you would like to discuss the results (data) and what might contribute to them and what may help improve the items that are causing concern/not changing.

### Presenting the HoNOS data

- **Start by making sure the person understands the graph(s).** Point out that there are twelve (fifteen for the HoNOSCA) items and show them where the labels for the items are. Point out that the scoring goes from 0, which means no problem to 4, which means a severe problem. State that the taller the bar the more severe the issue. Point out the levels between and how levels 2, 3 and 4 all indicate issues about which you might have some concern.

- **Indicate the items that you see as being of most concern.** Start off with the items with the highest scores. Ask the service user if they also see these things as being a concern. Listen carefully and engage constructively on their responses. Where you agree over a particular issue being a problem, ask the person how they think this could be improved**.**

- **Repeat the above process for items down to level 2** or even 1, if it feels relevant to do so.

- **Indicate the items that you do not see as being of concern** (scores of 0 and possibly 1). Ask the service user if they agree with your assessment of these items. Draw attention to these as areas of strength to enhance the person's sense of mastery.

- **If results from more than one time are available, talk about change over time.** Briefly point out both the positive and negative changes. Ask the service user if they see the same changes. Ask how they think the changes have occurred - things that have contributed to making things better or worse. If you have ideas about what has contributed to change, be prepared to share these also.

## A STRATEGY FOR PRESENTING THE INDIVIDUAL'S DATA TO A MULTIDISCIPLINARY TEAM.

Standard measures can structure thinking about the care and recovery needs of an individual service user in presenting to multidisciplinary teams or supervision. Standard measures form a part of, rather than an alternative to, discussion in these settings but can help make discussion more productive.

In these settings presentation of individual data such as HoNOS, HoNOS 65+ or HoNOSCA data, preferably in graphical form, can give colleagues a good starting point for understanding the following:

- the nature of strengths and difficulties the service user has,
- particular concerns with this person at present and, if standard measures from more than one time are available, how the situation has changed over time for this person.

Additional detail about the situation and context can be added to the discussion but standard measures provide a good foundation for understanding this information and its importance.

On the following page is a series of questions that can be used to analyse and report back on individual HoNOS, HoNOS 65+ or HoNOSCA data. These can be used to:

- think about the data for this individual prior to meeting with colleagues,
- review the interventions currently in place and
- identify different or additional interventions that may be particularly relevant at this time.

A one-page review form to assist with recording the results of this thinking is included below. The same series of questions can then be used to present the information back to the multidisciplinary team or other colleagues.

Once this approach has been practiced a little, it becomes quite a quick and efficient way of organising your thinking about an individual's situation to aid planning for recovery. This can make for a more effective presentation of the individual to the multidisciplinary team.

For more information about the implementation of this approach, see www.tepou.co.nz [2]

[2] Refer to "How to use HoNOS in Clinical Multidisciplinary Team (MDT) Meetings", a presentation by Dr Malcolm Stewart, in the Information Utility section of the Information page

Figure 10: Using HoNOSCA to Help Plan Care and Recovery

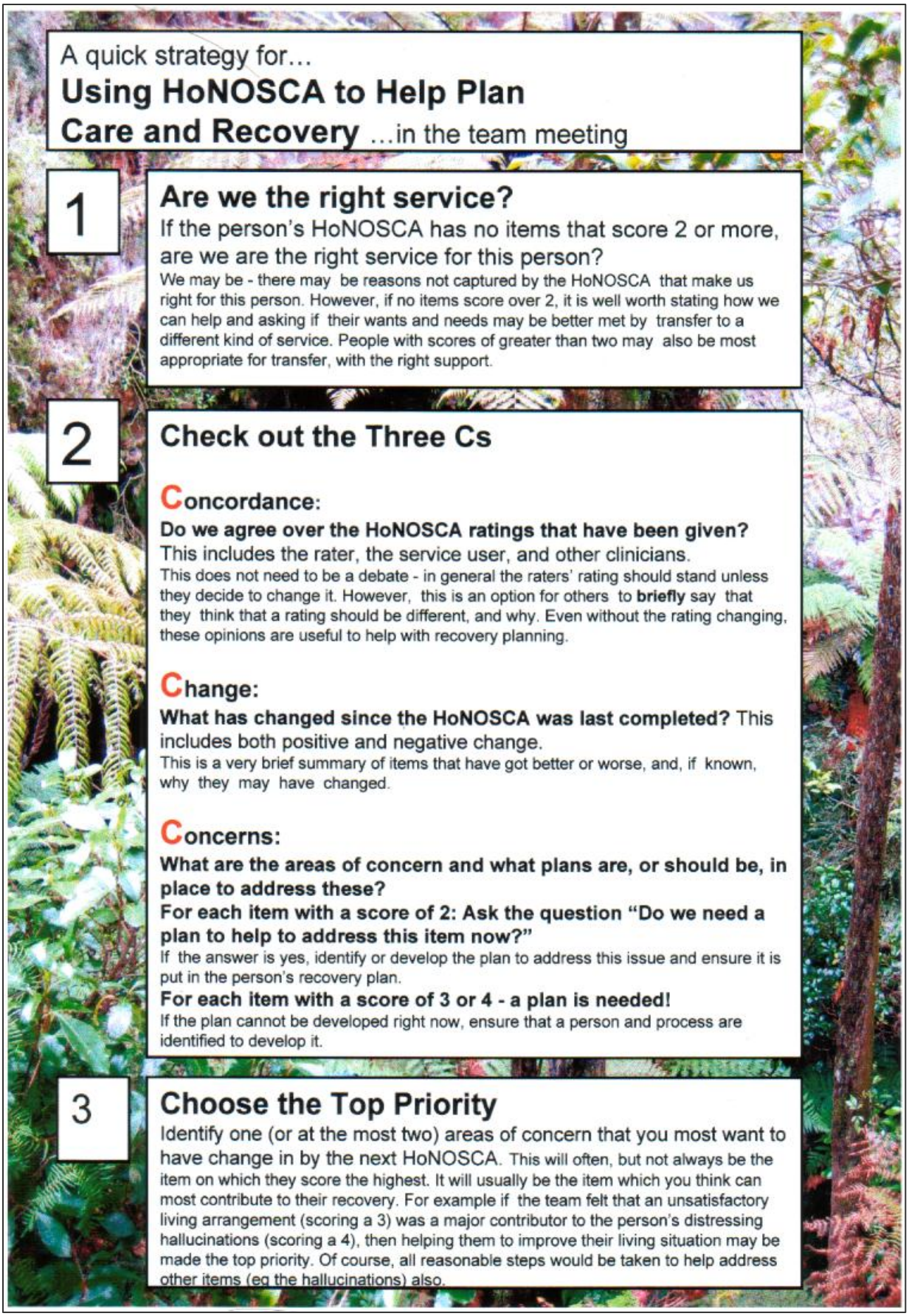


Figure 11: MDT Review Form

**MDT Review**

| **Name:** | **NHI:** | **Review Date:** |
|---|---|---|
| **Reason for Review / Desired Outcomes** | | |
| **Brief Background** (Maximum of 3 minutes to present) | | |
| **Current Situation** (Maximum of 2 minutes to present) | | |
| **Diagnoses** | **Formulation** | |
| **Changes in Last 3 Months** | | |
| **Concerns:** (If no HoNOS items score higher than 1, why are we still the right service for this person?) | | |
| ***HoNOS Items Scoring 3 and 4*** | | |
| **Item Name** | **Plan(s) to Address this Issue** | |
| ***HoNOS Items Scoring 2*** | | |
| **Item Name** | **Plan(s) to Address this Issue** | |
| ***Other Concerns/Goals not reflected by HoNOS Items*** | | |
| **Concern/Goal** | **Plan(s) to Address this Issue** | |
| **Top Priority for the Next 3 Months** | | |
| **Comments/Plans from Discussion** | | |

# Individual data - first collection

## Question(s) addressed

What is the service user's current status and where can we best focus our efforts for recovery.

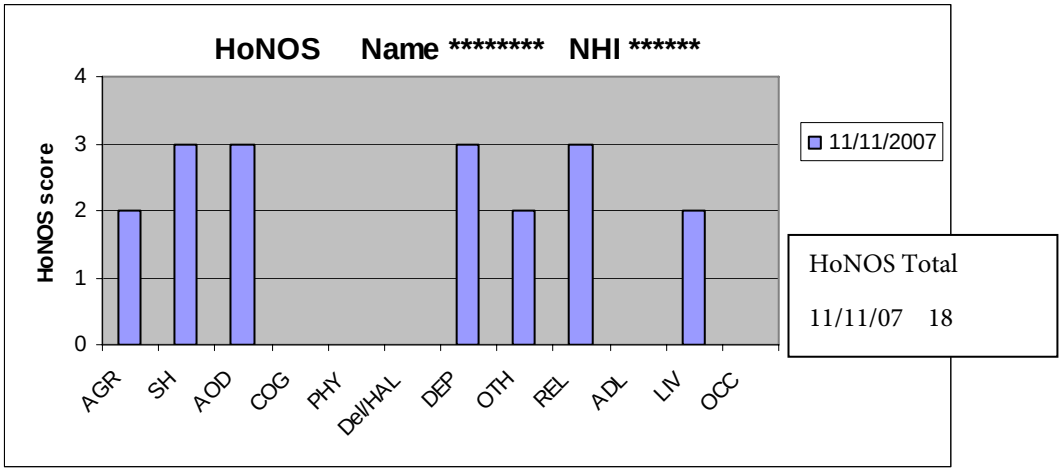
## Description

This is the report which is produced when a service user is first admitted to the service and their first HoNOS data is entered. It can be used as part of care and recovery planning. The strategy described in the previous section can be used to analyse and discuss the situation with the service user and colleagues in the multidisciplinary team to develop recovery plans.

## Report

This report is for a 33 year old female who was referred to a community mental health service when she was experiencing a period of deep depression marked by agitated and aggressive behaviour and strong suicidal ideation and intent. She also suffered considerable anxiety. She had a significant ongoing pattern of alcohol abuse that created considerable difficulties for her. While functioning reasonably well, despite these difficulties, in most areas of her life, she reported considerable relationship difficulties. Her living situation caused her additional stress as it exposed her to risk of criminal action. The report produced from her HoNOS data was as follows.

### Figure 12: Example of an Individual HoNOS Graph



### Interpretation

*Using the strategy described in the section above, immediate plans were made to address her risk of self-harm, alcohol use, depression and her relationship difficulties. Consideration was also given to whether to address her agitation, anxiety and living situation at this stage or whether to focus on these later. The decision was made that dealing with the anxiety along with the depression at this stage would be helpful, but as the agitation appeared to be secondary to depression and alcohol use no direct focus on this would be taken at this time. She was not interested in addressing her living situation at this stage, so this was not prioritised. The top priorities were reducing her risk of self-harm and her depression.*

### Comments

The total HoNOS scores can be useful for tracking change and giving some indication of overall severity of the situation but are not recommended for use as a cut-point criteria for service entry (i.e., it is not recommended that a service admit people with more than a particular score and do not admit people with a total score of less than that score).

The HoNOS results (data) could be linked more closely to care planning/recovery pathways. A list of biological, social and psychological strategies that may be of relevance to elevated levels on each item could be used as part of recovery planning. More complex approaches, for example, actioning of particular care pathways in response to particular combinations of elevated items, have also been proposed.

## Individual Data - Second and Subsequent Collections

### Question(s) addressed

What is the service user's current status and where can we best focus our efforts for recovery.

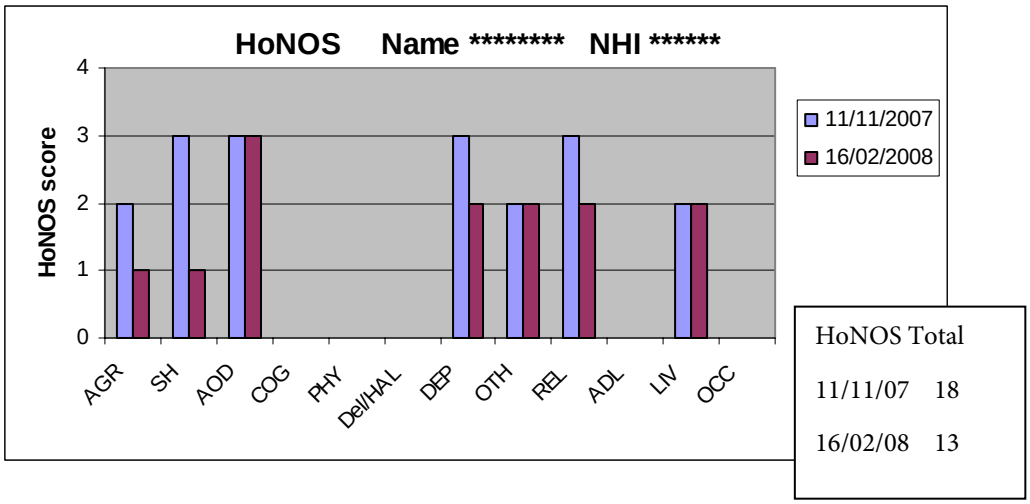
### Description

This report has the same format as the one above drawn from the first entry of data at admittance. The new HoNOS data is added to the right of each item from the previous collection. This data can now be used for recovery planning with the additional benefit that changes over time can be considered. The strategies described above to analyse and discuss the situation with the service user and the multidisciplinary team can still be used.

### Report

Following on with the HoNOS data for the 33 year old woman described in the previous section, this report records her initial HoNOS, and her three month review data.

### Figure 13: Example of an Individual HoNOS Graph (assessment & 1st review)




#### Interpretation

- *She has shown some improvement in her depression and a marked reduction in her self-harm risk. Her agitation is no longer in the clinical range (score 1). Her relationship difficulties have also shown some improvement. There has been no change in her alcohol use. Her living situation continues to place some stress on her. Anxiety also remains a problem for her.*
- *After discussion it was decided to increase efforts to address her alcohol use while continuing to focus on her depression and anxiety difficulties. She remained unwilling to address her living situation. A plan of waiting and working with her until she was ready to address this was implemented.*
- *She reported being comfortable with her relationship status at present, so no further work on this was planned.*

#### Comments

This shows how the HoNOS data can help track areas of change and no change and areas of ongoing concern. This can then be used to support recovery planning. The service user and, as appropriate their family/whanau, could be included in this discussion.

Additional collections can be added to this graph for each subsequent collection. To avoid being overwhelmed with data, preparation of such graphs should allow for selection of the time period over which data will be presented for (e.g. for the first data collection and for all data collections in the last year). However, it is also worth being able to plot all previous data collections on one graph so patterns over time can be recognised.

## Aggregated data

Aggregated data is data from more than one individual (e.g. more than one service user) that is grouped together. For example, the admission HoNOS scores for all service users using a respite service might be aggregated so the mean HoNOS score could be calculated. Aggregated data can be used at the team, organisation, regional or national level. When data is aggregated, a wide range of statistical techniques can be used to help understand what it means. Below are examples of some reports using aggregated data that may be useful.

## Summary of a team's outcomes over time - Example[3]

### Question(s) addressed

- What is the outcome of the team's activities over this period of time?
- What types of difficulties are the team addressing successfully (more and less) at the moment?
- What types of difficulties face people who use the team?

### Data used

This report shows the proportion of service users of the service who have a score in the clinical range (2, 3 and 4) for each item on the HoNOS.

This report uses HoNOS individual item data. It uses only matched data (i.e. data from both the beginning and end of a time period).
The period covered by this type of graph could be a specified period (e.g. three months) or defined by the service user's pathway (e.g. admission to discharge)

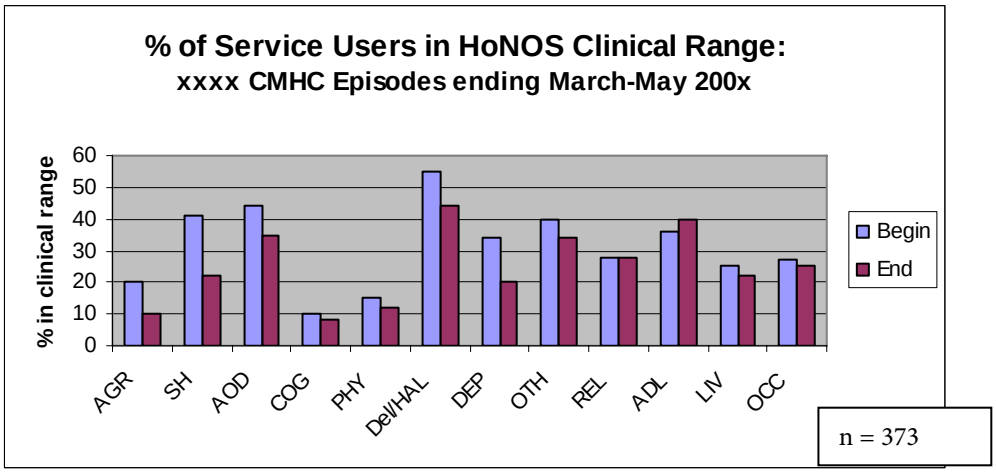
### Report

This shows the outcome for all matched pairs of HoNOS scores for:

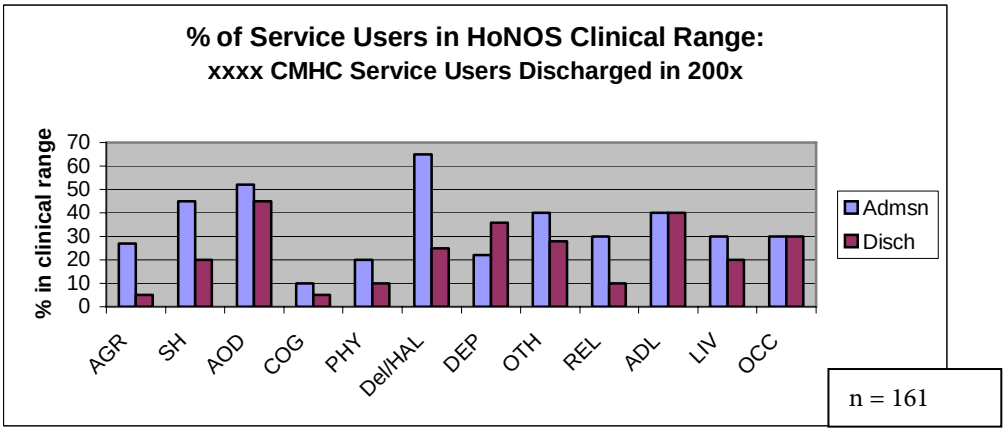
- all service users with an episode that finished in the last three months and
- from admission to discharge

[3] This data is for demonstration purposes and does not represent real service results.

### Figure 14: Percentage of service users in HoNOS clinical range for episodes ending March - May 200x




### Figure 15: Percentage of service users in HoNOS clinical range for service users discharged in 200x




#### Interpretation

- *The greater the reduction in the length of the bar from beginning to end of the period, the stronger the positive outcome for that item.*
- *If the lengths of the bars do not change much from beginning to end, then the team is probably not influencing the outcome on this item very much.*
- *The length of the bar at admission indicates how frequently that item is a problem for service users of that team.*
- *The "n=" box below the right hand bottom corner signifies the number of cases included in the analysis. For the first graph, 373 episodes of care for which there were matched pairs were collected in the three months. In the case of the second graph, 161 service users, for whom admission and discharge collections were available, were discharged during the full year.*

### Comments

This report shows most clearly *clinically significant* change, in that it highlights the proportion (%) of people who go from scoring in the clinical range to scoring in the non-clinical range for each HoNOS item.

## Use of HoNOS/HoNOSCA data to develop consumer and service profiles[4]

### Question(s) addressed

What are the characteristics of the people who access our services? What issues are they struggling with? How are these the same or different for different groups of service users? And many other similar questions.

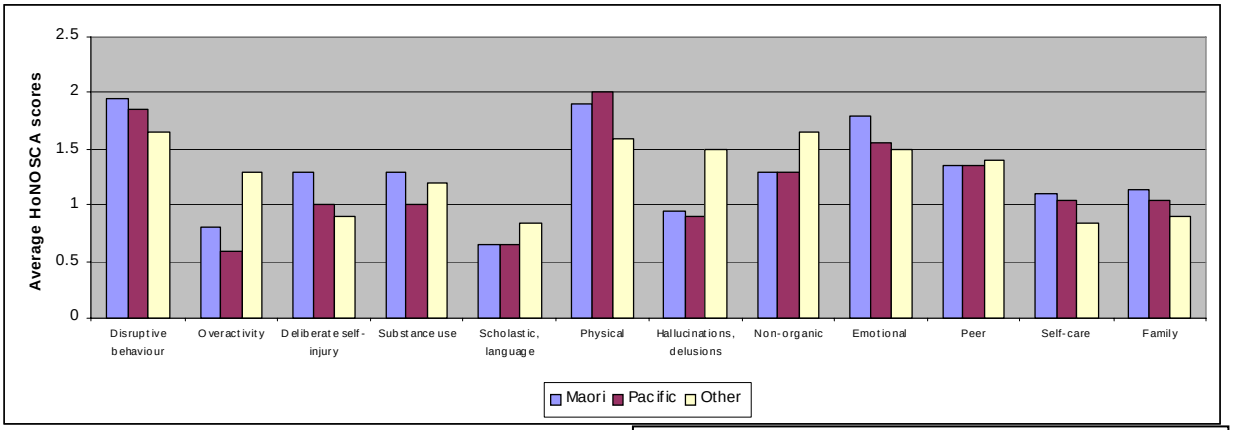
### Description

Utilising data from service, district, regional or national levels, HoNOS or HoNOSCA data can be used to develop service user profiles and can be used to compare different consumer groups, e.g. ethnic groups.

### Report

The following is an example of comparing the average inpatient admission data for the first 12 items on the HoNOSCA for different ethnic groups (Maori, Pacific Island, and Other, which includes Pākehā).

**Figure 16: Inpatient Admission Data - average total HoNOSCA scores (1st 12 items) by ethnicity**




n: Maori = 30, Pacific =26 Other = 75

[4] This data is for demonstration purposes and does not represent real service results.

## Interpretation

- *This graph does not include confidence intervals, so we cannot be sure about the statistical significance of differences between the groups. However some trends may be observed.*
- *This graph shows that the most severe difficulties overall for young people entering an inpatient unit were disruptive and antisocial behaviour (Item 1), physical illness and disability (Item 6) and emotional difficulties (Item 9).*
- *The "n=" box below the left hand bottom corner signifies the number of cases in each of the three ethnic groups included in the analysis.*
- *The pattern of difficulties appeared to be reasonably consistent over all three groups, and particularly, between Maori and Pacific people. Difficulties with over-activity (Item 2) and hallucinations and delusions (Item 7) seemed to be particularly common/severe in the Other group. There are some other differences between the three groups and these differences could be further explored.*

### Comments

The differences and similarities observed can now be used as a starting point for discussing why these differences might exist, and then using this discussion to attempt to cater better for all groups.

## Comparing outcomes of different groups of consumers[5]

### Question(s) addressed

How do the outcomes of different groups of consumers as defined by casemix compare.

### Description

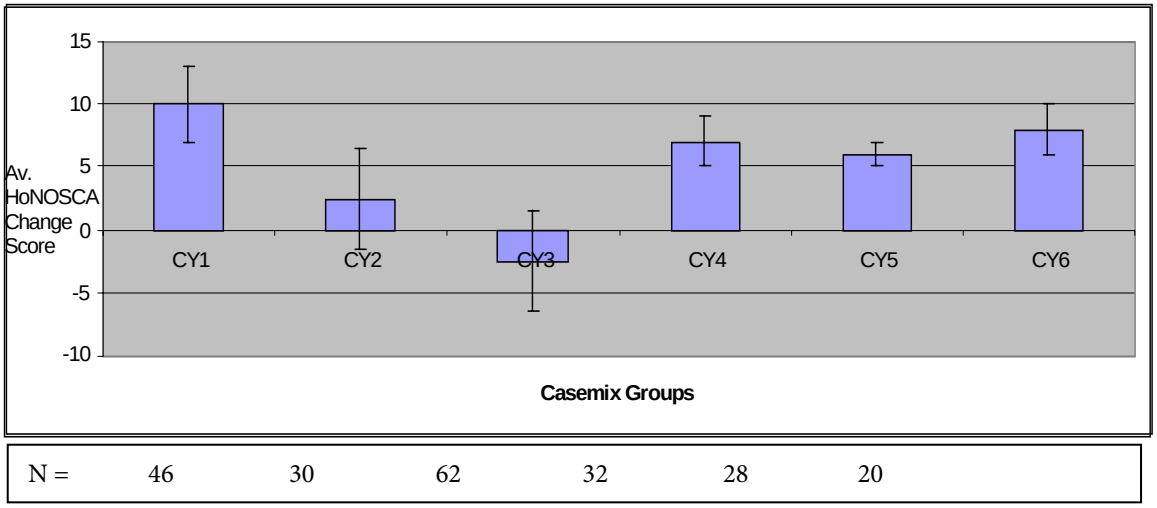
Casemix is frequently used to adjust data to allow comparisons across different services. However, it can also be used to define groups of service users within a service to assess how their needs are met.

### Report

A Child and Adolescent Mental Health Service was interested in the outcomes achieved by different groups of service users. To assess this, the average change in HoNOSCA total scores from admission to the most recent collection for different Casemix groups were calculated (Admission score - most recent score). It was found that there were sufficient service users in six Casemix groups for satisfactory analysis to be carried out. The codes CY1-CY6 represent these six different Casemix groups. The results are presented below.

[5] This data is for demonstration purposes and does not represent real service results

### Figure 17: Change in average HoNOSCA scores by Casemix class




| N = | 46 | 30 | 62 | 32 | 28 | 20 |
|---|---|---|---|---|---|---|

**Note:** Scores below zero indicate deterioration, scores above zero indicate improvement

95% Confidence intervals are shown on graph

#### Interpretation:

- *A larger positive number indicates a greater improvement in total HoNOSCA score. As their confidence intervals overlap, we conclude that the outcomes for groups CY1, CY4, CY5 and CY6 are not different. As their confidence intervals are above and do not include the zero line, we conclude that they show significant change (improvement).*
- *As the confidence limits for CY1 does not quite overlap with the confidence limits for CY2 we conclude that the outcome of CY1 is significantly different from CY2. Using these confidence intervals we can conclude that CY3 achieves a significantly poorer outcome than CY1, CY4, CY5 and CY6.*
- *As the confidence limits for CY2 and CY3 include 0 (no change), we conclude that these groups have shown no overall improvement.*
- *Taken all together this graph shows that there were markedly different outcomes for different Casemix groups with some benefiting more from service provision and some receiving little or no benefit.*
- *The "n=" line under the graph indicates the number of people in each group included in the analysis.*
- *Knowing that these different Casemix groups had substantially different outcomes, the team now looked at the characteristics of the Casemix categories and discussed why these categories may do worse. On the basis of this discussion, the team determined some of the likely barriers to clients achieving good outcomes. Plans were then put in place to reduce these barriers.*

### Comment:

This is a good example of how standard measures may raise questions rather than provide straightforward answers. However, raising questions and giving data to evaluate possible answers helps mental health service users, clinicians and managers find the best possible solutions to assist quality improvement.

## Comparing outcomes for different types of intervention[6]

### Question addressed

How do the outcomes of service users who are receiving different types of intervention or service compare?

### Description

A very significant issue for mental health services is identifying what kinds of interventions work best. Standard measures data can inform evidence based practice that helps services evaluate the impact of different types of interventions for particular service users and their service context. When the outcomes of groups receiving different kinds of services are compared, we can learn more about what does and does not work in our context. Over time the ability to delineate particular intervention types from the PRIMHD data set will strengthen as additional information is integrated into the data set.

It must be remembered that before the outcome of different interventions is evaluated, it is necessary to casemix the data to correct for differences in the type of people that receive services.

### Report

Ten northern DHBs were interested in understanding the impact of more explicitly Maori culture-based services on outcomes for Maori tangata whaiora and their whanau using the DHB's Child and Adolescent Services. A service audit tool was utilised to categorise CAMHS teams into three categories:

1. kaupapa Maori services,
2. mainstream services with strong Maori cultural input (e.g. Kaumatua or cultural advisors routinely involved as part of the assessment and treatment), and
3. mainstream services without strong Maori cultural input

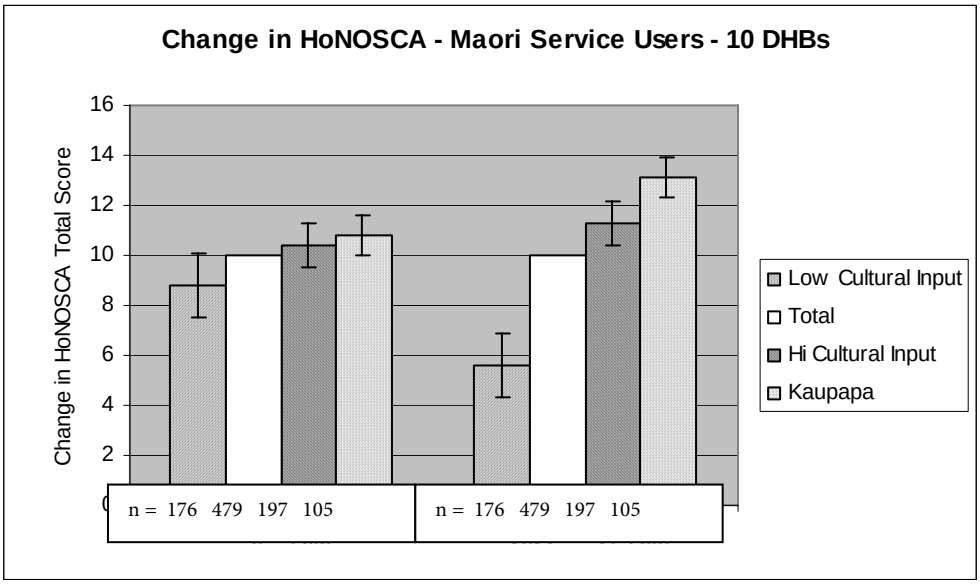
[6] This data is for demonstration purposes and does not represent real service results

The outcomes for all Maori clients of these services were evaluated using the following:

1. the change in total HoNOSCA score from admission to discharge, and
2. the number of unplanned Crisis Team contacts or similar need for crisis input during the time of contact and in the six months following discharge (recorded as average number of contacts per year).

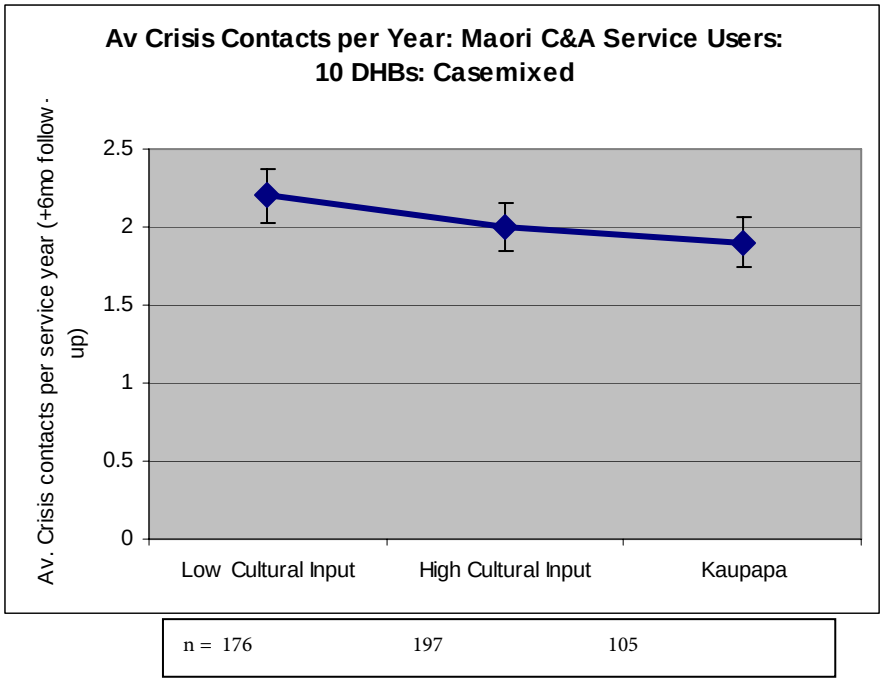
The following graph shows the data before and after casemixing for the three types of services. The data before casemixing (raw data) showed relatively little difference in the overall change in HoNOSCA between the three types of services - the confidence intervals for three overlap. However, the three types of services saw quite different types of clients so when casemix adjusting was carried out to compensate for these differences the relative outcomes looked quite different. The second set of bars on the graph shows relative change in HoNOSCA score after casemix adjustment. This indicates that Kaupapa Maori services and mainstream services with high cultural input achieved much better outcomes than mainstream services with low cultural input. Kaupapa Maori services showed somewhat better outcomes (that were just statistically significantly different) than high-cultural input mainstream services.

### Figure 18: Change in HoNOSCA - Maori service users, 10 DHBs




The average number of unplanned crisis contacts per year of service (including the period of six months after discharge) was also evaluated as an indicator of outcomes. This is based on the assumption that fewer crisis contacts meant better outcomes and/or a service that was being more pro-active about preventing crisis situations arising. The casemix adjusted data is presented below. These results show a trend towards lower crisis contacts by Maori service users who receive services with a stronger cultural component. Overlapping confidence intervals tells us that we cannot regard the differences between the groups as being statistically significant.

**Figure 19: Average crisis contacts per year - Maori C&A service users, 10 DHBs casemixed**




## Interpretation

*The interpretation of these findings is described in the previous report section. These results indicate better outcomes (as measured by their HoNOSCA total scores) for service users who have a higher level of cultural input, particularly those in kaupapa Maori services, but show no clear evidence of a similar result in crisis service usage.*

## Comment

This example shows the use of standard measures to explore the outcome of different types of interventions. Here the different levels of intensity of culturally-relevant input was explored. The same approach could be used to compare the outcomes of different models, and types of interventions such as:

- different types of clinical interventions (for example, those participating in a therapeutic group compared with those not participating or those receiving different types of medication).
- different models of service delivery within NGO services.
- different styles of service delivery (e.g. less or more assertive case management).

Remember that the outcomes shown here are not purely service outcomes (i.e. not purely the impact of the service on the service user's status) because other factors that are not related to the service (such as individual actions and the impact of societal changes) also contribute to this change. However, from a recovery-focused and client centred approach, the total change is of great importance.

It is not unusual to see, as in the example above, that some measures (in this case the HoNOSCA total score) show significant changes while others do not (in this case the average number of crisis contacts). There are many different reasons why there may be this type of variability in results including:

- the intervention may affect some potential outcomes but not others;
- some measures are more sensitive to therapeutic change than others; or
- other statistical reasons.

Therefore, finding different results for different measures does not invalidate your findings.

## Using PRIMHD data other than the HoNOS family of measures to evaluate outcomes[7]

### Question(s) addressed

What are the outcomes of services, using PRIMHD data other than HoNOS measures?

### Description

As previously discussed, PRIMHD is establishing collection of standard data for mental health services and service users in New Zealand. This data set incorporates and extends data collected for several years through MHINC and data collected in some parts of the country in recent years through the MH-SMART initiative. Eventually, additional standard measures will be added to the PRIMHD data set.

Most PRIMHD data is concerned with describing, in broad terms, characteristics of the service user, their interactions with the mental health system and the context of the data that is collected. However, another PRIMHD variable - Referral End Code - may be useful for helping services understand their outcomes. An example of the use of this is presented below.

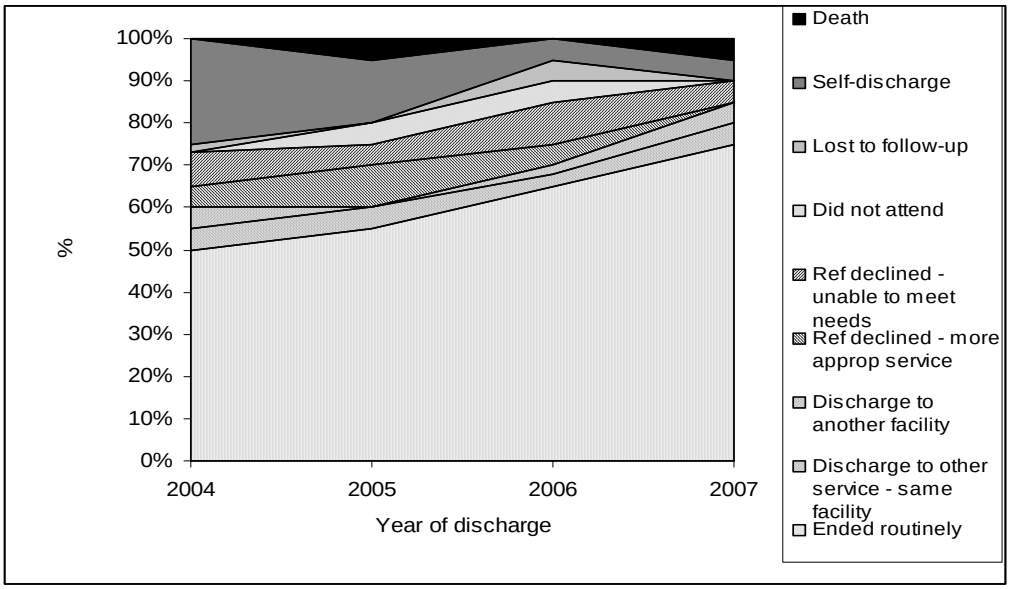
### Report

This report describes the percentage of service users discharged from (or referred but not admitted to) a service for different reasons, as recorded by the Referral End Code. It shows trends in the percentage of each Referral End Code over time. Some reasons for discharge can be regarded as more positive than others and the report shows Referral End Code categories ordered from (in general) most favourable at the bottom to least favourable at the top.

The percentage of discharges for each Referral End Code from a residential adult mental health rehabilitation facility for people with high and complex needs over a four year period is shown. The descriptors for each code value can be seen in the key at the right of the report.

[7] This data is for demonstration purposes and does not represent real service results

### Figure 20: Referral End Code by Year - ***** Residential Rehabilitation Facility




| N = | 193 | 141 | 187 | 132 |
|---|---|---|---|---|

## Interpretation

- *In this report, the outcomes are shown as bands that are loosely ranked from most favourable outcome (Ended routinely) at the bottom to least favourable outcome (Death) at the top. The broader the band, the more common that type of outcome.*
- *If the service was showing an improving pattern of outcomes over time, as rated by this measure, you would expect to see the lower bands becoming wider and moving closer to the right-hand top corner (as seen in this example).*
- *In this service, in 2004, only 50% of discharges occurred routinely with what staff regarded as treatment being completed as planned. Staff used the data above to identify that 25% of discharges occurred when service users self-discharged against medical advice (the dark grey band towards the top of the graph).*
- *Using this data, in early 2005 the staff developed a program to identify what lead to these self-discharges. This lead to a number of initiatives that reduced the proportion of self discharges to 15% in 2005 and 5% by 2006.*
- *At the end of 2006 the staff noted that inappropriate referrals (Referrals declined because the service was unable to meet needs requested and Referrals declined because other services were more appropriate) - represented on the graph by the two bands with strong diagonal stripes) typically accounted for about 15% of their discharges and had changed little over the last 3 years.*

- *In response to this data, in early 2007 the staff developed a program to educate likely referrers about their service and its admission criteria. Data for the end of 2007 suggests this has been successful as the proportion of inappropriate referrals (from around 15% for the previous three years) has reduced to about 5%.*
- *Any deaths that occurred in the service were intensively reviewed using Special Incidents procedures to minimise the risk of future similar events.*
- *The proportion of service users for whom "Ended routinely" was the Referral End Code had risen from 50% for 2004 to 75% for 2007, suggesting that more people were leaving the service with a successful outcome.*
- *The "n =" box indicates the number of discharges from the service during the year it is under.*

### Comments

This is an example of how PRIMHD information other than the HoNOS family of measures can be used to track service outcomes and to identify targets for service improvement.

It should be noted that the categories in the Referral End Code variable are quite broad so interpretation of this data needs to be considered carefully. For example, the "Deaths" category does not differentiate suicide from death from natural causes. Similarly, the ranking of codes by their desirability as an outcome is somewhat arbitrary.

Different types of services will have distinctly different Referral End Code profiles. Casemixing would be necessary to compare the profiles of different types of services, and the results may still be difficult to interpret validly. Analyses of a single service over time, as shown above, is likely to be more interpretable than comparisons between services.

## Using outcomes data to benchmark against other services[8]

### Question(s) addressed

How does this service compare with similar services nationwide?

### Description

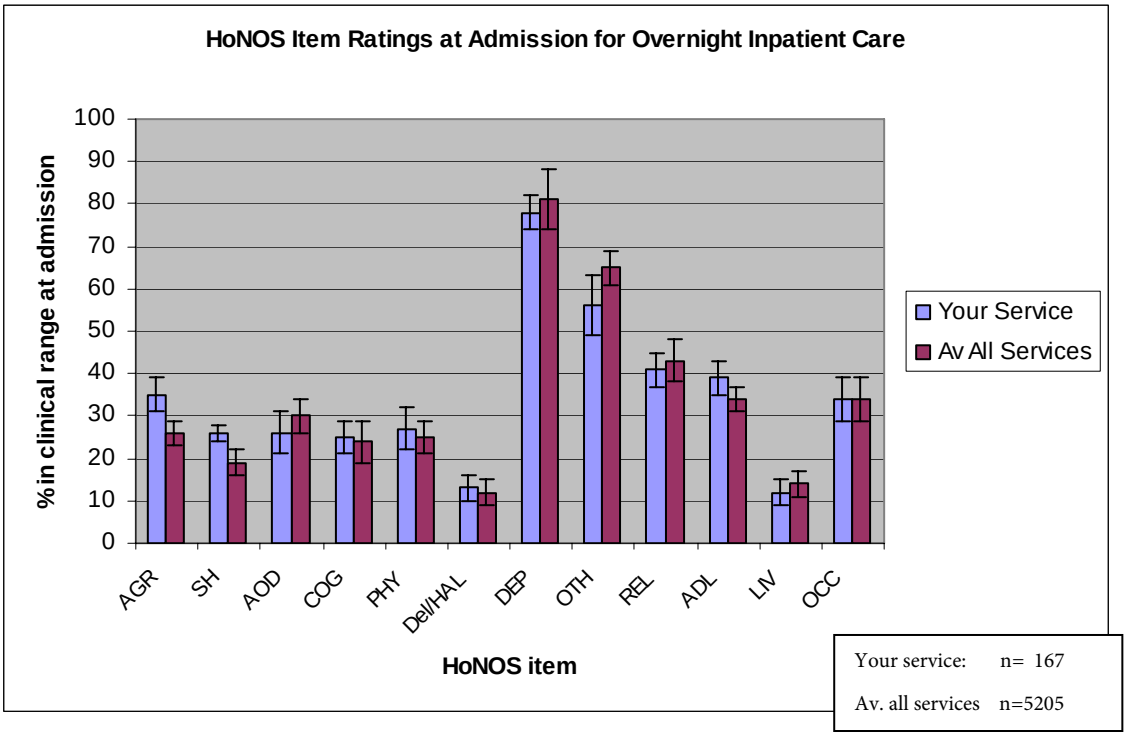
Benchmarking involves comparing characteristics of one service with other services or set standards. Set standards may include prescribed levels of expected performance (e.g. more than 90% of service users have a discharge planning meeting). They can be used to compare the service to others. For example, casemix adjusted outcomes of a service may be compared to casemix adjusted outcomes of a similar service that is identified as providing best practice and achieving good outcomes.

[8] This example is partly based on real service data

## Report

The following example shows a benchmarking report of the proportion of clients scoring in the clinical range on the HoNOS items at admission in a six month period for a single inpatient unit compared with the average proportion for all inpatient units.

### Figure 21: HoNOS Item Ratings at Admission for Overnight Inpatient Care




## Interpretation

- *We can see that the confidence intervals for all items except for Items 1 (Aggressive/overactive behaviour) and 2 (Non-accidental self-injury) overlap. Therefore, we conclude that for all items except 1 and 2 the pattern of admissions to this unit are quite similar to the general pattern.*
- *However, it is noted that the proportion scoring in the clinical range on the "overactive, aggressive, disruptive or agitated behaviour" item is significantly higher than average (shown by the non-overlapping error bars). The proportion in the clinical range is 35% as opposed to 26% for the average of all hospitals.*
- *The percentage in the clinical range on Item 2 (Non-accidental self-injury) is also significantly different, about a third higher than the average (25% as opposed to 19%).*

- *Understanding this lead to exploration about how the service could be modified to better address service users' needs. Examples of this exploration, and subsequent action, were:*

    1. *Given the higher than average level of aggressive/overactive behaviour at admission for this unit, careful consideration was given as to how to make the admission environment and processes as de-escalating as possible, resulting in subtle but significant changes in practice.*
    2. *Further analysis showed that 56% of those entering with clinical levels on Item 1 were intoxicated on alcohol or other drugs at the time of admission. Changes in process were made to address this issue in a de-escalating way prior to focusing on other aspects of the formal admission process.*
    3. *Feedback sought from service users and service user advisors also identified aspects of the Community Crisis Team approach when admitting people to the inpatient unit that could usefully be modified to reduce the level of aggressive/overactive behaviour for some service users.*
    4. *Additional training for self-harm recognition and prevention strategies in the inpatient setting was established to up skill staff in minimising the risk of self-harming behaviour in the unit. Some of these strategies "dovetailed" closely with strategies that were likely to be helpful in further assisting de-escalation of aggressive/overactive behaviour.*

### Comments

Benchmarking can be used in many ways to compare services. However, effective benchmarking requires an open mind and provision of accurate information from all services.

Benchmarked data does not take into account the contexts where data is collected. However, one of its major purposes is to pose questions rather than answer them. Taking this kind of data and using it as the starting point to understand what contextual, practice or other factors might contribute to what is found can be a complex but very useful task.

## Casemix-adjusted comparison of the outcomes of services

### Question(s) addressed

How do the outcomes achieved by different services compare after taking into account the difference between the populations served by those services?

This question may apply to the comparison of individual teams or whole DHBs.

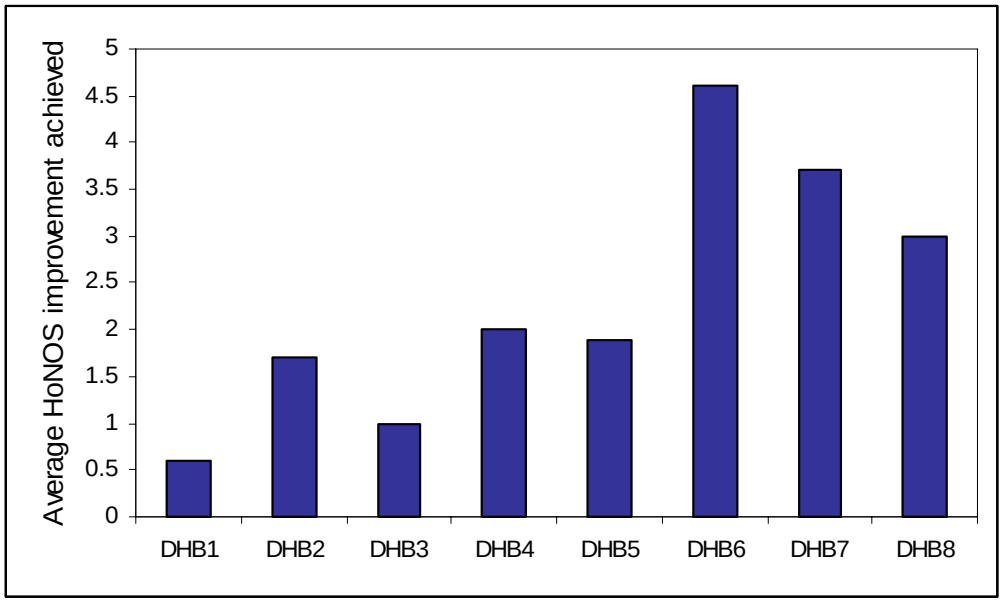
### Data Used

This approach may be used with many different outcome measures. An example would be the total score of the HoNOS.

### Description

The following example (based on work by Kathy Eagar and Tom Trauer) used data from the CAOS study. The figure below shows the change in average HoNOS total score achieved by the 8 DHBs. DHB 6 achieved the most change, with service users at this DHB improving by an average of over 4.5 HoNOS points. DHB 1 achieved the least change, with the average improvement being just 0.6 of a HoNOS point.

**Figure 22: Average HoNOS total score improvement for 8 DHBs (not Casemix adjusted)**




From this it would appear that DHB 6 is overall offering the most effective services. However, this analysis does not take into account the difference between the groups served by the different DHBs and the impact this may have on the outcomes. There are two possible reasons for the difference in the change scores achieved DHB1 and DHB6. One possible reason is service users at this DHB achieve better outcomes because DHB 6 delivers more effective mental health care. Another possible reason service users at DHB 6 achieve better outcomes because, as a group, they are different to those seen by DHB 1 and are in some sense easier to treat effectively.

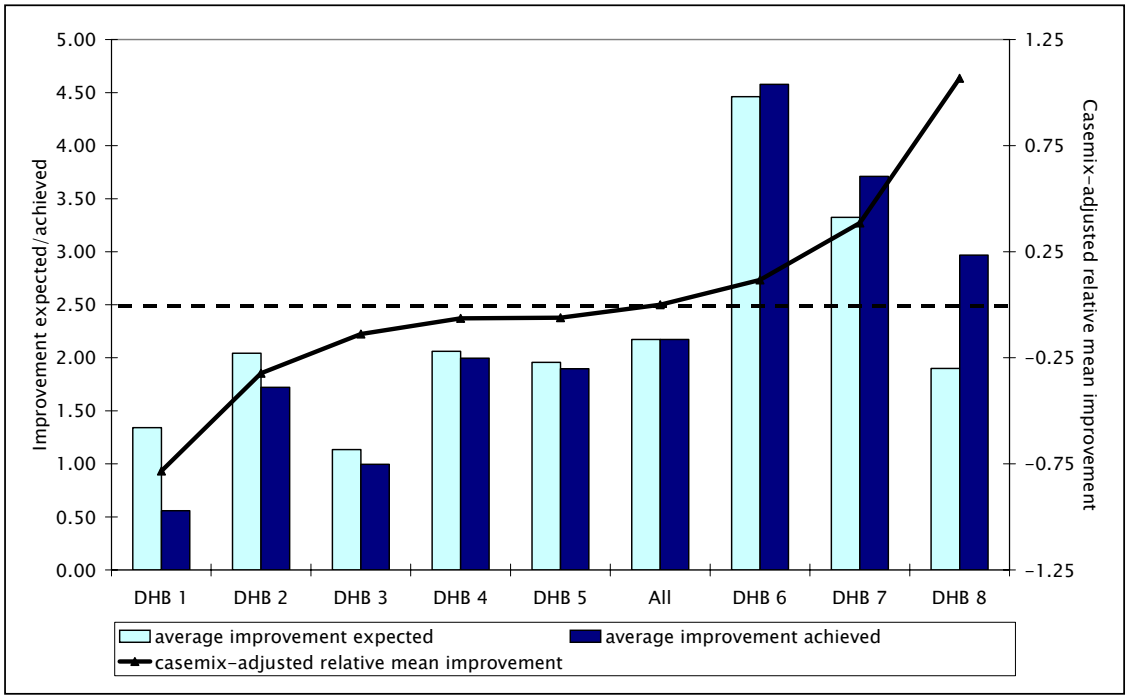
Therefore, comparison of the performance of the various DHBs can only be meaningful if the types of service users using the different services are taken into account. One obvious difference between the DHBs is that they have a different mix of community and inpatient services. But there are also other possible differences such as differences in the age profile, ethnicity profile or problem severity profile. Factors such as these needs to be taken into account to make meaningful comparisons.

To control (compensate) for differences in the mix of service users, the results must be casemix adjusted. Casemix adjustment assigns data to a Casemix class based on the characteristics of the service user. Service users with similar characteristics are assigned to the

same class. When outcome results are standardised in this way, any remaining differences can be attributed to differences between the services provided.

### Report

**Figure 23: Casemix Adjusted HoNOS Total Score Improvement for 8 DHBs**




### Interpretation

- *The first bar for each DHB shows the average improvement that would be expected after casemix adjustment. The second bar shows the actual average improvement achieved by the DHB (the same as the graph above).*
- *The black line shows the difference between the expected results and the achieved results (i.e. the difference of the two bars for each DHB). The scale for this is on the right.*
- *Those DHBs for which the black line is below the centre level (dotted line) are performing more poorly than expected on average after casemix adjustment. Those DHBs for which the black line is above the centre line are performing better than average after casemix adjustment.*
- *The higher the black line above the centre line, the more positive the outcomes of the service after casemix adjustment. The further the black line is below the centre line, the poorer the outcomes of the service.*
- *A negative casemix adjusted relative mean HoNOS improvement score indicates that service users achieved less improvement than similar service users in the national study did.*
- *Thus, based on the non-casemixed data it appeared that DHB 6 was offering the most effective service, after casemix adjustment it is clear that DHB 8 is actually offering the most effective service when the different mix of clients for the different DHBs are taken*

*into account. DHB 8 showed the biggest positive difference between its actual performance and its expected performance.*

- *After casemix adjustment, DHB 1 was still shown to be achieving the lowest level of performance. Its actual performance was well below what would have been expected given its mix of cases.*
- *Although their change in average HoNOS scores are quite different, from the second graph it can be seen that DHB 5 and DHB 6 were both performing very close to the level that would have been expected given the mix of cases they service.*
- *These results now become a starting point for trying to understand why the different services are performing differently - for instance what it is about the service provided by DHB 7 and DHB 8 that was more effective, and what it was about the service provided by DHB 1 and 2 that was less effective. The answers to this would clearly be complex, but potentially very fruitful. The value that Casemix has added to this is that now we more accurately know where to look to explore how to make services that work more effectively.*

### Comments

This example shows how Casemix is used to adjust for the differences in the mix of cases served by the different services. This prevents services that work with service users who have a more difficult path to recovery being disadvantaged if their outcomes are compared with other services with service users who typically recover more easily.

# Examples of information systems/tools to support data use

## Mental Health Outcomes System (MHOsys)

### Question(s) addressed

Standard measures data has great potential to support answers to many specific questions that information users may wish to ask in particular contexts. These may include (but are not limited to) questions such as what are the characteristics of the users of this service? What outcomes do we achieve for a specific kind of service user or a specific part of our service? What trends are occurring in our outcomes? Or what is the clinical profile of a specific kind of service user on admission? There will be many similar types of questions that a service could benefit from being able to answer. Some of these questions may be specific to a particular service or issue that occurs at a set time, so having a way of asking these questions other than through standard reports would be valuable.

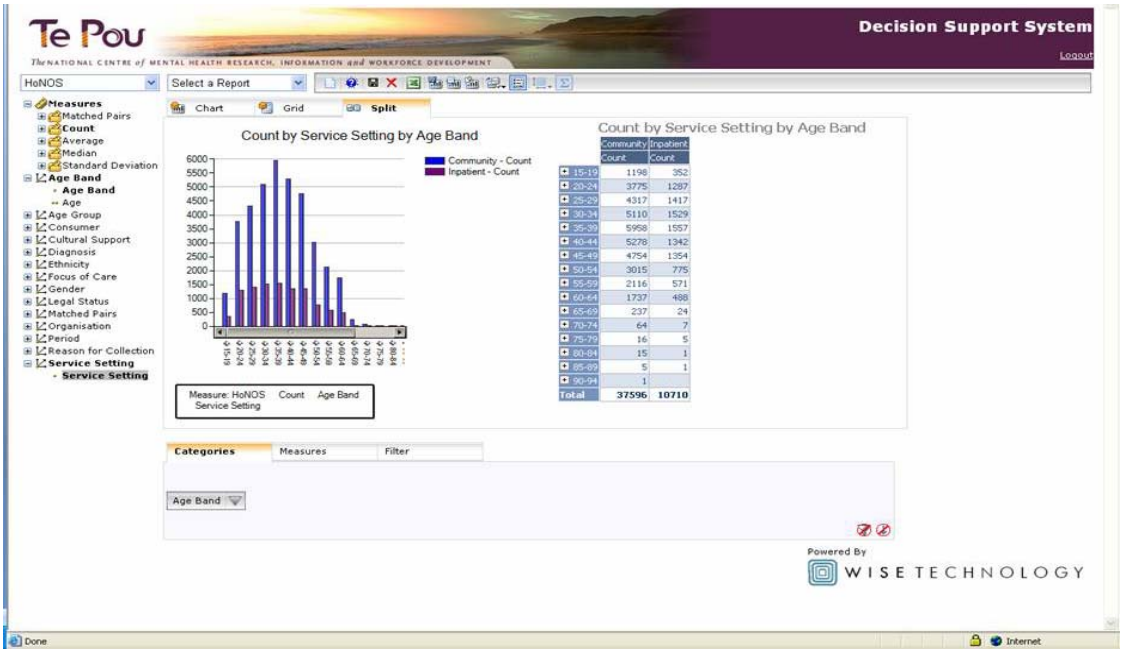
### Description

**MHOsys** is a web-based computer application that allows users to select and analyse standard measures information so that many sorts of questions can be answered. The information user can select cases on the basis of particular variables, and then undertake a range of analyses by selecting the variables of interest for the question being asked. It is able to present data from a single team or DHB up to national data. Australia has recently developed a similar tool that gives information users access to all national data in an aggregated and de-identified form. Matched or unmatched cases can be used. **MHOsys** is currently being piloted using data collected by several DHBs as part of Te Pou's Information Programme Test Project. Participating DHBs can look at their own data and compare this to aggregated data from all the DHBs. It is internet based and is accessed by accredited users with a username and password.

### Report

An example of an output from **MHOsys** that addresses the question "How many HoNOS do we have for service users in different age bands in our community services and our inpatient services?" follows. The column on the left allows selection of the variables of interest, the type of analysis to be presented and the variables to be used as the dependent variable or for defining sub-groups or selecting the data. A series of standard reports can also be selected from the menu on top. Data is presented as numbers and/or graphically.

## Figure 24: Mental Health Outcomes System (MHOsys)




### Interpretation

- *This data table and graph show the distribution, by age, of the number of completed collections of HoNOS data in some community and inpatient services.*
- *By selection of different variables from the column on the left, a great many different analyses could be undertaken, providing answers to a wide variety of questions. These variables can also be used to "filter" the data so particular groups (e.g. males or people with a particular type of diagnosis) can be analysed on their own. These variables can also be used to divide up the data set so that groups with different characteristics can be compared.*

### Comments

**MHOsys** is currently under development with plans to include casemix adjusted data once the casemix grouper has been developed. The flexibility of **MHOsys** makes it very useful for answering a wide variety of questions that could not be addressed in standard reports. However, it also increases the risk of using data inappropriately by drawing conclusions from an analysis that are not justifiable. It is for this reason that access is restricted to users who are fully conversant with the protocol under which standard measures data is collected and who have received training in the use of **MHOsys**.

Great care has been taken to ensure the confidentiality of individual service users data within **MHOsys**.

# The Mental Health "Dashboard"[9]

## Question(s) addressed

How is this service performing over time and compared with other mental health services on particularly important indicators?

## Description

The Mental Health Dashboard is a visual display that is designed to regularly give an organisation or service a quick snap-shot showing how it is functioning, based on a number of particularly important indicators. This information is often provided alongside that from similar services to allow comparison of performance (benchmarked dashboard).

Counties Manukau DHB (CMDHB) has been developing a Mental Health Dashboard which adapts the Balanced Scorecard approach (Kaplan & Norton, 1992) and measures a small number of variables that are considered to reflect service performance across four broad domains:

- service users,
- outcomes,
- services, and
- quality.

The Mental Health Dashboard presents information in the same way each time so that information users become familiar with the format and understand and interpret the meaning of information and trends.

## Dashboard Report

The dashboard report is in its first year of use by the Community Living Service (CLS) providers funded by CMDHB, but will be extended over 2009 to include a set of dashboards which will be customized for use by other mental health and addiction non-government organisations (NGOs) funded by CMDHB, the DHB provider arm and the DHB funder.

The Dashboard report is currently supplied to CLS providers every three months and forms a core part of the CLS provider benchmarking forum. In the future, the dashboard report will be provided for various levels of the organisation from the whole DHB down to individual services or teams within that DHB.

## Interpretation

The Dashboard will provide a basic set of important information that can be compared across time and between services. This information can then be used to identify areas for service improvement and development.

---

[9] Thanks To Phillipa Gaines for information on the Mental Health Dashboard

***Proposed measures for the dashboard report:*** The dashboard report will consist of approximately 16 measures that focus on some of the most important indicators of how a service is functioning, organised in the four domains described above. The preliminary range of indicators for DHB provider arm services are shown in the following table.

**TABLE 2: PROPOSED LIST OF DHB (PROVIDER ARM) DASHBOARD VARIABLES**

| | |
|---|---|
| **Consumers**<br>***(who used the service?)***<br>• Diagnosis and duration of care<br>• Average number of clients per actual FTE<br>• % of current clients seen, by gender & ethnicity compared to DHB population | **Outcomes** ***(impact of service?)***<br>• % of current clients in regular housing (*casemix*)<br>• % of current clients in open employment or education (*casemix*)<br>• % of current clients with deterioration in housing and employment since registration (*without a diagnosis of a psychotic disorder only*)<br>• Improvement based on HoNOS ratings - using all 12 items of the HoNOS |
| **Services** ***(what was done?)***<br>• Average monthly cost for clients using services prior to this team (*casemix*)<br>• Average monthly cost over last 12 months compared with DHB average for this type of team (*casemix*)<br>• Average annual cost<br>• Exited current clients - total cost of services compared with DHB average (for consumers with '*no psychotic disorder' only*) | **Quality** ***(how good was it?)***<br>• % of current clients in services >3 months who have used crisis, respite or inpatient services (incl average LOS for inpatients) in last 3 months<br>• % of current clients in service > 3 months who are currently under CTO<br>• Average weekly time spent in direct service delivery (SMO and other staff)<br>• Pre-admission community care<br>• Post-admission community care |

### Comments

To make it as meaningful as possible, collection of the information included in the Dashboard needs to be as consistent as possible over time and between different organisations and sites. This approach encourages stakeholders to focus on priority improvement goals in each area, instead of focusing exclusively on only one or two areas (such as the cost of services) to the exclusion of others. It also allows for the monitoring of dynamic relationships amongst the domains as they are obviously inter-connected and improvements in one part of the system may not necessarily translate to improvements in another. In fact, sometimes the reverse happens.

The Mental Health Dashboard is being developed for use with NGO providers and for DHB clinical teams. There are slight differences in the report being developed for NGO and DHB providers, reflecting their different roles within the recovery process. Routinely comparing the Dashboard results of different services at all levels will make it as useful as possible in benchmarking service processes. All services will have access to regular reports on their performance relative to everyone else. Sharing these reports will be a starting point for teams/providers to discuss, identify and build on best practice to improve the overall quality of services.

## Future Analysis and Reporting Tools

As further systems and tools are developed, these will be highlighted and accessible through Te Pou. Information Utility Training will incorporate these new developments.

# Glossary

This section defines some of the key terms that are used in this document. Where statistical and similar terms are defined, the definitions are designed to make the meaning as clear as possible to the average reader and are not formal statistical definitions of the terms.

**Aggregated data:** Data is aggregated when data from more than one individual (e.g. more than one service user) is grouped together. For example, the admission HoNOS scores for all service users using a respite service might be aggregated so the mean HoNOS score of users could be calculated.

**Benchmarking:** Benchmarking is a means of comparing aspects of different services (such as performance, processes or cost) with each other. Benchmarking often involves:

1. comparing different services with each other on the same measures, or
2. identifying best practice in a certain area and using this as a standard which other services can be compared with. For benchmarking to be meaningful, it is often necessary to casemix adjust (see below) the results for different services to account for differences in the types of people served by different services and programmes. For a detailed discussion of the role of benchmarking in mental health services see Bainey (2005).

**Casemix** is a term used to describe a system for grouping service users together in categories that have similar characteristics.

**Casemix adjustment** is the process of using a casemix system to statistically correct for the difference between the types of people assisted by different services so that the impacts of differences in what the service does can be seen more clearly (Herz, et al., 2006). Casemix adjustment allows us to compare services which have a different mix of cases. If the outcomes of two services were compared without casemixing the data, we could not be sure if better outcomes in one service were due to differences in what the team did or in who it served. If one team showed a better outcome that the other after the data had been casemix adjusted, we would have more confidence that the better outcome was due to that team doing something better. We might then study the teams to try and identify what that team was doing that was working better.

**Clinical significance:** See significance

**Confidence intervals** are very useful when you have aggregated data. Confidence intervals are used to help us:

1. determine if two groups are statistically significantly different (e.g. the difference between the HoNOS score for one group of consumers and another group),
2. determine if a group mean (or another parameter such as the slope of a line on a graph) is significantly different from a meaningful value; e.g. to assess if the HoNOS change score for a group is significantly different from zero).

Although it is not completely accurate in every situation (Wolfe & Hanley, 2002), a very useful rule of thumb that is very much in common practice is to make the confidence intervals two standard errors on either side of the sample mean. When we set up the confidence interval in this way, we can be 95% sure that the true mean for the whole population will be within this confidence interval. Confidence intervals are used to assess the significance of a difference as follows:

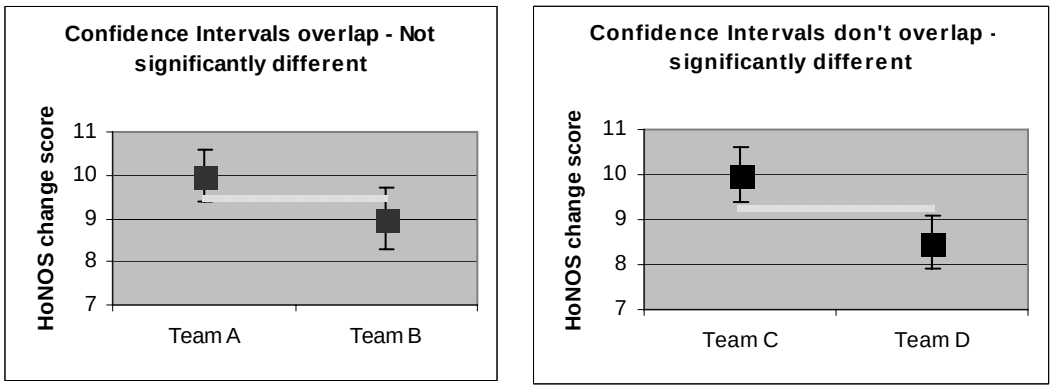
- **Comparing the parameters (e.g. two means) for two groups:** If the confidence intervals around the parameters for the two samples do not overlap then we can assume that the parameters are truly different. If the confidence limits do overlap, we assume that there is no real difference between the group means.

  *For example*: Ward A in an inpatient unit had an average admission HoNOS score of 28 (Confidence interval 24-32), while Ward B in the same unit had an average HoNOS score of 25 (Confidence interval 23-27). Staff wondered if the differences in means meant that Ward A tended to get more acutely unwell clients. However, when they looked at the confidence intervals for the two wards they saw that the confidence intervals overlapped (scores from 24-27 were in the confidence interval of both Wards). Because of this overlap, they concluded that there was no evidence of a significant difference in the HoNOS scores. Without looking at the confidence intervals they may have drawn a false conclusion that there was a significant difference.

- **Comparing a group parameter with a particular value:** If the confidence limit does not include the particular value of interest then we assume the true mean is different from this.

  *For example*: A particular service had a mean (average) HoNOS change score over three months of 5 (an improvement of 5) with a confidence interval of 3-7. As this interval does not include 0 we can assume that consumers in this service showed a real improvement. However, if the mean score had been 5 and the confidence interval was -1 to 11 then the score of 0 is within the confidence interval range so we could not validly conclude that the consumers showed a real improvement.

### Figure 25: Examples of Confidence Intervals




**Focus of Care:** Focus of Care is a measure by which the clinician indicates what has been the primary goal of care over the last period of time. The Focus of Care variable that is used in the PRIMHD data set has four types the clinicians could choose from:

1. **Acute** (main goal - the short term reduction in severity of symptoms and/or distress),
2. **Functional Gain** (main goal - to improve personal, social or occupational function or promote psychosocial adaptation),
3. **Intensive Extended** (main goal - to prevent or minimise deterioration and reduce risk for people with a stable pattern of severe symptoms, frequent relapses or severely limited ability to function independently), and
4. **Maintenance** (main goal - to maintain current status of a consumer who is stable and functioning relatively independently).

The goal of care for consumers may change over time as they move through different phases of illness. Clinicians are asked to identify what has been the main focus of care in the last period, even though there may have been occasions in which the consumer required care of a different type.

**He Korowai Oranga**: A cloak of health/wellbeing - Maori Health Strategy (Ministry of Health/Manatū Hauora NZ, 2002)

**Karakia**: Prayer, incantation, blessing

**Korowai**: ornamented cloak, a metaphor to represent that which envelopes tangata whaiora in a sense of well-being in the context of whanau ora. (Whanau ora is our korowai of recovery).

**Kaumatua:** elder or older person

**Outcomes:** Strictly, *service outcomes* means any change in the users of a service that occurs as a result of contact with that service. However, the term *outcomes* is often used more informally to describe any changes observed over time in users of a service. To determine whether changes in standard measures result purely from the activities of the service often requires rigorous research methods. However, if the changes observed are not necessarily the result of service activities, comparing standard measures over time as a proxy for outcomes is often useful to better understand services.

**Significance:** In the context of data analysis, significance means that the result seen is sufficiently large or strong to be "meaningful" and/or "real". When analysing standard measures in mental health, we may be particularly concerned about two types of significance:

> **Statistical significance** means that the result of a statistical analysis is sufficiently unlikely to have occurred by chance due to random variations and is likely to reflect a real effect (e.g. a real difference between means or a real correlation between variables). *For example*: The casemix adjusted HoNOS change scores for Ward A of an inpatient unit were somewhat larger than the casemix adjusted HoNOS change scores for Ward B. If the difference between the two was statistically significant, we could assume that service users in Ward A were achieving better outcomes. However, if the difference was not significantly different, we would conclude that there was no real difference between the HoNOS change scores for the two wards.
>
> **Clinical significance** means that the result of an analysis reflects a clinically meaningful effect for the people being analysed. Sometimes it is possible to find statistically significant differences (i.e. we have reason to assume that an effect observed is real) but the difference is not clinically significant (i.e. it is real, but too small to make an overall appreciable difference to the people concerned). For instance in the example above, if the difference in total HoNOS change scores for Ward A and Ward B were statistically significant, but the change in the total score was only 0.2, we would probably say that it was not clinically significant, because it shows very little difference in the average clinical improvement for people on the two wards.

**Standard measures:** Standard measures are assessment tools that allow consistent quantification of some aspect of the consumer's experience or status. They are often standardised rating scales and other types of measures (such as biological measures and demographic information) that are completed using consistent evaluation criteria by or for all consumers who use services. They are completed by or for the consumer at consistent times during the consumer's contact with a service (e.g. on admission, at three month periods after that and on discharge). Standard measures are usually chosen only if they have shown evidence of **reliability** (they yield the same result in the same situation if completed by different people or at different times) and **validity** (they measure what they are purported to measure).

**Statistical Significance:** See significance

**Tangata whaiora:** a person seeking wellbeing, mental health service user

**Tuituia**: paper of the MH-SMART initiative written by MH-SMART Outcomes Initiative tangata whaiora/motuhake rōpu (2005)

**Whanau ora**: Maori families supported to achieve their maximum health and wellbeing. Whanau ora is identified by tangata whaiora whanau as their korowai for recovery, it locates tangata whaiora within their whanau.

**Te Taha Hinengaro:** refers to mental and emotional wellbeing of a person.

**Te Taha Tinana**: the physical wellbeing of a person, which cannot be dealt with separately from the whanau, wairua, hinengaro.

**Te Taha Wairua**: refers to the spiritual wellbeing of a person. It determines one's identity. It provides a direct link with one's tūpuna (ancestors) and whanau group and strengthens the taonga (treasure, special possessions) and tikanga (customs, values) of one's cultural system.

**Te Whare Tapa Whā:** Maori model which describes the four holistic dimensions of Maori wellbeing - Taha Whanau, Taha Hinengaro. Taha Wairua, Taha Tinana

**Te Taha Whanau:** involves acknowledging the importance of whanau in providing sustenance, support and an environment that is important to good health

**Tino Rangatiratanga**: (personal power), self determination, self-management

**Whakawatea:** a process or ritual used for clearing the way for a gathering or meeting

**Whanau**: extended family, family group

# REFERENCES


Bayney, R. (2005) Benchmarking in mental health: an introduction for psychiatrists. *Advances in Psychiatric Treatment, 11,* 305-314.

Brown,C.S., Burlingame,G.M., Lambert,M.J., Jones,E., & Vaccaro,J. (2001). Pushing the quality envelope: a new outcomes management system. *Psychiatric Services, 52,* 925-934.

Britz, G. C., Emerling, D.W., Hare, L.B., Hoerl, R.W., Janis, S.J., & Shade, J.E. (2000). *Improving Performance Through Statistical Thinking.* Milwaukee, WI, ASQ Quality Press.

Buckingham, B. & Eager, K. (1999). *Developing a Casemix Classification for New Zealand Mental Health Services: Report to the New Zealand Health Research and Development Strategy.* Auckland: Health Research Council.

Burdett, J. (unpublished). (2005) *Prove To Me You're Making A Difference: Consumer Use of Outcome Measurement To Leverage Better Services*. Mind and Body Consultants, New Zealand.

Callaly, T., Hyland, M., Coombs, T., Trauer, T. (2008). Routine outcome measurement in public mental health - results of a clinician survey. *Australian Health Review, 30,* 164-173.

Cheung, G., & Strachan, J. (2007). Routine 'Health of Nations Otcome Scales for elderly people' (HoNOS65+) collection iin an acute psychogeriatric inpatient unit in New Zealand. *New Zealand Medical Journal, 120,* U2660.

Chipps, J., Mellsop, G. & Stewart, M. (unpublished). *The Use of MH-SMART Standard Measures Data: A Guide to the Improvement of Mental Health Services.* Auckland: Mental Health Research and Development Strategy.

Coombs, T. (2001*) NSW Mental Health Outcomes and Assessment Training (MH-OAT): Facilitators Manual.* Sydney: New South Wales Ministry of Health.

Copeland, M.E. (2002). *Wellness Recovery Action Plan.* Los Angeles: Peach Press

Covey, S.R. (1989), *The Seven Habits of Highly Effective People.* New York: Simon and Schuster.

Deegan, P. (1996). Recovery as a journey of the heart. *Psychiatric Rehabilitation Journal, 19,* 81-97.

Durie, M. (1998). *Whaiora - Maori health development* (2nd Edition). Auckland: Oxford University Press,

Eager, K., Gaines,P., Burgess,P., Green,J., Bower,A., Buckingham,B., & Mellsop, G. (2004). Developing a New Zealand casemix classification for mental health services. *World Psychiatry, 3,* 172-177.

Felton, J..S.. (1998). Burnout as a clinical entity – its importance in health care workers. *Occupational Medicine 48,* 237–250.

Few, S. *Information Dashboard Design: The Effective Communication of Data.* Sebastopol CA: O'Reilly Media

Gaines, P., Bower A, Buckingham B, Eager K, Burgess P, Green J. (2003) *New Zealand Mental Health Classification and Outcomes Study: Final Report*. Health Research Council of New Zealand: Auckland.

Ganju, V. (2006). Mental health quality and accountability: The role of evidence based practices and performance measurement. *Administration and Policy in Mental Health, 33,* 659-665.

Herz,L., Euseb,S., Breckenridge,J.N., & Rosen,A.K. (2006). Development and validation of a psychiatric case-mix system. *Medical Care, 44,* 568-580.

Hunter,C.E. & McFarlane,A.C. (1994). DRGs and Australian psychiatry. *Australian and New Zealand Journal of Psychiatry, 28,* 114-120.

Lakeman, R. (2004). Standard routine outcome measurement: Pot-holes in the road to recovery. *International Journal of Mental Health Nursing, 13,* 210-215.

Lambert,M.J., Harmon,C., Slade,K, Whipple,J.L., & Hawkins,E.J. (2005) Providing feedback to psychologists on their patients' progress: Clinical results and practice suggestions. Journal of Clinical Psychology/In Session, B (2), 165-174.

Levy M, Mar M, van Altvorst S, Sokratov A, Eade L, Meihana Te P, Winiata Te P. (2005). *Using HoNOS/HoNOS 65+/HoNOSCA to Enhance Outcomes for Tangata Whaiora/Motuhake and their Whanau: Principles to Guide Practice for Te Pou.* Auckland.

Lloyd-Evans, B., Johnson, S., & Slade, M. (2007). Assessing the content of mental health services: A review of measures. *Social Psychiatry and Psychiatric Epidemiology, 42,* 673-682.

Mellsop, G.W, & O'Brien, C.G. (2000). *Outcomes Summary Report.* Auckland: New Zealand Health Research Council

Mental Health Commission. (1998). *Blueprint for Mental Health Services in New Zealand.* Wellington: Author.

MH-SMART Outcomes Initiative tangata whaiora/motuhake rōpu. (2005) *Tuituia.* Wellington.

Miller W.R., & Rollnick,S. (1991) *Motivational Interviewing: Preparing People to Change Addictive Behaviour*. New York, Guilford.

Ministry of Health (2008) *Let's get real: Real Skills for people working in mental health and addiction.* Wellington.

Ministry of Health Manatū Hauora. (2002) *He Korowai Oranga - Maori Health Strategy*. Wellington.

Mirin S.M. & Namerow M.J. (1991). Why study treatment outcomes*? Hospital and Community Psychiatry, 42,* 1007 - 1013

New Zealand Privacy Commissioner (1994). *Health Information Privacy Code 1994.* Wellington: Author.

Onyett, S. (1992). *Case Management in Mental Health.* London: Chapman and Hall.

Rapp, C. (1998) *The Strength Model: Case Management with People Suffering from Severe and Persistent Mental Illness.* New York: Oxford University Press.

Te Pou MH-SMART suite of measures report to the Ministry of Health, 2009. (un-publish report)

Trauer, T., Eager, K., & Mellsop, G. (2006) Ethnicity, deprivation, and mental health outcomes. *Australian Health Review, 30,* 310-321.

Tufte, E.R. (1983). *The Visual Display of Quantitative Information.* Cheshire: Graphics Press.

Wing, J. K., Beevor, A. S, Curtis, R. H., Park, S .B. G., Hadden, S., & Burns, A. (1998). Health of Nations Outcome Scales (HoNOS): Research and Development. *British Journal of Psychiatry, 172,* 11-18.

Wolfe, R. & Hanley, J. (2002). If we're so different why do we keep overlapping: When I plus 1 doesn't make 2. Canadian Medical Association Journal, 166, 65-66.

# Appendix A: HoNOS User Guide

| | | | | Monitor? | Active treatment or management plan |
|---|---|---|---|---|---|
| Not clinically significant | 0 | No problem | Problem not present | × | × |
| | 1 | Minor problem | Requires no formal action. May or may not be recorded in clinical file. | **Maybe** | × |
| Clinically significant | 2 | Mild problem | Warrants recording in clinical notes. May or may not be incorporated into care plan. | √ | **Maybe** |
| | 3 | Moderate problem | Warrants recording in clinical notes. Should be incorporated into care plan. | √ | √ |
| | 4 | Severe to very severe problem | Most severe category for patient's with this problem. Warrants recording in clinical notes. Should be incorporated into care plan. *Note: patient can get worse* | √ | √ |