RC PSYCH
ROYAL COLLEGE OF PSYCHIATRISTS

CR240

Outcome measures in psychiatry

June 2024

COLLEGE REPORT

Contents

How to cite this publication:
Royal College of Psychiatrists (2024) College Report CR240: Outcome measures in psychiatry

Introduction

Overview

Measuring what medical care does to improve the lives of patients must be at the heart of what we do (Ryland, Carlile and Kingdon, 2021). A crucial part of achieving the ambition of parity between physical and mental health is for psychiatry to embrace, use and promote outcome measures as a way of demonstrating the positive impact that well-resourced and structured mental health services can have for patients, their families, and society more widely.

The Royal College of Psychiatrists strongly endorses and recommends the routine use of patient- and clinician-rated outcome measures in psychiatric practice (Tracy *et al.*, 2022). Outcome measurement can improve care planning, progress-tracking, quality improvement, service evaluation and research.

This report is intended to support clinicians and services to meet the needs and circumstances of the patients they are treating. It sets out some principles governing patient- and clinician-rated outcome measurement in mental health services and then provides more detailed guidance from the College faculties, covering the specialties within psychiatric care.

We take a principles-based approach that ensures this document will remain clinically useful in support of, and relevant to, wider national policy for the short, medium, and long term across all four nations of the United Kingdom. However, there are specific national policy drivers, for example the NHS England Long Term Plan (NHS England, 2019).

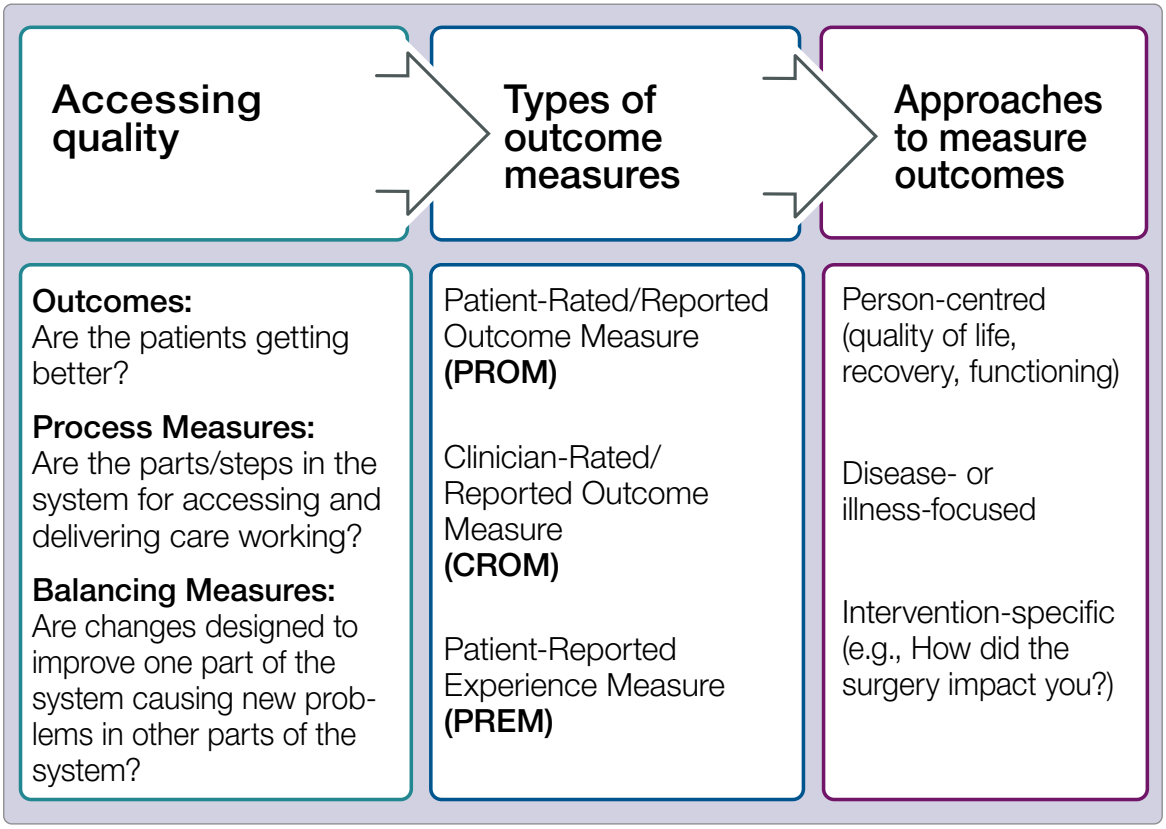
Outcome v process measurement

The Institute for Health Improvement (IHI) has defined what is meant by patient- and clinician-rated outcome and process measurement (IHI, 2022).

An outcome measure should help answer the question: 'How does the system impact the values of patients, their health and wellbeing?' –e.g., a patient's goal-based outcome.

A process measure, on the other hand, should identify "the parts/steps in the system performing as planned," e.g., time to treatment (NHS England and NHS Improvement, 2016).

Figure 1: Understanding outcomes



Patient- and clinician-rated outcome measurement is an area where there has been much research and a plethora of relevant tools for measurement developed. We list and reference some of the more widely used and tested instruments, however the list should not be taken as exhaustive or as indicating preference, except where explicitly indicated.

Choice of patient- and clinician-rated outcome measures will depend on the purpose for which they are being used. For patient- and clinician-rated outcome measures to be valuable they should be:

- Clinically relevant
- Reflect what people want
- Culturally appropriate
- Aligned with the wider system
- Used measures with established reliability and validity (NHS England and NHS Improvement, 2016).

For engagement with such measurements to be sustained and optimised, there needs to be agreement between service commissioners (funding bodies), services (provider organisations), clinicians, patients, and carers (where involved) about the outcomes which are important for them and for mental health services.

Principles informing the use of patient- and clinician-rated outcome measures.

The Royal College of Psychiatrists supports the development, selection, implementation, routine use, and interpretation of patient- and clinician-rated outcome measures, and sets out the following principles:

1. Patient-and clinician-rated outcome measures should support patient care.

- **a** The primary use of outcome data should support quality improvement.
- **b** When interpreting outcome data, one should be transparent about the limitations of such data and mindful of unintended consequences of implementation.
- **c** Comparison and benchmarking numerical changes in outcome data-gathering should be avoided where a 'good outcome' has not been established.
- **d** Measures may allow comparisons between teams and services, however this should be done with reflection, motivated by improving care and learning.
- **e** When data quality is poor and or there is lack of clarity of desired outcome, premature benchmarking or comparison of services should be avoided.
- **f** Enable reflection of data by all key stakeholders and used to support quality improvement.

2. Patient- and clinician-rated outcome measures need to be clinically meaningful.

- **a** Measures should be relevant to patients and clinicians and co-produced with patients.

3. Patient- and clinician-rated outcome measures needs to be clinically valid.

- **b** Patient- and clinician-rated outcome measures must be backed by scientific research and valid for the patient and wider population.
- **c** Focus on what is important to patients and carers.
- **d** Measures should be clear and unambiguous.
- **e** Measures should be validated for the purpose for which they are used.
- **f** Patient- and clinician-rated outcome measures are supported by digital enablers.
- **g** Digital systems should simplify the data collection and analysis and ensure maximum use of data already collected.
- **h** The system should optimise the reliability of outcomes data and data quality.

i Data should be used at the individual clinical, team and organisational level.

j There should be timely feedback of the data to patients, carers, and clinicians so that outcomes can influence the treatment process.

k Digital enablers must minimise the clinical burden for outcomes measurement.

While approaches around patient- and clinician-rated outcomes have been constantly evolving over the past years, we are confident that these principles still apply and, through their adoption, will help patients and services alike.

Many areas in the United Kingdom are considering mental health care provision within the context of integrated systems. Measuring patient- and clinician-rated outcomes across acute and mental health care, and social care, provided by a wide range of organisations, is complex, but it is important to achieve a truly integrated care pathway. Given the impact of physical morbidity and mortality among people with mental illness, and the lack of engagement of some individuals with severe mental illness with primary care services, physical health indices should be included in outcome measures for mental health services. Outcome measures used need to be culturally appropriate.

Individual outcomes

Documenting patient- and clinician-rated outcome measurements in clinical records is valuable to help identify what works and what does not work for that individual, and what other factors or circumstances may affect the impact of interventions. Many people who need mental health services will need their input over extended periods. Therefore, the timescales for assessing the effectiveness of interventions may be long and measurement should be tailored to each individual. It is important that at each stage, patients are offered treatments of established effectiveness. Sound approaches to applying outcome measurements could help to safeguard patients from inappropriate treatments that may be theoretically useful but are not effective in their individual case. Any treatment may also have side-effects, including psychosocial interventions, so these must also be measured.

Service and population level

Collecting patient- and clinician-rated outcome measures allows some comparison between services which may help optimise care and treatment, and provides opportunities for learning and sharing good practice. It is particularly important in this regard to consider possible moderating or mediating factors. For example, it might be expected that services in areas of high socio-economic deprivation will have greater difficulties establishing effective rehabilitative pathways and this must be considered when comparing outcomes, such as hospital admission rates or lengths of stay. Nonetheless, comparisons may help to identify aspects of well-performing services which could be modelled to enhance the treatment experience and outcomes for people in less well performing services. Pooled outcome data for population can also allow analysis of sub-groups or population with protected characteristics to allow planning for equitable

care provision and minimise health inequalities. Holistic and person-centred measures may offer insight into areas of discontent and need in a local population.

We recognise that there has been some interest in linking funding for services to improvements in patient- and clinician-rated outcome measures, but we suggest that the field is insufficiently developed for this to be a valid approach at the current time. At the initial stages, it might be advisable that financial drivers support good quality and routine outcome data-gathering and reflection on the data rather than focusing on numerical improvement. However, we do very much support the concept that transparency of outcomes, taking account of the limitations referred to above, is important in providing confidence that the investment of resources is being used effectively and efficiently. Therefore, RCPsych supports the routine collection of patient- and clinician-rated outcome measures.

Role of psychiatrists

The primary role of psychiatrists and mental health teams is to support the patient to achieve their goals while ensuring safety, promoting health, and keeping the patient's best interests in mind. Patient- and clinician-rated outcome measures for some conditions and individuals might be recovery-focused, while, for others, the aim might be stability. Goals are dependent on patient factors, such as their motivation for change, and are influenced by the natural course of conditions being treated. Treatment success varies by individual and is not always going to be about 'cure' or numerical improvement in scores. Outcomes, especially patient-reported outcome measures, can be a powerful driver for patient-centred care and co-production. Outcome measurement should encourage reflection to support the patient achieve their personal goals and allow services to improve the quality of care they offer.

The range of outcome measures available both for individual and service level measurement is considerable, so selection of a manageable number and type of measures to fit with information needs, skills, and resources available is essential. Having a small number of well-conducted and regularly viewed measurements is preferable to having large numbers of outcomes that might be inconsistently measured, need sophisticated analysis or that no one has the interest or time to examine. Measures should be chosen collaboratively, with input from all relevant stakeholders, including patients, carers, and clinicians, wherever possible. The choice of measures should be kept under review to maintain their relevance, as both patient and service needs change over time.

Interpretation of outcome measurements

While individual patient- and clinician-rated outcome measures inform treatment goals, anonymised, pooled outcome data digitally captured centrally can serve several functions. At a team level, it allows reflection and opportunity to use that data for quality improvement (QI). Visibility of outcome measures can inform and steer patient care, and team-level data and feedback on the benefits of work done with patients who have shown improvement have huge benefits for staff morale and wellbeing. Data on

improvement in mental health conditions might be pooled data over time and inform us about what a 'good outcome' is if captured routinely over time.

Whilst brief and focused interventions or hospital admissions often offer a clear start and end point for measuring impact of treatment through routine outcome measures captured at assessment and discharge for the episode, this model might not fit all mental health services or, in fact, conditions. We suggest routine outcome data-gathered should be captured and understood in the context of the course of the disease or condition and service provision. This might be of relevance in longer-term care provision for long-term conditions or chronic relapsing-remitting conditions. Routine outcome data-gathering can offer critical intelligence to researchers, clinicians and all organisations to compare, reflect and analyse data to offer clarity on desired outcomes, where such information is lacking.

The use of digital technology can help improve the efficiency of collecting outcome measures, empower patients in their own care as they can gather and review their own improvement, enable seamless data entry, and minimise the burden on front-line clinicians, as well as enabling feedback to individual clinicians and teams. There are likely to be real opportunities for learning in this area that emerged during the Covid-19 pandemic, where digital technology became crucial to maintaining levels of care for some patients, with some useful evidence as to which approaches work best with patients and at what time in the patient journey. Best practice guidelines should be followed when digitising patient- and clinician-reported outcome measures to ensure fidelity to the original scales and maintain the integrity of their psychometric properties.

Transparency and selection of outcomes

Transparency regarding the achievement of patient- and clinician-rated outcomes and quality levels supports patient choice, enables benchmarking of care services across equivalent services, and aids workforce planning and effective resource use.

The following chapters of the report offer specific advice and guidance on the use of patient- and clinician-rated outcomes across the range of specialties involved in psychiatry. Each faculty was asked to provide their own section focusing on outcome measurement, e.g., clinical outcomes, rather than process measurements, e.g., waiting times. Each faculty has therefore developed their section individually, rather than following a framework. The strengths of doing this include: prompting each faculty within RCPsych to consider patient- and clinician-rated outcome measurement, and giving each faculty ownership of its section. This also would, in future, allow sections to be individually updated, as needed. However, we recognise there are also limitations to this faculty-specific approach, such as a lack of standardisation between faculties in their approaches to outcomes, the potential for conflicting advice between faculties, and the fact that faculties do not precisely map to clinical services. Additionally, while there has been a lot of interest and progress in routine outcome data-gathering in mental health in recent years, this has meant that different mental health services have adapted different routine outcome data-gathering and reporting at different times and the systems are at different levels of maturity.

This report was prepared over a period of time based on RCPsych's stance on outcomes. In the future, we intend to mitigate the document's limitations by making regular updates to it. One of the issues we do acknowledge is the document in its current form does not place sufficient emphasis on patient-experience measures. In future iterations, we will need to consider the use of such measures, e.g., the care measure which the National Collaborating Centre for Mental Health (NCCMH) supports. We will also take emerging guidance, on both a national and international level, into account – for example, the NCCMH's Patient Reported Outcome Measures (PROMs) for People with Severe Mental Illness in Community Mental Health Settings.

References

Institute for Healthcare Improvement (2022) Science of Improvement: Establishing Measures. Available at: https://www.ihi.org/resources/Pages/HowtoImprove/ScienceofImprovementEstablishingMeasures.aspx

NHS England and NHS Improvement and (2016) Delivering the Five Year Forward View for Mental Health: Developing quality and outcomes measures. Available at: https://www.england.nhs.uk/mentalhealth/wp-content/uploads/sites/29/2016/02/mh-quality-outcome.pdf

NHS England (2019) 'NHS Long Term Plan', NHS Improvement

Ryland H, Carlile J and Kingdon D (2021) 'A guide to outcome measurement in psychiatry'. *BJPsych Advances*, **27(4)**: doi: 10.1192/bja.2020.58

Tracy DK et al (2022) 'Digital literacy in contemporary mental healthcare: electronic patient records, outcome measurements and social media', *BJPsych Advances.* Royal College of Psychiatrists, 1–8. doi: 10.1192/bja.2022.74

Academic psychiatry

The development, use and interpretation of validated outcome measures is of particular interest to the Academic Faculty of the Royal College of Psychiatrists. This faculty draws together psychiatrists who work at the forefront of clinical research, as well as teaching typically within medical school and higher education settings. The faculty seeks to promote and facilitate high-quality research that informs clinical training and evidence-based practice and to ensure the availability of clinical academic career pathways in psychiatry. Academic psychiatrists combine their research and educational activity with clinical service jobs and frequently take on roles within other College faculties and contribute to national and international scientific bodies. This can provide a contextualised perspective on appropriate use of outcome measures.

Quantitative research in mental health relies on careful measurement of psychological, biological and social factors either to test hypotheses for the understanding of mechanistic processes (e.g. that a treatment exerts an effect on a symptom) or to mine for non-random patterns in data that may yield novel insight into processes of interest (e.g. linking a condition to a particular pattern of genetic vulnerabilities). Outcome measures represent one subset of these quantitative tools for recording the impact and effectiveness of clinical activity on the health and wellbeing of patients. Biopsychosocial aspects of the clinical experience ought thus to be captured by the range of subjective (e.g. patient, carer or clinician-rated measures) and objective (e.g. service measures and, if possible, disease biomarkers) outcomes to enable a comprehensive picture. The advancement of psychiatric prevention and care through robust evidence-based practice fundamentally necessitates the use of outcome measures at a clinical service level. Many of these measures had been developed and implemented as research tools. Academic psychiatrists can take some responsibility for the shaping, selection and validation of clinical outcome measures and in guiding the appraisal, interpretation and revision of outcome data.

While one purpose of outcome measures is in the auditing of performance (potentially linked to assignment of resources or more explicitly 'payment by results'), many of the same principles, and reservations, apply to data for both audit and research. Generic outcome scores offer an opportunity for different interventions and services to be compared across different areas of (mental) healthcare, but there are recognised risks in generalised comparisons: For example, outcome measures of recovery do not fit well with neurodevelopmental disorder or dementia. These issues are considered in the sections below.

Similarly, there are potential pitfalls in the assimilation and interpretation of general data, where ratings scales may carry implicit assumptions of a linear continuum: Following an intervention, a one- or two-point change in outcome scores across a cohort of patients may represent a general improvement in wellbeing of the majority, or a major change/recovery for a small subgroup. Big data/data-mining approaches can often blur out nonparametric effects. The validated interpretation of outcome measures is critical to their effective and meaningful use. Across many areas of science, concerns are being raised regarding the replication of observed effects, where apparent breakthroughs turn out to be artefactual noise. These concerns are driving changes to registration, reporting, and statistical inference in science (e.g. Bell, 2017; Amrhein *et al.*, 2019).

In this context, the perspective of research-active academic psychiatrists is likely to be valuable to ensure inferences from quantitative outcome data are scrutinised, for example, when used as evidence for policy decisions.

Consensus on generalisable outcome measures can offer something valuable to clinical research. Experimental and observational research, including clinical trials, is often conducted in selected populations, and the effects are judged against fine-grained tailored metrics of change that are rarely used in standard clinical settings. Inclusion of easy-to-use outcome and experience measures can help frame and translate the research in broadly accessible terms. Moreover, they can be used to demonstrate the recognised benefits of research participation for patients and services.

Outcome measures are dependent on the expectations of the clinical workforce, stakeholders and service users. The teaching and training of future psychiatrists and allied mental health professionals is thus an important consideration. For example, the rating of a positive outcome may depend on the extent to which one was taught that schizophrenia, once established, is unremitting.

What we teach therefore should be accurate: High-quality research and clinical outcome data should provide continually updated information for the lifelong learning and professional development of the workforce. However, there is some circularity if what we teach biases the collection of information, and there is a further potential link to 'policy' when selecting and implementing outcome measures. Access to data is also important: Clinical academics, as both researchers and educators, need to communicate reliable and current information. Full and open access to outcome measures will remain fundamental to this. If outcome data become a key metric for performance monitoring, and funding of organisations, competitive, even commercial, interests might impede optimal use of this potentially valuable resource.

With increasing discussion of, and some progress towards, personalised care pathways and interventions in mental health, meaningful quantification of what is being achieved for different people necessitates measures that can be generalised. For the most part, biological psychiatry has yet to furnish markers and measures that will contribute to outcomes in the way an oncologist may read a scan.

However, steady progress is being made and step-changes are plausible. Still, current objective measures of outcome exist at the level of service use, occupational activity and social care or, more crudely, life expectancy. Patient and clinical rate outcomes, including both general and fine-tuned measures of change in symptoms and psychosocial functioning, are critical to capture the impact of mental health provision, so that it can be optimised for better patient care.

The Academic Faculty fully support the College's recommended use of easy-to-implement, validated and meaningful outcome measures in psychiatric practice. To maximise potential clinical and educational benefits, data from these measures must be accessible for open appraisal. Advances in the implementation of digital health records systems that allow for the extraction of anonymised data can support this goal at local and national levels.

References

Amrhein V, Greenland S and McShane B (1999) Scientists rise up against statistical significance. *Nature*, **567(7748)**: 30507

Bell, V (2017) Open science in mental health research, *Lancet Psychiatry,* 4(7): 525-26

Addiction psychiatry

Summary of outcome measures

Service	CROM	PROM	PREM
Required (Treatment Services)	Treatment Outcomes Profile (TOP) Alcohol Outcomes Record (AOR) Young People's Outcomes Record (YPOR) TOP and YPOR in secure settings	Treatment Outcomes Profile (TOP) Alcohol Outcomes Record (AOR) Substance Use Recovery Evaluator (SURE)	
Optional/ Other setting	AUDIT-C/AUDIT DUDIT	SUSS AUDIT-C/AUDIT DUDIT LDQ EQ-5D-5L	Friends and Family test Patient Satisfaction Scale

Principles informing the development of outcome measures

Monitoring of addiction treatment outcome is required of all addiction services in receipt of public funding for services in England, both those provided by NHS trusts and the voluntary sector. The nature of that monitoring is currently specified by the Office for Health Improvement and Disparities (OHID) (formerly Public Heath England) and is usually required as a contractual obligation by Local Authority Commissioners, as the statutory bodies commissioning publicly funded addiction treatment services.

The nature and timing of treatment outcome monitoring is specified by OHID in detail and is collected and collated centrally through the National Drug Treatment Monitoring System. Local and national healthcare professionals, commissioners and policy makers can access and use the data to monitor and improve performance of services and the treatment system (which can include a range of services) as a whole. Outcomes monitoring is effectively mandated, which has the advantage that it has a high level of implementation. The strong link to performance management and contract monitoring means however that the focus can be on system and financial outcomes which are of more interest to politicians (such as abstinence) rather than some of the more subtle factors (e.g., wellbeing or quality of life) which may be of more interest to individual patients.

Outcomes monitoring for patients in mental health services with co-morbid substance misuse has different requirements and is less well developed. The range of potential outcome measures was comprehensively reviewed by the CLAHRC Consensus Group on Outcome Measures for Addiction, and published as part of the RESULT project (Routine Evaluation of the Substance Use Ladder of Treatments).

The Government Drug Strategy, *From Harm to Hope,* published in 2021 followed a previous review of services by Dame Carol Black. There is an emphasis on comorbid mental health problems and substance misuse which will hopefully result in more focus on outcomes for service users affected by this.

Streams of outcome measurement development

Patient-informed measures

Of service:
Patient Satisfaction Scale

Of personal change:
a. Commissioned addiction services:

- Treatment Outcomes Profile (TOP) *self–completion sections*
- Alcohol Outcomes Record (AOR) *self–completion sections*

b. Comorbid substance use/dependence in other settings:

- SUSS: Substance Use in Specialist Services designed for routine clinical practice and research
- AUDIT-C: Alcohol Use Disorders Identification Test (consumption) designed as screening tool
- DUDIT: Drug Use Disorders Identification Test designed for screening and research
- LDQ: a 10-item measure of substance dependence designed to inform treatment planning in routine clinical practice or a self-directed intervention
- ASSIST-Lite: Alcohol, Smoking and Substance Involvement Screening Tool Lite is an alcohol, tobacco, cannabis, stimulants, sedatives, opioids and psychoactive substances screening tool. It has been modified and licensed for use in health and social care settings in the UK and included in the mental health services dataset. ASSIST-Lite screening tool: how to use (gov.uk)

Relative-/carer-informed measures

Of service:
Friends and Family Test

Of personal change:
None recommended

Staff-informed measures

Of service (set by OHID and used for performance management):

- Drug and alcohol treatment completion and drug misuse deaths.
- Adults with substance misuse treatment need who successfully engage in community-based structured treatment following release from prison.
- Smoking status at time of delivery in pregnant women.

Of personal change:
a. Commissioned addiction services:

- Treatment Outcomes Profile (TOP)
- Alcohol Outcomes Record (AOR)
- Young People's Outcomes Record (YPOR)
- TOP and YPOR in secure settings.

Research-informed measures

Of service:
None recommended

Of personal change:
The range of outcome measures that may be used in addiction research is large and broadly divides into:

- measures of psychological or physical dependence (e.g., LDQ, SDS, SDSQ, SDSS)
- measures of levels of substance use (frequency and quantity) (e.g., SUSS, AUDIT, DUDIT, ASSIST-Lite)
- recovery measures (ARQ, ARC)
- general psychological health, wellbeing, and quality of life (e.g., PHQ-9, GAD-7, EQ5D)
- Substance Use Recovery Evaluator (SURE).

Any requirement for a minimum data set

Of service:

- Drug and alcohol treatment completion and drug misuse deaths.

Of personal change:

- Treatment Outcomes Profile (TOP)
- Alcohol Outcomes Record (AOR)

Interpretation of outcome measurements

A range of outcome measures are required to assess the different domains of potential harm:

- levels of substance use and dependence
- physical and mental health
- social, housing, education and employment
- criminal activity
- quality of life.

The balance, as always, is to record person-specific outcomes while not over burdening with assessment.

In Scotland, NHS Health Scotland collates a series of core outcome measures for local Alcohol and Drug Partnerships (ADPs) (NHS Health Scotland, 2014). Many of these are high-level public health outcomes, whilst a number relate specifically to drug and alcohol treatment service outcomes – particularly measures benchmarking implementation of 'Medication Assisted Treatment' (MAT) standards.

The MAT standards aim to improve access, choice and care and to ensure that MAT is safe and effective:

1. All people accessing services have the option to start MAT from the same day of presentation.

2. All people are supported to make an informed choice on what medication to use for MAT, and the appropriate dose.

3. All people at high risk of drug-related harm are proactively identified and offered support to commence or continue MAT.

4. All people are offered evidence-based harm reduction at the point of MAT delivery.

5. All people will receive support to remain in treatment for as long as requested.

6 The system that provides MAT is psychologically informed (tier one); routinely delivers evidence-based low-intensity psychosocial interventions (tier two); and supports individuals to grow social networks.

7 All people have the option of MAT shared with primary care.

8 All people have access to independent advocacy and support for housing, welfare and income needs.

9 All people with co-occurring drug use and mental health difficulties can receive mental health care at the point of MAT delivery.

10 All people receive trauma-informed care.

Benchmarking for each standard occurs through numerical evidence, process evidence and experiential evidence. Examples of outcomes include: less than 2% unsupported discharges; evidence of standard operating procedures for delivery of standards; and, feedback from individuals with lived and living experience.

Other outcome measures include:

- percentage reduction in daily drugs spend during treatment
- reduction in the percentage of clients injecting in the last month during treatment
- proportion of clients who abstain from illicit drugs between initial assessment and 12-week follow-up
- proportion of clients receiving drugs treatment experiencing improvements in employment/education profile during treatment
- number of treatments drug service clients receive at 3 month and 12 month follow-up
- wait time targets.

The Welsh Government has been collecting TOP data on treatment outcome since 2009 but has a separate data system from England. There is no equivalent treatment outcome monitoring system in Northern Ireland although they are planning to introduce the TOP.

Glossary of relevant measures

1. Treatment Outcomes Profile (TOP)

This uses a validated data collection tool (Marsden *et al.*, 2008; Public Health England, 2016) to assist clinicians (typically keyworkers) to collect patient-level clinical information on:

- substance use
- injecting risk behaviour
- crime and health
- housing, education and employment
- quality of life.

It contains both clinician-rated and patient-rated outcomes. It is collected at various points in the treatment pathway, including at treatment entry, during regular clinical reviews, at discharge, and post-discharge.

The data can be used by clinicians to support treatment planning and setting treatment goals with clients. Managers and commissioners of services are able to use TOP quarterly outcomes reports to track service performance and assist in quality improvement and contract monitoring.

2. Alcohol outcomes record (AOR)

The AOR measures change and progress of patients in treatment for alcohol misuse. It includes four outcomes from the TOP identified as a minimum data set for this patient group:

- alcohol use
- tobacco use
- psychological health
- physical health.

Services can choose whether to use the AOR or the full TOP with their patients.

3. Young people's outcomes record (YPOR)

Keyworkers can use the YPOR to record outcomes for young people in treatment for substance misuse. It has a series of questions on alcohol use, drug use, health and wellbeing.

4. Measuring outcomes in secure settings

All secure settings should record a TOP or YPOR (depending on age) for any new detainees receiving substance misuse treatment. Secure settings include:

- prisons
- immigration removal centres
- secure children's homes
- welfare only homes
- youth offender institutions
- secure training centres.

5. Alcohol Use Disorders Identification Test (AUDIT)

The AUDIT is a 10-item screening tool (Maximum score = 40) developed by the World Health Organization (WHO) to assess alcohol consumption, drinking behaviours, and alcohol-related problems. Both a clinician-administered version and a self-report version of the AUDIT are provided. AUDIT-C is the first three questions of the full AUDIT recording frequency and quantity of alcohol use (Maximum score = 12).

6. ASSIST-Lite

The ASSIST-Lite (Alcohol, Smoking and Substance Involvement Screening Tool - Lite) is a shortened version of the ASSIST screening tool which was developed for the World Health Organization (WHO) by an international group of researchers. Its function is to help detect and manage substance use and related problems in healthcare settings.

The ASSIST-Lite has been modified and licensed for use in health and social care settings throughout the UK.

Two versions of the ASSIST-Lite have been developed. One version is specifically adapted for use in mental health settings, the other is for use in all other health and social care settings.

Both versions of the tool include the alcohol use disorders identification test – consumption (AUDIT-C) for identifying health risk from alcohol consumption.

7. Substance Use Recovery Evaluator (SURE)

SURE is a psychometrically valid, quick and easy-to-complete outcome measure, developed with unprecedented input from people in recovery. It can be used alongside, or instead of, existing outcome tools. There are 21 questions and each question scores 1, 2 or 3. This means it is possible to score between 21 and 63.

Child and adolescent psychiatry

Introduction to CYPMH outcome measures

The Child and Adolescent Faculty of the Royal College of Psychiatrists strongly supports the use of outcome measures. Their use improves effectiveness and the quality of services, and they are valued by both clinicians and service users. As set out in the introduction to this report, the College believes that the principles set out in its Occasional Paper: *Outcomes Measures Recommended for Use in Adult Psychiatry* (2011) remain relevant.

In 2016, NHS England noted that local areas would need to ensure a suite of quality and outcome measures were utilised and that these should be:

- clinically relevant, so that they are seen to add value for clinicians as a routine part of their clinical practice and as part of a continuous quality improvement process
- reflect what people who use the service (and their families) want
- culturally appropriate and culturally reliable
- aligned with system-wide objectives
- measurable using metrics with established reliability and validity.

Specialist CYPMH and the whole system

The NHS has a Mental Health Services Data Set (MHSDS) for England to assist in the storage, management and dissemination of patient information. The MHSDS has been in use since January 2016 and all NHS funded mental health services in England are required to send data to this collection as a priority. However, many providers report difficulties in doing this.

The Child and Adolescent Faculty has had extensive involvement in developing outcome metrics since 2012, spanning four areas:

- Development of the new database.
- The development and piloting of a CAMHS currency.
- Steering the development and selection of outcomes metrics suitable for national roll-out.
- The use of outcome measures within CYP IAPT.

The NHS has set a clear direction to prioritise the recording and use of data based on outcomes. By the year 2020/21, it was intended that there would be "national metrics to support improvements in children and young people's mental health outcomes" and that "all services should routinely collect and publish outcomes data."

The NHS Mental Health and Dementia Programme Board, which covers all aspects of mental health delivery in the NHS, has agreed a metric to capture reliable improvement in the presenting problems of children and young people (CYP) which will take into account movement towards goals. This metric will be tested out, which will involve gathering feedback from clinicians and managers about its strengths and weaknesses as a measure of improvement.

Reliable change index (RCI)

This has been agreed by NHSE and the Anna Freud Centre to be trialled as the outcome metric as it:

- prioritises the voices of children and young people
- considers the amount of change achieved and takes into account measurement error of the measure
- does not require expectation of complete recovery
- has been trialled within child IAPT data, and shows improvement of about 50%
- aligns with approaches being developed across all age MH services
- can be used with different measures allowing practitioners to use specific and appropriate measures
- supports clinical conversation with CYPs and carers and enhances informed choice and shared decision making.

Two data points are required to make an RCI. These can be collected at any point in the patient journey. RCI categorises outcomes into 'improved', 'no change' and 'worse'. Reliable change will be calculated centrally by NHS digital.

The definition of reliable change is taken from this the report by Wolpert and colleagues (2016):

> *"Reliable change considers the amount of change from one time point to another, relative to the properties of the measure used, thereby counting as reliable change that is unlikely to be attributable to measurement error alone.*
>
> *– Jacobson & Truax, 1991*

> ***1. Reliable improvement:*** *where a score on at least one measure changed enough for it to be considered statistically reliable and no other score reliably deteriorates.*

> ***2. No reliable change:*** *where a score on all completed measures did not show any statistically reliable change.*

3. Reliable deterioration: *where a score on one or more measures changed enough in a negative direction for it to be considered statistically reliable.*

Context is important, and the authors suggest adaptations to the criteria are made in the light of the child context; to take into account the much wider range of self- and parent-reported measures. To be classed as reliably improved, at least one scale has to have reliably improved and no scale can have reliably deteriorated."

RCIs can be calculated on all measures that follow; however, central calculation of RCI by NHS digital will be done on those measures that have a current licence.

Context

In specialist CYPMH, service-based outcomes are linked to a whole system approach that includes (not an exhaustive list):

- education
- social services
- youth services
- youth justice services
- voluntary organisations
- parent support groups
- cultural and religious support networks.

Outcomes of specialist CYPMH depends on this interlinked system; service resources and service function affect outcomes within and between systems. A whole system approach to CYPMH contains a number of models for assessment, formulation and measurement, including the Choice and Partnership Approach (CAPA), and I-THRIVE. All these are person-centred, and focus on values of the CYP and carers, as well as promoting shared decision making, and routine use of clinically meaningful outcome measures (Law & Wolpert, 2014).

NHS digital does collect service-based measures (activity levels, waiting times, DNAs), and this gives a measure of service provision against service resource. There is a flow of this data to NHS digital, and public dashboards are available that show monthly service data.

The main focus here is on clinician- and patient-reported outcome measures (CROMs and PROMs). The principles in the introduction apply to this data and should be borne in mind throughout.

Aim of this chapter

- To provide information about the Mental Health Services Data Set (MHSDS) and the outcome measures within this. The MHSDS applies only to England. Other parts of the United Kingdom have their own datasets.
- To provide a set of general measures (applicable to all disorders) based on recommendations from members of the Faculty of Child and Adolescent Psychiatry, and CYPMHS academics.

To suggest, where possible, additional condition specific measures (that may not necessarily be present within the current MHSDS). This is not intended as a definitive recommendation and will be updated as required via feedback from clinicians and other stakeholders. The guidance offered in this chapter is not binding for services.

The advice here may be relevant to other national groups (Wales, Scotland, Northern Ireland). However, the MHSDS applies to England only, and the information on the RCI applies to CYPMH services in England, and their flow of data into the MHSDS. 'Parent' in the document refers to parent, carer, or guardian.

Wales

There is an agreed minimum dataset, which will be integrated with the Health and Social Care database system over the next three years. To start with, the baseline outcomes data will include the Strength and Difficulties Questionnaire (SDQ), Experience of Service Questionnaire (ESQ) and the Goal Based Outcomes (GBOs) tool.

Scotland

At present there is no national agreement about outcomes, and each Health Board has its own requirements for outcome data returns.

Northern Ireland

At present there is no national agreement about outcomes. At a regional level, a minimum dataset is agreed and available focusing on demographics, activity and counts of disorders present. This dataset is currently being revised and may incorporate outcome recommendations. At a trust level, agreed outcomes have developed independently; for example, in the Northern Health and Social Care Trust, agreed outcomes are similar to those recommended within this document.

Child- or young-person-informed measures

The guidance about questionnaires detailed below, and in the summary table, were chosen via integration of discussions with the following: clinicians, College executive meeting attendees, researchers, experts in specific fields, and College representatives leading on outcome measurements.

A glossary of terms (if not defined in text), and hyperlinks to questionnaires is included in this chapter's Appendix, together with the current digital licences, and link to version 4.0 of the MHSDS (which contains a list of questionnaires that can be submitted to MHSDS).

The general measures are in use and often multiple measures are used with any one family. There is a general desire to reduce the numbers of measures in order to reduce administrative overheads on both clinicians and admin staff and to collect a more complete dataset on individuals and families.

A Quality Improvement (QI) approach is recommended to address difficulties in outcome measure collection. With sufficient QI work and allocation of resources, more complete RCI calculations are possible as well as a clearer focus between clinicians, CYPs and their families, about the importance and meaning of using outcome measures (Law & Wolpert, 2014).

Of service:

Activity data is collected and sent (or 'flowed') to MHSDS in England and equivalents in Northern Ireland, Wales and Scotland.

- **Current View questionnaire**
  This is a multiple-domain summary document, being trialled in the CYPMH currency pilot (payment by results). It provides data about complexity, comorbidity, and severity of conditions. It is a summary clinician-reported outcome measure (CROM) and does not have known psychometric properties, so RCI generation is not possible.

- **Experience of Service Questionnaire (ESQ)**
  12-item Likert scale satisfaction questionnaire. Its psychometric properties have been researched (Brown *et al.*, 2014); however, using it before and after service use would not be particularly useful, as it is, by definition, an 'experience' questionnaire. Thus, RCI is unlikely to be a useful measure. However, it can be used to compare satisfaction between services.

Of personal change:

- **Child Outcome Rating Scale (CORS) and Outcome Rating Scale (ORS)**
  Four-item continuous-type scale that can be completed by parent of child including under 5s (YCORS). Designed as potentially usable session by session, or as an outcome measure across time points.

- **Goal Based Outcomes (GBOs)**
  This is an idiographic measure of change. It is a 10-point Likert scale and allows personalisation of goal setting, and is well liked and used by CYPs and clinicians. Currently GBO is being reported alongside the metric of reliable change, with movement of three points being considered reliable. This method is being tested further.

- **SCORE-15 (Family functioning)**
  15-item Likert scale questionnaire, with additional qualitative/descriptive detail. The 15 items generate three subscales of family functioning.

- **Strength and difficulty scale (youth, parent, teacher versions)**
  25-item questionnaire with three points, generates total score and five subscales. The impact scales are useful in clinical situations where symptoms may stay present, but impact of those symptoms lessen.

Parent-informed measures

Of service:

Activity data as collected and flowed to MHSDS in England and equivalents in Northern Ireland, Wales and Scotland.

- **Experience of Service Questionnaire (ESQ)**
  See above in the 'Child or young person informed measures' section under the subheading, 'Of service'.

Of personal change:

- **Brief Parental Self-Efficacy Scale (BPSES)**
  Five-item Likert scale. There is no psychometric data available, so it is difficult to generate RCI.

- **SCORE-15 (Family functioning)**
  15-item Likert scale questionnaire, with additional qualitative/descriptive detail. The 15 items generate three subscales of family functioning.

- **Strength and difficulty scale (Youth, parent, teacher versions)**
  See above in the 'Child or young person informed measures' section under the subheading, 'Of personal change'.

Staff-informed measures

Of service:

- **Trust staff survey questionnaires**
  Complaints, staff sickness and retention rates, significant and near-miss events. Trust reports of service and inspection carried out by external bodies such as CQC.

Of personal change:

- **Children's Global Assessment Scale (CGAS)**
  A single scale (0 to 100) measure.

- **The Health of the Nation Outcome Scales for Children and Adolescents (HoNOSCA)**
  This group of scales (HoNOS family) consist of 12 domains (Likert scales) covering symptoms and social functioning. There are parent and self-rating versions available, although these are not used often.

- **Condition specific/research outcome measures**
  There is considerable overlap between research outcome measures and use of outcome scales to improve patient outcomes, including service-based quality improvement initiatives. Many questionnaires in quality improvement are used as either improvement measures or are used to support research. What follows is a list of outcome measures that are in common use. The measures detailed here have known psychometric properties.

In-patient care

Clinician-Reported Outcome Measures (CROMs):

The Current view for HoNOSCA and CGAS: NHSE stipulates the use of HoNOSCA and CGAS. The former is not particularly sensitive to change (according to clinician feedback during the preparation of this chapter). However, the HoNOSCA has subscales that are used clinically to measure change, and there are parent and youth self-report versions. At present, reliable change calculation for total scale is in development.

Patient-Rated Outcome Measures (PROMs):

Use condition-specific questionnaires.

Disorder specific

Depressed mood and anxiety disorders:

Children and young people often present with more than one mental health condition, e.g., anxiety and depression are often co-morbid.

Clinician-Reported Outcome Measures (CROMs):

NICE guidelines for depression suggest using the Kiddie-Schedule for Affective Disorders and Schizophrenia (K-SADS) as a possible CROM. Until further evidence is available, this is a possible option; however, this measure is relatively labour intensive. Other brief condition-specific CROMs include subscales of the HoNOSCA.

Patient-Rated Outcome Measures (PROMs):

The Revised Children's Anxiety and Depression Scale (RCADS) is a 47-item Likert questionnaire that produces depression and anxiety specific subscales as well as total scores. It is available in child and parent versions. It is in common use, is recommended for use by CYP IAPT and has good user acceptability. It also is available in symptom tracker formats (6-10 items per symptom category) that can be used to track symptoms over sessions while in treatment. It is not currently on the licensed list of NHS digital, but is within the recommended MHSDS list. Because it is not yet licensed, it is unclear whether RCIs will be calculated, though psychometric properties allow RCI calculation.

There are a number of other questionnaires that are not within the MHSDS at present, but are being used by specialist CYPMH services. These include the Screen for Child Anxiety Related Disorders (SCARED), Children's Obsessional Compulsive Inventory (CHOCI), Child PTSD Symptom Scale (CPSS) and the Mood and Feeling Questionnaire (MFQ). The MFQ is not in the MHSDS, nor licensed, though it is used clinically and in research trials.

Neurodevelopmental disorders

This covers generalised learning disability (LD), attention deficit hyperactivity disorder (ADHD) and autism spectrum disorder (ASD).

Clinician-Reported Outcome Measures (CROMs):

The LD version of the HoNOS, or the HoNOSCA.

Patient-Rated Outcome Measures (PROMs):

- **The Sheffield Learning Disability Outcome Measure (SLDOM)** is a parent measure of CYP (3–16) symptoms and their parental coping ability. Part one is a parent efficacy understanding 7-item Likert scale questionnaire. Part two is about engagement and collaboration with the specialist LD team (10-item Likert scale).

- **For ADHD**, services often have specific measures in use ranging from subscales of the SDQ, to Conners (various versions) and the Swanson, Nolan, and Pelham-IV (SNAP-IV). The recommendation is the SNAP-IV, though this is not in the MHSDS, whereas the SDQ is within the dataset.

- **The SNAP-IV** comes in various versions, though the 18-item teacher and parent version is commonly used. Each item is rated on a four-point scale, and the questionnaire generates three sub-scores: total, inattention, and hyperactivity-impulsivity.

- **For ASD**, services often have specific measures and use condition-specific measures for comorbidities (anxiety, obsessive-compulsive disorder (OCD), depressed mood). Suggested measures include the Repetitive Behaviours Questionnaire (RBQ-2), the Vineland Adaptive Behaviour Scale (teacher and parent version) and the Social Communication Questionnaire (SCQ).

Tics:
The Parent Tic Questionnaire (PTQ) consists of a list of 14 motor and vocal tics. It rates both frequency and intensity on five- and nine-point Likert scales respectively. It takes about 20 minutes to complete and can be easily completed by parents. There is an online version. For Tourette's an additional specific quality of life measure, Gilles Tourette Quality of Life (GTS-QOL), is suggested.

Sleep:
Two questionnaires covering sleep parasomnias and behaviours respectively: the Sleep Disturbance Scale for Children (SDSC), and the Children's Sleep Habits Questionnaire (CSHQ).

Eating disorders
See the next chapter of this document.

Emerging personality disorders
Basic symptom scales are not recommended as they are not sensitive to change.

Patient-Rated Outcome Measures (PROMs):
Personalised care and the use of questionnaires to help detail emotions, cognitions and behaviours can be helpful. Scales to do this include the Defeat Scale, and the Fear of Negative Emotions Scale (both brief measures). The Beck Youth Inventory can be used to identify a range of emotions when young people have difficulties expressing emotions verbally. A frequency count of self-harm is useful as a measure of progress. Two measures used more routinely are the Connor-Davidson Resilience Scale, and the Inventory of Interpersonal Problems.

Clinician-Reported Outcome Measures (CROMs):
The Recovery Star (a version of the Outcomes Star) and the HoNOS family of measures can be used as goal-based outcomes. Use other scales for comorbidity.

Psychoses

Specialist services such as EIS and Transition may use a range of adult outcome measures such as the Positive and Negative Syndrome Scale (PANSS) or Comprehensive Assessment of At-Risk Mental States (CAARMS). However, advisors to this document suggest that generic measures like the HoNOSCA are a first step.

Clinician-Reported Outcome Measures (CROMs):

The **PANSS** provides outcome scales for positive and negative symptoms of psychoses, a general scale, and an aggression scale. In total there are 30 items, rated on a seven-point scale. It is in common use.

A note on the **CAARMS**:
The Mental Health Services Data Set (MHSDS) lists the CAARMS. This is a CROM covering four areas, each rated on three items: frequency and duration, severity, and distress within the MHSDS. It is not regarded as an outcome measure and is used more to determine at-risk mental health states.

Patient Rated Outcome Measures (PROMs):

For YP with emerging difficulties including psychoses, services may have their own outcome measures in use, or use a range of condition specific measures. Guidance for delivering Early Intervention in Psychosis (EIP) services requires the use of DIALOG and QPR as a minimum. Although these are developed for over 18s, there may be some young people for whom the use of these tools is appropriate.

One measure used for Bipolar (mania) is the Young Mania Rating Scale (YMRS). This is an 11-item questionnaire rated on a five-point scale, giving subjective experience of mood and behaviour over the last 48 hours.

Summary table/glossary

Below is a summary table of CROMs and PROMs, and a note on whether the questionnaire/scale is free for use. Those marked with an asterisk are not currently within the MHSDS.

General measures		CROM	PROM	Free (Y/N)
All	Key measures	CGAS	CORS, YCORS ORS, GBO	Y
	Additional measures	HoNOSCA Current View	SDQ BPSES SCORE-15	N Y Y
Condition-specific measures		**CROM**	**PROM**	**Free (Y/N)**
In-patient Units	Key	HoNOSCA	CORS, ORS	Y
	Additional		BPSES SCORE-15	Y

Depressed mood and anxiety	Key		RCADS	Y
	Additional	K-SADS*	MFQ* SCARED*	Y Y
Anxiety conditions specified	Additional OCD		CHOCI* in conjunction with global measures	Y
	Additional PTSD		Child PTSD CPSS* Child revised impact of events scale CRIES	Y Y
Learning difficulties and neuro-develop-mental disorders	Key	HoNOS-LD	Sheffield SLDOM	Y
	Key Tics and Tourette's		Parent Tic Questionnaire (PTQ)	Y
	Additional Tourette's		Tourette's quality of life (GTS-QOL)	Y
	Key ADHD		SNAP* in conjunction with quality of life measure	Y
	Additional ADHD		Conners* SDQ	N N
	Additional Sleep		Sleep Disturbance Scale for Children (SDSC)* Children's Sleep Habits Questionnaire (CSHQ)*	Y Y
	Additional ASD		Repetitive Behaviours Questionnaire (RBQ-2)* Social Communication Questionnaire (SCQ)* Vineland questionnaire, parent and teacher*	Y N N
Emerging Personality Disorders	Key	Recovery (outcomes) Star*	Connor-Davidson Resilience Scale * Inventory of interper-sonal problems*	N to both
Psychoses	Key	HoNOS family PANSS*	YMRS*	Y except PANSS
	Additional	Recovery (Outcomes) Star*	Side effect scales * (AIMS, Barnes, Simpson-Angus)*	Y except for Star

Appendix

Abbreviation	Name	Link
CGAS	Children's Global Assessment Scale	https://www.corc.uk.net/outcome-experience-measures/childrens-global-assessment-scale-cgas/
CORS	Child Outcome Rating Scale	https://www.corc.uk.net/outcome-experience-measures/outcome-rating-scale-ors-child-outcome-rating-scale-cors/
ORS	Outcome Rating Scale	https://www.corc.uk.net/outcome-experience-measures/outcome-rating-scale-ors-child-outcome-rating-scale-cors/
HoNOSCA (or the wider HoNOS family of measures)	Health of the Nation Outcome Scales Children and Adolescents	https://www.corc.uk.net/outcome-experience-measures/health-of-the-nation-outcome-scales-for-children-and-adolescents-honosca/
BPSES	Brief Parental Self-Efficacy Scale	https://www.corc.uk.net/media/1279/brief-parental-self-efficacy-scale.pdf
SDQ	Strength and Difficulties questionnaire	https://www.sdqinfo.org/a0.html
Current View	Current view questionnaire	https://www.corc.uk.net/outcome-experience-measures/current-view/
SCORE-15	Systemic Clinical Outcome and Routine Evaluation	https://www.corc.uk.net/outcome-experience-measures/systemic-clinical-outcome-and-routine-evaluation-score-15/
CHI-ESQ	Experience of Service Questionnaire	https://www.corc.uk.net/outcome-experience-measures/experience-of-service-questionnaire-esq/
RCADS	Revised Children's Anxiety and Depression Scale	https://www.corc.uk.net/outcome-experience-measures/revised-childrens-anxiety-and-depression-scale-rcads/
MFQ	Mood and Feeling questionnaire	https://www.corc.uk.net/outcome-experience-measures/mood-and-feelings-questionnaire-mfq/
SCARED	Screen for Child Anxiety Related Disorders	https://www.ohsu.edu/sites/default/files/2019-06/SCARED-form-Parent-and-Child-version.pdf
ChOCI	Children's Obsessional Compulsive inventory	https://novopsych.com.au/assessments/diagnosis/obsessional-compulsive-inventory-ocd-child-self-report/
CPSS	Child PTSD Symptom Scale	https://istss.org/clinical-resources/assessing-trauma/child-ptsd-symptom-scale-for-dsm-5-(cpss-5)
CRIES	Child Revised Impact of Events Scale	https://www.corc.uk.net/outcome-experience-measures/child-revised-impact-of-events-scale-cries/

Abbreviation	Name	Link
SLDOM	Sheffield Learning Disabilities Outcome Measure	https://www.corc.uk.net/outcome-experience-measures/sheffield-learning-disabilities-outcome-measure-sldom/
PTQ	Parent Tic Questionnaire	https://fcon_1000.projects.nitrc.org/indi/enhanced/assessments/PTQ.html
SNAP-IV	Swanson, Nolan, and Pelham-IV (26 item or 18 item)	http://www.shared-care.ca/files/Scoring_for_SNAP_IV_Guide_26-item.pdf
SDSC	Sleep Disturbance Scale for Children	https://www.med.upenn.edu/cbti/assets/user-content/documents/Sleep%20Disturbance%20Scale%20for%20Children%20(SDSC).pdf
CSHQ	Children's Sleep Habits Questionnaire	https://njaap.org/wp-content/uploads/2016/04/Childrens-Sleep-Habits-Questionnaire.pdf
Conners	Conners family of questionnaires	https://www.pearsonclinical.co.uk/
RBQ-2	Repetitive behaviours' questionnaire	https://research.ncl.ac.uk/cargo-ne/RBQ.html
SCQ	Social communication questionnaire	https://www.hogrefe.co.uk/shop/social-communication-questionnaire-85199.html
Vineland-3	Parent and teacher forms	https://www.pearsonclinical.com/
PANSS	Positive and Negative Syndrome Scale	https://www.pearsonclinical.co.uk/store/ukassessments/en/Store/Professional-Assessments/Personality-%26-Biopsychosocial/Positive-and-Negative-Syndrome-Scale/p/P100029000.html (purchase required)
CDRS	Connor-Davidson Resilience Scale	http://www.connordavidson-resiliencescale.com/
IIP	Inventory of interpersonal problem	https://www.mindgarden.com/113-inventory-of-interpersonal-problems
YMRS	Young Mania Rating scale	https://dcf.psychiatry.ufl.edu/files/2011/05/Young-Mania-Rating-Scale-Measure-with-background.pdf
Recovery STAR	Outcomes star family	https://www.corc.uk.net/outcome-experience-measures/outcomes-star/
AIMS	Abnormal involuntary movement scale	https://www.ohsu.edu/sites/default/files/2019-10/%28AIMS%29%20Abnormal%20Involuntary%20Movement%20Scale.pdf
Barnes Akathisia	Barnes Akathisia scale	https://simpleandpractical.com/wp-content/uploads/2014/09/Barnes-Akathisia-Rating-Scale-BARS.pdf
Angus	Simpson Angus Extrapyramidal scale	https://medilib.ir/uptodate/show/105378

Licensing by NHS digital

The NCCR tools and measures library lists copyrighted clinical assessment tools and outcome measures for use within health and social care, plus how to access them. Most are available via the Copyright Licensing Service.

Scales relevant to CAMHS

Available tools (November 2018)

- Brief Parental Self-Efficacy Scale (BPSES)
- Child Outcome Rating Scale (CORS)
- Child Session Rating Scale (CSRS)
- Children's Global Assessment Scale (CGAS)
- Comprehensive Assessment of At-Risk Mental States – CAARMS
- Current View
- Experience of Service Questionnaire (ESQ)
- Goals Based Outcome
- Group Session Rating Scale (GSRS)
- HoNOSCA
- HoNOSCA Self
- HoNOSCA-P
- HoNOS-LD
- HoNOS Working Age Adults
- How are things? (ODD-p)
- Me and My School
- Me and My Feelings
- Outcome Rating Scale (ORS)
- SCORE-15
- Session Feedback Questionnaire
- Session Rating Scale (SRS)
- Strengths and Difficulties Questionnaire (SDQ)

Questionnaires that can be submitted to MHSDS

NHS Digital provides tools and guidance to help mental health service providers implement the MHSDS.

The list of questionnaires (whether licensed or not) can be found on pages 84–86 of the **MHSDS version 6.0** (updates occur at intervals).

Eating disorders

There is consensus in the College, through its Eating Disorders Faculty, that outcome data collection needs to be sustainable and clinically meaningful, so that it can be used for benchmarking, service development and research. Many services struggle with consistent data collection owing to insufficient staffing. Where it works better is in those organisations that have a dedicated administrator or psychology assistant support. Dedicated staffing is required for consistent data collection and analysis.

There are significant differences between children and young people's (CYP) and adult eating disorder services, with respect to service configuration, care pathways and process. These differences are reflected in QNIC and QED standards, and it would be helpful to achieve consistency across the age range where possible.

There is general support for including International Classification of Diseases (ICD) diagnoses across the age range, including primary diagnosis and physical and psychiatric comorbidity, as these influence outcomes. Furthermore, some of these comorbidities also reflect outcomes: for example, the presence of osteoporosis or extreme malnutrition is an important outcome for the patient.

Process outcomes, such as waiting times for assessment and treatment, etc., would also be helpful, as there are variations between CYP and adult services, and geographically. Recent investment in CYP ED services has partly addressed geographical differences in this age group, but significant discrepancies in adult services remain.

Adult eating disorders services

The most widely used patient-rated outcome measure is the EDE-Q, which has good psychometric properties, is available free of charge and has also been included in various apps and online data collection systems. Furthermore, EDE-Q also has a child and adolescent version, and hence it can be used across the age range.

The clinical impairment assessment questionnaire (CIA) has been developed to measure the impact of the eating disorder on a patient's general functioning, and it is sensitive to change. However, it does not capture functional impairment due to comorbidity.

The HoNOS (clinician-rated) was not developed with eating disorders in mind; however, the Eating Disorder Faculty of the Royal College of Psychiatrists has developed specific guidance on scoring to support its use for eating disorders. Additional training is needed for reliable scoring, and there is very little research evidence using HoNOS in eating disorders. Only a few providers have found it to be a helpful outcome measure – mainly to record psychiatric comorbidity.

A few services use the CORE-OM as a patient-rated measure to reflect general psychopathology, but there is very little research using CORE-OM in eating disorders. Some services use the CORE-10.

Clinical research often uses various additional measures, given the high level of psychiatric comorbidity. The most commonly used measure is the Beck Depression Inventory (BDI) as depression is common in eating disorders. However, the BDIis not freely available. Adopting the IAPT outcome measures (PHQ and GAD-7) would provide a good alternative and they would also help with benchmarking.

BMI is an important physical outcome measure for underweight patients.

Outcome measures commonly used in adult eating disorder services

		CROM	PROM	Parent/carer
General functioning	Recommended		CIA	N/A
	Optional	HoNOS/GAF/ CGI	CORE-OM CORE-10	
Mental health symptom tracker	Recommend-ed		EDE-Q	N/A
	Optional	EDE Morgan-Rus-sell	PHQ, GAD-7	
Physical health symptom tracker	Recommend-ed	BMI		N/A
Patient experi-ence	Optional		Friends and family tests	Friends and family tests

The most commonly used clinician-rated outcome measure in research is the Morgan-Russell Outcome Measure, and it is free to use. It is not mandatory, but could be retained as an optional clinician-rated measure specific to eating disorder symptomatology. It can be used in adolescents and adults.

Children and young people's community eating disorder services

There are several self-rated and parent-completed measures in this area which are suggested for use in the NHS England 'Access and waiting time standards'. They are outlined in the table below.

Outcome measures proposed for CAMHS eating disorder services

Type	Recommended or optional	Clinician-reported	Child-/young person- reported	Parent-/carer-reported
General functioning	Recommended	CGAS	ORS	
	Optional	HoNOSCA	HoNOSCA-YP SDQ	HoNOSCA-P Parent SDQ
Mental health symptom tracker	Recommended		EDE-Q Other symptom measures as relevant	N/A
Physical health symptom tracker	Recommended	Percentage of median BMI for patients who are underweight	N/A	N/A
Family functioning	Optional	Current view tool	Score 15 – family functioning measure	Score 15 – family functioning measure BPSES
Alliance	Optional	N/A	SFQ	SFQ
Goals	Optional	N/A	GBOs ORS (13+) ORS (6–12)	GBOs
Satisfaction		N/A	PREM: Chi-ESQ	Chi-ESQ

Key:
BPSES = Brief Parental Self-Efficacy Scale; BMI = body mass index; CGAS = Children's Global Assessment Scale; CHI-ESQ = Commission for Health Improvement Experience of Service Questionnaire; EDE-Q =Eating Disorder Examination Questionnaire; GBOs = Goals Based Outcomes; HoNOSCA = Health of the Nation Outcome Scales; MHSDS = Mental Health Services Data Set; ORS = Outcomes Rating Scale; RCADS = Revised Children's Anxiety and Depression Scale; SDQ = Strengths and Difficulties Questionnaire; YP = young person

For underweight patients, height (stunting is a risk in children) is important alongside BMI percentile.

For comorbidity in children, most CYP services adhere to the CORC (Child Outcomes Research Consortium, www.corc.uk.net) suite of outcome measures, which include the Revised Children's Anxiety and Depression Scale (and Subscales) (RCADS) for depression and anxiety.

Summary

The following outcome measures have been most consistently used in eating disorder services, and therefore can be recommended as common outcome measures across the age range:

1. Patient-rated outcome measures:
    - EDE-Q (eating disorder examination questionnaire)
    - CIA (clinical impairment assessment questionnaire)

2. Clinician Reported Outcome Measures:
    - CGAS/GAF

3. Physical health symptom:
    - BMI (BMI percentile for <18-year-olds) is a good measure of severity and outcome for patients who are underweight, as severity of malnutrition is associated with poor outcomes and high morbidity.

4. Patient and carer experience:
    - Chi-ESQ
    - Friends and Family test

Additional outcome measures may be used by individual services. There is a consensus in the faculty that the HoNOS/HoNOSCA is not very helpful in this patient population. If, however, the College recommends these for continuing use in the NHS, it would be important to emphasise the need for additional training and the use of a specialised glossary to capture psychopathology and related abnormal behaviours.

Forensic psychiatry

People who become patients in forensic mental health services share many of the same problems and desired outcomes with patients in all other mental health services. They may differ in that a greater proportion may have complex presentations, very high rates of experienced trauma, and more offending with consequent contact with multiple agencies of care and control beyond health services. They also tend to be in services for a long period of time. It is therefore important to monitor individual outcomes with, as far as possible, similar recognised structured measures of mental state and of risk throughout their time within mental health services. Staff who make and interpret the measures must be fully trained in those measures to ensure accuracy of recording and appropriate interpretation.

Outcome measures need to address therapeutic progress as well as risk reduction. However, a recent review found that despite the large number of instruments potentially available, mostly related to the assessment of risk, evidence to support their use as outcome measures in forensic mental health services is limited (Ryland *et al.*, 2021d). The authors recommend that future research and instrument development must involve patients and carers, alongside clinicians, to ensure adequate content validity, so that the outcomes measured are those most valued by the key stakeholders (Terwee *et al.*, 2018).

As reflected in the introduction to this report, while recovery is possible in some cases, the severity and chronicity of some patients who are engaged with forensic psychiatry means that a stabilisation and limiting of deterioration may be the optimum outcome. It is important to set realistic outcomes for each individual and, if necessary, review and revise the initial target outcomes as treatment progresses. However, in recognising the potential limitations and that full recovery is unlikely for some patients, it is important that this does not become the default for all patients, resulting in therapeutic nihilism.

The Forensic Faculty agrees with the principles outlined by RCPsych in relation to the development of outcome measures and highlight the need for outcomes to be important for both patients and clinicians, and for the measures to be culturally appropriate, valid, reliable and clinically useful. For patients in forensic mental health services, outcomes reported by significant others should be given appropriate weight depending on the individual circumstances and ongoing relationship with the patient, given that interpersonal dissonance and violence has commonly been a precursor to admission to services. Attempts should be made wherever feasible to involve carers in measuring outcomes. Forensic patients tend to have relatively long in-patient admissions where discharge for restricted patients may only be achieved with relevant ministerial or tribunal approval depending on jurisdiction.

This means that some outcomes for patients within forensic services are therefore beyond the direct control of those services. Once patients are discharged from forensic services, many agencies are likely to be involved with that individual's ongoing care. Forensic services should endeavour to set people up for the best possible results on discharge, however many other factors will inevitably contribute to a former forensic patient's outcomes in the community.

Consultation beyond mental health professionals on outcome measures is particularly important for forensic psychiatry, where work with other agencies is routine and safety of others as important as that of the patient. Further research is needed to develop consensus between agencies in this area. Outcome measures related to the use of coercion, as well as levels and types of security, are also relevant areas to include as these may be useful measures of a lack of onward care package (including social care package).

Research in the United Kingdom and internationally involving patients, carers and clinicians has identified important areas in which to measure outcomes (Ryland *et al.*, 2021b, Wallang *et al.*, 2018, Livingston, 2018, Morrissey *et al.*, 2017). Patients should be able to lead lives that are meaningful to them, have hope for the future and enjoy a good quality of life, with positive experiences of their care. Risk to self and others should be managed in a sustainable way, so that patients feel safe, others feel safe around them and negative events, such as self-harm, suicide, acts of aggression and violent offending are reduced or eliminated. Improving mental and physical health outcomes is also essential, with reduction of symptoms, side effects, morbidity and premature mortality all being important goals. The development of life skills is highly valued, and this can range from simple occupational activities within a ward setting, to paid employment and creative achievement in the community, depending on the needs, desires and aptitudes of each patient. Finally, progress towards greater independence is a commonly cited aspiration, which is again specific to individual context.

Outcome measures are important in forensic mental health services to:
1. measure individual patients' progress, identify unmet needs and set the pathway
2. know whether a service is delivering as designed, guide service improvement and support research.

Instruments should be simple to complete, meaningful for patient care and recognised as useful. It needs also to be recognised that planned (e.g. service redesign) or co-incidental changes (e.g. as a result of the COVID-19 pandemic) are likely to affect outcomes so measurements before and after the event may be useful to understand the effects of that event or change. It should also be recognised that some outcomes may be difficult to measure directly so some process measures may need to be used as proxies to quantify the change in outcome. In this case it is important to be clear that these are process measures, and clearly articulate how these are thought to relate to the outcome of interest. Some outcome measures are specific to forensic psychiatry while others, for example related to health and illness, will be measures used in mental and physical health services. A mixture of both types is likely to be needed to obtain a comprehensive understanding of relevant outcomes.

Service-level outcomes

The following are some suggested parameters that could be measured to assess outcome. It is important to recognise that most outcomes following discharge from forensic mental health services will be subject to factors beyond the control of those services. Patient and carer satisfaction with services is important to quantify, however measures need to be selected carefully, given that few would choose to be detained

in forensic mental health services (comparison-based tests, such as the Friends and Family test may therefore be inappropriate). Measures of ward atmosphere, such as the EssenCES, may be useful proxies to gauge the quality of the therapeutic milieu (Schalast *et al.*, 2008).

Whilst in forensic inpatient services:

- No. of deaths
- No. of serious untoward events (SUI)
- No. of patients in out of area placements (OAPs)
- No. of seclusions
- No. of risk events (as reported by systems such as DATIX)
- No. and type of complaints
- No. of unplanned admissions
- No. of patients involved in occupation, education, leisure within the unit
- Patient and carer satisfaction with services
- EssenCES

Following discharge:

- Degree of recovery from mental illness
- No. of readmissions within a specified period
- No of patients who re-offend and the nature of that reoffending
- No of patients in employment, education or engaged in regular voluntary work
- Quality of life in the community
- Life expectancy and physical morbidity

Patient-level outcomes

Patient-level outcomes should be measured at regular intervals throughout the patient pathway. Outcomes should ideally be rated by both patients and clinicians, although the domains reported by each are not necessarily identical (patients will be uniquely placed to evaluate their own quality of life, for example). Some patients may choose not to participate in the process of rating outcomes, but clinicians should try to facilitate their involvement wherever possible.

While no dedicated carer-reported scales have been identified, carers should still be involved in evaluating outcomes, for example by participating in care planning meetings where appropriate. Patient and clinician reported outcome measures can potentially be used at an aggregate level to gauge the operation of services, however this should be done only with caution as linking incentives, such as funding, may alter the way that respondents complete questionnaires (Ryland *et al.*, 2021a).

Useful forensic specific outcome measure include:

- Health of the Nation Outcome Scale Secure Version (HONOS-Secure)
- Camberwell Assessment of Need Forensic Version (CANFOR)
- Forensic Outcome Measure (FORUM)

HONOS secure, though primarily designed as an outcome measure was not identified as particularly useful by clinicians (Ryland *et al.*, 2021c), however it is currently widely used as it is mandated by some commissioners . It only has a clinician-rated scale and does not have a patient equivalent. The CANFOR is a well-established, extensively validated instrument looking at need in a forensic population developed with extensive stakeholder input (Thomas *et al.*, 2008). It is both patient and clinician reported and has several versions, including long, short and research. FORUM is a more recently developed, validated instrument to measure individual patient outcomes. It has patient and clinician scales based on empirical research with stakeholders, which are designed to be used together to support patients' care planning and treatment (Ryland *et al.*, 2021b).

The Dangerousness, Understanding, Recovery and Urgency Manual (DUNDRUM) quartet is designed to assess specialist dependency needs, helpful to audit the appropriateness of placements, and is not designed to be an outcome measure. The DUNDRUM 3 and 4 scales concern programme completion and recovery, and patient-reported versions were subsequently developed, with limited patient involvement, to mirror the clinician versions (O'Dwyer *et al.*, 2011, Davoren *et al.*, 2015).

Illness-specific

Patients with a mental disorder, especially those with a severe and chronic mental illness, have significant physical morbidity and mortality and, therefore, outcome measures looking at physical health are as relevant and important as those related to mental health.

There are a number of well-established instruments to assess and measure change and outcome of specific physical and mental illnesses. Instruments should be selected according to the health needs of each patient and used to measure progress and change in that individual. Again, instruments that are both clinician and patient reported should be used wherever possible.

Risk and offending

Risk, especially to others, is afforded particular prominence by forensic mental health services, given the profiles of patients using these services, including high rates of previous violent offending. Numerous risk assessment instruments exist to help predict the likelihood of negative outcomes such as violence and criminal recidivism. There are a large number of instruments that assess risk in forensic mental health services, which vary in their format, length, purpose and evidence base (Ramesh *et al.*, 2018). Risk assessment instruments with dynamic components may be inappropriately used in practice to retrospectively measure if a risk has reduced over time.

Similarly, measures of risk outcome may be inappropriately used as predictive measures to plan an individual's care pathway. It is important to be clear what the purpose of measurement is and select an appropriate instrument designed and validated for the intended use (De Vet *et al.*, 2011).

The Historical-Clinical-Risk Management-20 (HCR-20) is the most widely researched and used structured professional judgement instrument employed by clinicians in forensic mental health services. It is designed to help develop formulations and plan scenarios in order to understand and manage risk (Douglas *et al.*, 2014). Only the clinical and risk scales are dynamic and the authors emphasise the need for users to focus on the scenario planning, rather than the scoring of particular items. Caution should therefore be used before utilising the HCR-20 as a risk outcome measure. Other popular dynamic or partially dynamic risk assessments, which have been well validated for the purpose of prediction, such as the Short-Term Assessment of Risk and Treatability (START) (Webster *et al.*, 2006) and Structured Assessment of Protective Factors for violence risk (SAPROF) (de Vogel *et al.*, 2011) could also potentially be repurposed as outcome measures, however further research is needed for this to happen and caution is necessary in interpreting scores (Ryland *et al.*, 2021d).

Generic measures

Other well-validated instruments developed for use in general psychiatric contexts may be helpful in forensic mental health services to measure parameters such as quality of life and satisfaction with care. For example, the Recovering Quality of Life (ReQoL) instrument has been developed to assess the quality of life for people with different mental health conditions (Keetharuth *et al.*, 2018). DIALOG+ is an outcome measure and patient–clinician communication tool designed for community mental health care which may be useful for discharged forensic patients (Priebe *et al.*, 2015).

Conclusion

Outcome measurement in forensic mental health services is important. The outcomes measured must be those that are important for key stakeholders including patients, their carers, clinicians and the public. Valued outcomes include better physical and mental health, reduced risk and criminal justice involvement, building life skills, achieving greater independence and a sustained improvement in quality of life. Outcomes should be routinely measured at both the individual patient level and service level, using a mixture of administrative data and validated instruments reported by patient and clinicians.

References

Davoren M, Hennessy S, Conway C, Marrinan S, Gill P and Kennedy, HG (2015) Recovery and concordance in a secure forensic psychiatry hospital - the self rated DUNDRUM-3 programme completion and DUNDRUM-4 recovery scales. *BMC Psychiatry*, **15**

De Vet, H C, Terwee CB, Mokkink LB and Knol DL (2011) Measurement in medicine: a practical guide, Cambridge University Press.

De Vogel V, De Vries Robbé M, De Ruiter C And Bouman, Yha (2011) Assessing Protective Factors in Forensic Psychiatric Practice: Introducing the SAPROF. International Journal of Forensic Mental Health, 10: 171–77.

Douglas KS, Hart Sd, Webster CD, Belfrage H, Guy LS And Wilson CM (2014) Historical-Clinical-Risk Management-20, Version 3 (HCR-20V3): Development and Overview. International Journal of Forensic Mental Health, 13: 93–108.

Keetharuth AD, Brazier J, Connell J, Bjorner JB, Carlton J, Taylor Buck E, Ricketts T, Mckendrick K, Browne J, Croudace T And Barkham M (2018) Recovering Quality of Life (ReQoL): a new generic self-reported outcome measure for use with people experiencing mental health difficulties. Br J Psychiatry, 212: 42–49.

Livingston JD (2018) What Does Success Look Like in the Forensic Mental Health System? Perspectives of Service Users and Service Providers. International Journal of Offender Therapy and Comparative Criminology, 62: 208–28.

Morrissey C, Geach N, Alexander R, Chester V, Devapriam J, Duggan C, Langdon Pe, Lindsay B, Mccarthy J And Walker D (2017) Researching outcomes from forensic services for people with intellectual or developmental disabilities: a systematic review, evidence synthesis and expert and patient/carer consultation. Health Services and Delivery Research, 5.

O'Dwyer S, Davoren M, Abidin Z, Doyle E, Mcdonnell K And Kennedy HG (2011) The DUNDRUM Quartet: validation of structured professional judgement instruments DUNDRUM-3 assessment of programme completion and DUNDRUM-4 assessment of recovery in forensic mental health services. BMC Research Notes, 4: 229.

Priebe S, Kelley L, Omer S, Golden E, Walsh S, Khanom H, Kingdon D, Rutterford C, Mccrone P and Mccabe R (2015) The Effectiveness of a Patient-Centred Assessment with a Solution-Focused Approach (DIALOG+) for Patients with Psychosis: A Pragmatic Cluster-Randomised Controlled Trial in Community Care. Psychother Psychosom, 84: 304–13.

Ramesh T, Igoumenou A, Vazquez Montes M and FazeL S (2018) Use of risk assessment instruments to predict violence in forensic psychiatric hospitals: a systematic review and meta-analysis. European Psychiatry, 52: 47–53.

Ryland H, Carlile J and Kingdon D (2021a) A guide to outcome measurement in psychiatry. BJPsych Advances, 27, 263-271.

Ryland H, Cook J, Ferris R, Markham S, Sales C, Fitzpatrick R and Fazel S (2021b) Development of the FORUM: a new patient and clinician reported outcome measure for forensic mental health services. Psychology, Crime & Law, 1–18.

Ryland H, Cook J, Fitzpatrick R and Fazel S (2021c) Ten outcome measures in forensic mental health: A survey of clinician views on comprehensiveness, ease of use and relevance. Criminal Behaviour & Mental Health.

Ryland H, Cook J, Yukhnenko D, Fitzpatrick R and Fazel S (2021d) Outcome measures in forensic mental health services: A systematic review of instruments and qualitative evidence synthesis. European Psychiatry, 64: e37.

Schalast N, Redies M, Collins M, Stacey J and Howells K (2008) EssenCES, a short questionnaire for assessing the social climate of forensic psychiatric wards. Crim Behav Ment Health, 18, 49–58.

Terwee CB, Prinsen CAC, Chiarotto A, Westerman Mj, Patrick Dl, Alonso J, Bouter LM, De Vet HCW and Mokkink LB (2018) COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. Qual Life Res, 27, 1159-1170.

Thomas SD, Slade M, Mccrone P, Harty Ma, Parrott J, Thornicroft, G and Leese M (2008) The reliability and validity of the forensic Camberwell Assessment of Need (CANFOR): a needs assessment for forensic mental health service users. International Journal of Methods in Psychiatric Research, 17, 111–20.

Wallang P, Kamath S, Parshall A, Saridar T and Shah M (2018) Implementation of outcomes-driven and value-based mental health care in the UK. Br J Hosp Med (Lond), 79: 322–27.

Webster Cd, Nicholls Tl, Martin Ml, Desmarais Sl And Brink J (2006) Short-Term Assessment of Risk and Treatability (START): The case for a new structured professional judgment scheme. Behav Sci Law, 24: 747–66.

General adult psychiatry

Streams of outcome measurement development

Patient-informed measures

Of service:
Friends and family test

Of personal change:
For routine measurement we recommend choosing one out of the following five measures:

- Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS / SWEMWBS)
- Euro Quality of Life Questionnaire (EQ-5D)
- DIALOG
- Process of recovery questionnaire (QPR)
- Recovering quality of life (REQoL)

Relative/carer informed measures

Of service:
Friends and family test

Of personal change:
None recommended

Staff-informed measures

Of service:
None recommended

Of personal change
- Clinical Global Impression – severity scale (CGI-S) at initial contact
- Clinical Global Impression – improvement scale (CGI-I) subsequently
- Global Assessment of Functioning (GAF)
- Health of the Nation Outcome Scale (HoNOS)

Research-informed measures

Of service:
None recommended

Of personal change:
None recommended

Any requirement for a minimum dataset

Of service:
Type of contact:
- Initial contact/follow-up contact
- Community setting/Hospital in-patient setting
- Patient's home/Clinical base/other setting

Ratings interval:
- At least twice in an episode
- At least every six months

Of personal change:
Demographics:
- Basic demographics including age, sex, ethnicity and disability status along with recording of deaths in people under the care of mental health services.

Diagnostic information:
- We recommend that all patients should have a working diagnosis. However, we recognise that in many general adult psychiatric settings, formal standardised assessment within a service contact episode by a psychiatrist is neither realistic nor desirable. Further, diagnosis is a process which is longitudinal and relies on observation and collateral data collection which may not be available at any one cross-sectional time point. Fuller assessment over time can often reveal comorbid conditions, such as neurodevelopmental disorders and comorbid diagnoses which have an impact on prognosis, so diagnosis is inevitably a dynamic process. We propose using a simple grid as shown below to summarise diagnostic information. Working diagnoses should be made by adhering to current ICD criteria and not on the basis of 'gut feeling' or a clinician's view distorting patients' reported information.

  The grid should be simple and record who is making both diagnosis and grid statements.

Suggested simple grid for summarising diagnostic information				
	Definite	Probable	Possible	Insufficient information
Psychotic disorder				
Mood or anxiety disorder				
Neurodevelopmental disorder				
Personality disorder				
Substance misuse disorder				
Other: specify				

In addition, it is important to record the time aspects of the disorder as these can also influence the type of intervention and prognosis.

Other episode and illness information:

- Episodic (i.e. with wellness for >6 months) Yes/no
- Date of onset of symptoms
- Date of first diagnosis
- Start of current episode (from wellness)

Diagnosis

A working diagnosis should be made at initial assessment by the assessing clinician and adjusted as clinically indicated at subsequent reviews. We expect any registered mental health professional to be able to make a working diagnosis, but more junior staff should routinely seek support from senior staff, including doctors. In many cases, but not always, a diagnostic assessment from a psychiatrist will be needed during the episode of care.

HoNOS, for all its shortcomings, captures key aspects germane to diagnosis and management plans. However, there are important diagnostic aspects that it ignores – for these, systematic enquiry is an important tool in achieving a greater predictive validity in terms of outcomes; notable omissions (not a comprehensive list) include the historical presence of hypomania, anxiety and neurodevelopmental disorders.

Further, particular care needs to be exercised in enquiring about symptoms that change a treatment plan but may not be obvious at first and not volunteered by service users, e.g. the presence of mood-congruent delusions (reference, guilt, punishment) and derogatory hallucinations in major depressive disorders.

CGI-S and CGI-I

These scales have been used for about 40 years and, although criticised because they are more subjective than others, they correlate well with all the more specific symptom-based scales, include an appreciation of functional as well as clinical change, and when pooled

across large numbers of assessments are a more robust and sensitive index of change than symptom-based disease-specific scales. They are very simple to complete as there are only two questions where one out of seven possible responses are ticked.

GAF

The Global Assessment of Functioning (GAF) is a numeric scale (1– 100) used to rate globally the social, occupational, and psychological functioning of adults. It has been in use for over 20 years but is no longer included in the current version of DSM (DSM-V). However, it is easy to score as a number between 0 and 100 is chosen that fits in best with the person's current state. While WHO DAS now takes the place of GAF in DSM-V, GAF remains preferable as a clinician instrument because it is quick. WHO DAS is gaining traction and there are patient and carer versions that are likely to be preferable in the future.

HoNOS

HoNOS was developed over 20 years ago by the Royal College of Psychiatrists with the specific aim to capture what is important in mental health care delivery in the UK's NHS. Most trusts already collect data using it and its use is evaluated by comparing it with other instruments by The UK Routine Clinical Outcomes in Mental Health Group. HoNOS is currently collected through the use of the Mental Health Clustering Tool (MHCT) which has 6 additional questions.

The MHCT is not an outcome measure as such, but has been mandated for use for a proposed 'payment by results' system since 2011. This has had a dramatic effect in enhancing the initial recording of CROMs but with serious concerns about the use of professional time and quality assurance. In particular, we are concerned that HoNOS should not be used as a triage service threshold. It was designed as an outcome measure, not as a triage tool.

WEMWBS and SWEMWBS

These are quick-to-use measures – whether using the 14-item scale or the shorter 7-item version. They record psychological wellbeing rather than overall wellbeing. However, psychological wellbeing is a hallmark of mental health and, as such, it is a measure which is relevant to successful management of conditions from the patient's perspective and, therefore, preferable to scales that are exclusively based on symptom measures.

EQ-5D

The EQ-5D comprises the following five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. The patient is asked to indicate their health state for each dimension on one of 3 levels: no problems, some problems, and extreme problems. It has been adopted in the UK for routine outcome measurement and is preferred by NICE to calculate quality-adjusted life-years (QALYs) for use in cost-effectiveness analyses. It has been widely used in physical healthcare and the advantage of using it in mental healthcare would be to allow comparisons of quality-of-life improvement to be made between physical health and mental health interventions.

However, although it has been shown that the EQ-5D is valid and responsive for depression, the results in other psychiatric disorders have been less convincing.

DIALOG

This is a service user-rated outcome measure, which focuses on quality of life, care needs and treatment satisfaction with 11 items rated on a seven-point scale. It has been recommended for use in NHS early intervention in psychosis services by NICE as part of the access and waiting time standard guidance.

QPR

This measure was developed in collaboration with service users and asks about key aspects of personal recovery including connectedness, hope, identity, meaning to life, and empowerment. There are 11 items to be rated on a five-point scale. It has also been recommended for use in NHS early intervention in psychosis services by NICE as part of the access and waiting time standard guidance.

ReQoL

This is a newer measure which has been designed specifically for use in mental health populations. There are two versions, a 10-item and a 20-item questionnaire both rated on a five-point scale. Recent findings have shown it to be a more sensitive and responsive measure than the EQ-5D.

Intellectual disability

Introduction and background

Measuring outcomes is essential to improving quality of services. While the public health approach to healthcare measures have evolved around both process and outcomes monitoring, it is the general consensus that some of the important properties for patient-based outcome measures are appropriateness, reliability, validity, responsiveness, precision, interpretability, acceptability and feasibility (Thornicroft and Tansella, 2013). They are also used as evidence of the contribution of health services in providing quality assurance, and value for money. While the aforementioned seven criteria are key to examining outcomes, it is precision, generalisability and appropriateness that are the most important criterion for such patient-related outcome measures to be effective (Fitzpatrick et al., 1998). These are also of significant relevance to the field of Intellectual Disabilities (ID).

Porter (2010) proposed an outcome measure hierarchy when determining the group of relevant outcomes for any condition or patient population. He stated that for any condition or population, multiple outcomes collectively define success and proposed a three-tiered hierarchy, in which the top tier is generally the most important and lower-tier outcomes involved a progression of results contingent on success at the higher tiers.

- **Tier 1:** Health status achieved or retained
- **Tier 2:** Process of recovery
- **Tier 3:** Sustainability of health

Tier 1: Health status achieved or retained

The top tier measures survival and the degree of health or recovery achieved or retained at the peak or steady state. This is achieved by clinician-rated outcome measures and includes measures for functional assessment. To measure survival, services could measure outcomes in terms of mortality data, including suicide data. Learning from Lives and Deaths (LeDeR) provides a good level of information regarding potential contributory factors and actions that would need to be taken to improve outcomes.

The second level of Tier 1 is the degree of health or improvement achieved or retained. There are a number of measures that could be used..One option is the Health of Nation Outcome Scale – Learning Disability (HoNOS-LD), which measures the overall outcome for patients and for overall evaluation of services, examining overall health of an individual rather than focusing on one set of symptoms. It could be used with people with an intellectual disability with mental health needs irrespective of the degree of their disability (Roy et al., 2002). It has been tested for inter-rater reliability, convergent reliability, validity to change and acceptability in a national pilot study.

Tier 2: Process of recovery

The two levels at Tier 2 are:

- Time required to achieve recovery and return to normal or best attainable function.
- The disutility (adverse effects) of the care or treatment process in terms of discomfort, retreatment, short-term complications, and errors and their consequences.

It is important to acknowledge that while process measurements are key to measuring health system performance, there is wide variability in these. There are no standard models of care, and specifications as well as the needs of services are governed by the policy context or aspirations laid down by NHS frameworks. When measuring the time required to achieve recovery, teams should take into account various figures including waiting times for first contact, time to start of treatment and overall length of time to achieving desired outcome, etc).

For measuring disutility of the care or treatment, there are a number of measures that could be used depending on the type of service provision. For services focusing on challenging behaviours, using Periodic Service Review (PSR) recommended by the British Institute of Learning Disability would be a one approach. Organisational measures like the number of Serious Untoward Incidents (SUI) and number of DATIX entries are proxy measures of quality. Unplanned admissions, delayed discharges, delayed transfers of care, complaints and disengagement from services are useful ways of measuring overall outcomes at this level. It may be noted that some of this level is a heterogenous mix of service level parameters, outputs and blended process derivatives, a large number of which has not been validated and needs careful consideration, despite being widely accepted as useful.

Tier 3: Sustainability of health

This tier includes measures for sustainability of health or recovery and nature of recurrence as well as long-term consequences of therapeutic interventions. This tier should include measures for relapse rates, re-referrals, placement breakdowns, re-admissions etc.

Quality of life measures would also give a good indication of outcomes at this tier. The Maslow Assessment of Needs Scale (MANS-LD) represents a value-driven approach to assessing outcomes of services for people with intellectual disabilities. Another tool to consider would be EQ-5D developed by EuroQol Group as a measure of health-related quality of life that can be used in a wide range of health conditions and treatments; it measures outcomes on five dimensions, namely mobility, self-care, usual activities, pain/discomfort and anxiety/depression.

Scope of this guidance

The scope of the this guidance is to examine the key characteristics of the validated outcome measures which are validated and adapted in intellectual disability population. While the disease or individual health condition verses the systematic tools for measurements have their own specific utility, we have attempted to examine them together in this guidance.

Research needs

It is important to emphasise that the focus of such validity measures has conventionally rested squarely with either patient-related outcomes measures (PROMs) or clinician-reported outcome measures (CROMs), where certainly most of the research is currently located. Within intellectual disability, there is a pressing need to develop adapted measures that capture the essence of a validated tool design. These measures need to be both emancipatory and appropriate to use in the ID population. Some of the important domains to consider are: Patient Reported Experience Measures (PREMs), Carer Reported Outcome Measures (CAROMs) and Carer Reported Experience Measures (CAREM). While there are limitations of using PREMs across the spectrum of intellectual disability, such as limitations of use across moderate to severe/profound disability, CAROM has the potential to improve service-related outcome measures by making them more participatory and needs based.

The Clinical Global Impression scale (CGI) emerged as the consensus choice of PROM scale (clinician rated as well as carer rated). While not validated across the ID population, it is established as a baseline for overall clinical research.

CGI is a brief, stand-alone assessment of the clinician's view of the patient's global functioning both before and after intervention. It provides an overall clinician-determined summary measure that considers all available information, including knowledge of the patient's history, psychosocial circumstances, symptoms, behaviour, and the impact of the symptoms on the patient's ability to function.

PREMs, while not robustly validated, are considered to be practical tools. The PREM-LD, for example, is a one-item questionnaire that asks the patient to answer the question: "How do you feel about your treatment in this service? This retains simplicity while capturing experiential reality. The essence of these tools is to capture experiential measures across intellectual disability groups of patients, regardless of level of disability. A similar tool, the CAREM-LD, is proposed to capture carer-reported experiential measures.

Measuring utility

Measuring utility is important in outcomes research because it allows for a more comprehensive evaluation of the value of healthcare interventions. Utility measures capture the impact of an intervention on a patient's quality of life, considering both the length and the quality of life. Utility measures can be used to compare the effectiveness of different interventions and to inform decisions about resource allocation.

One commonly used utility measure is the Quality Adjusted Life Year (QALY) which is designed to measure the healthy years lived by an individual. It is a standard measure of disease burden that jointly looks at quality of life (morbidity) and survival (mortality); QALY is one of the health indicators that looks at both quantity and quality of life. Because of this feature, QALY provides an estimation of the amount of quality time experienced by an individual due to any health interventions/programs. A score of 1 represents perfect health while 0 represents poor health status of an individual. It is calculated by multiplying the time spent in a particular health state by a utility score that reflects the patient's

preference for that health state. QALYs provide a way to compare the effectiveness of different interventions in terms of the amount and quality of life gained.

Another important utility measure is the Disability Adjusted Life Year (DALY), which measures the burden of disease by combining the years of life lost due to premature death with the years of life lived with disability. The DALY is a useful measure for comparing the impact of different diseases and conditions on population health. While DALY has significant challenges, including allocation of preferential value-added weights, this has provided indices for comparison for landmark studies such as global estimates of burden of diseases.

Measuring utility is essential for outcomes research because it provides a way to assess the effectiveness of healthcare interventions in terms of their impact on patients' quality of life. Utility measures can help healthcare providers, policymakers, and researchers make informed decisions about resource allocation and treatment choices based on the outcomes that matter most to patients.

The most widely accepted utility measure for allocation decision is the use of QALY, the data for which can be generated from EQ-5D. However, currently, there is rare application of EQ5D data which is beginning to make headway. Raczka, Theodore and Williams (2020) have used EQ5D in part. Currently, lack of such data in intellectual disability implies that we are limited to using cost–impact, budget–impact and cost–consequence analysis and hence limiting a rational voice in resource allocation decision.

Short Form Survey (SF-36)

Another useful global measure is the Short Form Survey SF-36 which is a universally accepted measure and has been validated in the intellectual disability population with some adaptations. The SF-36 shows good internal consistency $\kappa = 0.92$ and a moderate inter-rater reliability of 0.63 (Jones, Dagnan and Ruddick, 1997). There is, however, limited use of SF-36 data, which has limited further utility values reports.

Policy context

The national service model (NHS England 2015) sets out a co-produced list of the expectations of people with intellectual disabilities and sets the service priorities for commissioners and providers.

- People should have a good and meaningful everyday life.
- Care and support should be person-centred, planned, proactive and coordinated.
- People should have choice and control over how their health and care needs are met.
- People should be supported to live in the community with support from and for their families and carers.
- People should have a choice of where and with whom they live and with a choice of housing.
- People should get good care and support from mainstream health services.

- People should be able to access specialist health and social care support in the community.
- People should be able to get support to stay out of trouble.
- When their health needs cannot be met in the community, people should be able to access high-quality assessment and treatment in hospital.

Within this chapter, we have attempted to present validated outcomes measures in intellectual disability psychiatry that are patient-reported or system-wide. Below, we aim to map the evidence in key individual domains.

Challenging behaviours

Challenging behaviours are one of the core presentations of intellectual disability and continues to be of key importance to this population. There is wide variability in measurement, recording and interpretation of severity in this context. This may include a range of behaviours which often are contextual and may not always be rated or measured by any specific outcome measures.

The British Association of Psychology commissioned a study of appropriate measures and validated results for people in the UK with intellectual disability (Morris, Bush and Joyce, 2014). The report focussed on three key domains of measurement: Generic measures, measures of frequency and impact, and measure of quality of life. Following the results of validation, the following key outcome measures emerged:

Generic measures:

- HoNOS-LD (Health of Nations Outcome Scales – Learning Disability (Roy *et al.*, 2002)

Frequency intensity and impact of behaviours:

- Challenging Behaviour Interview (Oliver *et al.*, 2003)
- Behaviour Problems Inventory (Rojahn *et al.*, 2012)

Measures of quality of life:

- Life Experiences Checklist (Ager, 2002)
- Maslow Assessment of Needs Scales (MANS) for Learning Disability (Skirrow and Perry, 2009)

Difference between pre and post scores on each measure
(Morris, Bush and Joyce, 2014)

HoNOS-LD	CBI	BPI-01	MANS-LD
P=0.0013	P=0.3371	P=0.0002	P = 0.0121

The following is a brief description of the tools considered in the pilot.

HoNOS-LD

Well established for use with people with a learning disability and mental health needs, regardless of the degree of their disability, The HoNOS-LD is a clinician-rated scale, which grades severity on a 5-point scale. The scale has been well validated in populations and reports moderate to good interrater reliability with κ values of (0.56–0.86). (Roy et al., 2002). HoNOS-LD is a generic health tool and not an individual health-state tool.

Challenging Behaviour Interview (CBI)

A 14-item scale which assesses the severity of challenging behaviour in children and adults with intellectual disability, relating to self-injury verbal aggression, inappropriate vocalisation and disruption to environment. The validated tool reports excellent κ values for Part 1 (0.7–0.9) and moderate for Part 2 (0.6–0.8) (Oliver et al., 2003).

Behaviour Problems Inventory (BPI)

A 52-item respondent-based behaviour rating instrument for self-injurious, stereotypic, and aggressive/destructive behaviour. Items are rated on a frequency and severity scale. It has good consistency and excellent inter-rater reliability (r = 0.76) with good test-retest reliability (r = 0.76). It is useful for challenging behaviour at risk as an outcome for intervention studies. However, this may not be used for overall assessment of challenging behaviours.

Maslow Assessment of Needs Scales (MANS) for Learning Disability (Skirrow and Perry, 2009)

A value-driven scale, quantified and adapted to the needs of intellectual disability subjects. It uses higher categories of needs as actualised in Maslow's hierarchy of needs. The scale's 19-item questionnaire is designed for use with a five-point scale, with symbols to help the person decide on their response. Questions are focussed on mapping satisfaction to well-known needs from Maslow's hierarchy, which range from basic to higher-level needs. The MANS-LD is supplemented by an eight-item questionnaire adapted from the World Health Organization Quality of Life WHO-QOL assessment (which is not adapted to ID specific population). The validation results show moderate to good validity and reliability data.

An adaptation of **Maslow Assessment of Needs Scales (MANS) for Learning Disability, Mini-MANS-LD** has been developed by Raczka, Theodore and Williams (2020) which reports acceptable psychometric properties, including moderate congruent validity and acceptable internal consistency (α =0.74).

Interestingly, the authors mapped relationship between Mini-MANS-LD and EQ 5D-Y (which they used in conjunction) and report that it was significantly correlated with one health state. This is a significant development in quality-of-life measurement scales in intellectual disability subjects, and once expanded and validated, further paves the way for mapping the quality-of-life data which would hopefully in future allow for calculations of much more precise ICER (incremental cost effectiveness ratios) and hence aid in robust decision making.

Following this is a compilation of other tools and outcome measures which are useful in measuring challenging behaviours in various context and are validated in the ID population.

Modified Overt Aggression Scale

In addition, the **Modified Overt Aggression Scale** (MOAS) (Ratey and Gutheil, 1991) has been shown to have excellent results in specific trials (κ= 0.65–0.9). The informant-rated tool considers severity of verbal aggression, aggression towards property, self-harm, and physical aggression over a week period. However, there are some limitations owing to the implicit nature of ratings which compromise reliability, and we have not been able to find more recent validations from this measure.

Diagnostic Assessment for the Severely Handicapped (DASH-II)

DASH-II measures mental ill health and challenging behaviours in severely intellectually disabled people. It is informant based and has 84 items rated on a 3-point Likert. The scale shows good test–retest reliability of (0.8-0.9), but poor to moderate internal consistency (α =0.53–0.84). This is considered a good measure for screening. It is important to note that Raitasuo, Taiminen and Salokangas (1999) have used DASH and BPRS (Brief Psychiatric Rating Scale) in a Finnish study. While BPRS has been used more as a proxy measure to measure improvement to intervention, the study establishes the potential of BPRS to be used in intellectually disabled subjects. They reported DASH-II to be non-discriminatory, but also attributed this to a minimal degree of potential change.

Aberrant Behaviour Checklist (ABC)

The Aberrant behaviour checklist ABC (Aman *et al.,* 1985) is one of the most widely used instruments to assess challenging behaviour among children and adults with intellectual disability. This 58-item checklist has five subscales for agitation, lethargy, social withdrawal, stereotypies, hyperactivity or noncompliance, and inappropriate speech. ABC has been found to have good to excellent internal consistency (Cronbach α = 0.8–0.9) and moderate inter-rater reliability. In monitoring changes, ABC has been reported to be better than DASH-II.

Dementia

Assessment of dementia is one of the challenging areas in intellectual disability. There are some scales specifically designed and dedicated to diagnosis of dementia. The following is a compiled list of key validated measures which have an evidence base for use in ID populations.

The Dementia for Learning Disability (DLD) scale was developed in the 1980s in The Netherlands, and its intended use was for adult with intellectual disability. Since then, it has been used widely in Europe and in the UK, both in clinical practice as well as in research (Strydom and Hassiotis, 2003). This is completed by a family member or carer who knows the person well. It has 50 items giving two main scores: cognitive scores (SCS) and social scores (SSC). Even though Evenhuis (1996) reported that the DMR has a sensitivity of 100%, other studies have shown that the DLD has lower sensitivity and specificity (0.61/0.63) despite its wide usage.

Temporal stability of DLD for dementia diagnosis: DLD test–retest reliability was moderate overall but was stronger for the cognitive subscale than the social subscale (Koehl *et al.*, 2020). When comparing the Rapid Assessment of Dementia in Developmental Disabilities (RADD)' sensitivity to dementia in Down Syndrome, Walsh *et al* (2015) report that RADD exhibited high sensitivity (0.87) and specificity (0.81) in discriminating among individuals with and without dementia. Results of the Severe Impairment battery (SIB) test showed (ROC) curves for groups varying in pre-morbid severity of ID, and the RADD exhibited high sensitivity (0.87) and specificity (0.81) in discriminating among individuals with and without dementia. It should be noted that SIB and Brief praxis tests are not intellectual disability specific.

Dementia Screening Questionnaire for Individuals with Intellectual Disability (DSQIID)

DSQIID is 53-item dementia screening questionnaire for people with ID. The scale reported internal consistency α =0.91 inter-rater consistency of 0.9 and test–retest validity of 0.95 (O'Caoimh, 2013; Deb et al., 2007). The scale has a fixed cut off score which limits its applicability in people with more severe forms of intellectual disability.

Modified Cambridge Cognitive Examination with Down's Syndrome (CAMCOG-DS)

This modified version of a Cambridge cognitive scale (CAMCOG) has been adapted from CAMDEX (Cambridge Examination for Mental Disorders of the Elderly) and has been validated in the Down's Syndrome population (Hon et al., 1999). The validated tool reports a κ (inter-rated reliability) of 0.8 for 91 percent of domains rated (Ball et al., 2004), hence making it a very useful tool.

Geydes Dementia Scale for Down's Syndrome

The Geydes Dementia Scale for Down's Syndrome (Jozsvai, Hewitt and Gedye, 2018) for assessing severity in people with ID reports a good sensitivity and specificity of 85%, and is one of the NICE recommended rating scales for assessment of dementia in Down's syndrome.

Quality Outcome Measures in Dementia (QOMID)

QOMID is one of few validated tools to measure quality outcomes in dementia in ID population (Dodd, Bush and Livesey, 2015). The tool shows robust psychometric properties. Principal component analysis (PCA) shows that QOMID tool components have good eigen values and are, hence, factorizable. It has good face validity. The QOMID has good internal reliability (Cronbach α 0.84) suggesting that all domains contribute equally towards the construct of quality outcome. Further adaptions of QOMID are currently being piloted and it promises to be important predictor in an otherwise neglected area of research.

Outcomes in forensic and inpatient intellectual disability

It is widely acknowledged that people with intellectual disabilities who have additional mental health needs, including forensic needs, have relatively long stays with significant financial and health implications. To address that, a review by Morrissey *et al.* (2017) systematically examined the area. The review focussed on three key outcome super domains: Effectiveness, patient safety and patient and carer experience.

The review found that length of stay, while a good service-level outcome measure, had wide variability in terms of methods of reporting. The case is similar for discharge outcomes – they are useful for defining a service-level framework and outcome measures, but are challenging to validate. It should be kept in mind that these outcomes reflect more on health system responsiveness and impact than on patient-related outcome measures.

Examining outcomes of interest in forensic inpatient settings, a systematic review by Morrissey *et al.* (2017) identify 'effectiveness' within their final framework as a key superordinate domain. The sub-domains included under this were those that captured aspects of the care pathway, along with a focus on clinical symptoms, recovery and a reduction in reoffending. Related variables such as length of stay, discharge, and need for security were included, but these were not always directly correlated with clinical need. This reflects on wider variability encountered elsewhere in outcomes research in reporting such variables.

Further sub-domains which clustered around safety and the patient and carer experience were incorporated into the final framework as proxy indicators of the quality of forensic services. While the overall review provides key evidence in this area, determination of evidence purely from the perspective of outcome measures needs further validation. The review has explicitly outlined that there is a gap in understanding of the recovery processes and meaning in this context and that this is a relevant research need. The review considers it important to consider individual-level outcomes to understand recovery.

A newly reported tool specifically developed to reflect the measures of forensic mental health services called FORUM (Forensic outcome measures) which has clinical- and patient-reported measure variants, has reported robust properties. The FORUM-P is a patient-reported outcome measure consisting of 20 items, while the FORUM-C is a clinician-reported outcome measure with 23 items. Twelve items in the FORUM-P

correspond to 13 items in the FORUM-C. The psychometric characteristics have reasonably good internal consistency, Cronbach α for FORUM-P was 0.87 (0.80–0.93) and for FORUM-C was 0.93 (0.91–0.96). The test retest reliability κ for FORUM-P was low at 0.44 (0.24–0.63), while it was very good for FORUM-C 0.78 (0.73–0.85) (Ryland *et al.*, 2022).

The remit of this guidance is to understand which measures would be most appropriate to recommend or consider and, hence while it is acknowledged that the above systematic review fills a large evidence gap of system effectiveness, there is a strong research need to validate patient-related outcome measures in the ID population. There is also a need for systematic examination of process level outcomes and for robust measures to be designed and validated. This is one of the key areas where research is recommended.

Inpatient outcome measures in intellectual disability

Another important systematic review in this area (Melvin *et al.*, 2022) has examined effectiveness, safety and experiences in inpatient settings and has reported that, overall, services were associated with improvements in mental health for this population. The review reported that n=16 studies reported clinical outcomes measures which included the Brief Symptom Inventory, Emotional Problem Scales, Mini Psychiatric Assessment Schedules for Adults with Developmental Disabilities (mini PAS-ADD), Health of the Nation Outcome Scale (HoNOS)-secure, Reiss Screen for Maladaptive Behaviour and the Clinical Global Impressions Scale. Some of these scales are not validated in the intellectual disability population but have been used in both inpatient and forensic ID inpatient services.

Within Forensic ID, inpatients service parameters such as length of stay were reported in n=19 studies, and clinical outcomes scales such as HoNOS-LD (validated in ID population), the Historical Clinical and Risk Management (HCR-20) assessment, the psychopathy checklist screening, emotional problems scale and self-report inventory were also used. Again, some of these scales have not been validated or reported in ID populations. The review showed that admission to in-patient services, and that is to either general or specialist intellectual disability in-patient services, was associated with improvements in symptoms during the stay. Patient experience was reported in n=8 studies from specialist ID inpatient services, but with no comparators. Quality of life was reported in only n=2 of these.

Overall, from the perspective of outcome monitoring, it is important to note that while such reviews are crucial in determining the health system responsiveness and key priorities, there is a research need to better understand which patient-related outcome measures are more sensitive to change in the ID population, in both forensic and inpatient settings. There is a pressing need to map utility data so that cost-effectiveness data can be gathered.

Importance of developing outcome measures for inpatient services in intellectual disability

The current policy context and the direction of future of inpatient services in intellectual disability is radically evolving. However, there are significantly higher levels of psychiatric morbidity in people with intellectual disabilities. This means that the identification and treatment of their comorbid mental health problems require specialist expertise, both in generic and specialist settings. Hence, it is crucial that we, at the College, revisit the need for monitoring long-term needs by formulating and monitoring outcome measures, both for informing robust, data-driven approaches, as well as for making a case for appropriate resource allocation.

The College's Faculty of Intellectual Disability produced a faculty report 'People with learning disability and mental health, behavioural or forensic problems the role of in-patient services' (2013) which lays down summative guidance for the minimum data set for outcome variables (see this chapter's Appendix).

This is very diverse area and, while we hope that newer validated outcome measures like FoRUM will provide a concrete recommendation, there is an imminent research need to develop more process-driven outcome-mapping.

ADHD: Outcome measures

ADHD is co-morbid in up to 20% of people with intellectual disabilities (Miller, Perera and Shankar, 2020). Suitable diagnosis allows for early treatment; therefore, outcomes are directly linked to diagnosis. To facilitate suitable diagnosis and, by extension, the best outcomes, structured assessments are important.

For the general population, the Diagnostic Interview for ADHD in adults (DIVA) – which is a structured assessment on the principles set out by the DSM-5 criteria for ADHD – has evolved. The DIVA has been modified to encapsulate the presentation of ADHD in people with ID. To do this, it has been simplified by providing concrete clinical examples (in child and adults) of the 18-symptom criteria. The clinical examples are common presentations, which are described so carers and patients can compare and see how similar their individual experiences are. It includes a range of social influencers such as work/school, relationships, recreational and social activities etc.

People without ID usually take around one and a half hours to complete the DIVA. It can take longer in people with ID. The DIVA-ID has not been separately validated in the ID population. There is now a proposed screening tool for ADHD in people with ID (Sawhney *et al.*, 2021). Using multiple logistic regressions, three questions were identified and proposed for screening. Larger-scale replication is needed to generate generalisable results.

Epilepsy: Outcome measures

It is recognised that 22.2% of people with ID have epilepsy (Robertson *et al.*, 2015). Of these, 70% are pharmaco-resistant. Outcomes in epilepsy for people with ID need to consider a wide range of issues.

1. Seizure issues: The traditional measure is for the patient to keep a seizure diary, which allows for correlations to be identified between interventions and their possible effect on seizures. Details of what needs to be recorded in seizure diaries are usually individual to the person but broadly cover common domains. The seizure types, when they each happen and last for, any triggers, how they are associated with medication change or other change (such as sleep, stress, periods etc) are basic data which help provide information on seizure-related wellbeing. In people with ID, there might be a need for this to be provided in easy-to-read diaries or to be managed by a carer or family member.

2. Safety issues: People with ID and epilepsy are at significant risk of harm from seizures. The SUDEP and Seizure Safety Checklist is a good-practice, evidence-based tool (Shankar *et al.*, 2018). It allows for understanding and communication change in at-risk people with ID. An evidenced, patient-facing digital tool, EpSMon, is available to provide support and feedback on risk change (Newman *et al.*, 2020).

3. Drug treatment: Measuring the impact of anti-seizure medication (ASMs) is important (Watkins *et al.*, 2020). There is good guidance on the suitable best practice in prescribing ASMs to people with ID. The use and removal of rescue medication, specifically midazolam, is a good indicator of improved or worse outcomes.

4. Holistic presentation: The HoNOS-LD captures seizures as a category and, in addition to other issues, links it to the general presentation of the individual.

5. Regional outcomes: The purple light tool kit (developed with suitable co-production) is intended to examine service and community outcomes for people with ID and epilepsy. It is recommended by RCPsych and the NHSE (England). The annual LeDeR report provides insights into local mortality outcomes, including epilepsy in ID.

Constipation in people with intellectual disability

Constipation is a major cause of morbidity and premature mortality for people with an intellectual disability. Over a third of people with intellectual disabilities encounter chronic constipation across their lifetimes (Maslen *et al,* 2022). A systematic review (Robertson *et al,* 2018) found that while only 0.1% of the general population, the equivalent rate was 25% for people with an intellectual disability. Whilst there are general measures of constipation, including the Bristol Stool Chart and the Rome IV Criteria, these can be difficult to report for people with intellectual disabilities (Maslen *et al,* 2022).

One simple outcome measure used is defining constipation as fewer than three bowel movements per week and/or taking laxatives three or more times per week (Laugharne *et al,* 2024). Using this definition may be used as an outcome measure to measure change, such as after the implementation of individualised bowel care plans.

Rating scales of global diagnostic value

There are multiple key rating tools which are standardised for diagnosis. Most have detailed psychometric properties reported for them (we have mentioned the tools here and excluded any reported psychometric properties as this is beyond the scope of the review). Some of those scales include the Wechsler Adult Intelligence Scale (WAIS), the Adaptive Behaviour Assessment System (ABAS), the Vineland (Pearson) and the Global Assessment of Functioning (GAF).

There is also a wider variability of scales in autism diagnosis and management compared to the scope of the current review. There are good and comprehensive systematic reviews and recommendations in this area (Brugha *et al.*, 2015; McConachie *et al.*, 2015).

Similarly, for Functional Performance in Autism, the Spectrum Star (Triangle Consulting Social Enterprise Limited) is a self-completion measure for quantifying discussions between the individual and their support-worker. Although becoming more widely used since its development 9 years ago, there is little published in terms of patient-related outcome measures.

Other key outcome measures used for mental ill health

Psychiatric Assessment Schedules for Adults with Developmental Disabilities PAS-ADD (Moss *et al* 1988) & mini PAS-ADD *(*Moss *et al.,* 1998)

The PAS-ADD is designed to be used by community health workers and can give cumulative scores that cluster in affective/neurotic disorder, psychotic disorders and possible organic disorders among adults with intellectual disability. The tool has a good sensitivity of 66% and specificity of 70%. It has a reasonably robust factor structure, with the first factor accounting for 20% of the variance, and the subsequent eight factors accounting for only 4–8% of the variance.

Mini-PAS ADD

The Mini PAS-ADD has seven subscales for depression, anxiety, bipolar disorder, psychosis, obsessive compulsive disorder, unspecified disorders and autism. Other than in anxiety and bipolar disorder, the scale shows good internal consistency. However, overall in ID (Prosser et al., 1998) there are reports that it has a good inter-rater reliability with a κ of 0.74. There are reports of good validity (a positive predictive value owing to high specificity). However, an independent community sample validation reported a relatively reasonable sensitivity of 66% and a specificity of 70% (Sturmey et al., 2005). But overall, both PASS-ADD and Mini-PAS ADD continue to be reliable and valid instruments to measure outcomes in intellectual disability.

Glasgow Depression Scale for People with a Learning Disability (GDS-LD)

The Glasgow Depression Scale for People with a Learning Disability (GDS-LD) is a self-reporting instrument that measures depression among people with ID. It's a 20-item scale and, despite being a short rating scale, it has a good internal consistency with α of 0.9. It has a good test–retest reliability and has a very good inter-rater reliability of 0.98 (Cuthill, Espie and Cooper, 2003).

Reiss Profile for Mental Retardation/Development Disability

The Reiss Profile for MR/DD is an adaptation of the Reiss profile and has 15 subscales and motivational profiles. The tool reports a good internal consistency (average alpha=0.84), significant variability in the inter-rater reliability (average intraclass correlation coefficient = 0.52), and excellent validity (95% of the correct profiles were chosen) (Lecavalier and Havercamp, 2004).

Appendix (Taken from Faculty Report FR/ID/03)

Minimum dataset of outcome variables for in-patient beds in categories 1 and 4	
Measures at baseline	
Desirable	• Diagnoses using ICD-10 criteria or equivalent: include degree of learning disability, pervasive developmental and other developmental disorders, personality disorders, mental illnesses, substance misuse or dependence and physical disorders (Gray 2007; Alexander et al, 2011) • IQ score on WAIS–IV or equivalent (Wechsler, 2008) • Coded forensic history: index offence, nature of detention, • Past convictions for offences of violence, sex, arson and other offences, history of aggression towards other people, property and self (Alexander et al, 2006, 2011; Gray et al, 2007) • HoNOS-secure score (Dickens et al, 2007)
Desirable	• PCL:SV score (Hart et al, 1995; Morrissey, 2003, 2007, 2011; Gray et al, 2007; Fitzgerald et al, 2011) • HCR-20 (Webster et al, 1995; Gray et al, 2007; Fitzgerald et al, 2011) • VRAG score (Gray et al, 2007; Quinsey et al, 2006; Fitzgerald et al, 2011) • START score (Webster et al, 2004)
Measures of effectiveness	
Essential	• Global measures or measures of symptom severity: HoNOS secure, yearly and at discharge (Dickens et al, 2007) • Progress measures: community leave status (no leave/escorted leave/ unescorted leave) • Progress measures: length of stay • Progress measures: direction of care pathway (whether moved to a less restrictive setting)
Desirable	• Symptom-specific assessment scales (e.g. measures of anger, • depression/anxiety, other psychopathology) • HCR-20: yearly and at discharge • START score: regular intervals (e.g. 2-monthly and at discharge) • Clinical Global Impression (CGI) scale (Guy, 1976)
Measures of patient safety	
Essential	• Proxy measures of aggression: index of the number of restraints and seclusions (total number divided by length of stay) (Alexander et al, 2010) • Proxy measures of self-injury/self-harm: index of the number of incidents (total number divided by length of stay) • Number of alerts regarding patient safety • Any ‘never’ incidents: escapes, suicide
Measures of patient experience	
Essential	• Evidence of patient participation in treatment planning: My Shared Pathway (NHS Networks; Esan et al, 2012) • Patient satisfaction surveys • Evidence of carer/family participation in treatment

Desirable	• Measures of social climate: Essen Climate Evaluation Schema or equivalent (Schalast et al, 2008) • Quality of Life measure: EQ-5D-3L or equivalent, yearly and at discharge (EuroQol Group, 1990
Minimum dataset of outcome variables for in-patient beds in categories 2, 3 and 5	
Desirable	• Comorbid diagnoses on ICD-10 criteria or equivalent: include degree of learning disability, pervasive developmental and other developmental disorders, personality disorders, mental illnesses, substance misuse or dependence and physical disorders • IQ score on WAIS–IV or equivalent (Wechsler, 2008) • HoNOS learning disability score
Desirable	• Measure of symptom severity using TAG (Slade et al, 2000) • Reiss Screen Test (Reiss, 1988), PIMRA (Matson, 1988), PASADD checklist (Moss et al, 1998), MOAS (Oliver et al, 2007) and symptom-specific assessment scales (e.g. measures of anger, depression/anxiety, other psychopathology)
Measures of effectiveness	
Essential	• Global measures or measures of symptom severity: HoNOS learning disability, on admission, discharge and at regular intervals • Progress measures: community leave status (no leave/escorted leave/unescorted leave) • Progress measures: length of stay • Progress measures: direction of care pathway (whether moved to a less restrictive setting)
Desirable	• Measure of symptom severity using TAG, Reiss Screen Test, PIMRA, PASADD checklist, MOAS and symptom-specific assessment scales (e.g. measures of anger, depression/anxiety, other psychopathology) • CGI scale
Measures of patient safety	
Essential	• Proxy measures of aggression: index of the number of restraints • and seclusions (total number divided by length of stay) (Alexander et al, 2010) • Proxy measures of self-injury/self-harm: index of the number of incidents (total number divided by length of stay) • Number of alerts regarding patient safety • Any 'never' incidents: escapes, suicide
Measures of patient experience	
Essential	• Evidence of patient participation in treatment planning: My Shared Pathway (NHS Networks; Esan et al, 2012) • Evidence of community participation: education, work experience and leisure • Patient satisfaction surveys • Evidence of carer/family participation in treatment
Desirable	• Measures of social climate: Essen Climate Evaluation Schema or equivalent (Schalast et al, 2008) • Quality of Life measure: EQ-5D-3L or equivalent, yearly and at discharge (EuroQol Group, 1990)

References

Ager A (2002) 'Quality of Life' Assessment in Critical Context. *Journal of Applied Research in Intellectual Disabilities*, **15**:369–76.

Aman M, Singh N, Stewart A and Field C (1985) The Aberrant Behavior Checklist: A behavior rating scale for the assessment of treatment effects. American journal of mental deficiency, 89:485–91

Ball SL, Holland AJ, Huppert FA, Treppner P, Watson P and Hon J (2004) The modified CAMDEX informant interview is a valid and reliable tool for use in the diagnosis of dementia in adults with Down's syndrome. Journal of Intellectual Disability Research, 48(6): 611–20.

Brugha TS, Doos L, Tempier A, Einfeld S and Howlin P (2015) Outcome measures in intervention trials for adults with autism spectrum disorders; a systematic review of assessments of core autism features and associated emotional and behavioural problems. Int J Methods Psychiatr Res, 24(2) 99–115.

Cuthill FM, Espie CA and Cooper SA (2003) Development and psychometric properties of the Glasgow Depression Scale for people with a Learning Disability, Individual and carer supplement versions. Br J Psychiatry, 182: 347–53

Deb S, Hare M, Prior L and Bhaumik S (2007) Dementia Screening Questionnaire for Individuals with Intellectual Disabilities. British Journal of Psychiatry, 190(5): 440–44

Dodd K, Bush A and Livesey A (2015) Developing and piloting the QOMID – quality outcome measure for individuals with intellectual disabilities and dementia. Advances in Mental Health and Intellectual Disabilities, 9(6):298–311

England N (2020) Rightcare: Epilepsy toolkit otimising a system for people living with epilepsy. Available at: https://www.england.nhs.uk/rightcare/wp-content/uploads/sites/40/2020/03/rightcare-epilepsy-toolkit-v2.pdf (Accessed 22 February 2024).

Evenhuis HM (1996) Further evaluation of the Dementia Questionnaire for Persons with Mental Retardation (DMR). *J Intellect Disabil Res*, **40 (Pt 4)**:369–73.

Fitzpatrick, Davey, Buxton and Jones (1998) Evaluating patient-based outcome measures for use in clinical trials. Health Technology Assessment, 2(14)

Hon J, Huppert FA, Holland AJ and Watson P (1999) Neuropsychological assessment of older adults with Down's Syndrome: An epidemiological study using the Cambridge Cognitive Examination (CAMCOG). British Journal of Clinical Psychology, 38(2): 155–65.

Jones J, Dagnan D and Ruddick L (1997) A pilot study of the use of the SF-36 to assess health status of adults with learning disabilities living in small community based homes. British Journal of Developmental Disabilities, 43: 27–35

Jozsvai E, Hewitt S and Gedye A (2018) 'Gedye Dementia Scale for Down Syndrome', in Prasher, V.P. (ed.) Neuropsychological Assessments of Dementia in Down Syndrome and Intellectual Disabilities. Cham: Springer International Publishing, 57–71

Koehl L, Harp, J, Van Pelt KL, Head E and Schmitt, FA (2020) Longitudinal assessment of dementia measures in Down syndrome. Alzheimers Dement (Amst) 12(1): e12075.

Laugharne R, Sawhney I, Perera B, et al (2024) Chronic constipation in people with intellectual disabilities in the community: cross-sectional study. BJPsych Open, 10(2):e55 doi:10.1192/bjo.2024.12

Lecavalier L and Havercamp SM (2004) Are caregivers reports of motivation valid? Reliability and validity of the Reiss Profile MR/DD'. J Intellect Disabil Res, 48(Pt 3): 217–24.

Maslen C, Hodge R, Tie K, Laugharne R, Lamb K, & Shankar R (2022) Constipation in autistic people and people with learning disabilities. The British journal of general practice: the journal of the Royal College of General Practitioners, 72(720): 348–51 https://doi.org/10.3399/bjgp22X720077

McConachie H, Parr J, Glod M, Hanratty J, Livingstone N, Oono I, Robalino S, Baird G, Beresford B, Charman T, Garland D, Green J, Gringras P, Jones G, Law J, Couteur A, Macdonald G, McColl E, Morris C and Williams K (2015) Systematic Review of Tools to Measure Outcomes for Young Children with Autism Spectrum Disorder. *Health technology assessment (Winchester*, England), **19:** 1–506

Melvin CL, Barnoux M, Alexander R, Roy A, Devapriam J, Blair R, Tromans S, Shepstone L and Langdon PE (2022) A systematic review of in-patient psychiatric care for people with intellectual disabilities and/or autism: effectiveness, patient safety and experience. *BJPsych Open*, **8(6)**: e187

Miller J, Perera B and Shankar R (2020) Clinical guidance on pharmacotherapy for the treatment of attention-deficit hyperactivity disorder (ADHD) for people with intellectual disability. *Expert Opinion on Pharmacotherapy*, **21(15)**: 1897–1913.

Morris J, Bush A and Joyce, T (2014) Outcome measures of challenging behaviour interventions London: The British Psychological Society

Morrissey C, Langdon PE, Geach N, Chester V, Ferriter M, Lindsay WR, McCarthy J, Devapriam J, Walker, D-M, Duggan C and Alexander R (2017) A systematic review and synthesis of outcome domains for use within forensic services for people with intellectual disabilities. BJPsych Open, 3(1):41–56

Moss S, Prosser H, Costello H, Simpson N, Patel P, Rowe S, Turner S and Hatton C (1998) Reliability and validity of the PAS-ADD Checklist for detecting psychiatric disorders in adults with intellectual disability. Journal of Intellectual Disability Research, 42(2): 173–83

Newman C, Ashby S, McLean B and Shankar R (2020) Improving epilepsy management with EpSMon: A Templar to highlight the multifaceted challenges of incorporating digital technologies into routine clinical practice, Epilepsy & Behavior, 103: 106514.

O'Caoimh, R (2013) Screening for Alzheimer's Disease in Downs Syndrome. Journal of Alzheimer's Disease & Parkinsonism, 03

Oliver C, McClintock K, Hall S, Smith M, Dagnan D and Stenfert-Kroese B (2003) Assessing the Severity of Challenging Behaviour: Psychometric Properties of the Challenging Behaviour Interview. Journal of Applied Research in Intellectual Disabilities, 16(1): 53–61

Porter ME (2010) What Is Value in Health Care?. New England Journal of Medicine, 363(26): 2477–81

Prosser H, Moss S, Costello H, Simpson N, Patel P and Rowe S (1998) Reliability and validity of the Mini PAS-ADD for assessing psychiatric disorders in adults with intellectual disability. J Intellect Disabil Res, 42 (Pt 4): 264–72

Raczka R, Theodore K. and Williams J (2020) An initial validation of a new quality of life measure for adults with intellectual disability: The Mini-MANS-LD'. Journal of Intellectual Disabilities, 24(2):177–93.

Raitasuo S, Taiminen T and Salokangas RK (1999) Inpatient care and its outcome in a specialist psychiatric unit for people with intellectual disability: a prospective study. J Intellect Disabil Res, 43 (Pt 2): 119–27.

Ratey JJ and Gutheil CM (1991) The measurement of aggressive behavior: reflections on the use of the Overt Aggression Scale and the Modified Overt Aggression Scale. J Neuropsychiatry Clin Neurosci, 3(2):S57–60

Royal College of Psychiatrists (2013) Faculty Report FR/ID/03: People with learning disability and mental health, behavioural or forensic problems the role of in-patient services, London: Royal College of Psychiatrists' Faculty of Psychiatry of Intellectual Disability. Available at: https://www.rcpsych.ac.uk/docs/default-source/members/faculties/intellectual-disability/id-fr-id-03.pdf?sfvrsn=cbbf8b72_2. (Accessed 22 February 2024)

Robertson J, Hatton C, Emerson E and Baines S (2015) Prevalence of epilepsy among people with intellectual disabilities: A systematic review, Seizure, 29:46–62.

Robertson J, Baines S, Emerson E, Hatton C (2018) Prevalence of constipation in people with intellectual disability: a systematic review. J Intellect Dev Disabil 2018; 43: 392–406

Rojahn J, Rowe EW, Sharber AC, Hastings R, Matson JL, Didden R, Kroes, DBH and Dumont ELM (2012) The Behavior Problems Inventory-Short Form for individuals with intellectual disabilities: Part II: reliability and validity. *Journal of Intellectual Disability Research*, **56(5)**:546–65.

Roy A, Matthews H, Clifford P, Fowler V and Martin DM (2002) Health of the Nation Outcome Scales for People with Learning Disabilities (HoNOS–LD). The British Journal of Psychiatry, 180(1):61–66

Ryland H, Cook J, Ciobanasu O, Oluwabamise O, Cornish R, Al-Taiar H, Chris-Okoro E, Hoggart C, Vallakalil A, Fitzpatrick R and Fazel S (2022) Reliability and validity of the FORUM-P and FORUM-C: two novel instruments for outcome measurement in forensic mental health. Psychology, *Crime & Law:* 1-16.

Sawhney I, Perera B, Bassett P, Zia A, Alexander RT and Shankar R (2021) Attention-deficit hyperactivity disorder in people with intellectual disability: statistical approach to developing a bespoke screening tool,. BJPsych Open, 7(6)

Shankar R, Henley W, Boland C, Laugharne R, McLean BN, Newman C, Hanna J, Ashby S, Walker MC and Sander JW (2018) Decreasing the risk of sudden unexpected death in epilepsy: structured communication of risk factors for premature mortality in people with epilepsy. European Journal of Neurology, 25(9): 1121–27

Skirrow P and Perry E (2009) The Maslow Assessment of Needs Scales (MANS): An Outcome Measure and Planning Tool for People with Intellectual Disabilities.

Sturmey P, Newton JT, Cowley A, Bouras N and Holt G (2005) The PAS–ADD Checklist: independent replication of its psychometric properties in a community sample. British Journal of Psychiatry, 186(4): 319323.

Thornicroft G and Tansella M (2013) The balanced care model for global mental health. Psychological Medicine, 43(4): 849–63.

Walsh DM, Doran E, Silverman W, Tournay A, Movsesyan N and Lott IT (2015) Rapid assessment of cognitive function in down syndrome across intellectual level and dementia status. Journal of Intellectual Disability Research, 59(11):1071–79

Watkins L, O'Dwyer M, Kerr M, Scheepers M, Courtenay K and Shankar R (2020) Quality improvement in the

management of people with epilepsy and intellectual disability: the development of clinical guidance. *Expert Opinion on Pharmacotherapy*, **21(2)**: 173–81

Liaison psychiatry

Background: Outcome measurement development in liaison psychiatry

The increasing focus upon outcome and performance measurement over recent years, including the need to establish the collection of outcomes data as a matter of routine, has been of particular importance within liaison psychiatry. Although there is evidence for the economic benefit of liaison psychiatry services, there has been a relative lack of evidence relating to clinical outcomes.

Attempts have been made in the past to identify which measures should be recommended for use across liaison psychiatry services, but this has been challenging due to the variety of service settings and types of intervention which characterise liaison psychiatry work.

Liaison psychiatry service provision includes input to general hospital emergency departments and medical or surgical in-patient wards, provision of specialist out-patient services (generic or single-condition/service area), and in some cases designated liaison psychiatry in-patient beds. There are also a small number of liaison services within primary care. Within these various settings, contacts and interventions may include single or multiple assessments, diagnosis and/or formulation, guidance and advice, changes to current treatment, brief interventions, triage and signposting or longer term psychotherapeutic or biopsychosocial interventions.

The aim: Developing a framework

The first Framework for Routine Outcome Measurement in Liaison Psychiatry (FROM-LP) was published in 2015. It was intended that FROM-LP would be adopted by liaison psychiatry services throughout the NHS, to provide consistency in the collection of outcome measures.

FROM-LP drew on the Centre for Mental Health's report: Outcomes and Performance in Liaison Psychiatry: developing a measurement framework (Fossey and Parsonage, 2014), which provided a clear and structured account of the issues faced in attempting to measure outcomes consistently in liaison psychiatry services and suggested possible ways forward. No single instrument existed that could be universally applied across all liaison psychiatry services, given their complexity and heterogeneity. It was recommended that different groups of outcomes measures be used for different contexts (but overall creating a 'balanced scorecard'), so a working group was set up by the Faculty of Liaison Psychiatry to create an effective approach, which should be simple, easy to apply and consistently deliverable.

Progress since FROM-LP

Since its launch, there has been considerable uptake of FROM-LP. It was included in the recent National Institute for Health and Care Excellence's (NICE) and NHS England's *Achieving Better Access to 24/7 Urgent and Emergency Health Care Part 2: Implementing the Evidence-Based Treatment Pathway for Urgent and Emergency Liaison Mental Health Services for Adults and Older Adults – Guidance* (NICE, 2016). FROM-LP is also endorsed by the RCPsych Psychiatric Liaison Accreditation Network4 (PLAN) in its standards for services. Whilst the majority of these policy drivers are England-centric, there is acknowledgement from the Liaison Faculty and devolved nations that FROM-LP is relevant to all liaison services.

Positive feedback on the use of FROM-LP has been gathered from liaison psychiatry services, through the annual Liaison Psychiatry Survey of England (LPSE), and from clinicians through the Liaison Psychiatry Faculty Outcome Measure working group. However, there is recognition of the need to improve and optimise the measures used. FROM-LP has therefore been revised and FROM-LP (II) aims to build on the original approach, incorporating feedback and modifying the recommended measures where necessary.

Next steps: Development of FROM-LP (II)

The original aims of FROM-LP were to provide a robust and effective approach to outcome measurement, to enable liaison psychiatry services to demonstrate their clinical outcomes and effectiveness. The development and expansion of liaison psychiatry services remains a key aim of the NHS Long Term Plan, but robust data are needed to justify and support investment. These remain the key aims of FROM-LP (II). It further standardises and improves outcome measurement, incorporates new NHS England Access standards, and considers the implementation of such measures, ensuring that they are practical, cost-effective, validated and relevant to patient and carers. Feedback from faculty members in all the constituent nations has been listened to and acted upon for the benefit of all stakeholders in liaison psychiatry services.

FROM-LP (II) builds on the earlier version to again focus on brief, simple, easy, and deliverable systems of data collection, including measures of progress and outcomes spanning clinical-related outcome measures (CROMs), patient-related outcome measures (PROMs), and patient, carer and referrer-rated satisfaction scales. In a change from FROM-LP, case-types (single clinical contact/series of clinical contacts) are not defined, rather measures are rated as 'highly recommended' or 'optional', depending on the needs of individual liaison psychiatry services. This can be viewed as a 'menu of choices' e.g., the collection of response times may be highly relevant for emergency department (ED) cases, but less so for outpatients.

With increasing use of electronic record systems in acute hospital and psychiatry settings, FROM-LP II can be incorporated into electronic record templates for ease of use after each clinical encounter but is adaptable enough to be used across a range of record systems.

CR241: FROM-LP(II) is now available from the College website.

Evaluation and interpretation of results

FROM-LP (II) will encourage clinicians to embed evidence-based clinical and performance outcome measurement in routine practice and will aid evaluation of the impact this subsequently has on service development, clinical effectiveness, and patient care. It will be evaluated by the collection of feedback from liaison faculty members and liaison psychiatry services. Whilst the tools included may be used for comparison between services, this is not the main intention of the framework.

FROM-LP (II) includes several validated, evidence-based tools, but there needs to be flexibility and an understanding of local needs when a service decides which measures to adopt.

Some services may have some additional local data collection requirements, beyond those recommended in the framework, e.g:

- Minimum Data Set, i.e., patient demographics
- referral source, referral profile etc.
- structure (resources and inputs)
- process in a broader sense (e.g., number of patients seen/treated)
- education and training of general hospital staff/teams
- impact on local health service use.

Challenges with feedback for liaison services

It has been recognised by clinicians and the Psychiatric Liaison Accreditation Network that it is often challenging for liaison psychiatry services to collect feedback from patients and carers, especially in relation to older adults.

Challenges include:

- Patients may be too physically unwell to give feedback.

- Patients presenting in mental health crisis may be too mentally unwell and/ or anxious and distressed to give feedback at the time of the assessment.

- Patients may be too cognitively impaired to give feedback and/ or consent to doing so

- following their discharge from the hospital, the rate of response to requests for feedback is often low, possibly because patients do not wish to be reminded of their period of illness.

- Patients and carers often find it difficult to distinguish the care provided by liaison psychiatry from their overall care and experience within the general hospital.

Suggested ways of maximising feedback include:

- Seeking feedback from carers when patients are unable to give this.
- Asking patients to provide feedback shortly before they are discharged from the hospital, e.g., by asking them to complete a form and sealing this in an envelope to increase anonymity.
- Asking a liaison psychiatry team member who has not been directly involved in a patient's care to seek feedback, either verbally, or in writing. This can help to clarify uncertainties, e.g., about which aspects of care feedback is being sought.
- Providing a stamped addressed envelope for return of a written feedback form.
- Build feedback into e-systems for automatic feedback gathering.

There will be consideration of recommending the incorporation of FROM-LP (II) into NHS Digital as a mandated data set for NHS liaison psychiatry services.

Conclusion

The Framework for Routine Outcome Measurement in Liaison Psychiatry (II) (FROM-LP II) will encourage clinicians to embed evidence-based clinical and performance outcome measurement in routine practice and will aid evaluation of the impact on service development, clinical effectiveness, and patient care. The intention is that this, like FROM-LP, will be adopted by liaison psychiatry services across the NHS.

References

Baugh C, Blanchard E, Hopkins I (2020) *Psychiatric Liaison Accreditation Network (PLAN).* Quality Standards for Liaison Psychiatry Services, Sixth Edition, Royal College of Psychiatrists. Available at: https://www.rcpsych.ac.uk/improving-care/ccqi/quality-networks-accreditation/psychiatric-liaison-accreditation-network-plan/plan-standards

Department of Health (2012) The NHS Friends and Family Test Implementation Guidance. *London: Department of Health*

Education and Welfare, Public Health Service, Alcohol, Drug Abuse, and Mental Health Administration National Collaborating Centre for Mental Health (NICE) (2016) Achieving Better Access to 24/7 Urgent and Emergency Mental Health Care – Part 2: Implementing the Evidence-based Treatment Pathway for Urgent and Emergency Liaison Mental Health Services for Adults and Older Adults

Fossey M and Parsonage M (2014) Outcomes and performance in liaison psychiatry: developing a measurement framework. *London: Centre for Mental Health*

Guy W (1976) *ECDEU assessment manual for psychopharmacology.* Rockville, MD, U.S. Dept of Health,

Persaud A, Pearson S, Oates M, Strathdee G, et al (2008) Client Satisfaction Questionnaire. *Outcomes Compendium. London: National Institute for Mental Health in England*

Trigwell P (2014a) IRAC: Identify and Rate the Aim of the Contact. *Personal communication.*

Trigwell P (2014b) Referrer satisfaction scale (After Persaud A et al, 2008). *Personal communication*

Neuropsychiatry

Types of neuropsychiatric disorders

Neuropsychiatric disorders are varyingly defined and potentially cover a wide range of disorders at the interface of neurology and psychiatry.

The disorders addressed in this document are as follows:

1. Functional Neurological Disorder (FND): Also known as conversion disorder or dissociative (neurological) disorder. FND is where neurological symptoms such as seizures, motor or sensory disturbance occur in the absence of neurological disease and have been historically assumed to be psychological in origin. FND is considered the neurological variant of a range of potentially overlapping functional disorders affecting other bodily systems, varyingly referred to as medically unexplained symptoms (MUS) or somatoform, psychosomatic or dissociative disorders.

2. Brain Injury (BI): Either caused by physical trauma, i.e. Traumatic Brain Injury (TBI), or other mechanisms such as hypoxia, intra-cerebral bleed or cerebrovascular accident (CVA) or resulting in significant psychiatric and/or behavioural disturbance.

3. Alcohol Related Brain Disorders (ARBD): This encompasses various neurological, cognitive and psychiatric conditions that are associated with long-term alcohol misuse and related vitamin deficiencies.

4. Neurodegenerative Disorders: This includes a wide range of disorders from those with complex neurological and psychiatric manifestations, e.g. Parkinson's and Huntington's diseases, to classical dementias with less overt neurological symptoms presenting mostly with cognitive disorders such as Alzheimer's disease, vascular and fronto-temporal dementias.

5. Neurodevelopmental disorders: e.g. Tourette's Syndrome.

6. Neuropsychiatry services also regularly cover psychiatric aspects of neurological disorders such as epilepsy and multiple sclerosis. There are many such disorders and evidence is accumulating for particular outcome measures for psychiatric symptoms in these groups. It is envisaged that these will be added to this framework as and when there is sufficient evidence to make specific recommendations.

General measures

There is a need to collect universal measures across all disorders to allow comparison/baselines regarding changes in clinical status. Given the above breadth of disorders that fall under neuropsychiatry, this is a challenge but one which is important to achieve, even if captured by relatively simple/basic measures.

The measures in this section are therefore recommended for all neuropsychiatric conditions. Additional measures, specific to particular conditions, are detailed in the following sections.

Clinician-Rated Outcome Measures

Clinical Global Impression (CGI)
Perhaps the simplest and most intuitive general measure, which is intrinsically applicable to all disorders, is the Clinical Global Impression as it rates an impression of severity and improvement. It can be applied to the wide variety of disorders seen in neuropsychiatry. The CGI has two subscales – the CGI-severity (CGI-S) which measures severity, and the CGI-improvement (CGI-I) which measures improvement.

The Health of the Nation Outcome Scale (HoNOS)
This scale is a more detailed option. It consists of 12 domains covering symptoms and social functioning – there is one item on physical illness but this is not meant to capture functional symptoms. The HoNOS also has several variants designed for different disorders – including one for Acute Brain Injury (the HoNOS-ABI; see below in brain injury section) which might be more suitable than the original HoNOS for neuropsychiatric disorders in general, as well as Acute Brain Injury patients.

Patient-Rated Outcome Measures

The CGI-I has also been used for ratings by patients and carers as well as clinicians. The advantage of using the CGI here is that its use – together with a clinician-rated CROM – provides a complimentary pair of assessments and discrepancies and this may be of clinical value. This might be of particular value in functional disorders when a significant discrepancy is often found between objective and subjective symptoms.

CORE-10 is a validated instrument used for a wide variety of psychiatric disorders. However, it focuses just on mental symptoms and does not cover physical/functional symptoms.

Patient-Rated Experience Measures

- Friends and family test (FFT)
- Patient Satisfaction Scale (PSS)

Recommendations

We recommend the following minimum:

- CROM: CGI-I, rated by the clinician
- PROM: CGI-I, rated by the patient (and/or carer – especially when a patient report either is not possible or likely to be unreliable)
- PREM: Friends and family test.

With the following for extra optional measures:

- CROM: HoNOS
- PROM: CORE-10
- PREM: Patient Satisfaction Scale.

Condition-specific measures

1. Functional neurological disorder (FND)

Clinician-reported outcome measures:

There are only three validated CROMs for FND in adults; two are for movement disorders, one of which is a very complex scale (the PMDRS, (Hinson *et al.*, 2005) that requires blinded expert neurologist rating using video), and a newer simplified version scale S-FMDRS (Nielsen *et al.*, 2017) that still requires blinded video rating, but can be rated by other MDT members (e.g. physiotherapists).

The PMDRS has also been adapted for functional seizures (aka non-epileptic seizures (NES) or dissociative seizures), but this scale is similarly complex and has not been adopted by researchers, let alone clinicians. This collection of related scales is considered too detailed for routine clinical use and is intended for research purposes.

Clinician-rated strength (using MRC 0–5 scale) on clinical examination has been used but is theoretically problematic as strength can usually be increased with specific clinical manoeuvres that recruit automatic movement (e.g. Hoover's sign). In the future, actigraphy devices might provide meaningful objective measurement of movement disorders and seizures. Specific scales for all FND symptoms are also currently in development by international collaborations and will hopefully be available in the coming years.

Patient-rated outcome measures:

Patient report: Subjective strength ratings (e.g. 0–10 or 0–100%) have been used in clinical trials and may be meaningful and are simple and quick to collect. Research studies of NES have mostly focused on patient (and/or carer report) of seizure frequency, but with wide variation in how this is recorded. There is some recent evidence that seizure clusters might be a meaningful measure.

EQ-5D-5L: This scale consists of five descriptions (Mobility, Self-care, Usual activities, Pain/discomfort, Anxiety/depression), each rated at five levels (no, slight, moderate, severe, unable/extreme). It also has a visual analogue scale requesting patients to rate their health from 0 (worst you can imagine) to 100 (best you can imagine). The original scale rates health on the day of interview. It has been recommended by FROM-LP and has been used in the large multicentre CODES trial of CBT for NES.

PHQ-15: this is a 15-item somatic symptom severity scale. It rates 15 common symptoms throughout the body on a scale of 'not bothered at all', 'bothered a little' and 'bothered a lot'. However, it has not been found to perform well with FND, especially for identifying cases, and is therefore not recommended.

SF-12 and SF-36: The Short Form (SF) functional disability measure has a 12- and 36-item version; both have mental and physical health domains and are consequently widely used in medicine in mental and physical disorders. They have been regularly used in both longitudinal observational studies and RCTs in FND.

WHO DAS 2.0: The World Health Organization (WHO) has recently developed a second iteration of the Disability Assessment Scale (DAS) which, similarly to the SF scales, has a short (12-item) and longer (36-item) version, both of which are freely available on the WHO website. Both are available in interviewer, patient and proxy versions, have been extensively validated in many conditions, and have the advantage (compared to the SF scales) of being free to use and have free online/electronic support – e.g. downloadable electronic scoring sheets.

BIPQ: The Brief illness perception questionnaire is a 9-item scale to assess illness perceptions and has been used frequently in functional patients (both in patients with neurological symptoms and those with symptoms in other systems).

Recommendations:

- CROM: Nil at present; new measures are in development
- PROM: EQ-5D-5L

With the following for extra optional measures:

- CROM: S-FMDRS – but only applicable for movement disorders
- PROM: WHO DAS 2.0 12/36 (or SF12/36)

2. Brain Injury (BI)

Outcome measures for brain injury rehabilitation have been specified since 2012 by the specialist UK Rehabilitation Outcomes Collaborative (UKROC), which is now used by NHS England to provide the commissioning dataset for all specialist neurorehabilitation services across England. UKROC provides monthly activity reports and quarterly benchmarking reports. It also provides costing information on case mix and treatment costs which is updated annually to inform the development and updating of complexity-weighted tariffs. As well as outcome specification it provides information on rehabilitation requirements, the inputs provided to meet them and cost-benefits of rehabilitation for patients with different levels of need. The outcome measures recommended also depend on the complexity of patient need and the level of rehabilitation

service that they require. The CROMs detailed below are those required for patients with the highest complexity of need who are receiving rehabilitation in specialised Level 1c neuropsychiatric rehabilitation services for brain injury.

Clinician-reported outcome measures

UKROC dataset

- The Northwick Park (Nursing) Dependency Scale (NPDS): this is a measure of patient care needs that incorporates outcomes for activities of daily living, safety awareness, behavioural management and communication. The Barthel Index of Activities of Daily Living can be derived from the outcome domains of the NPDS.

- The Northwick Park Therapy Dependency Assessment (NPTDA): this provides an assessment of therapy dependency. It includes 30 items of therapy dependency across seven domains which include physical handling; basic function; activities of daily living; cognitive, psychosocial and family support; discharge planning; indirect interventions and additional activities; special facilities; investigations and procedures. It is completed in high intensity specialised neurorehabilitation services on a fortnightly basis by a multidisciplinary team.

- FIM/FAM: The Functional Independence Measure (FIM) is a measure of disability and can be scored alone or with the additional 12-item Functional Assessment Measure (FAM). The FIM-FAM is a 30-item measure scored between one (complete dependence) and seven (complete independence).

- Rehabilitation Complexity Scale Extended (RCS-E): this provides an overall measure of five domains: care, nursing, medical, therapy and equipment needs, and provides a banding of complexity.

Other measures

- HoNOS-ABI: This version of the HoNOS has been specifically adapted for brain injury patients. It is not part of the UKROC dataset and is an additional optional measure that is used, in particular, in acquired brain injury populations with neuropsychiatric symptoms.

- SASNOS: The St Andrews-Swansea Neurobehavioural Outcome Scale was developed to measure neurobehavioural disability across five major domains (interpersonal behaviour, cognition, aggression, inhibition and communication) with 49 rated items.

- Mayo-Portland Adaptability Inventory (MPAI): This measure was designed to evaluate people during the post-acute period following acquired brain injury. There are three subscales (ability index, adjustment index and participation index).

- Brain Injury Needs Indicator (BINI): This scale was developed by the Department of Health and Brain Injury Rehabilitation Trust (BIRT) and has four sections. Section 1 gathers information about the brain injury history (e.g. nature of the brain injury

and the level of initial recovery). Section 2 assesses the patient and relatives view of the effect it currently has in everyday life (pre-injury/currently). Section 3 compares perceptions from the patient and support network and uses results from Sections 1 and 2 to estimate the level of risk (classifying it as high, medium or low). Some CCGs request this scale to be completed.

Patient-rated outcome measures

The European Brain Injury Questionnaire (EBIQ) is a self-reported and carer-report measure for subjective outcomes in social, cognitive and emotional domains after brain injury. It has been validated in the brain injury population as being reliable in demonstrating treatment outcomes.

Recommendations

- CROM: For Level 1c neuropsychiatric rehabilitation services, the UKROC dataset is recommended: NPDS, NPTDA, FIM/FAM, RCS-E
- PROM: EBIQ

With the following for extra optional measures:

- CROM: HoNOS-ABI, SASNOS, MPAI-4, BINI

3. Alcohol-Related Brain Damage (ARBD)

ARBD outcomes can be divided into the two key domains of cognitive and behavioural assessment.

Cognition

- Brief: M-ACE or Montreal Cognitive Assessment (MoCA)
- Detailed: Addenbrooke's Cognitive Examination-III (ACE-III)
- Frontal: Frontal Assessment Battery – as ARBDs frequently present with frontal lobe dysfunction additional specific assessments may be helpful.

Behaviour

HoNOS-ABI: This has been recommended after withdrawal from alcohol, after three months of abstinence and subsequently every six months until an optimum level of independence is achieved. If there is residual cognitive impairment, then six-monthly cognitive assessments are recommended over a follow-up period of up to three years.

Recommendations

- CROM: M-ACE (or MoCA)
- PROM: HoNOS-ABI

With the following for extra optional measures:

- CROM: ACE-III, FAB

Neurodegenerative disorders

- Clinician-Rated Outcome Measures

- Mini Mental State Examination (MMSE): This 30-point scale has been in widespread use since its publication in 1975, despite several weaknesses – particularly its lack of items focusing on executive function.

- Addenbrooke's Cognitive Examination: This 100-point scale was designed to build on the 30-point MMSE to capture a more detailed and global measure of cognitive function. In the latest iteration, the ACE-III, the MMSE has been replaced with the authors' own 30-point subscale, the M-ACE which has the advantage of being free to use as well as having several theoretical advantages over the MMSE and promising supportive data on its validity in most common types of dementia.

- Montreal Cognitive Assessment MoCA: This 30-point scale has been validated for use in most dementias, including those associated with Parkinson's disease.

Patient-rated outcome measures

Quality of Life in Alzheimer's disease (QOL-AD): This 13-point quality of life scale has been developed for Alzheimer's disease and can be completed by patients or, if necessary, a suitable proxy or carer.

Dementia Quality of Life (DEMQOL): This larger (29-point) scale has been validated for use in all common dementias and also has a proxy/carer version.

Recommendations

- CROM: M-ACE (or MoCA)

With the following for extra optional measure:
- CROM: ACE-III
- PROM: QOL-AD (for Alzheimer's disease cases only) or DEMQOL

Neurodevelopmental disorders

Tourette's syndrome

Clinician-reported outcome measures

Yale Global Tic Severity Scale (YGTSS): A comprehensive scale for assessing the extent and severity of motor and phonic tics and their impact on self-esteem and social, educational and occupational function.

Measuring outcomes in autism spectrum disorder (ASD)

This is a substantial topic which is well reviewed for childhood (McConachie *et al.*, 2015) but the equivalent adult overview is more superficial (Henninger and Taylor, 2013). As a long-term neurodevelopmental condition, ASC shares many of the issues raised in *An Intellectual Disability Outcomes Framework for improving the quality of services for people with intellectual disability* (Royal College of Psychiatrists, 2015). However, the majority of people with ASD do not have ID.

Outcome implies a repeated measure of change. It would exclude instruments that are purely diagnostic although there are some that, measuring current symptomatology, can serve both purposes (e.g. the Autism Diagnostic Observation Schedule [ADOS]).

Areas of change

There are a large number of areas relevant to ASD. Measures addressing these might focus on change in:

- autism symptomatology – this might include social and communication skills, repetitive behaviours, and sensory symptomatology/behaviour
- symptomatology associated with ASD – such as the presence of behavioural difficulties, sleep, and eating difficulties
- the symptoms of coexistent conditions (comorbidity) associated with ASD – such as other neurodevelopmental disorders (notably ADHD) and psychiatric disorders (notably anxiety, depression and OCD)
- the management of medical conditions such as epilepsy as well as general health
- independent living (which ranges from self-help skills and the ability to function independently through to employment), social participation, quality of life and happiness
- the impact of the person on those around them, notably their family mortality and those elements identified as being associated with this – notably factors contributing to suicidality (Hirvikoski, Mittendorfer-Rutz *et al.*, 2015).

Change is slow and it can take months to discover whether it is sustained or the more superficial, short-term temporary response to a new initiative or circumstance.

Instruments to measure change

These are not necessarily specific to the individual areas listed above: they may cover several of these and then not always comprehensively.

There are few instruments that are specific to ASC. Instruments that have been developed with other populations (e.g. children, adult psychiatric populations, offenders) are often used but, if they do not take into account the particular characteristics of ASD, their results can be very misleading.

There are some instruments that deserve particular mention, as follow.

The Autism Diagnostic Observation Schedule (ADOS):

An interview with the individual which aims to elicit autistic symptomatology as the basis for a systematic, detailed, description and coded rating of 28 symptoms (Lord *et al.*, 2000). Well-researched, its reliability and validity (achieved by an intensive training course) make it one of the main research tools, although it appears to be less sensitive to identifying symptoms in more able adults with ASC (Asperger syndrome). Its main shortcoming is that it is an observation of the individual over a brief (30–60 minutes) of time which may not be representative of their wider functioning. It is a tool that gives clinicians (psychiatrists and psychologists) the basis for individual interview to arrive at a systematic evaluation of someone's current state.

While its use in research requires formal training over several days, there is a self-teach training pack for clinical use. Although well-tested, its cost, requiring the time of trained interviewers, is a limitation.

The Vineland Adaptive Behavior Scales (2nd Edition): An interview with an informant (e.g. a caregiver, close relative or friend) that systematically assesses the person's functional ability across domains that include communication, daily living skills, and socialisation. There are also Maladaptive Behaviour Subscales. It has a solid research basis and its validity has been extended to cover the whole life-span. Its thorough research basis and its structure mean that an assessment can be achieved in 15–30 minutes.

The Spectrum Star:

This has been specifically designed for charting progress and measuring outcomes in autism. It covers nine domains, including social difficulties, communication, daily living skills, managing health issues, managing sensory issues, responsible behaviour, and leisure activities/hobbies. It is recommended for use on the pathway with people with mild LD through to normal intellectual functioning in association with ASD.

It is a range of Outcome Stars which have been developed for a variety of client groups and services including older people, people with an intellectual disability, and people with mental health problems. These measures focus on a number of key areas and engage the individual in progressing their life in these areas, using the concept of a ladder of change which stretches from the person being stuck, through to developing the skills and self-reliance to manage their difficulties in each area.

The Star is completed jointly by the individual with their worker/professional to serve as the basis for an action plan, drawn up after each reading, to help focus the service user on moving forward. They are designed, therefore, to be a repeated measures instrument to chart change and progress. The emphasis is on self-completion, limited only by the individual's objectivity, but it is relatively new and is only now getting into wider use.

Details are available at: https://www.outcomesstar.org.uk/using-the-star/find-your-star/autism-and-adhd/

The Adolescent/Adult Sensory Profile:

A 60-item questionnaire developed by Brown and Dunn 2002) to provide a systematic analysis of the individual's sensory thresholds and responsiveness. Although labelled a 'self questionnaire', it is widely held that reliability requires that it be administered as an interview by a clinician. This is an area that has only recently come to prominence (Horder *et al.*, 2014), particularly after DSM-5 identified it as symptomatic of ASD, and it is being identified as a significant area of difficulty.

Available from Pearson.

Disorder		CROM	PROM	PREM
All	Recommended	CGI-I	CGI-I	Friends and Family test
	Optional	HoNOS*	CORE-10	Patient Satisfaction Scale
Functional neurological disorder (FND)	Recommended	-	EQ-5D-5L**	
	Optional	S-FMDRS***	WHO DAS 12/36 (or SF-12/36) BIPQ	
Brain injury	Recommended	NPDS NPTDA FIM/ FAM RCS-E	EBIQ	
	Optional	HoNOS-ABI SASNOS MPAI-4 BINI		
Alcohol-related brain disorders (ARBD)	Recommended	M-ACE (or MoCA) HoNOS-ABI		
	Optional	ACE-III FAB		
Neurodegenerative Disorders	Recommended	M-ACE (or MoCA)		
	Optional	ACE-III	QOL-AD or DEMQOL	
Tourette's syndrome	Recommended	-		
	Optional	YGTSS		

*HoNOS-ABI is potential alternative once validated for neuropsychiatric disorders beyond brain injury
**Adapted to rate last week (not last 24 hrs)
*** S-FMDRS – is only applicable for movement disorders

References

Brown C and Dunn W (2002) Adolescent-adult sensory profile: user's manual, Therapy Skill Builders

Henninger N A and Taylor JL (2013) Outcomes in adults with autism spectrum disorders: A historical perspective. *Autism 17(1): 103–16*

Hinson VK, Cubo E, Comella CL, Goetz CG and Leurgans S (2005) Rating scale for psychogenic movement disorders: Scale development and clinimetric testing. *Movement Disorders*, **20(12)**: 1592–97 https://doi.org/10.1002/mds.20650

Hirvikoski T, Mittendorfer-Rutz E, Boman M, Larsson H, Lichtenstein P, and Bölte S (2016) Premature mortality in autism spectrum disorder. The British Journal of Psychiatry. 232 232..238 (cambridge.org)

Horder J, Wilson CE, Mendez MA, and Murphy DG (2014) Autistic traits and abnormal sensory experiences in adults. Journal of autism and developmental disorders 44(6): 1461–69

Lord C, Risi S, Lambrecht L, Cook Jr EH, Leventhal BL, DiLavore PC, Pickles A, and Rutter M (2000) The Autism Diagnostic Observation Schedule--Generic: A standard measure of social and communication deficits associated with the spectrum of autism. Journal of Autism and Developmental Disorders 30(3): 205–23

McConachie H, Parr JR, Glod M, Hanratty J, Livingstone N, Oono IP, Robalino S, Baird G, Beresford B, and Charman T (2015) Systematic review of tools to measure outcomes for young children with autism spectrum disorder

Royal College of Psychiatrists (2015) *An Intellectual Disability Outcomes Framework for improving the quality of services for people with intellectual disability.* London, Royal College of Psychiatrists. Avaliable at: https://www.rcpsych.ac.uk/docs/default-source/members/faculties/intellectual-disability/id-fr-id-07.pdf?sfvrsn=de6623a3_4

Old age psychiatry

Principles informing the development of outcome measures

Measuring outcomes – the changes, benefits, learning or other effects that actually occur as a result of what is done is increasingly important to service users, commissioners and providers of health care. People want to understand not only the inputs and interventions provided by the NHS but also the end results and improvements delivered through those interventions.

Most mental health trusts already collate information relating to outcomes, but this is usually dispersed throughout their information systems, reported to trust boards and commissioners in various sections, and does not allow front-line staff to have a shared understanding of differences they make for service users and carers, who in turn do not have a sense of the changes they expect to see from the care they receive.

Basics

The principles that should guide development of outcome measures are that the measures should be relevant to patients and clinicians, simple and easy to use without burdening both parties with more form filling, clear and unambiguous, validated for the purpose for which they are used, simple for IT systems to support data collection and analysis and finally allow comparison between teams and services locally and nationally.

Musts

All service lines would need to agree on a CROM, PROM and PREM at the minimum.

Streams of outcome measurement development

Patient-informed/-reported measure (PROM)

- SWEMWEBS or ReQoL can be used as PROM for **cluster 1–17**
- QoL-AD (Carer proxy version) can be used as a PROM for **cluster 19–21**
- Efficacy of Post Diagnostic Support for dementia via a simple questionnaire can be used as a PROM for **cluster 19–21**

Patient-reported experience measure

Of service:
Friends & Family Test

Optional measures

- Do patients and carers report being treated with dignity and respect?
- Do patients receive well-coordinated easily accessible care & feel safe and protected from avoidable harm?

For **cluster 19-21**, the following optional measures can also be considered:

- Reduction in avoidable, unscheduled care for people with dementia in A&E
- % referred to Post Diagnostic Support Group in 8 wks
- % referred to Cognitive Stimulation Therapy in 12 wks

Relative-/carer-informed measures (cluster 19-21)

Carer's quality of life in patients suffering from dementia (EQ-5D)

Staff-/Clinician-reported outcome measure (CROM)

We recommend HoNOS for **all clustered patients**; both functional and organic mental disorder.

Other optional measures **(cluster 19-21):**

- Rate of progression from one organic cluster to other.
- Median survival from first assessment per organic cluster.

ICHOM has suggested looking at seven domains of outcomes in dementia. These domains cover several of the highlighted streams of outcome measurement development but the Old Age Faculty believe that it would be unrealistic to expect services to measure so extensively. Dementia is a particularly difficult condition to measure outcomes as, by its very nature, it is a progressively deteriorating condition.

Research-informed measures

Of service
None identified

Of personal change
None identified

Any requirement for a minimum data set

Of service
None identified

Of personal change
None identified

Interpretation of outcome measurements

- Changes in **total HoNOS score and/or categorical change by item (either CNWL or SWYPFT approach)**.
- Change in **total SWEMWBS or Re-QoL or QoL-AD scores**.
- Effect size calculation, if adequate numbers available.

Transparency and data protection issues
None identified

Innovative use of outcomes [for example in modifying treatment programmes, in conjunction with technology]
None identified

Old Age Psychiatry Faculty outcome measure summary table			
Cluster	**CROM**	**PROM**	**PREM**
1-17	HoNOS	SWEMWEBS or ReQoL	FFT
19-21	HoNOS	QoL- AD (carer proxy version) Efficacy of post diagnostic support	FFT

Glossary of relevant measures
N/A

Perinatal psychiatry

Across the UK, there has been an increase in the provision of perinatal mental health services – both in-patient Mother and Baby Units (MBU) and specialist perinatal community mental health services – as part of the Five Year Forward View. Measures have been selected which can be used to assess changes in both maternal mental health and mother–infant interaction quality.

The extent of the evidence base in support of the reliability and validity of measures designed to assess the quality of mother–infant interaction, as well as their suitability for use in routine clinical practice, is very limited. Consequently, those recommended below have some evidence supporting their use, but they require further evaluation as outcome measures. The decision as to which measures to recommend involved discussion with colleagues from the National Collaborating Centre for Mental Health and NHS England.

Summary of recommended perinatal outcome measures

In order to assess the effectiveness of perinatal interventions, paired data should be collected, i.e. at two time points – at the beginning and end of episode of treatment. Further ratings may be made at significant time points, e.g. following delivery.

The following outcome measures are recommended by the College as a minimum for use in perinatal services. Table 1 outlines the full range of outcome measures that can be used in perinatal in-patient and community services. Table 1 also sets out outcome measures to consider for use in common mental health disorders within primary care.

Minimum outcomes measures for use in perinatal services

1. Generic measures of maternal mental health for use in MBUs and specialist community perinatal teams:

- HoNOS (CROM)
- HoNOSCA (CROM)
- CORE-10 (PROM)
- CORE-OM (PROM for those receiving psychological therapies)

2. Mother–infant measures:

The two measures below should be used as the minimum standard:

- Postpartum Bonding questionnaire (PROM) – Can be used in MBU or community setting
- Bethlem Mother–infant interaction Scale (CROM) – MBU setting only

The two measures below focus on the quality of mother–infant relationship.

- PIIOS (2–7 months) (CROM) – Use if concerns raised by clinician or NICHD (3-15 months) (CROM)
- Crittenden Care Index (0–24 months) (CROM)

3. Infant measure:

Alarm Distress Baby Scale (CROM) – Used as minimum in MBU setting and if concerns raised by the clinician or mother regarding the baby's interactions.

4. Patient outcomes and experience measure:

POEM (PROM/PREM)

Specific conditions for use in MBUs and Specialist Community Teams:

Perinatal outcomes measures		
Common mental health disorders	**Type**	**Stage**
*Edinburgh Postnatal Depression Scale (EPDS): A 10-item measure for screening and measuring the severity of postnatal depression.	PROM	Consider if woman responds positively to Whooley questions or clinical concern at booking, during pregnancy and first year after birth.
*Patient Health Questionnaire (PHQ-9): A nine-item measure for screening, monitoring and measuring the severity of depression based on each of the 9 DSM-IV related diagnostic criteria.	PROM	Consider if woman responds positively to Whooley questions or clinical concern at booking, during pregnancy and first year after birth.
*Generalised Anxiety Disorder Scale (GAD-7): A seven-item measure for assessing the presence and severity of generalised anxiety.	PROM	Consider if woman responds positively to GAD-2, Whooley questions or clinical concern at booking, during pregnancy and first year after birth.
Generic measures	**Type**	**Stage**
**Health Of the Nation Outcome (HoNOS): A 12-item scale measuring behaviour, impairment, symptoms and social functioning.	CROM	At initial assessment, CPA review and discharge.
**Health of the Nation Outcome Scales Children and Adolescent Mental Health (HoNOSCA): A 15-item scale measuring behaviour, impairment, symptoms and social functioning.	CROM	At initial ssessment, CPA review and discharge.
***Recovering Quality of Life (ReQoL): An 11-item measure developed to measure improvement in quality of life with different ental health conditions.	PROM	At initial ssessment, CPA review and discharge.
**CORE-OM: Is a measure of Global Distress with 34 -items and four subscales, including well-being symptoms, function and risk.	PROM	At initial assessment, CPA review and discharge. Consider for inpatient and psychological services.
**CORE-10: Is a 10-item measure taken from the CORE-OM.	PROM	At initial assessment, CPA review and discharge. Consider for Perinatal Community Services.
***Camberwell Assessment of Needs CAN-M: A semi-structured interview schedule, for assessing the needs of pregnant women and mothers with severe mental illness.	CROM/ PROM	At initial assessment, CPA review and discharge. Once fully assessed, it may be appropriate to complete one-page summary at review.

Mother–infant measures	Type	Stage
**Postpartum Bonding Questionnaire (PBQ) Brockington: A 25-item measure, using a six-point Likert scale, to identify mother-infant isorders.	PROM	Initial assessment, review as appropriate and discharge.
**Bethlem Mother-Infant Interaction Scale (BMIS): A seven-item scale, measuring the quality of mother-infant interaction in the mother and baby setting (MBU setting).	CROM	Repeated weekly ratings.
***Mothers' Object Relations Scale (MORS-SF): A 14-item measure, scored on a six-point Likert scale, used to measure parental representation of the baby	PROM	Initial assessment, CPA review, and discharge.
**Care Index: Detailed measure of mother–infant interaction with parental sensitivity central. Seven aspects of interactional behaviour are observed. Three maternal scales: sensitivity, control and unresponsiveness. Four infant scales: cooperativeness, compulsivity, difficultness and passivity. Valid from birth to 24 months.	CROM	If problem identified then use at initial screening and repeat at discharge.
**The Parent–Infant Interaction Observation Scale (PIIOS): A 13-domain observer-rated measure, scored on a three-point Likert scale to evaluate parent-–infant relationship with parental sensitivity and responsiveness central. Valid for infants aged 2–7 months.	CROM	If problem identified then use at Initial assessment and repeat at discharge.
**NICHD: A measure of parent–infant interaction quality. Five core maternal scales: sensitivity/responsiveness, intrusiveness, detached/disengagement behaviours, positive and negative disregard for infant. Four core infant scales: positive mood, negative mood, activity and sustained attention. Valid for infants aged 3–15 months.	CROM	If problem identified then use at Initial assessment and repeat at discharge.
Infant measures	**Type**	**Stage**
**The Alarm Distress Baby Scale (ADBB): An eight-item measure, The Modify ADBB (M-ADBB) is the shorter version and is a five-item measure. These scales are designed to assess infant social withdrawal behaviours in interaction with health professional conducting the assessment.	CROM	Used if there are concerns regarding the baby's interactions– e.g. limited eye contact, little facial expression and withdrawal in baby.

Patient experience measures	Type	Stage
**Patient Outcome and Experience Measure (POEM): A measure of patient satisfaction. Two forms of measure for MBU and community service.	PREM/ PROM	At discharge from in-patient MBU or community perinatal team.
***Perinatal VOICE (Views On In-patient CarE): A 19-domain measure of patient satisfaction; for use in in-patient MBU setting.	PREM/ PROM	At discharge from in-patient MBU.
Specific conditions	**Type**	**Stage**
***Brief Psychiatric Rating Scale (BPRS): A 24-item clinician-rated scale used as part of a clinical interview, measuring positive, negative and affective symptoms of people with psychotic disorders, especially schizophrenia.	CROM	Assessment, CPA review and discharge.
***Young Mania Rating Scale (YMRS): An 11-item scale used to assess manic symptoms based on the person's subjective report of his or her clinical condition over the previous 48 hours.	CROM	Assessment, CPA review and discharge.
***Difficulties in Emotional Regulation Scale (DERS and DERS-SF): A 36-item and 18-item scale, respectively, for assessing emotion regulation problems in adolescents and adults.	PROM	Assessment, CPA review and discharge.
***Health Anxiety Inventory (short version: SHAI): A 14-item plus four-item inventory. The scores can be combined, and a cut-off score of 15 indicates a mixture of people who are hypochondriacal and health-anxious. A score of 18 or above fulfils the criteria for the DSM-IV diagnostic criteria for hypochondriasis.	PROM	Assessment, CPA review and discharge.
***Yale-Brown Obsessive Compulsive Scale: A 10-item scale to assess the severity and type of symptoms in patients with OCD.	PROM	Assessment, CPA review and discharge
***The Impact of Events Scale Revised (IES-R): A 22-item scale primarily used for the provisional diagnosis of post-traumatic stress disorder.	PROM	Assessment, CPA review and discharge.
***Panic Disorder Severity Scale (PDSS): A seven-item scale with a cut-off score of eight that is an indicator of panic disorder.	PROM	Assessment, CPA review and discharge.
***Agoraphobia-Mobility Inventory (MI): A 27-item scale used for provisional diagnosis of agoraphobia. The total score indicates the severity of the agoraphobia.	PROM	Assessment, CPA review and discharge.

* Consider use in Primary Care, e.g. midwives, health visitors, GPs and IAPT services
** Recommended as a minimum in Community/MBU Perinatal Mental Health Services
*** May be used in Community/MBU Perinatal Mental Health Services

Most measures for assessment of parent–infant interactions require further validation. The Crittenden Care index has been evaluated but, due to time taken to train to reliability, it is less suited to routine use. NICHD has evidence in support of its predictive validity to a range of child development outcomes. PIIOS is newly developed as a screening tool for use between two and seven months of age and has only been validated against the care index to date.

Perinatal Faculty outcome measures summary table

General measures		CROM	PROM	PREM
All	Recommended	Health of the Nation Outcome Scale (HoNOS)	CORE-10 CORE-OM (inpatient and psychology input)	
	Optional	Health of the Nation Outcome Scales-Children and Adolescents (HoNOS-CA) Camberwell Assessment of Needs (CAN-M) (Also PROM)	Recovering Quality of Life (ReQol)	
Mother-infant measures		**CROM**	**PROM**	**PREM**
MBU	Recommended	Bethlem Mother–Infant Interaction Scale (BMIS)		
All	Recommended		Postpartum Bonding Questionnaire (PBQ)	
	Recommended when problem identified	CARE Index Parent–Infant Interaction Observation Scale (PIIOS) NICHD		
	Optional		Mothers' Object Relations Scale (MORS-SF)	
Infant measures		**CROM**	**PROM**	**PREM**
All	If problem identified	The Alarm Distress Baby Scale (ADBB)		
Patient experience		**CROM**	**PROM**	**PREM**
All	Recommended		Patient Outcome and Experience Measure (POEM)	
	Optional		Perinatal VOICE (Views On In-patient CarE) Questionnaire	

For disorder-specific measures, please see:
College report CR216: Framework for Routine Outcome Measures in Perinatal Psychiatry

Psychological treatment and psychological services

Patient-informed measures

Of service
Patient Satisfaction Scale

Of personal change
- Clinical Outcomes in Routine Evaluation – Outcome Measure (CORE-OM)
- Inventory of Interpersonal Problems (IIP)
- Work and Social Adjustment Scale (WSAS)

Relative-/carer-informed measures

Of service
Friends and family test

Of personal change
None recommended

Staff-informed measures

Of service
None recommended

Of personal change
Global Assessment of Functioning (GAF)

Research-informed measures

Of service
Service Engagement Scale (SES)

Of personal change
The range of outcome measures that may be used in research trials of psychological treatments is large and will vary according to the specific modalities and specific mental conditions under investigation. Clinician-rated scales, e.g. the Hamilton Rating Scale for Depression (HRSD-21), may be used to augment and triangulate information

obtained from patient-rated scales, e.g. the Patient Health Questionnaire-9 (PHQ-9), and to identify discrepancies between therapist and patient-determined outcomes. Outcome measures for personality disorders may include those assessing relevant personality traits (such as impulsivity, measured by the Barratt Impulsiveness Scale (BIS)) and measures of comorbid conditions (such as substance misuse disorder, measured by the Alcohol Use Disorders Identification Test (AUDIT) and the Drug Use Disorders Identification Test (DUDIT)).

Studies of specific psychological treatment modalities may include specific outcome measures of particular psychological mechanisms relevant to the specific theory of change underpinning that psychological therapy. For example, in psychodynamic psychotherapy studies, measures of defence mechanisms, such as the Defense Mechanism Rating Scale (DMRS) may be of interest. In trials of cognitive-behaviour therapy (CBT) for depression, the Behavioural Activation for Depression Scale (BADS) may be used to measure changes in avoidance and activation over the course of behavioural activation; and in trials of mentalisation-based treatment (MBT), assessment of reflective function with the Reflective Functioning Questionnaire (RFQ-46) will be of relevance.

Any requirement for a minimum data set

- Setting (e.g., IAPT service, secondary mental health community setting, specialist psychotherapy service, personality disorder service, day hospital, therapeutic community, inpatient setting)

- Modality of therapy (e.g. CBT, psychodynamic psychotherapy, MBT)

- Length of therapy

- Ratings interval
  - For IAPT services, session-by-session outcome monitoring
  - For other services, baseline measurement and follow-up at least six-monthly

- Demographics
  - Basic demographics including age, sex, ethnicity and disability status

- Diagnostic information

Interpretation of outcome measurements

Psychological therapies' services span a wide range of different psychological modalities, settings, patient groups and mental disorders. Therefore, it is difficult to recommend a minimum data set of common outcome measures that are appropriate for all services. Services vary in the complexity and severity of the patient population treated. First-line services in primary care, such as IAPT, offer treatment for less severe conditions, such as mild depression and anxiety. Secondary and tertiary out-patient services offer specific treatments for personality disorder or other rarer conditions, such as severe obsessive-compulsive disorder, body dysmorphic disorder, medically unexplained symptoms, paraphilias, etc in day hospitals and therapeutic communities. In-patient

and secure settings treat patients with serious mental illnesses or forensic patients. Therefore, type and delivery of outcome measures will need to be chosen accordingly to take into account the patient population and setting in which they are delivered.

Moreover, psychotherapy services offer different modalities of psychological therapy – cognitive behavioural therapy (CBT) is the most common treatment offered in IAPT services (although some services also offer interpersonal psychotherapy (IPT), dynamic interpersonal therapy (DIT) and other modalities), whereas psychotherapy services in mental health settings may offer a broader range, including psychoanalytic or psychodynamic psychotherapy, systemic therapy, and other evidence-based psychological therapies originally developed for particular disorders, such as dialectical behaviour therapy (DBT), cognitive analytic therapy (CAT), mentalisation-based treatment (MBT), schema therapy, mindfulness-based therapies and others. Specific measures may be recommended for each of these different modalities of psychological therapy, reflecting the theoretical framework on which they are based which informs their theory of change and corresponding therapeutic techniques. For example, outcome measures chosen for patients treated with CBT may focus on changes in thought patterns, symptoms and problematic behaviours, whereas psychodynamic psychotherapists may wish to measure modifications of the patient's personality traits that underlie their overt symptomatology and behaviour, or specific aspects of the therapeutic relationship. Quality of life (QOL) is increasingly being measured, especially in patients with personality disorder, for whom QOL may be severely impaired but improves with treatment.

We recognise many patients may present for psychological treatment with ill-defined complaints, such as problems in relationships, low self-confidence, somatic complaints, or disturbances in their sense of self or identity, that do not neatly fit into or fulfil criteria in recognised diagnostic categories. However, we recommend, nevertheless, that all patients should have a working diagnosis, adhering to ICD-10 criteria. It is also important to note comorbidities, such as substance misuse disorders, personality disorders or neurodevelopmental disorders.

Caution should be exercised in the interpretation of changes in the scores of particular outcome measures at different points in therapy. For example, while increased scores on the CORE-OM, PHQ-9 or GAD-7 indicate increased feelings of depression or anxiety early in treatment, they could also denote an improvement in the patient's awareness of, and hence ability to work with, their feelings I – rather than indicating a relapse of their illness or a failure of treatment.

The following recommendations are informed by the results of a national survey that was sent to all members of the Medical Psychotherapy Faculty to determine what measures were being used at the time in psychotherapy services across the country and, individually, what the therapist's experience was of using them. In choosing these measures, we were mindful of practicalities and ease of administration. Ideally, all services should measure the patient's health and wellbeing, symptoms, satisfaction with the service, employment, disability, social inclusion status, access to services, attrition/non-adherence with treatment, and adverse events. We recommend measures that should be used for all patients treated with psychological therapies and further, more specific, symptom measures that are recommended depending on the particular clinical condition that is being treated (problem descriptor) in line with those administered in IAPT services.

DISORDER		CROM	PROM	PREM
All	Recommended	GAF	CORE-OM IIP WSAS	Family and friends test Patient Satis-faction Scale
	Optional	PFS WAI-SR	BSI EQ-5D SFQ WAI-SR WEMWEBS	
Depression	Recommended		PHQ-9 GAD-7	
General anxiety disorder	Recommended		PHQ-9 GAD-7	
Mixed anxiety/ depression	Recommended		PHQ-9 GAD-7	
Social anxiety	Recommended		SPIN	
Post-traumatic stress disorder (PTSD)	Recommended		IES-R	
Agoraphobia	Recommended		MI	
Obsessive compul-sive disorder (OCD)	Recommended		OCI or YBOCS	
Panic disorder	Recommended		PDSS	
Body dysmorphic disorder (BDD)	Recommended		BDD-YBOCS	
Irritable bowel syndrome (IBS)	Recommended		Francis IBS scale	
Chronic fatigue syndrome (CFS)	Recommended		CFQ-11	
Chronic pain (in context of anxiety/ depression)	Recommended		GAD-7	
Medically unex-plained symptoms (MUS)	Recommended		BIPQ PHQ-15	
Borderline personality disorder	Recommended	ZAN-BPD	WSAS EQ-5D SFQ	
Antisocial personality disorder	Recommended		OAS-M STAXI-2 EQ-5D SFQ	

Glossary of relevant measures

Alcohol Use Disorders Identification Test (AUDIT)

The AUDIT is a 10-item screening tool developed by the World Health Organization (WHO) to assess alcohol consumption, drinking behaviours, and alcohol-related problems. Both a clinician-administered version and a self-report version of the AUDIT are provided.

Barratt Impulsiveness Scale (BIS)

The BIS is a widely used measure of impulsiveness. It includes 30 items that are scored to yield six first-order factors (attention, motor, self-control, cognitive complexity, perseverance, and cognitive instability impulsiveness) and three second-order factors (attentional, motor and non-planning impulsiveness).

Behavioural Activation for Depression Scale (BADS)

The BADS is a 25-item questionnaire measuring changes in avoidance and activation over the course of behavioural activation.

Brief Symptom Inventory (BSI)

The BSI is a 53-item self-report instrument designed as a shorter alternative to the complete Symptom Checklist-90-Revised (SCL-90-R). It is composed of nine primary symptom dimensions (somatisation, obsessive-compulsive, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation and psychoticism). It includes three global indices of distress (Global Severity Index, Positive Symptom Distress Index and Positive Symptom Total), which measure the overall psychological distress level, the intensity of symptoms, and the number of self-reported symptoms.

Brief Illness Perception Questionnaire (BIPQ)

The BIPQ assesses patients' perceptions of illness and consists of nine items rated on a scale from 0 to 10 assessing the patient's perceptions and beliefs about the following: the effect of their illness on life, duration of illness, control over illness, effectiveness of treatment, experience of symptoms, concern about illness, mood, and degree of understanding of the illness. The final item is open-ended, asking respondents to rank the three most important factors causing their illness.

Chalder Fatigue Questionnaire (CFQ-11)

The CFQ-11 is a self-administered questionnaire for measuring the extent and severity of fatigue within both clinical and non-clinical, epidemiological populations, and is widely used to measure physical and mental fatigue in chronic fatigue syndrome patients.

Clinical Outcomes in Routine Evaluation – Outcome Measure (CORE-OM)

The CORE-OM is the most common measure used in psychological therapy services. It is a 34-item generic measure of psychological distress that is pan-theoretical (i.e. not associated with a school of therapy), pan-diagnostic (i.e., not focused on a single presenting problem) and draws upon the views of what practitioners consider to be the most important generic aspects of psychological wellbeing health to measure. The CORE-OM comprises four domains: Wellbeing (four items), Symptoms (12 items), Functioning (12 items) and Risk (six items). It takes 5–10 minutes to complete and is free to use. There are briefer versions that can be used for repeated monitoring or quick initial assessment.

Defense Mechanism Rating Scale (DMRS

The DMRS is an observer-rated measure of defence mechanisms comprising 27 individual defences organised into seven levels from low adaptive to high adaptive: (1) action, (2) major image distorting, (3) disavowal, (4) minor image distorting, (5) other neurotic, (6) obsessional, and (7) high adaptive. Each utterance or phrase in an interview or therapy session that is considered a defence can be rated at one of the seven levels. Higher scores indicate better adaptive functioning and lower scores indicate poorer adaptive functioning.

Drug Use Disorders Identification Test (DUDIT)

DUDIT was developed as a parallel instrument to the AUDIT for identification of individuals with drug-related problems.

EuroQOL five dimensions Health-Related Quality of Life Questionnaire (EQ-5D)

The EQ-5D is a self-report measure comprising the following five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. The patient is asked to indicate their health state for each dimension on one of three levels: no problems, some problems, and extreme problems. It has been adopted in the UK for routine outcome measurement and is preferred by NICE to calculate quality-adjusted life-years (QALYs) for use in cost-effectiveness analyses.

Generalised Anxiety Disorder Scale – 7 items (GAD-7)

The GAD-7 is a self-administered, seven-item patient questionnaire used as a screening tool and severity measure for generalised anxiety disorder (GAD). It is based on the DSM-IV diagnostic criteria for GAD but is also sensitive to severity of symptoms of social phobia, post-traumatic stress disorder and panic disorder.

Global Assessment of Functioning (GAF)

The GAF is a clinician-rated measure comprising a numeric scale (1–100) used to rate globally the social, occupational and psychological functioning of adults. It has been in use for over 20 years but is no longer included in the current version of the DSM (DSM-5). However, it is easy to score as a number between zero and 100 is chosen that fits in best with the person's current state.

Hamilton Rating Scale for Depression (HRSD-21)

The HSRD-21 is a clinician-rated 21-item questionnaire used to provide an indication of depression and as a guide to evaluate recovery. The questionnaire is designed for adults and is used to rate the severity of their depression by probing mood feelings of guilt, suicide ideation, insomnia, agitation or retardation, anxiety, weight loss, and somatic symptoms.

Impact of Events Scale – Revised (IES-R)

The IES-R is a 22-item self-report measure that assesses subjective distress caused by traumatic events. Items correspond directly to 14 of the 17 DSM-IV symptoms of PTSD. Respondents are asked to identify a specific stressful life event and then indicate how much they were distressed or bothered during the past seven days by each 'difficulty' listed.

Inventory of Interpersonal Problems (IIP)

The IIP is a self-report instrument that identifies distress arising from interpersonal difficulties. It has been validated for use with psychotherapy populations and can track the level of interpersonal distress before, during and after therapy. There is a long (64-item) and a short (32-item) version. The derived scales are: Domineering/Controlling, Vindictive/Self-Centred, Cold/Distant, Socially Inhibited, Non-assertive, Overly Accommodating, and Self-Sacrificing.

Patient Health Questionnaire-9 (PHQ-9)

The PHQ-9 is a nine-item depression scale assessing symptoms and functional impairment to make a tentative diagnosis of depression, and deriving a severity score to help select and monitor treatment.

Psychodynamic Functioning Scale (PFS)

The PFS measures clinician-rated psychodynamic changes in treatment. The six scales have the same format as the Global Assessment of Functioning Scale and measure psychological capacities over the last 3 months. The scales comprise quality of family relationships, quality of friendships, quality of romantic/sexual relationships, tolerance for affects, insight, and problem-solving capacity.

Reflective Functioning Questionnaire (RFQ-46)

The RFQ is self-report measure of mentalising. It was developed to assess severe impairments or imbalances in mentalising as typically observed in patients with borderline personality disorder features and may not be particularly suitable for use in normal community samples.

Service Engagement Scale (SES)

The SES is a clinician-rated measure comprised of 14 questions to measure service user engagement with mental health services.

Social Functioning Questionnaire (SFQ)

The SFQ is a self-report measure assessing the patient's perceived social function. It has eight questions covering the 'essential' aspects of social interaction: work and home tasks, financial concerns, relationships with family, sexual activities, social contacts, and spare time activities.

Social Phobia Inventory (SPIN)

The SPIN is a self-reported questionnaire for screening and measuring the severity of social anxiety disorder. It consists of 17 items, which cover the main spectrum of social phobia, such as fear, avoidance, and physiological symptoms.

State Trait Anger Expression Inventory-2 (STAXI-2)

The STAXI-2 is a 57-item inventory which measures the intensity of anger as an emotional state (State Anger) and the disposition to experience angry feelings as a personality trait (Trait Anger). It consists of six scales measuring the intensity of anger and the disposition to experience angry feelings.

Working Alliance Inventory – Short Revised (WAI-SR)

The WAI-SR is a measure of the therapeutic alliance that assesses three key aspects of the therapeutic alliance: (a) agreement on the tasks of therapy, (b) agreement on the goals of therapy, and (c) development of an affective bond.

Warwick–Edinburgh Mental Wellbeing Scale (WEMWBS)

The WEMWBS is a measure that records psychological wellbeing rather than overall wellbeing. However, psychological wellbeing is a hallmark of mental health and, as such, it is a measure which is relevant to successful management of conditions from the patient's perspective and, therefore, is preferable to scales that are exclusively based on symptom measures.

Work and Social Adjustment Scale (WSAS)

The WSAS is a simple, five-item self-report scale. It is a reliable and valid measure of functional impairment attributable by the person to an identified problem and offers the potential for readily interpretable comparisons across studies and disorders, as well as before and after therapy. The areas assessed are: Ability to work, Home management, Social leisure activities, Private leisure activities, Close relationships.

Yale-Brown Obsessive Compulsive Scale (YBOCS)

The YBOCS is a clinician-rated, 10-item scale, each item being rated from zero (no symptoms) to four (extreme symptoms). The scale includes questions about the amount of time the individual spends on obsessions, how much impairment or distress they experience, and how much resistance and control they have over these thoughts.

Yale-Brown Obsessive-Compulsive Scale for BDD (BDD-YBOCS)

The BDD-YBOCS is a 12-item observer-rated scale to assess the severity of body dysmorphic disorder (BDD) symptoms.

Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD)

The ZAN-BPD is a nine-item, validated, clinician-based, diagnostic interview assessing the severity of DSM-IV-based borderline personality disorder symptoms and measuring meaningful changes in symptoms over time. Each of the nine criteria for BPD is rated on a five-point anchored rating scale of 0–4, yielding a total score of 0–36.

Rehabilitation and social psychiatry

Patient-informed measures

Of service:
Friends and Family Test

Of personal change:
- Quality of life and recovery – DIALOG, Recovering Quality of Life (ReQoL), Questionnaire about the Process of Recovery (QPR)

- Wellbeing – Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS)

Relative-/carer-informed measures

Of service:
Friends and Family Test

Staff-informed measures

Of service quality (completed by the service manager)
Quality Indicator for Rehabilitative Care (QuIRC) for inpatient rehabilitation units and QuIRC-SA for supported accommodation services

Of service user progress
- Clinical status – Health of the Nation Outcome Scale (HoNOS), Clinical Global Impression (CGI)
- Social and everyday function – Life Skills Profile (LSP) (MOHOST also used routinely by OTs)
- Needs – Camberwell Assessment of Needs Short Appraisal Scale (CANSAS)

Research measures

The outcome measures used in research studies of mental health rehabilitation services depend on the research question(s) being investigated. Service quality is assessed using QuIRC and QuIRC-SA. Service user characteristics may include symptoms (e.g. Brief Psychiatric Rating Scale [BPRS] and Positive and Negative Syndrome Scale [PANSS]) and function (LSP, Social Function Questionnaire [SFQ]) alongside measures of additional

comorbid problems (e.g., substance misuse – Alcohol Use Disorders Identification Test [AUDIT], Clinician Alcohol and Drug Scale [CADS]). Measures of side effects from medication include Barnes Akathisia Scale and AIMS. For health economic evaluation, EQ-5D is the preferred option.

Any requirement for a minimum data set

Service level:

- Specific rehabilitation service type: High Dependency Rehabilitation Unit (HDRU), Longer Term HDRU, Highly Specialist HDRU, Community Rehabilitation Unit.
- Number of admissions/discharges in last 12 months.
- Average (mean and median) length of stay in the current service/unit (benchmarked against the expected length of stay for the specific service type as set out in the service typology published by the Rehabilitation Faculty and adopted by CQC).
- Number (%) SUs participating in work, education, or leisure activities in the community.
- Number (%) SUs who have received annual physical health check.
- Number adverse/risk events in last 12 months.
- Number of re-admissions within 12 months of discharge.

Service user:

- Demographics (age, sex, ethnicity, section status).
- ICD Diagnosis.
- Number (%) SUs who achieve successful (sustained) move-on to more independent settings (65% will move on successfully over five years; 10% will achieve independent living by 5 years).

Placements in out of area mental health rehabilitation units (CCGs and trusts would need to provide these data for their local population)

- Number of rehab out-of-area placements (OAPs).
- Total, mean and median cost of rehab OAPs.

Glossary

CANSAS

This is a brief measure completed by a clinician who knows the patient well. It covers 22 domains that are each rated as 0 = no need, 1= met need, 2= unmet need. Total met and unmet need scores are generated. It is useful for identifying areas of the person's life that require further input and for showing change in met/unmet needs over time at the service level (collated data).

CGI

The CGI provides a quick 'global impression' based on symptoms and function. The scale is very simple to complete as there are only two questions (each with seven possible response options).

DIALOG

This is a service user-rated outcome measure, which focuses on quality of life, care needs and treatment satisfaction with 11 items, each rated on a seven-point scale. It has been recommended for use in NHS early intervention in psychosis services by NICE and is available as a digital application that provides a useful mechanism for feedback to patients so that scores can be discussed and feed into care planning.

EuroQOL Health-Related Quality of Life Questionnaire (EQ-5D)

The EQ-5D is a self-report measure assessing mobility, self-care, usual activities, pain/discomfort and anxiety/depression. The patient is asked to indicate their health state for each dimension on one of three levels: no problems, some problems, and extreme problems. It has been adopted in the UK for routine outcome measurement and is preferred by NICE to calculate quality-adjusted life-years (QALYs) for use in cost-effectiveness analyses.

HoNOS

HoNOS is a brief, universal clinical status tool for all mental health conditions. As such, it does not capture all relevant aspects of symptoms and function relevant in rehabilitation (e.g. manic symptoms, neurodevelopmental disorders, specific cognitive problems associated with complex psychosis). It is not particularly sensitive to change in clinical status for people receiving mental health rehabilitation services (items 10 and 11 are probably the most useful), but it is included as it is still widely used across mental health services.

LSP

The Life Skills Profile is completed by a clinician who knows the service user well. There are 39 items providing an overall score and five sub-scores. This measure is included in the national dataset in Australia for all mental health service users and has been used widely in research and as a routine outcome measure in mental health rehabilitation services in the UK.

QPR

This measure was developed in collaboration with service users and asks about key aspects of personal recovery including connectedness, hope, identity, meaning to life, and empowerment. There are 11 items to be rated on a five-point scale. It has also been recommended for use in NHS early intervention in psychosis services by NICE as part of the access and waiting time standard guidance.

QuIRC and QuIRC-SA

The QuIRC is the only tailor-made quality assessment tool for mental health rehabilitation services. The tool has also been adapted (QuIRC-SA) for mental health supported accommodation services. Both are free and completed online (www.quirc.eu) by the manager of the service and take around 45 minutes. They cover: service provision (e.g. number of beds/places, average length of stay, staffing, staff turnover, training, supervision, treatment and support ordered); links with community organisations (e.g. colleges, employment agencies, sport and leisure facilities); the therapeutic milieu (e.g. collaborative care planning, service user involvement, promotion of service users' independent living skills); and the protection of service users' human rights (e.g. their privacy and dignity, their legal rights, access to advocacy and the use of restraint and seclusion).

Seven domain scores are calculated: Living Environment; Therapeutic Environment; Treatments and Interventions; Self-Management and Autonomy; Social Integration; Recovery Based Practice; Human Rights and a report produced showing the service's performance against the average scores for similar services.

ReQoL

This is a newer measure which has been designed specifically for use in mental health populations. There are two versions, a 10-item and a 20-item questionnaire both rated on a five-point scale. Recent findings have shown it to be a more sensitive and responsive measure than the EQ-5D.

Process measures are important in rehabilitation outcomes as it can help improve care delivery and patient experience.

SFQ

The SFQ is a self-report measure assessing the patient's social function. It comprises eight questions covering the 'essential' aspects of social interaction: work and home tasks, financial concerns, relationships with family, sexual activities, social contacts, and spare time activities.

WEMWBS (and SWEMWBS)

These are quick-to-use measures – whether using the 14-item scale or the shorter 7-item version. They record psychological wellbeing rather than overall wellbeing. However, psychological wellbeing is a hallmark of mental health and, as such, it is a measure which is relevant to successful management of conditions from the patient's perspective and, therefore, preferable to scales that are exclusively based on symptom measures.