



Australian Mental Health Outcomes and Classification Network

'Sharing Information to Improve Outcomes'

An Australian Government funded initiative

**A review of clinician-rated instruments that could be used
to assess adults' levels of functioning in specialised public
sector mental health services**

Philip Burgess, Meredith Harris, Tim Coombs, Jane Pirkis

June 2016

Table of contents

Abstract	2
Background	2
Method	2
Results	2
Conclusions	3
Method	7
Results	9
Overview of identified articles and instruments	9
Hierarchical, criterion-based assessment of the instruments	12
Criterion 1: Is brief (<50 items) and simple to score	12
Criterion 2: Is not made redundant by more recent instruments	13
Criterion 3: Relevant version has been scientifically scrutinised	14
Criterion 4: Considers functioning in a contemporary way	14
Criterion 5: Demonstrates sound psychometric properties	15
Discussion	20
References	24

Abstract

Background

Functioning is one of the key domains that has been emphasised in the routine assessment of outcomes that has been occurring in specialised public sector mental health services across Australia since 2002, via the National Outcomes and Casemix Collection (NOCC). For adult consumers (aged 18-64), the Life Skills Profile (LSP-16) has been the instrument of choice to measure functioning. The continued use of the LSP-16 has been questioned, however. The current review set out to evaluate, against a set of pre-determined criteria, the most suitable existing clinician-rated instruments for routinely measure functioning for adult consumers. It did not rule out the LSP-16, but explored the potential of other candidate instruments.

Method

We used two existing reviews of functioning measures as our starting point, and conducted a search of Medline and PsycInfo to identify articles relating to additional clinician-rated instruments. We used a hierarchical, criterion-based approach to assess whether given instruments showed promise for measuring functioning in the Australian context.

Results

We identified 20 relevant instruments, five of which met our criteria. The LSP-16 remained in contention, along with the Health of the Nation Outcome Scales (HoNOS), the Illness Management and Recovery

Scale – Clinician Version (IMRS-C), the Multnomah Community Ability Scale (MCAS), and the Personal and Social Performance Scale (PSP).

Conclusions

The current review can help to inform decisions about which clinician-rated instruments hold promise for assessing whether functioning improves, deteriorates or remains the same for adult consumers of Australian mental health services. Further work is required to determine which, if any, of these instruments satisfy further criteria relating to appropriateness, acceptability and feasibility.

Background

The International Classification of Functioning, Disability and Health (ICF) recognises functioning as an essential component of health and wellbeing.¹ The ICF emphasises functioning over disability, focusing on what people have the potential to do and actually do, irrespective of their mental (or physical) health conditions. The ICF stresses two key elements of functioning: ‘activity’ (the execution of tasks) and ‘participation’ (involvement in life situations).¹

Functioning is one of the key domains that has been emphasised in the routine assessment of outcomes that has been occurring in specialised public sector mental health services across Australia since 2002, via the National Outcomes and Casemix Collection (NOCC).² Under the NOCC protocol, various outcome measurement instruments are administered for all consumers in ambulatory and community residential settings at set points in their episode of care. For adults (aged 18-64), who are the focus of the current report, the main instrument used to assess functioning to date has been the Life Skills Profile (LSP-16).³⁻⁷ The Health of the Nation Outcome Scales (HoNOS),⁸⁻¹² which is primarily used to assess symptoms, also contains a small number of items that relate to functioning. The LSP-16 and the HoNOS are both clinician-rated. More information about the full NOCC suite of instruments and the framework that guides their administration can be found elsewhere.¹³

In 2013, NOCC was reviewed by the National Mental Health Information Development Expert Advisory Panel (NMHIDEAP) which gathered information from a variety of sources, including multi-modality stakeholder consultations and analysis of NOCC data.¹⁴ With respect to adults, the review recommended that the instruments be rationalised, and that a simple clinician-rated instrument be developed that assesses functioning and symptomatology and, potentially, other relevant domains. Such an instrument might take the form of a single existing instrument, or alternatively it might be a composite of several instruments, but either way it should be brief. The HoNOS was promoted as the starting point for this

instrument, largely because it is relatively well accepted by clinicians, and is completed in the majority of episodes of care. The continued use of the LSP-16 was questioned, however, because of criticisms that it uses outmoded, inappropriate language and is not strengths-based, and because it is not optimally completed.¹⁴

Since the review, NMHIDEAP has given some thought to the development of the instrument. In particular, it has proposed a 'domain framework' that should guide the endeavour.¹⁵ This emphasises personal recovery, social recovery and clinical recovery. It has also suggested several options for how the new instrument should be developed: augmenting the HoNOS with a measure of functioning that replaces the LSP-16 and, if necessary, some additional clinically-relevant items; or constructing a new instrument that is purpose-designed to cover off all of the areas in the domain framework (again, this might have the HoNOS at its core). In addition, NMHIDEAP has proposed various sets of criteria which could be used to identify potentially useful existing instruments.¹⁵

We conducted the current systematic review in order to inform the debate about how functioning should be captured in the new instrument. We did this as part of our role with the Australian Mental Health Outcomes and Classification Network (AMHOCN), which has been responsible for data management, training and service development, and analysis and reporting related to NOCC since 2003.¹⁶ Our starting point was two reviews of functioning measures that had been conducted for different purposes. One of these looked at instruments that might be used in community managed organisations in Australia,¹⁷ and the other considered instruments that might be used in clinical services in New Zealand.^{18,19} Once we had considered the instruments that were shortlisted in these reviews, we conducted our own systematic review of the academic literature. We sought to identify articles that had been published since the original reviews, as well as any articles that might have been missed by these reviews. Our review aimed to answer the following question: *Is there an existing clinician-rated instrument that might be used to routinely measure consumer functioning in Australian specialised public sector mental health services?*

The current review set out to evaluate, against a set of pre-determined criteria, the most suitable existing clinician-rated instruments for routinely measure functioning for adult consumers.

Method

We conducted an iterative search of Medline and PsycInfo from their respective years of inception to April 2016 for journal articles that described relevant functioning instruments. In the first iteration, we searched titles and abstracts using the following search string: ('mental' OR 'psychiatr*') AND ('social function*' OR 'personal function*' OR 'community function*' OR 'social abilit*' OR 'personal abilit*' OR 'community abilit*' OR 'social perform*' OR 'personal perform*' OR 'community perform*' OR 'occupation* function' OR 'occupation* perform*' OR 'community participat*' OR 'community involve*' OR 'work' OR 'leisure' OR 'educat*' OR 'personal relationship*' OR 'interpersonal relationship*' OR 'social inclusion' OR 'living skill*' OR 'life skill*' OR 'self-care'). In the second iteration, we searched titles only for the names of identified instruments, in order to ensure that we picked up as many relevant articles on each as possible. We also searched the reference lists of key review papers and articles on individual instruments. Our search was restricted to English-language articles.

At the title and abstract screening and full-text screening stages, we excluded articles that made reference to instruments that could not be readily rated by a clinician without recourse to other information (e.g., consumer-rated instruments, instruments that required structured or semi-structured interviews with consumers or other informants, instruments that involved a systematic extraction of information from case notes). We also excluded articles on instruments that were designed for use with non-adult populations or clinically defined sub-populations (e.g., instruments designed for use with children and adolescents or older persons, instruments designed for use with people with intellectual disabilities, instruments designed for use in forensic mental health settings). In addition, we excluded articles on instruments that assessed only a limited aspect of functioning (e.g., instruments that were exclusively about activities of daily living, instruments that focused only on work performance).

Once we had identified our pool of relevant articles, we assessed whether each of the given instruments they described might be candidates for routinely assessing changes in functioning of consumers in Australian public sector mental health services. We did this using a hierarchical, criterion-based approach based on one that we used for a previous review of recovery instruments.²⁰ Under this approach, we progressively excluded instruments from further consideration if they did not meet a specific criterion. The criteria were:

1. Is brief (<50 items) and simple to score;
2. Is not made redundant by more recent instruments;
3. Relevant version has been scientifically scrutinised;
4. Considers functioning in a contemporary way; and
5. Demonstrates sound psychometric properties.

Results

Overview of identified articles and instruments

An overview of the identified articles is provided in Figure 1. In total, our search identified 5,908 journal articles. Removal of duplicates and screening titles and abstracts left 336 full text articles, a further 81 of which were excluded when the full text was reviewed. The remaining 255 articles provided information about 20 clinician-rated instruments designed to measure functioning.

Figure 1: Article selection

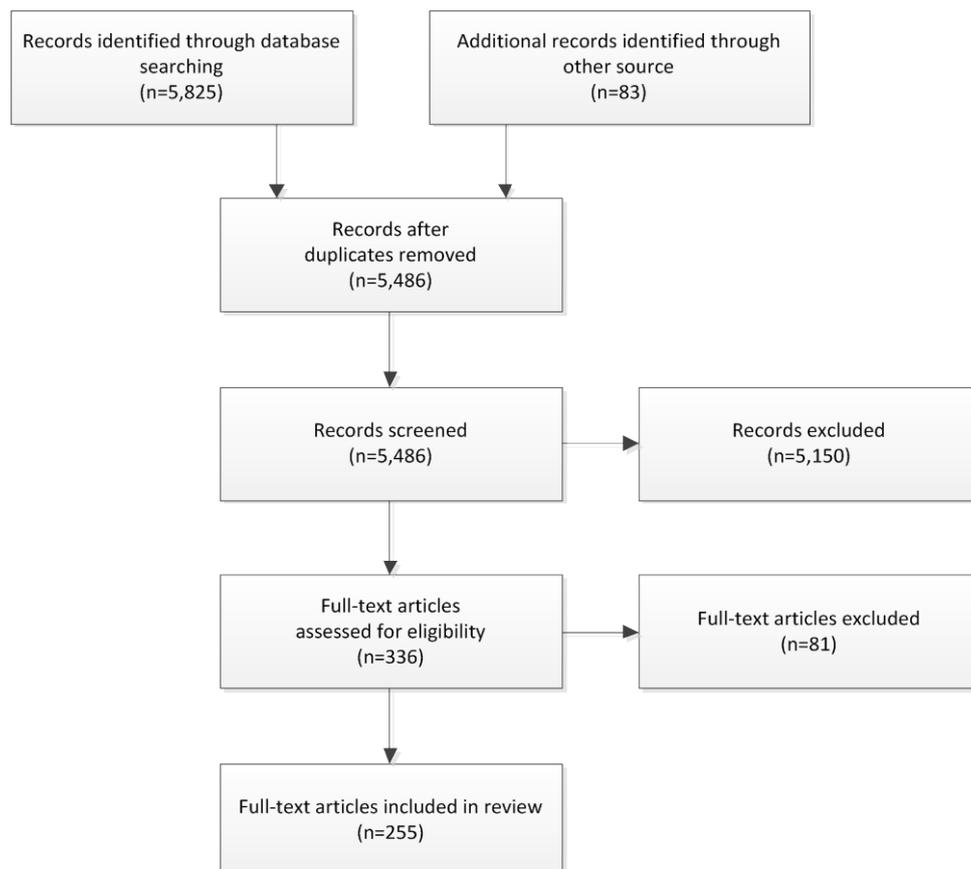


Table 1 profiles the 20 instruments, describing them in terms of when and where they were developed, the domains they assess, and their item structure. It should be noted that the LSP-16 is included among these instruments. We focused on the LSP-16 because this is the version of the instrument that is in current use in specialised public sector mental health services in Australia. Reference is made to its parent instrument, the LSP-39, as appropriate, however.

Table 1: Profile of clinician-rated instruments designed to assess functioning

INSTRUMENT	DATE ^a	COUNTRY	DESCRIPTION
Disability Rating Form (DRF)	1992	United States	The Disability Rating Form (DRF) was designed to rate disabilities associated with severe mental illness. It consists of five items that equate to five areas of disability: <i>activity of daily living; social functioning; concentration and task performance; adaptation to change; and impulse control.</i> ^{21,22}
FACE Core Assessment	1994	United Kingdom	The FACE Core Assessment was developed for use in adult mental health services. It went through several iterations during its development, but the final version contains 50 items that cover the following domains: <i>behaviour; cognitive; mental health; physical wellbeing; activities of daily living; interpersonal relationships; social circumstances; and response to care.</i> In addition, it includes a global rating of the impact of consumers' problems on their quality of life and/or functioning in the past month. ²³⁻²⁷
Global Assessment of Functioning (GAF)	1987	United States	The Global Assessment of Functioning (GAF) scale was introduced in the revised version of the third edition of the Diagnostic and Statistical Manual (DSM-III-R) as a means of assessing 'adaptive functioning'. ²⁸ The GAF is a 100-point, single item scale that assesses three different dimensions of functioning (<i>psychological; social; and occupational</i>). It can yield a single score (where only the most severe of the symptom and functioning values are recorded) or separate scores for symptoms (GAF-S) and functioning (GAF-F).
Health of the Nation Outcome Scales (HoNOS)	1998	United Kingdom	The Health of the Nation Outcome Scales (HoNOS) was designed as an instrument that could be used routinely by clinicians to measure outcomes for consumers with a mental illness. ⁹⁻¹² It can be regarded a general measure of mental health and social functioning in people with a mental illness. ¹² It has 12 items which roll up into four subscales (<i>behaviour, impairment, symptoms, and social</i>). The <i>social</i> subscale contains four items (Items 9-12), each of which relates to functioning. These cover: <i>relationships; activities of daily living; living conditions; and occupation and activities.</i>
Illness Management and Recovery Scale - Clinician Version (IMRS-C)	2004	United States	The Illness Management and Recovery Scales (IMRSs) were developed to assess outcomes for participants the widely-used IMR program (a program designed to teach people with schizophrenia illness self-management strategies). ^{29,30} There are clinician-rated and consumer-rated IMRSs, both of which use the same anchor points. The clinician-rated IMRS (IMRS-C) has 15 items which aggregate into three scales: <i>recovery (progress towards goals; knowledge; contact with people outside of family; relapse prevention planning; involvement with self-help activities); management (symptom distress; impairment of functioning; relapse of symptoms; psychiatric hospitalisations; coping); and biology (using medication effectively; alcohol use; drug use).</i> ³¹
Level of Functioning Scale (LFS)	1989	United States	The Level of Functioning Scales (LFS) was developed through a factor analysis of 73 items on the 79-item Missouri Level of Care (MLC) that measure functioning. The resultant factors on the LFS are: <i>community skills; self-care skills; nuisance behaviour; sociability; skilled nursing; proclivity for violence; and control of anger.</i> ³²
Life Functioning Assessment	2013	United States	The Life Functioning Assessment Inventory (L-FAI) was specifically designed to measure functioning in consumers with psychosis. It assesses four life domains:

INSTRUMENT	DATE ^a	COUNTRY	DESCRIPTION
Inventory (LFAI)			<i>work; social; relationships; leisure; and homemaking</i> . Each domain is given a status score (reflecting general performance) and a grade score (reflecting a more specific performance level within the grade). ³³
Life Skills Profile (LSP-16)	1989	Australia	The 16-item Life Skills Profile (LSP-16) was developed as a short-form version of the 39-item LSP-39. The original instrument was designed to measure constructs relevant to survival and adaptation in the community for individuals with schizophrenia and chronic mental illness, ^{6,7} and the modified version was created to minimise the rating burden on clinicians participating in the Australian Mental Health Classification and Service Costs (MH-CASC) Project. ³⁻⁵ The LSP-16 measures the following aspects of functioning: <i>withdrawal; self-care; compliance; and anti-social behaviour</i> .
Mini-ICF-APP	2009	Germany	The Mini-ICF-APP was created with reference to the International Classification of Functioning, Disability and Health (ICF), initially in German ³⁴ and then in several other languages, including English. ³⁵ It assesses 13 domains: <i>adherence to regulations; planning and structuring of tasks; flexibility; competency; endurance; assertiveness; contact with others; public exposure; intimacy; non-work activities; self-maintenance; mobility; and competence to judge and decide</i> .
Multi-Function Needs Assessment (MFNA)	1982	United States	The 134-item Multi-Function Needs Assessment (MFNA) was developed to assess the service needs and general functional performance of consumers in a single psychiatric hospital. ³⁶ A briefer, clearer instrument was later created which contains 118 items covering 13 areas of functioning: <i>physical self-maintenance; physical health; substance abuse; motor behaviour; psychiatric symptoms; attitude and motivation; attention and memory; verbal communication; family interaction; social interaction; independent living skills; public behaviour; and work/school/leisure</i> . ³⁷
Multnomah Community Ability Scale (MCAS)	1994	United States	The Multnomah Community Ability Scale (MCAS) is a 17-item instrument that assesses the level of functioning of consumers with chronic mental illness living in the community. The 17 items aggregate into four subscales: <i>interference with functioning; adjustment to living; social competence; and behaviour problems</i> . ³⁸⁻⁴⁰
Need of Support and Service Questionnaire (NSSQ)	2005	Sweden	The Need of Support and Service Questionnaire (NSSQ) was developed for a study examining whether mental health staff and social services staff were consistent in their judgements of consumers' needs. The instrument has 33 items, 23 of which relate to three domains of need: <i>need of support in activities of daily living; need of service provided by the public health and social service sectors; and need of assisted care living and need of work</i> (the remaining 10 items relate to socio-demographic information). ⁴¹
Personal and Social Performance Scale (PSP)	2000	Italy	The Personal and Social Performance Scale (PSP) was developed as part of a package for planning and evaluating psychiatric rehabilitation. It is a 100-point, single item rating scale that yields a rating based on four main areas: <i>socially useful activities, including work and study; personal and social relationships; self-care; and disturbing and aggressive behaviours</i> . ⁴²
Profile of Community Psychiatry Clients (PCPC)	1998	Australia	The Profile of Community Psychiatry Clients (PCPC) was developed as a tool to measure common problems and probable needs experienced by consumers in the community. It was designed to meet a range of purposes, including use in screening, service quality assurance, and research. It contains 35 items which measure the following domains: <i>coping limitations; behavioural problems; levels of social support; and organic problems</i> . ⁴³
Rehabilitation Evaluation Hall and Baker (REHAB)	1984	United Kingdom	Rehabilitation Evaluation Hall and Baker (REHAB) is a measure of socially appropriate or adaptive behaviour, designed for use with people with chronic mental illness. It contains 23 items, seven of which form a <i>deviant behaviour</i> subscale and 16 of which form a general behaviour subscale. It covers the following areas: <i>social activity; disturbed speech; communication skills; self-care skills; and community skills</i> . ^{44,45}
Residential Competency Scale (RCS)	1989	Canada	The Residential Competency Scale (RCS) was designed to assess the community living skills of consumers residing in community residential support services following deinstitutionalisation. It contains 82 items that measure skills in the following areas: <i>community skills; self-care; friendship; consideration; social competence; clarifying communication; money management; self-control; meal preparation; leisure; time routine; independence; and time planning</i> . ⁴⁶

INSTRUMENT	DATE ^a	COUNTRY	DESCRIPTION
Social Adjustment Behavior Rating Scale (SABRS)	1962	United States	The Social Adjustment Behavior Scale was designed to measure two aspects of consumers' social adjustment – work level and socialisation level. It contains 61 items, 29 of which relate to <i>work</i> and 33 of which relate to <i>socialisation</i> , with one overlapping item. ⁴⁷
Social Functioning Index (SFI)	1983	United States	The Social Functioning Index (SFI) was designed to measure social skills in consumers receiving community-based care, post-hospitalisation. It contains 51 items that relate to: <i>energy; self-control; hygiene; communication; and awareness of the environment</i> . ⁴⁸
Social Occupational Functioning Assessment Scale (SOFAS)	1994	United States	The SOFAS is a one-item rating of consumer functioning scored 0-100, and is provided in DSM-IV as an Axis V measure. ⁴⁹ In contrast to the GAF, it is intended to assess social and occupational functioning independently of the severity of psychological symptoms.
Uniform Client Data Instrument (UCDI)	1982	United States	The Uniform Client Data Instrument was designed to enable standardised information to be collected for consumers of mental health services in the United States. ^{50,51} It covers a range of domains, two of which have been used as explicit measures of functioning: <i>community living skills</i> (10 items); and <i>social activities</i> (8 items). ⁵²

a. Refers to the date of published information on the original version of the instrument.

Hierarchical, criterion-based assessment of the instruments

Criterion 1: Is brief (<50 items) and simple to score

Figure 2 shows that 13 of the 20 instruments meet the first criterion. The exceptions are the FACE Core Assessment, the Level of Functioning Scale (LFS), the Life Functioning Assessment Inventory (LFAI), the Multi-Function Needs Assessment (MFNA), the Residential Competency Scale (RCS), the Social Adjustment Behavior Rating Scale (SABRS) and the Social Functioning Index (SFI). The LFAI is complex to score because it the domains it assesses are given status scores (reflecting general performance) and grade scores (reflecting more specific performance levels within the grade). The remaining exceptions range in length from 50 items (the FACE Core Assessment) to 134 items (the MFNA), making them too unwieldy for use in routine outcome measurement. These instruments are excluded from further analysis.

Figure 2: Summary of instruments meeting criteria at each level of the hierarchy

All instruments	Disability Rating Form (DRF) FACE Core Assessment Global Assessment of Functioning (GAF) Health of the Nation Outcomes Scales (HoNOS) Illness Management and Recovery Scale - Clinician Version (IMRS-C) Level of Functioning Scale (LFS) Life Functioning Assessment Inventory (LFAI) Life Skills Profile 16 (LSP-16) Mini-ICF-APP Multi-Function Needs Assessment (MFNA)	Multnomah Community Ability Scale (MCAS) Need of Support and Service Questionnaire (NSSQ) Personal and Social Performance Scale (PSP) Profile of Community Psychiatry Clients (PCPC) Rehabilitation Evaluation Hall and Baker (REHAB) Residential Competency Scale (RCS) Social Adjustment Behavior Rating Scale (SABRS) Social and Occupational Functioning Assessment Scale (SOFAS) Social Functioning Index (SFI) Uniform Client Data Instrument (UCDI)
1. Is brief (<50 items) and simple to score	Disability Rating Form (DRF) Global Assessment of Functioning (GAF) Health of the Nation Outcomes Scales (HoNOS) Illness Management and Recovery Scale - Clinician Version (IMRS-C) Life Skills Profile 16 (LSP-16) Mini-ICF-APP	Multnomah Community Ability Scale (MCAS) Need of Support and Service Questionnaire (NSSQ) Personal and Social Performance Scale (PSP) Profile of Community Psychiatry Clients (PCPC) Rehabilitation Evaluation Hall and Baker (REHAB) Social and Occupational Functioning Assessment Scale (SOFAS) Uniform Client Data Instrument (UCDI)
2. Is not made redundant by more recent instruments	Disability Rating Form (DRF) Health of the Nation Outcomes Scales (HoNOS) Illness Management and Recovery Scale - Clinician Version (IMRS-C) Life Skills Profile 16 (LSP-16) Mini-ICF-APP	Multnomah Community Ability Scale (MCAS) Need of Support and Service Questionnaire (NSSQ) Personal and Social Performance Scale (PSP) Profile of Community Psychiatry Clients (PCPC) Rehabilitation Evaluation Hall and Baker (REHAB) Uniform Client Data Instrument (UCDI)
3. Relevant version has been scientifically scrutinised	Health of the Nation Outcomes Scales (HoNOS) Illness Management and Recovery Scale - Clinician Version (IMRS-C) Life Skills Profile 16 (LSP-16)	Multnomah Community Ability Scale (MCAS) Personal and Social Performance Scale (PSP) Rehabilitation Evaluation Hall and Baker (REHAB)
4. Considers functioning in a contemporary way	Health of the Nation Outcomes Scales (HoNOS) Illness Management and Recovery Scale - Clinician Version (IMRS-C) Life Skills Profile 16 (LSP-16)	Multnomah Community Ability Scale (MCAS) Personal and Social Performance Scale (PSP)
5. Demonstrates sound psychometric properties	Health of the Nation Outcomes Scales (HoNOS) Illness Management and Recovery Scale - Clinician Version (IMRS-C) Life Skills Profile 16 (LSP-16)	Multnomah Community Ability Scale (MCAS) Personal and Social Performance Scale (PSP)

Criterion 2: Is not made redundant by more recent instruments

Figure 2 shows that the majority of the remaining 13 instruments remain in contention when this criterion is examined. The two exceptions are the Global Assessment of Functioning (GAF) and the Social and Occupational Functioning Assessment Scale (SOFAS). The GAF was introduced in the revised version of the third edition of the Diagnostic and Statistical Manual (DSM-III-R) as a means of assessing ‘adaptive functioning’,²⁸ but was eliminated from subsequent versions of the DSM because it was regarded as being inadequate for assessing a construct like functioning that may be volatile and may not operate

independently of symptomatology, and because of the training required for it to be used appropriately.⁵³ The GAF was replaced by the SOFAS, on the grounds that the SOFAS assessed social and occupational functioning independently of symptom severity.⁵⁴ In turn, the SOFAS has been superseded by the Personal and Social Performance Scale (PSP), which demonstrates stronger psychometric performance.⁴² This sequence of instrument development and replacement led us to eliminate the GAF and the SOFAS from further consideration.

Criterion 3: Relevant version has been scientifically scrutinised

We considered whether each of the remaining 11 instruments had been subjected to scientific scrutiny. To satisfy this criterion, the given instrument had to have been assessed by investigators who were independent of the original instrument developers, and the results of that assessment had to have been published in the peer reviewed literature. Figure 2 indicates that six instruments satisfied this criterion. Those which have not been subjected to scientific scrutiny are the Disability Rating Form (DRF), the Mini-ICF-APP, the Need of Support and Service Questionnaire (NSSQ), the Profile of Community Psychiatry Clients (PCPC) and the Uniform Client Data Instrument (UCDI). These were excluded from further examination.

Criterion 4: Considers functioning in a contemporary way

We evaluated whether the remaining six instruments consider functioning in a contemporary way. Figure 2 shows that we removed the Rehabilitation Evaluation Hall and Baker (REHAB) at this point. This instrument was developed in 1984, in the era of deinstitutionalisation, and was designed for use with residents of long term psychiatric facilities who were being relocated to community residential support settings. It takes a limited view of functioning, and not one that recognises the capacity of people with mental illness to lead contributing lives. It primarily deals with activities of daily living, and only includes

relatively few items on other aspects of functioning. Most of these are framed negatively, falling into the 'deviant behaviours' subscale of the instrument. It should be noted that we retained the Life Skills Profile (LSP-16) at this juncture, despite criticisms that it uses outdated language. The decision was a borderline one; a number of the items on the LSP are consistent with a contemporary view of functioning, but there are certainly some that are outmoded.

Criterion 5: Demonstrates sound psychometric properties

Table 2 summarises the psychometric properties of the five remaining instruments. Specifically, it considers:

- Validity, or the extent to which the instrument measures what it intends to measure. Three types of validity are examined: construct validity; concurrent validity; and predictive validity.
- Reliability, or the extent to which the instrument gives stable, consistent results. Three aspects of reliability are considered: internal consistency; inter-rater reliability; and test-retest reliability;
- Sensitivity to change, or the extent to which, assuming the instrument is valid and reliable, it demonstrates the capacity to pick up change over time.

All five of the remaining instruments have been subject to independent psychometric testing by investigators other than the original developers. Figure 2 shows that all five have relatively sound psychometric properties, although some caveats are worth noting here. For example, the Health of the Nation Outcome Scales (HoNOS) has been extensively examined in its entirety, but less attention has been paid to the *social* subscale which contains the four items (Items 9-12) that relate to functioning that are relevant here. When this subscale and its component items have been assessed, they have sometimes performed less well than other elements of the instrument (particularly Items 11 and 12, which relate to *living conditions* and *occupation and activities*, where functioning is not independent of opportunities). Similarly, the Life Skills Profile (LSP-16) has been fairly extensively tested, but some of the

information on its psychometric properties comes from assessments of its parent instrument, the LSP-39. In addition, the Illness Management and Recovery Scale – Clinician Version (IMRS-C) has undergone less psychometric testing than the other four instruments; further information on its reliability and sensitivity to change would be desirable.

Table 2: Psychometric properties of instruments meeting Criteria 1-5

INSTRUMENT	PSYCHOMETRIC PROPERTIES	
Health of the Nation Outcome Scales (HoNOS) ^a	Construct validity	Various studies have examined the subscale structure of the HoNOS, typically using confirmatory factor analysis or principal components analysis. ⁵⁵ These studies have produced varying results, with some finding the four factor model defined by the original subscales had a good fit, ⁵⁶ and others suggesting that an alternative structure might be more appropriate. ⁵⁷⁻⁶⁴ Generally speaking, support has been found for the structure of the <i>social</i> subscale which comprises four items relating to functioning: Item 9 (<i>problems with relationships</i>), Item 10 (<i>problems with activities of daily living</i>), Item 11 (<i>problems with living conditions</i>), and Item 12 (<i>problems with occupation</i>). However, analyses that have considered the relative contribution of the different HoNOS items to its total score Items 11 and 12 make a lesser contribution than other items. ^{58,59,61}
	Concurrent validity	Numerous studies have considered the concurrent validity of the HoNOS. ⁵⁵ Some of these have shown that it correlates well with other clinician-rated instruments that are designed to measure functioning, including the Camberwell Assessment of Need Short Appraisal Schedule (CANSAS), ⁶⁵ the Role Functioning Scale (RFS), ¹² the Global Assessment of Functioning (GAF), ^{61,66-72} the Life Skills Profile (LSP), ⁶⁹ and the Disability Assessment Schedule (DAS), ⁶⁶ although some studies have found low correlations between the functioning-related items on the HoNOS and other relevant measures of functioning (e.g., one study found no correlation between Item 9 (<i>problems with relationships</i>) and the social communication scale on the Beeinträchtigungs-Schwere-Score (BSS), and no correlation between Item 10 (<i>problems with daily activities</i>) and the GAF ⁵⁸). The HoNOS has been found to discriminate between groups of consumers on the basis of where they receive treatment, which may in turn be a reflection of their levels of functioning. For example, it has been shown to distinguish between those receiving standard case management versus those assertive case management, ⁷³ those in residential/nursing home, day patient, outpatient and inpatient settings, ^{59,68,70,74,75} those in acute and sub-acute settings, ⁷¹ and those in long-stay settings with low, medium and high expectations of discharge. ⁷⁶
	Predictive validity	Several studies have examined the predictive validity of the HoNOS. Most have found it to have reasonably good predictive validity, explaining a significant proportion of the variance in resource use (e.g., as measured by service contacts, length of stay and costs) and treatment outcome (e.g., as measured by readmission rates, retention in the community, treatment response and death). ^{69,74,77-79} , and/or demonstrating its ability to act as a flag for discharge or transfer decisions. ⁸⁰ There have been exceptions to this rule, however, with some studies finding no correspondence or only a small association between the HoNOS and treatment outcome ⁷² or resource use. ⁸¹ In both cases, the findings have applied not only to the total score but also to scores on the subscale relating to functioning.
	Internal consistency	In studies examining the internal consistency of the HoNOS, Cronbach's alpha has ranged from 0.59 to 0.76, indicating that the instrument has a moderately high level of internal consistency and low levels of item redundancy, although it should be noted that the <i>social</i> subscale does not always perform as strongly as the <i>behaviour</i> and <i>symptoms</i> subscales. ^{12,57-59,61,68,70,71,82-84}
	Inter-rater reliability^b	Most of the studies that have assessed the inter-rater reliability of the HoNOS have found that the overall agreement between pairs of raters is fair to moderate, ^{70,85,86} or even moderate to good. ^{12,66,68,71,74,87} There are exceptions to this, ⁸⁸ however, and agreement tends to be poor on particular items, including three of the functioning items: Item 9 (<i>problems with relationships</i>), ⁶⁸ Item 11 (<i>problems with living conditions</i>) ^{68,89} and Item 12 (<i>problems with occupation</i>). ^{12,85,89}
	Test-retest	Several studies have examined the test-retest reliability of the HoNOS. These generally report

INSTRUMENT	PSYCHOMETRIC PROPERTIES	
	reliability^b	fair to moderate, or even good to very good, overall reliability score ^{68,70,72,86,87} although some items stand out as having lower reliability scores than others, including Item 10 (<i>problems with activities of daily living</i>).
	Sensitivity to change	The sensitivity to change of the HoNOS has been assessed in a number of studies. The simplest of these have examined change in HoNOS over time in given settings, hypothesising that there should be a decrease in severity as the consumer nears the end of an episode and generally finding support for this, ^{58,61,72,75,79,81,87,89-91} albeit with some variability on individual items, with some studies shedding doubt on the ability of Item 11 (<i>problems with living conditions</i>) and Item 12 (<i>problems with occupation</i>) to measure change. ^{58,61,79,87} Other studies have used clinician or consumer judgement as the 'gold standard' against which to judge the HoNOS's sensitivity to change. These studies have found correlations between HoNOS scores and both clinicians' assessments of improvement or deterioration ^{73,92} and consumers' self-reports of the extent to which their goals had been met. ⁹³ Still other studies have compared the HoNOS's dynamic properties and capacity to detect change against other, more established measures of outcome, including measures of functioning like the GAF. ⁶¹ predominantly for the behaviour and symptoms subscales. ⁹⁴ Finally, one study sought clinical expert opinion about the ability of the HoNOS to detect change. It found that clinicians believed that the HoNOS was sensitive to change, but that the <i>social and impairment</i> subscales were less responsive than the <i>behaviour and symptoms</i> subscales. ⁹⁵
Illness Management and Recovery Scale - Clinician Version (IMRS-C)	Construct validity	Several studies have used Rasch analysis or confirmatory factor analysis to test the construct validity of the IMRS-C. These studies provide support the factors of <i>recovery, management and biological vulnerability</i> , ⁹⁶ although one study suggests that the first two of these perform better than the third ⁹⁷ and another suggests some modifications, including replacement of the <i>biological vulnerability</i> factor with a <i>substance use</i> factor. ⁹⁸
	Concurrent validity	The IMRS-C has demonstrated good concurrent validity, showing strong negative correlations with clinicians' assessments of symptomatology, ⁹⁹ and strong positive correlations with clinicians' contemporaneous assessments of progress towards employment, education and housing goals ⁹⁸ and other clinician-rated instruments, including the Multnomah Community Ability Scale (MCAS), ¹⁰⁰ the Psychosis Evaluation Tool for Common Use by Caregivers (PECC), ⁹⁹ and the Substance Abuse Treatment Scale-Revised (SATS-R) ^{98,100} . The IMR-C also correlates well with the consumer-rated version of the instrument, ^{96,99,100} and with other consumer-rated instruments like the Coping Efficacy Scale (CES), ⁹⁶ the Multidimensional Scale for Perceived Social Support (MSPSS), ⁹⁶ the Manchester Short Assessment of Quality of Life (MANSA), ⁹⁹ the Recovery Assessment Scale (RAS), ⁹⁹ and the Modified Colorado Symptom Index (MCSI). ⁹⁹
	Predictive validity	The IMRS-C has demonstrated good predictive validity, showing strong positive correlations with clinicians' subsequent assessments of progress towards employment, education and housing goals. ⁹⁸
	Internal consistency	The IMRS-C has generally demonstrated good internal consistency, with Cronbach's alpha ranging from 0.70 to 0.82 in various studies. ^{96,98-100}
	Inter-rater reliability^b	No information available.
	Test-retest reliability^b	The IMRS-C has demonstrated strong test-retest reliability over a period of two weeks (r=0.81-0.88). ^{99,100}
	Sensitivity to change	There is limited information about the ability of the IMRS-C to measure change, but it has been used in evaluations of the IMR program, ^{101,102} and demonstrated outcomes in the expected direction.
Life Skills Profile (LSP-16)	Construct validity	The original five subscales of the LSP-39, as identified by the instrument's developers, were <i>communication, social contact, non-turbulence, self-care and responsibility</i> . These have been tested in subsequent studies using methods like principal components analysis and confirmatory factor analysis, with the result that alternative subscale structures have been proposed. One study, for example, led to the suggestion that the subscales should be <i>bizarre, withdrawal, self-care, compliance and anti-social behaviour</i> , ¹⁰³ and another resulted in the recommendation that the subscales could be further divided into the two dimensions of <i>general impairment and difficulty</i> . ¹⁰⁴ The LSP-16 has only four subscales – <i>self-care, antisocial, withdrawal and compliance</i> . These have also been subjected to testing, with one study using multilevel confirmatory factor analysis to demonstrate that the four-factor model was imperfect, leading to a proposal for a 15-item version of the LSP that fitted the data better. ¹⁰⁵

INSTRUMENT	PSYCHOMETRIC PROPERTIES	
	Concurrent validity	The LSP-39 has been shown to perform well against the Health of the Nation Outcome Scales (HoNOS), ^{69,83,106-108} the Katz Adjustment Scale (KAS), ⁶ the Multnomah Community Ability Scale (MCAS), ¹⁰⁹ the Strauss-Carpenter Levels of Functioning Scale (LOF), ¹⁰⁹ the Global Assessment of Functioning Scale (GAF), ^{69,109,110} the Role Functioning Scale (RFS), ⁸³ the Quality of Life Scale (QOL), ¹¹¹ the Interviewer-rated Quality of Life Scale (IQL), ¹⁰⁷ the Social Behaviour Schedule (SBS), ¹¹⁰ the Resource Associated Functional Level Scale (RAFLS) ¹⁰³ and the Global Assessment Scale (GAS). ¹⁰⁷ However, it has demonstrated poor or mixed performance against the BASIS-32 [®] , ⁸³ the Mental Health Inventory (MHI), ⁸³ the Short Form-36 (SF-36), ⁸³ the General Wellbeing Scale (GWB), ^{111,112} the Brief Psychiatric Rating Scales (BPRS), ^{7,103,107} the Dysexecutive Questionnaire (DEX), ^{107,110} Cantril's Ladder ¹⁰⁷ and the Affect Balance Scale (ABS). ¹⁰⁷ The LSP-16 has been shown to correlate well with the LSP-39, ^{3,113} and with the Health of the Nation Outcome Scales (HoNOS) ¹¹⁴ but has demonstrated poor or mixed performance against the Behaviour and Symptom Identification Scale (BASIS-32 [®]). ¹¹⁴ Both versions of the instrument have been shown to discriminate between consumers on the basis of the stability and independence of their living situations, ^{104,112,115-118} levels of social functioning (as measured by factors like unstable employment, welfare dependency, police contact and complaints by neighbours), ^{7,104,119} legal status, ¹⁰⁸ and diagnosis. ^{108,120}
	Predictive validity	Several studies have examined the predictive validity of the LSP-16 and LSP-39. In general, these studies have shown that the different versions of the instrument can predict outcomes relating to community tenure, ¹²¹ hospital readmission, ^{104,119} change in locus of care, ¹¹⁸ length of inpatient stay, ^{122,123} and overall costs, ^{112,123} although one study reported discrepant findings. ⁶⁹
	Internal consistency	A number of studies have examined the internal consistency of the LSP-39, from which the LSP-16 is derived. In these studies, the internal consistency has been reported as moderately high with subscale and total score Cronbach's alphas ranging from 0.64 to 0.88 ^{6,7,83,103} and 0.93 to 0.94, ^{83,103,109} respectively.
	Inter-rater reliability^b	A number of studies have assessed the inter-rater reliability of both the LSP-39 and the LSP-16. These studies found the overall agreement between pairs of raters on the LSP-39 to be fair to moderate ^{83,103,124} or moderate to good. ^{6,7,103}
	Test-retest reliability^b	The few studies that have examined the test-retest reliability of the LSP have reported a high overall reliability score, albeit for the LSP-39 only. ^{6,83,124} By way of example, one study established high test-retest reliability for case workers, residential carers and parents, each of whom were asked to rate the same person with relatively stable chronic schizophrenia at two points in time (one month apart). ⁶
	Sensitivity to change	A number of studies have compared the LSP's ability to detect change against more established measures of outcome, finding, for example, significant associations between changes on the LSP and changes on the Global Change Ratings Scale (GCRS), ⁸³ the Modified Clinical Global Impressions Scale (CGI), ⁸³ the Role Functioning Scale (RFS), ⁸³ the HoNOS, ^{69,83} and the Global Assessment of Functioning (GAF). ⁶⁹ Other studies have examined changes in LSP scores for different consumer groups that would be expected to show greater or lesser degrees of improvement depending on their treatment circumstances. These studies have typically found that the LSP demonstrates greater levels of improvement in those who participate in intensive case management than in those who undergo routine case management, ^{8,125-129} but there have been some exceptions. ^{130,131} Still other studies have used self-reported improvement or deterioration as the 'gold standard' against which to assess the sensitivity to change of the LSP. One such study involved an analysis of LSP-39 scores over time for groups showing differing levels and directions of self-reported change, and found that LSP scores worsened in the group who reported a decline in their levels of functioning, but there was no association between LSP change score and self-reported change for any other group. ⁸³
Multnomah Community Ability Scale (MCAS)	Construct validity	Tests of the factor structure of the MCAS have suggested that the four factor model proposed by the instrument's developers (<i>interference with functioning; adjustment to living; social competence; and behaviour problems</i>) ³⁸⁻⁴⁰ may not be the best solution, and alternative structures have been proposed. ^{54,132,133}
	Concurrent validity	The MCAS has been shown to perform well against clinicians' global assessments ³⁹ and consumers' levels of resource use, ^{39,40} and to discriminate between consumers in meaningful ways (e.g., on the basis of age-related factors, severity of symptoms and cognitive functioning). ^{39,40,134} It has also been shown to correlate with various measures of related constructs, including the Client Satisfaction Questionnaire (CSQ), ⁵⁴ the physical health scale of

INSTRUMENT	PSYCHOMETRIC PROPERTIES	
		the SF-36, ⁵⁴ the Lehman Quality of Life Scale (LQLS), ⁵⁴ the SOFAS, ⁵⁴ the Brief Psychiatric Rating Scale (BPRS), ¹³⁵ and the PANSS, ¹³⁶ but not the Recovery Assessment Scale (RAS). ¹³⁷
	Predictive validity	The MCAS has demonstrated good predictive validity, with poorer scores being associated subsequent hospitalisations. ^{39,138}
	Internal consistency	The internal consistency of the MCAS was found to be good by the instrument's developers, with Cronbach's alphas up to 0.90, ³⁹ but lower levels of internal consistency have been found by others using both the original structure and revised structures. ^{54,132,133}
	Inter-rater reliability^b	The MCAS has shown good inter-rater reliability, with an intra-class correlation coefficient of 0.85 found during the development of the instrument. ³⁹ Subsequent studies have reported a broader range (0.62-0.99). ^{136,139,140}
	Test-retest reliability^b	The MCAS performs well in relation to test-retest reliability, demonstrating an intra-class correlation coefficient of 0.83. ³⁹
	Sensitivity to change	The MCAS's sensitivity to change has not been formally tested, although it has been used to measure outcomes in several studies and improvements in the expected direction have been detected. ^{141,142} In at least one of these studies change on the MCAS was associated with commensurate levels of change on other relevant instruments, including the Brief Psychiatric Rating Scales (BPRS). ¹⁴³
Personal and Social Performance Scale (PSP)	Construct validity	The factor structure proposed by the original developers of the PSP (<i>socially useful activities, including work and study; personal and social relationships; self-care; and disturbing and aggressive behaviours</i>) has been confirmed in subsequent independent analysis. ¹⁴⁴
	Concurrent validity	The PSP has demonstrated strong correlations with various clinician-rated and consumer-rated instruments assessing functioning and related constructs, including the two instruments it has superseded (the GAF ¹⁴⁵⁻¹⁵⁰ the SOFAS ^{147,149,151}), the Mini-ICF-APP, ^{147,149} the Strauss-Carpenter Level of Functioning (SCLF) ¹⁵² , Activities of Daily Living Rating Scale II (ADLRS-II) ¹⁵³ and the Quality of Life Scale. ^{144,150} It has been found to show marginal correlations with a consumer-rated version of the PSP, although this was mediated by consumers' levels of insight. ¹⁵⁴ The PSP has been shown to discriminate between consumers on the basis of their treatment setting (e.g., inpatient versus community), ^{145,149,155} their diagnosis, ¹⁵¹ and their neurocognitive capacity and levels of symptomatology, as measured by instruments like the Wechsler Memory Scale-Revised (WMS-R), Continuous Performance Test (CPT), Wisconsin Card Sorting Test (WCST), Negative Syndrome Scale (PANSS) and Clinical Global Impressions-Severity (CGI-S). ^{144-153,155,156}
	Predictive validity	The PSP has been shown to predict relapse in consumers with schizophrenia. ¹⁵⁷
	Internal consistency	The internal consistency of the PSP has been found to be moderate to high, with Cronbach's alpha ranging from 0.64 to 0.87 across studies. ^{144-149,151,157}
	Inter-rater reliability^b	The inter-rater reliability of the PSP has varied across studies, with intra-class coefficients ranging from 0.43 to 1.0. ^{42,146-149,155,158}
	Test-retest reliability^b	Studies of the test-retest reliability of the PSP have reported intra-class coefficients of between 0.61 and 0.98. ^{147,148,151,152,155}
	Sensitivity to change	Various studies have demonstrated the PSP's capacity to detect change alongside more established outcome instruments like the PANSS and the CGI-S. ^{148,150-152,155,156} Several authors have explored thresholds for clinically significant change on the PSP, variously proposing that increases in total scores of 7-9 may indicate clinically significant improvement, ^{152,155} and that decreases of 10 may indicate clinically significant decline. ¹⁵⁷

- a. More detail on the psychometric properties of the HoNOS can be found elsewhere.⁵⁵ The information presented in this table relates primarily to the HoNOS items that are concerned with functioning (Items 9-12).
- b. The level of reliability of an instrument is traditionally measured by a kappa value. Kappas of ≤ 0.20 are regarded as poor, 0.21-0.40 as fair, 0.41-0.60 as moderate, 0.61-0.80 as good, and ≥ 0.81 as very good.

Discussion

We used a hierarchical, criterion-based approach to identify candidate instruments for measuring functioning among adult consumers of specialised public sector mental health services. By the end of the elimination process, we had reduced 20 potential instruments to five: the Health of the Nation Outcome Scales (HoNOS); the Illness Management and Recovery Scale – Clinician Version (IMRS-C); the Life Skills Profile (LSP-16); the Multnomah Community Ability Scale (MCAS); and the Personal and Social Performance Scale (PSP). The HoNOS, the MCAS and the PSP were all shortlisted in the two previous reviews that we drew upon, and the LSP-16 was shortlisted in the Australian review but not the New Zealand one.¹⁷⁻¹⁹ The IMRS-C was not identified in either of these reviews, so it did not feature in their shortlists.¹⁷⁻¹⁹

All five of the above instruments are recommended for consideration as clinician-rated instruments that might be used to routinely measure adult consumers' functioning in Australian mental health services. The consideration process should be systematic and structured. To a large extent, it should relate to the appropriateness of the candidate instruments for assessing functioning in the current context, their acceptability to clinicians and consumers, and the feasibility of using them in routine practice. The process should involve seeking stakeholders' opinions about, for example, the specific domains of functioning covered by each instrument, and the language used in individual items. Ideally, the process should also involve some real-world testing of clinicians' completion of the instruments in busy inpatient and community mental health settings. AMHOCN has conducted stakeholder consultations and run stakeholder surveys of this kind in the past. AMHOCN has also provided data on completion rates for the two instruments that are already part of the NOCC suite, with the most recent (2014-2015) data showing that the HoNOS was completed at 83% of review/discharge collection occasions, and the LSO-16 was

competed at 71%. AMHOCN could also assist with field testing the other three instruments, should this be viewed as desirable.

A key consideration in terms of appropriateness relates to the capacity of each of the instruments to be used to measure outcomes. Our review examined a range of psychometric properties, including sensitivity to change, and all of the instruments performed reasonably well on these. Other factors need to be taken into account too, however. For instance, the time period covered by the instruments needs to be short enough that the no two rating periods overlap. The PSP asks about the consumer's general functioning, without specifying time period, so this is not an issue for this instrument. The HoNOS covers the previous two weeks, which is sufficiently short that the issue of potentially overlapping assessment periods is minimised in most cases. The LSP-16 covers the last three months, as does the IMRS-C. The MCAS has rating periods of three months and one year, depending on the specific item. Consideration might be given to exploring whether these instruments can be modified to cover shorter time periods. Precedents exist for these sorts of modifications; an alternative version of the MCAS exists which has a rating period of one month.¹⁴⁰

The five shortlisted instruments each have items that address 'activities' and 'participation', identified as core elements of functioning in the ICF. In part this is because our initial exclusion criteria meant that instruments that only measured activities (or, more specifically, activities of daily living) were discarded before they reached the point of review. The various instruments placed differing emphasis on these two elements, however, and include divergent domains within each of them. The HoNOS assesses *relationships, activities of daily living, living conditions and occupation and activities*. The IMRS-C covers *recovery, management and biology*. The LSP-16 focuses on *withdrawal, self-care, compliance and anti-social behaviour*. The MCAS considers *interference with functioning, adjustment to living, social competence and behavioural problems*. The PSP provides a rating that is based on *socially useful activities, personal and social relationships, self-care and disturbing and aggressive behaviours*. When

the appropriateness, acceptability and feasibility of the five instruments is explored, consideration should be given to the precise domains that stakeholders believe should be assessed.

It is worth noting that a vast array of functioning instruments did not feature in our 'long list', and consequently had no opportunity to make our 'short list'. This was because we were specifically looking for clinician-rated measures of functioning for adult consumers that could be used in as part of the new instrument that was recommended by NMHIDEAP.¹⁴ As noted above, this meant that we excluded consumer-rated instruments and clinician-rated instruments that sought information via consumer interviews or case note reviews, doing so at the stage of screening the abstracts and full-text of identified journal articles. More than 80 additional instruments designed to measure functioning were eliminated at this pre-review stage. Some of these instruments undoubtedly have merit. For example, the Camberwell Assessment of Need Short Appraisal Schedule (CANSAS)¹⁵⁹ is popular and has sound psychometric properties, but was excluded because it involves a structured interview in which clinician, consumer and carer views of need can be recorded separately. If the examination of appropriateness, acceptability and feasibility of the five instruments does not yield positive findings, then consideration might be given to broadening the search criteria and identifying additional instruments (albeit ones that might need to be modified to be fit for purpose).

Decisions about whether or not to use one of the five identified instruments – or to seek alternatives – should not be made in isolation. The clinician-rated instruments in the current NOCC suite are complemented by various consumer-rated instruments. At present, these primarily relate to levels of distress and other psychological symptoms, but there is an appetite for broadening these to include constructs like social inclusion and recovery. AMHOCN has reviewed existing social inclusion and recovery instruments,^{20,160} and has developed and trialled a new social inclusion instrument (the Living in the Community Questionnaire).¹⁶¹ There is an argument that these constructs are closely related to functioning, particularly the 'participation' element of functioning. There is also an argument that whereas a consumer's level of functioning can be assessed by either a clinician or by the consumer

him/herself, social inclusion and recovery are more appropriately measured by the consumer, because of their experiential nature. Consideration should be given to how the selected clinician-rated measure of functioning complements proposed consumer-rated social inclusion and recovery instruments.

Identifying an appropriate clinician-rated functioning instrument should not stop with adult consumers. The current review excluded instruments that were designed for specific populations, including children and adolescents and older people. Norms around functioning are clearly age-related to some extent, so it makes sense that functioning instruments that have utility for adult consumers may not do so for younger and older consumers. For younger consumers, levels of maturity will impact upon functioning. For older consumers, physical and cognitive abilities may play a role. Age-specific functioning instruments are required for these groups, and we would recommend a similar process for identifying them.

We acknowledge that our review had some limitations. Despite our best efforts, we may have missed some relevant and potentially useful instruments designed to assess functioning (e.g., if our search terms did not pick up articles related to them, or if these articles were not indexed in the two academic databases we used). In addition, we may have missed some articles relating to the instruments we did identify, so our examination of the psychometric properties of the final five may not have been exhaustive. Finally, the articles we did retrieve did not always provide optimal detail on the instruments they described (particularly with respect to the specific items on these instruments), so it is possible that we misinterpreted information about some of them.

These limitations aside, we believe that the current review can help to inform decisions about which clinician-rated instruments hold promise for assessing whether functioning improves, deteriorates or remains the same for adult consumers of Australian mental health services. Further work is required to determine which, if any, of these instruments satisfy further criteria relating to appropriateness, acceptability and feasibility.

References

1. World Health Organization. International Classification of Functioning, Disability and Health (ICF). Geneva: World Health Organization, 2001.
2. Pirkis J, Callaly T. Mental health outcome measurement in Australia. In: Trauer T, ed. Outcome Measurement in Mental Health. Cambridge: Cambridge University Press; 2010.
3. Buckingham W, Burgess P, Solomon S, Pirkis J, Eagar K. Developing a Casemix Classification for Mental Health Services. Volume 2: Resource Materials. Canberra: Commonwealth Department of Health and Family Services, 1998.
4. Buckingham W, Burgess P, Solomon S, Pirkis J, Eagar K. Developing a Casemix Classification for Mental Health Services. Volume 1: Main Report. Canberra: Commonwealth Department of Health and Family Services, 1998.
5. Burgess P, Pirkis J, Buckingham W, Eagar K, Solomon S. Developing a casemix classification for specialist mental health services. *Casemix* 1999; **1**(4): 4-20.
6. Parker G, Rosen A, Emdur N, Hadzi-Pavlov D. The Life Skills Profile: Psychometric properties of a measure assessing function and disability in schizophrenia. *Acta Psychiatrica Scandinavica* 1991; **83**(2): 145-52.
7. Rosen A, Hadzi-Pavlovic D, Parker G. The life skills profile: A measure assessing function and disability in schizophrenia. *Schizophrenia Bulletin* 1989; **15**(2): 325-37.
8. Craig T, Doherty I, Jamieson-Craig R, Boocock A, Attafua G. The consumer-employee as a member of a Mental Health Assertive Outreach Team: (1) Clinical and social outcomes. *Journal of Mental Health* 2004; **13**(1): 59-69.
9. Wing J, and Lelliott, P. Reliability and validity of HoNOS. *Psychiatric Bulletin* 1999; **23**(6): 375.

10. Wing J, and Lelliott, P. Reliability and validity of HoNOS [Correspondence]. *Psychiatric Bulletin* 1999; **23**(6): 375.
11. Wing J, Curtis RH, Beevor A. Health of the Nation Outcome Scales (HoNOS). Glossary for HoNOS score sheet. *British Journal of Psychiatry* 1999; **174**: 432-4.
12. Wing JK, Beevor AS, Curtis RH, Park SB, Hadden S, Burns A. Health of the Nation Outcome Scales (HoNOS). Research and development. *British Journal of Psychiatry* 1998; **172**: 11-8.
13. Department of Health. Mental Health National Outcomes and Casemix Collection: Technical Specification of State and Territory Reporting Requirements, Version 1.90. Canberra: Department of Health, 2015.
14. National Mental Health Information Development Expert Advisory Panel (NHMIDEAP). Mental Health National Outcomes and Casemix Collection: NOCC Strategic Directions 2014-2024. Canberra: Commonwealth of Australia, 2013.
15. National Mental Health Information Development Expert Advisory Panel (NHMIDEAP). National Outcomes and Casemix Collection (NOCC): Domain Framework. Canberra: Commonwealth of Australia, 2015.
16. Burgess P, Coombs T, Clarke A, Dickson R, Pirkis J. Achievements in mental health outcome measurement in Australia: Reflections on progress made by the Australian Mental Health Outcomes and Classification Network (AMHOCN). *International Journal of Mental Health Systems* 2012; **6**: 4.
17. Australian Mental Health Outcomes and Classification Network (AMHOCN) and Community Mental Health Australia (CMHA). National Community Managed Organisation (CMO) Outcome Measurement Project: Final Report to the Mental Health Information Strategy Standing Committee. Sydney: AMHOCN, 2013.
18. Lutchman R, Thompson A, Tait H, et al. In search of a standardised, comprehensive assessment of functioning. *New Zealand Journal of Occupational Therapy* 2007; **54**(1): 33-8.

19. Waikato Evaluation Team. Functional Outcome Measures for Children and Youth, Adults and Older Adult Consumers in New Zealand Mental Health Services. Hamilton, NZ: University of Waikato, 2005.
20. Burgess P, Pirkis J, Coombs T, Rosen A. Assessing the value of existing recovery measures for routine use in Australian mental health services. *Australian and New Zealand Journal of Psychiatry* 2011; **45**: 267-80.
21. Hoyle RH, Nietzel MT, Guthrie PR, Baker-Prewitt JL, Heine R. The Disability Rating Form: A brief schedule for rating disability associated with severe mental illness. *Psychosocial Rehabilitation Journal* 1992; **16**(1): 77-94.
22. Hoyle RH, Nietzel MT, Guthrie PR, Baker-Prewitt JL, Heine R. The Disability Rating Form. *Psychosocial Rehabilitation Journal* 1993; **16**(3): 153-60.
23. Clifford P. The FACE recording and measurement system: A scientific approach to person-based information. *Bulletin of the Menninger Clinic*; **63**: 305-31.
24. Clifford P. The FACE Project: Final Report to the Department of Health. London: University College London, 1994.
25. Clifford P. Structuring the Clinical Record: The FACE Assessment and Outcome System (Main Report to the Department of Health). London: University College London, 1997.
26. Clifford P. Structuring the Clinical Record: The FACE Assessment and Outcome System (Appendices). London: University College London, 1997.
27. Clifford PI, Orbach G, Hobbins J, Pavyer J. Measuring disability and outcomes in routine practice with the FACE Core Assessment. *Bulletin of the Menninger Clinic* 1999; **63**(3): 332-45.
28. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (3rd edition, revised). Washington, DC: American Psychiatric Association, 1987.
29. Gingerich S, Mueser K. Illness Management and Recovery Implementation Resource Kit. Rockville, MD: Center for Mental Health Services, Substance Abuse and Mental Health Services Administration, 2002.

30. Mueser K, Meyer P, Penn D, Clancy R, Clancy D, Salyers M. The Illness Management and Recovery program: Rationale, development, and preliminary findings. *Schizophrenia Bulletin* 2006; **32**(1): 32–43.
31. Mueser K, Gingerich S, Salyers M, McGuire A, Reyes R, Cunningham H. The Illness Management and Recovery Scale (IMR). Lebanon, NH: New Hampshire-Dartmouth Psychiatric Research Center, 2004.
32. Massey OT, Pokorny LJ, Kramer BH. The development of factor-based level of functioning scales from a level of care instrument. *Journal of Clinical Psychology* 1989; **45**(6): 903-9.
33. Hui CL-M, Li Y-K, Leung K-F, et al. Reliability and validity of the Life Functioning Assessment Inventory (L-FAI) for patients with psychosis. *Social Psychiatry and Psychiatric Epidemiology* 2013; **48**(10): 1687-95.
34. Linden M, Baron S, Muschalla B. Mini-ICF-Rating für Aktivitäts- und Partizipationsstörungen bei Psychischen Erkrankungen (Mini-ICF-APP). Hogrefe AG, Bern: Verlag Hans Huber, 2009.
35. Molodynski A, Linden M, Juckel G, et al. The reliability, validity, and applicability of an English language version of the Mini-ICF-APP. *Social Psychiatry and Psychiatric Epidemiology* 2013; **48**(8): 1347-54.
36. Angelini DE. Functional needs of the chronically mentally ill: Implications for service delivery. *Psychosocial Rehabilitation Journal* 1982; **5**(1): 29-33.
37. Weiner HR. Multi-Function Needs Assessment: The development of a functional assessment instrument. *Psychosocial Rehabilitation Journal* 1993; **16**(4): 51-61.
38. Barker S, Barron N, McFarland B, Bigelow DA. Multnomah Community Ability Scale: User's Manual. Portland, OR: Western Mental Health Research Center, Oregon Health Sciences University, 1994.
39. Barker S, Barron N, McFarland BH, Bigelow DA. A community ability scale for chronically mentally ill consumers: I. Reliability and validity. *Community Mental Health Journal* 1994; **30**(4): 363-83.

40. Barker S, Barron N, McFarland BH, Bigelow DA, Carnahan T. A community ability scale for chronically mentally ill consumers: Part II. Applications. *Community Mental Health Journal* 1994; **30**(5): 459-72.
41. Jansson L, Sonnander K, Wiesel F-A. Needs assessed by psychiatric health care and social services in a defined cohort of clients with mental disabilities. *European Archives of Psychiatry and Clinical Neuroscience* 2005; **255**(1): 57-64.
42. Morosini PL, Magliano L, Brambilla L, Ugolini S, Pioli R. Development, reliability and acceptability of a new version of the DSM-IV Social and Occupational Functioning Assessment Scale (SOFAS) to assess routine social functioning. *Acta Psychiatrica Scandinavica* 2000; **101**(4): 323-9.
43. Cheah Y, Parker G, Hadzi-Pavlovic D, Gladstone G, Eysers K. Development of a measure profiling problems and needs of psychiatric patients in the community. *Social Psychiatry and Psychiatric Epidemiology* 1998; **33**(7): 337-44.
44. Baker R, Hall J. User's Manual for Rehabilitation Evaluation. Aberdeen: Vine Publishing, 1984.
45. Baker R, Hall J. REHAB: A new assessment instrument for chronic psychiatric patients. *Schizophrenia Bulletin* 1988; **14**: 97-110.
46. Kazarian SS, Cole JD, Barnes SM, Sussman SI. The Residential Competency Scale: A new measure of residential functioning skills. *Journal of Community Psychology* 1989; **17**(4): 297-303.
47. Aumack L. A social adjustment behavior rating scale. *Journal of Clinical Psychology* 1962; **18**(4): 436-41.
48. Peterson L. Social functioning assessment of aftercare psychiatric patients in socialization therapy. *Psychological Reports* 1983; **53**(3, Pt 2): 1123-30.
49. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (4th edition). Washington, DC: American Psychiatric Association, 1994.

50. Mulkern V, Manderscheid R. Characteristics of community support program clients in 1980 and 1984. *Hospital and Community Psychiatry* 1989; **40**(2): 165-72.
51. Tessler R, Goldman H. *The Chronically Mentally Ill: Assessing Community Support Programs*. Cambridge, Massachusetts: Ballinger; 1982.
52. Widlak PA, McKee D, Greenberg JR, Greenley JR. An assessment of client function scales in the Uniform Client Data Instrument (UCDI). *Psychosocial Rehabilitation Journal* 1992; **15**(4): 19-35.
53. Suzuki T, Uchida H, Sakurai H, et al. Relationships between global assessment of functioning and other rating scales in clinical trials for schizophrenia. *Psychiatry Research* 2015; **227**(2-3): 265-9.
54. Hendryx M, Dyck DG, McBride D, Whitbeck J. A test of the reliability and validity of the Multnomah Community Ability Scale. *Community Mental Health Journal* 2001; **37**(2): 157-68.
55. Pirkis JE, Burgess PM, Kirk PK, Dodson S, Coombs TJ, Williamson MK. A review of the psychometric properties of the Health of the Nation Outcome Scales (HoNOS) family of measures. *Health and Quality of Life Outcomes* 2005; **3**: 76.
56. Preston NJ. The Health of the Nation Outcome Scales: Validating factorial structure and invariance across two health services. *Australian and New Zealand Journal of Psychiatry* 2000; **34**(3): 512-9; discussion 20-1.
57. Trauer T. The subscale structure of the Health of the Nation Outcome Scales (HoNOS). *Journal of Mental Health* 1999; **8**(5): 499-509.
58. Andreas S, Harfst T, Rabung S, et al. The validity of the German version of the Health of the Nation Outcome Scales (HoNOS-D): A clinician-rating for the differential assessment of the severity of mental disorders. *International Journal of Methods in Psychiatric Research* 2010; **19**(1): 50-62.
59. Lovaglio PG, Monzani E. Validation aspects of the Health of the Nation Outcome Scales. *International Journal of Mental Health Systems* 2011; **5**: 20.

60. Lovaglio PG, Monzani E. Health of the nation outcome scales evaluation in a community setting population. *Quality of Life Research* 2012; **21**(9): 1643-53.
61. McClelland R, Trimble P, Fox ML, Stevenson MR, Bell B. Validation of an outcome scale for use in adult psychiatric practice. *Quality in Health Care* 2000; **9**(2): 98-105.
62. Newnham E, Harwood K, Page A. The subscale structure of the Health of the Nation Outcome Scales. *Journal of Mental Health* 2009; **18**: 326-34.
63. Speak B, Muncer S. The structure and reliability of the Health of the Nation Outcome Scales. *Australasian Psychiatry* 2015; **23**(1): 66-8.
64. Speak BL, Muncer SJ. Factorial structure of the Health of the Nation Outcome Scales: An ordinal confirmatory factor analysis using a national sample of clinician ratings in England. *International Journal of Mental Health Nursing* 2016; **25**(1): 87-98.
65. Issakidis C, Teesson M. Measurement of need for care: A trial of the Camberwell Assessment of Need and the Health of the Nation Outcome Scales. *Australian and New Zealand Journal of Psychiatry* 1999; **33**(5): 754-9.
66. Amin S, Singh SP, Croudace T, Jones P, Medley I, Harrison G. Evaluating the Health of the Nation Outcome Scales. Reliability and validity in a three-year follow-up of first-onset psychosis. *British Journal of Psychiatry* 1999; **174**: 399-403.
67. Browne S, Doran M, McGauran S. Health of the Nation Outcome Scales (HoNOS): Use in an Irish psychiatric outpatient population. *Irish Journal of Psychological Medicine* 2000; **17**(1): 17-9.
68. Orrell M, Yard P, Handysides J, Schapira R. Validity and reliability of the Health of the Nation Outcome Scales in psychiatric patients in the community. *British Journal of Psychiatry* 1999; **174**: 409-12.
69. Parker G, O'Donnell M, Hadzi-Pavlovic D, Proberts M. Assessing outcome in community mental health patients: A comparative analysis of measures. *International Journal of Social Psychiatry* 2002; **48**(1): 11-9.

70. Shergill SS, Shankar KK, Seneviratna K, Orrell MW. The validity and reliability of the Health of the Nation Outcome Scales (HoNOS) in the elderly. *Journal of Mental Health* 1999; **8**(5): 511-21.
71. Phuaphanprasert B, Srisurapanont M, Silpakit C, et al. Reliability and validity of the Thai version of the Health of the Nation Outcome Scales (HoNOS). *Journal of the Medical Association of Thailand* 2007; **90**(11): 2487-93.
72. Preti A, Pisano A, Cascio MT, et al. Validation of the Health of the Nation Outcome Scales as a routine measure of outcome in early intervention programmes. *Early Intervention in Psychiatry* 2012; **6**(4): 423-31.
73. Gallagher J, Teesson M. Measuring disability, need and outcome in Australian community mental health services. *Australian and New Zealand Journal Psychiatry* 2000; **34**(5): 850-5.
74. Hope JD, Trauer T, Keks NA. Reliability, validity and utility of the Health of the Nation Outcomes Scale (HoNOS) in Australian adult psychiatric services. *Schizophrenia Research* 1998; **29**(1,2): 9-10.
75. Kisely S, Campbell LA, Crossman D, Gleich S, Campbell J. Are the Health of the Nation Outcome Scales a valid and practical instrument to measure outcomes in North America? A three-site evaluation across Nova Scotia. *Community Mental Health Journal* 2007; **43**(2): 91-107.
76. Allan S, McGonagle I. A comparison of HoNOS with the Social Behaviour Schedule in three settings. *Journal of Mental Health* 1997; **6**(2): 117-24.
77. Broadbent M. Reconciling the information needs of clinicians, managers and commissioners: A pilot project. *Psychiatric Bulletin* 2001; **25**(11): 423-5.
78. Schneider J, Wooff D, Carpenter J, Brandon T, McNiven F. Service organisation, service use and costs of community mental health care. *Journal of Mental Health Policy and Economics* 2002; **5**(2): 79-87.
79. Kisely S, Campbell LA, Cartwright J, Cox M, Campbell J. Do the Health of the Nation Outcome Scales measure outcome? *Canadian Journal of Psychiatry* 2010; **55**(7): 431-9.

80. Prowse L, Coombs T. The use of the Health of the Nation Outcome Scales (HoNOS) to inform discharge and transfer decisions in community mental health services. *Australian Health Review* 2009; **33**(1): 13-8.
81. Goldney RD, Fisher LJ, Walmsley SH. The Health of the Nation Outcome Scales in psychiatric hospitalisation: a multicentre study examining outcome and prediction of length of stay. *Australian and New Zealand Journal of Psychiatry* 1998; **32**(2): 199-205.
82. Page AC, Hooke GR, Rutherford EM. Measuring mental health outcomes in a private psychiatric clinic: Health of the Nation Outcome Scales and Medical Outcomes Short Form SF-36. *Australian and New Zealand Journal of Psychiatry* 2001; **35**(3): 377-81.
83. Stedman T, Yellowlees P, Mellsop G, Clarke R, Drake S. Measuring Consumer Outcomes In Mental Health: Field Testing of Selected Measures of Consumer Outcome in Mental Health. Canberra: Department of Health and Family Services, 1997.
84. Oiesvold T, Bakkejord T, Sexton JA. Concurrent validity of the Health of the Nation Outcome Scales compared with a patient-derived measure, the Symptom Checklist-90-Revised in out-patient clinics. *Psychiatry Research* 2011; **187**(1-2): 297-300.
85. Bebbington P, Brugha T, Hill T, Marsden L, Window S. Validation of the Health of the Nation Outcome Scales. *British Journal of Psychiatry* 1999; **174**: 389-94.
86. Brooks R. The reliability and validity of the Health of the Nation Outcome Scales: validation in relation to patient derived measures. *Australian and New Zealand Journal of Psychiatry* 2000; **34**(3): 504-11; discussion 20-1.
87. Andreas S, Harfst T, Dirmaier J, Kawski S, Koch U, Schulz H. A psychometric evaluation of the German version of the 'Health of the Nation Outcome Scales, HoNOS-D': On the feasibility and reliability of clinician-performed measurements of severity in patients with mental disorders. *Psychopathology* 2007; **40**(2): 116-25.

88. Webster J, Bretherton F, Goulter NS, Fawcett L. Does an educational intervention improve the usefulness of the Health of the Nation Outcome Scales in an acute mental health setting? *International Journal of Mental Health Nursing* 2013; **22**(4): 322-8.
89. Trauer T, Callaly T, Hantz P, Little J, Shields R, Smith J. Health of the Nation Outcome Scales. Results of the Victorian field trial. *British Journal of Psychiatry* 1999; **174**: 380-8.
90. Audin K, Margison FR, Clark JM, Barkham M. Value of HoNOS in assessing patient change in NHS psychotherapy and psychological treatment services. *British Journal of Psychiatry* 2001; **178**: 561-6.
91. Egger ST, Weniger G, Prinz S, Vetter S, Muller M. Health of the Nation Outcome Scales in a psychiatric inpatient setting: assessing clinical change. *Journal of Evaluation in Clinical Practice* 2015; **21**(2): 236-41.
92. Taylor JR, Wilkinson G. HoNOS v. GP opinion in a shifted out-patient clinic. *Psychiatric Bulletin* 1997; **21**(8): 483-5.
93. Hunter R, McLean J, Peck D, et al. The Scottish 700 Outcomes Study: A comparative evaluation of the Health of the Nation Outcome Scale (HoNOS), the Avon Mental Health Measure (AVON), and an Idiographic Scale (OPUS) in adult mental health. *Journal of Mental Health* 2004; **13**(1): 93-105.
94. Bech P, Bille J, Schutze T, Sondergaard S, Wiese M, Waarst S. Health of the Nation Outcome Scales (HoNOS): Implementability, subscale structure and responsiveness in the daily psychiatric hospital routine over the first 18 months. *Nordic Journal of Psychiatry* 2003; **57**(4): 285-90.
95. Burgess P, Trauer T, Coombs T, McKay R, Pirkis J. What does 'clinical significance' mean in the context of the Health of the Nation Outcome Scales? *Australasian Psychiatry* 2009; **17**(2): 141-8.
96. Hasson-Ohayon I, Roe D, Kravetz S. The psychometric properties of the Illness Management and Recovery scale: client and clinician versions. *Psychiatry Research* 2008; **160**(2): 228-35.

97. McGuire AB, Kean J, Bonfils K, Presnell J, Salyers MP. Rasch analysis of the illness management and recovery scale-clinician version. *Journal of Evaluation in Clinical Practice* 2014; **20**(4): 383-9.
98. Sklar M, Sarkin A, Gilmer T, Groessl E. The psychometric properties of the Illness Management and Recovery scale in a large American public mental health system. *Psychiatry Research* 2012; **199**(3): 220–7.
99. Färdig R, Lewander T, Fredriksson A, Melin L. Evaluation of the illness management and recovery scale in schizophrenia and schizoaffective disorder. *Schizophrenia Research* 2011; **132**: 157–64.
100. Salyers M, Godfrey J, Mueser K, Labriola S. Measuring illness management outcomes: A psychometric study of clinician and consumer rating scales for illness self management and recovery. *Community Mental Health Journal* 2007; **43**(5): 459–80.
101. Hasson-Ohayon I, Roe D, Kravetz S. A randomized controlled trial of the effectiveness of the illness management and recovery program. *Psychiatr Serv* 2007; **58**(11): 1461-6.
102. Salyers MP, Godfrey JL, McGuire AB, Gearhart T, Rollins AL, Boyle C. Implementing the illness management and recovery program for consumers with severe mental illness. *Psychiatr Serv* 2009; **60**(4): 483-90.
103. Trauer T, Duckmanton RA, Chiu E. The Life Skills Profile: a study of its psychometric properties. *Australian and New Zealand Journal of Psychiatry* 1995; **29**(3): 492-9.
104. Andrews G, Teesson M, Stewart G, Houtt J. Follow-up of community placement in the chronic mentally ill in New South Wales. *Hospital and Community Psychiatry* 1990; **41**: 184-8.
105. Little J. Multilevel confirmatory ordinal factor analysis of the Life Skills Profile-16. *Psychological Assessment* 2013; **25**(3): 810-25.
106. Trauer T, Eagar K. New Zealand Mental Health Consumers and their Outcomes. Auckland: Health Research Council of New Zealand, 2004.

107. Wooff D, Schneider J, Carpenter J, Brandon T. Correlates of stress in carers. *Journal of Mental Health* 2003; **12**(1): 29-40.
108. Eagar K, Trauer T, Mellsop G. Performance of routine outcome measures in adult mental health care. *Australian and New Zealand Journal of Psychiatry* 2005; **39**(8): 713-8.
109. Dickinson D, Coursey RD. Independence and overlap among neurocognitive correlates of community functioning in schizophrenia. *Schizophrenia Research* 2002; **56**(1-2): 161-70.
110. Simon AE, Giaocomini V, Ferrero F, Mohr S. Dysexecutive syndrome and social adjustment in schizophrenia. *Australian and New Zealand Journal of Psychiatry* 2003; **37**(3): 340-6.
111. Norman RMG, Malla AK, McLean T, et al. The relationship of symptoms and level of functioning in schizophrenia to general wellbeing and the Quality of Life Scale. *Acta Psychiatrica Scandinavica* 2000; **102**(4): 303-9.
112. Trauer T, Duckmanton RA, Chiu E. Estimation of costs of public psychiatric treatment. *Psychiatr Serv* 1998; **49**(4): 440-2.
113. Rosen A, Trauer T, Hadzi-Pavlovic D, Parker G. Development of a brief form of the Life Skills Profile: The LSP-20. *Australian and New Zealand Journal of Psychiatry* 2001; **35**(5): 677-83.
114. Trauer T. Analysis of Outcome Measurement Data from the Four Victorian 'Round One' Agencies: Mental Health Branch, Department of Human Services, Victoria, 2003.
115. Browne G, Courtney M. Measuring the impact of housing on people with schizophrenia. *Nurs Health Sci* 2004; **6**(1): 37-44.
116. Keller S, Hayes R. The relationship between the Allen Cognitive Level Test and the Life Skills Profile. *American Journal of Occupational Therapy* 1998; **52**(10): 851-6.
117. Kirkby K, Daniels B, Jones I, McInnes M. A survey of social outcome in schizophrenia in Tasmania. *Australian and New Zealand Journal of Psychiatry* 1997; **31**(3): 405-10.
118. Trauer T, Duckmanton RA, Chiu E. The assessment of clinically significant change using the Life Skills Profile. *Australian and New Zealand Journal of Psychiatry* 1997; **31**(2): 257-63.

119. Parker G, Hadzi-Pavlovic D. The capacity of a measure of disability (the LSP) to predict hospital readmission in those with schizophrenia. *Psychological Medicine* 1995; **25**(1): 157-63.
120. Parker G, Rosen A, Trauer T, Hadzi-Pavlovic D. Disability associated with mood states and comparator conditions: Application of the Life Skills Profile measure of disability. *Bipolar Disorders* 2007; **9**(1-2): 11-5.
121. Preston NJ. Predicting community survival in early psychosis and schizophrenia populations after receiving intensive case management. *Australian and New Zealand Journal of Psychiatry* 2000; **34**(1): 122-8.
122. Ballesteros RJ, Martinez SM, Martin CM, Ibarra GN, Bulbena VA. Assessment of the Life Skills Profile and the Brief Psychiatric Rating Scale as predictors of the length of psychiatric hospitalization. *Actas Espanolas de Psiquiatria* 2002; **30**: 225-32.
123. Kisely S, Preston N, Rooney M. Pathways and outcomes of psychiatric care: Does it depend on who you are, or what you've got? *Australian and New Zealand Journal of Psychiatry* 2000; **34**(6): 1009-14.
124. Andrews G, Peters L, Teesson M. Measurement of Consumer Outcome in Mental Health: A Report to the National Mental Health Information Strategy Committee. Sydney: Clinical Research Unit for Anxiety Disorders, 1994.
125. Hambridge J, Rosen A. Assertive community treatment for the seriously mentally ill in suburban Sydney: A programme description and evaluation. *Australian and New Zealand Journal of Psychiatry* 1994; **28**: 438-45.
126. Hamernik E, Pakenham KI. Assertive community treatment for persons with severe mental disorders: A controlled treatment outcome study. *Behaviour Change* 1999; **16**(4): 259-68.
127. Johnston S, Salkeld G, Sanderson K, Issakidis C, Teesson M, Buhrich N. Intensive case management: A cost-effectiveness analysis. *Australian and New Zealand Journal of Psychiatry* 1998; **32**(4): 551-9.

128. Rosen A, Teesson M. Does case management work? The evidence and the abuse of evidence-based medicine. *Australian and New Zealand Journal of Psychiatry* 2001; **35**(6): 731-46.
129. Sanderson K, Issakidis C, Johnston S, Teesson M, Salkeld G, Buhrich N. Cost-effectiveness of Intensive Case Management for People with Serious Mental Illness. Sydney: South Eastern Sydney Area Mental Health Service, 1996.
130. Ford R, Barnes A, Davies R, Chalmers C, Hardy P, Muijen M. Maintaining contact with people with severe mental illness: 5-year follow-up of assertive outreach. *Social Psychiatry and Psychiatric Epidemiology* 2001; **36**(9): 444-7.
131. Ford R, Ryan P, Beadsmoore A, Craig T, et al. Intensive case management for people with serious mental illness - Site 2: Clinical and social outcome. *Journal of Mental Health* 1997; **6**(2): 181-90.
132. Corbiere M, Crocker AG, Lesage AD, Latimer E, Ricard N, Mercier C. Factor structure of the Multnomah Community Ability Scale. *Journal of Nervous and Mental Disease* 2002; **190**(6): 399-406.
133. Bassani DG, Dewa CS, Krupa T, et al. Factor structure of the Multnomah Community Ability Scale: Longitudinal analysis. *Psychiatry Research* 2009; **167**(1-2): 178-89.
134. Prouteau A, Verdoux H, Briand C, et al. Cognitive predictors of psychosocial functioning outcome in schizophrenia: A follow-up study of subjects participating in a rehabilitation program. *Schizophrenia Research* 2005; **77**: 343-53.
135. Brown E, Rush A, Biggs M, Shores-Wilson K, Carmody T, Suppes T. Clinician ratings vs. global ratings of symptom severity: A comparison of symptom measures in the bipolar disorder module, phase II, Texas Medication Algorithm Project. *Psychiatry Research* 2003; **117**: 167-75.
136. Trauer T. Symptom severity and personal functioning among patients with schizophrenia discharged from long-term hospital care into the community. *Community Mental Health Journal* 2001; **37**(2): 145-55.

137. Lavin D, Ryan P. Using quantitative research to measure recovery outcomes and correlates. *Irish Journal of Psychological Medicine* 2012; **29**(3): 157-62.
138. Zani B, McFarland B, Wachal M, Barker S, Barron N. Statewide replication of predictive validation for the Multnomah Community Ability Scale. *Community Mental Health Journal* 1999; **35**(3): 223-9.
139. Dickerson FB. Assessing clinical outcomes: the community functioning of persons with serious mental illness. *Psychiatr Serv* 1997; **48**(7): 897-902.
140. Dickerson FB, Origoni AE, Pater A, Friedman BK, Kordonski WM. An expanded version of the Multnomah Community Ability Scale: anchors and interview probes for the assessment of adults with serious mental illness. *Community Mental Health Journal* 2003; **39**(2): 131-7.
141. Hopkins M, Ramsundar N. Which factors predict case management services and how do these services relate to client outcomes? *Psychiatric Rehabilitation Journal* 2006; **29**(3): 219-22.
142. McDevitt J, Wilbur J, Kogan J, Briller J. A walking program for outpatients in psychiatric rehabilitation: pilot study. *Biological Research for Nursing* 2005; **7**(2): 87-97.
143. Adair CE, McDougall GM, Mitton CR, et al. Continuity of care and health outcomes among persons with severe mental illness. *Psychiatr Serv* 2005; **56**(9): 1061-9.
144. Kawata AK, Revicki DA. Psychometric properties of the Personal and Social Performance scale (PSP) among individuals with schizophrenia living in the community. *Quality of Life Research* 2008; **17**(10): 1247-56.
145. Apiquian R, Elena Ulloa R, Herrera-Estrella M, et al. Validity of the Spanish version of the Personal and Social Performance scale in schizophrenia. *Schizophrenia Research* 2009; **112**(1-3): 181-6.
146. Brissos S, Palhava F, Marques JG, et al. The Portuguese version of the Personal and Social Performance Scale (PSP): reliability, validity, and relationship with cognitive measures in hospitalized and community schizophrenia patients. *Social Psychiatry and Psychiatric Epidemiology* 2012; **47**(7): 1077-86.

147. Juckel G, Schaub D, Fuchs N, et al. Validation of the Personal and Social Performance (PSP) Scale in a German sample of acutely ill patients with schizophrenia. *Schizophrenia Research* 2008; **104**(1-3): 287-93.
148. Tianmei S, Liang S, Yun'ai S, et al. The Chinese version of the Personal and Social Performance Scale (PSP): Validity and reliability. *Psychiatry Research* 2011; **185**(1-2): 275-9.
149. Schaub D, Brune M, Jaspens E, Pajonk F-G, Bierhoff H-W, Juckel G. The illness and everyday living: Close interplay of psychopathological syndromes and psychosocial functioning in chronic schizophrenia. *European Archives of Psychiatry and Clinical Neuroscience* 2011; **261**(2): 85-93.
150. Nafees B, van Hanswijck de Jonge P, Stull D, et al. Reliability and validity of the Personal and Social Performance scale in patients with schizophrenia. *Schizophrenia Research* 2012; **140**(1-3): 71-6.
151. Garcia-Portilla MP, Saiz PA, Bousoño M, et al. Validation of the Spanish Personal and Social Performance scale (PSP) in outpatients with stable and unstable schizophrenia. *Revista de Psiquiatria y Salud Mental* 2011; **4**(1): 9-18.
152. Nasrallah H, Morosini P, Gagnon DD. Reliability, validity and ability to detect change of the Personal and Social Performance scale in patients with stable schizophrenia. *Psychiatry Research* 2008; **161**(2): 213-24.
153. Hsieh PC, Huang HY, Wang HC, et al. Intercorrelations between the Personal and Social Performance Scale, cognitive function, and activities of daily living. *Journal of Nervous and Mental Disease* 2011; **199**(7): 513-5.
154. Schaub D, Brune M, Bierhoff H-W, Juckel G. Comparison of self- and clinician's ratings of personal and social performance in patients with schizophrenia: The role of insight. *Psychopathology* 2012; **45**(2): 109-16.

155. Patrick DL, Burns T, Morosini P, et al. Reliability, validity and ability to detect change of the clinician-rated Personal and Social Performance scale in patients with acute symptoms of schizophrenia. *Current Medical Research and Opinion* 2009; **25**(2): 325-38.
156. Jelastopulu E, Giourou E, Merekoulias G, Mestousi A, Moratis E, Alexopoulos EC. Correlation between the Personal and Social Performance scale (PSP) and the Positive and Negative Syndrome Scale (PANSS) in a Greek sample of patients with schizophrenia. *BMC Psychiatry* 2014; **14**: 197.
157. Nicholl D, Nasrallah H, Nuamah I, Akhras K, Gagnon DD, Gopal S. Personal and social functioning in schizophrenia: defining a clinically meaningful measure of maintenance in relapse prevention. *Current Medical Research and Opinion* 2010; **26**(6): 1471-84.
158. Srisurapanont M, Arunpongpaisal S, Chuntaruchikapong S, et al. Cross-cultural validation and inter-rater reliability of the Personal and Social Performance scale, Thai version. *Journal of the Medical Association of Thailand* 2008; **91**(10): 1603-8.
159. Phelan M, Slade M, Thornicroft G, et al. The Camberwell Assessment of Need: the validity and reliability of an instrument to assess the needs of people with severe mental illness. *British Journal of Psychiatry* 1995; **167**(5): 589-95.
160. Coombs T, Nicholas A, Pirkis J. A review of social inclusion measures. *Australian and New Zealand Journal of Psychiatry* 2013; **47**(10): 906-19.
161. Australian Mental Health Outcomes and Classification Network (AMHOCN). Development of the Living in the Community (LCQ) Measure of Social Inclusion for Use in Mental Health: Final Report. Sydney: AMHOCN, 2015.