Psychometric properties of the Mental Health Recovery Star

Helen Killaspy, Sarah White, Tatiana L. Taylor and Michael King

Background
The Mental Health Recovery Star (MHRS) is a popular outcome measure rated collaboratively by staff and service users, but its psychometric properties are unknown.

Aims
To assess the MHRS's acceptability, reliability and convergent validity.

Method
A total of 172 services users and 120 staff from in-patient and community services participated. Interrater reliability of staff-only ratings and test–retest reliability of staff-only and collaborative ratings were assessed using intraclass correlation coefficients (ICCs). Convergent validity between MHRS ratings and standardised measures of social functioning and recovery was assessed using Pearson correlation. The influence of collaboration on ratings was assessed using descriptive statistics and ICCs.

Results
The MHRS was relatively quick and easy to use and had good test–retest reliability, but interrater reliability was inadequate. Collaborative ratings were slightly higher than staff-only ratings. Convergent validity suggests it assesses social function more than recovery.

Conclusions
The MHRS cannot be recommended as a routine clinical outcome tool but may facilitate collaborative care planning.

Declaration of interest
None.

Healthcare providers in many countries increasingly use standardised measures and routine data collection to monitor the effectiveness of health services.[1] In recent years, the use of Patient Reported Outcome Measures (PROMs) has also gained popularity.[2] In the UK, recent mental health policy has emphasised the need for services to adopt a 'recovery orientation' to improve service users' experience of care, social inclusion and recovery.[3,4] One measure currently under consideration by the Department of Health in the UK for recommendation for routine use in mental health services is the Mental Health Recovery Star (MHRS),[5] developed by the Mental Health Providers' Forum and Triangle Consulting. Already popular across England, this tool has also started to attract international interest, particularly in Australia.[6] It aims to assess a person's recovery from mental ill health, using the contemporary meaning of recovery as a personal and dynamic process of adjustment and growth following the development of a mental health problem.[7] The tool is described as valuing service users' perspectives, enabling empowerment and choice, assessing service users' progress and supporting recovery and social inclusion. Based on the Outcomes Star, which was developed for users of homelessness services, it was adapted for the wider mental health sector through an iterative piloting process with service users, managers and front-line staff informed by the published literature on recovery.[5]

The MHRS assesses ten life domains (managing mental health; self-care; living skills; social networks; work; relationships; addictive behaviour; responsibilities; identity and self-esteem; trust and hope) each of which are represented diagrammatically as a point on a ten-arm star. Each domain (or arm) is rated on a 'ladder of change' scale from 1–2 (stuck) through 3–4 (accepting help), 5–6 (believing), 7–8 (learning) to 9–10 (self-reliance). Detailed guidance on how to rate each of these for each domain is given in the user guide.[5] The MHRS ratings are agreed through a collaborative discussion between the service user and mental health worker that lasts approximately 1 h. This collaborative approach to rating is unusual in outcome measurement and has the advantage of providing a focus for service user–staff discussions.

The MHRS is currently in use in many voluntary, statutory and independent mental health services across England,[8] but its psychometric properties have not been established.[6,9] The collaborative component of the measure presents methodological difficulties in the assessment of its reliability because of the influence of staff and service user on each other in agreeing the final rating. It is also possible that the tool may be difficult to use with service users with more severe mental health problems that impair their ability to engage in the collaborative discussions necessary for rating. The aims of our study were therefore to assess the reliability and validity of the MHRS among people with severe and enduring mental health problems in order to inform its appropriateness for use as a clinical outcome tool and/or a clinical engagement tool in mental health services. The study protocol was approved by the Institute of Child Health/Great Ormond Street Hospital Research Ethics Committee (Ref 10/H0713/9).

Method

Sampling and recruitment

Participants were recruited from four study sites across England where staff were trained to use the MHRS by the Mental Health Providers' Forum. Participants were recruited from services suggested by our local collaborators (community day centres, community and in-patient rehabilitation units, and low secure in-patient wards). The researcher met with the service manager and staff team to explain the purpose and details of the study. Staff then introduced the researcher to any service users they considered potentially eligible for participation, i.e. those who had the capacity to provide informed consent and were able to participate in a collaborative discussion. The researcher explained the purpose of the study to potential participants and answered any questions. They were given a participant information sheet containing the same information and could take up to 7 days to

consider whether they wished to participate. If they were willing, she met with them again to gain their written informed consent. All contacts with service users were made at their local mental health service.

Sample size

Participant numbers were based on a pragmatic balance between ideal sample sizes and the limitations of the study budget. The target sample size for the majority of the analyses was 100, to ensure the estimates of intraclass correlation coefficients (ICCs) and correlation coefficients could be estimated with sufficient precision; the lower bound of an acceptable ICC of 0.7 would be estimated, with a width of the 95% confidence interval no more than s.d. = 0.1.

Data collection

To assess the influence of collaboration between service users and staff on the reliability of the MHRS, ratings were made with and without the collaborative discussion (i.e. by staff and service users together and by staff alone) as follows.

Analysis 1: test–retest reliability of staff-only ratings

A staff member who knew the service user well (i.e. somebody who had been working directly with them for at least a few months and knew enough about their current mental health needs to be able to make a rating) rated them using the MHRS without collaborative discussion with the service user. The ratings were repeated by the same staff member within 1 month to test their stability. Staff were asked their opinion of the MHRS's usability (ease of completion and time to complete) and usefulness. Data on staff demographics (age, gender, ethnicity), the type of post they held, their professional background and the length of time they had worked in mental health services were also gathered.

Analysis 2: staff interrater reliability

A second staff member who also knew the service user well, independently rated the service users rated in analysis 1 using the MHRS without discussion with the other staff member or with the service user.

Analysis 3: convergent validity of staff ratings

A staff member who participated in analysis 1 or 2 also rated the service user using the Life Skills Profile (LSP),[10] a well-established, standardised measure of social functioning that has been used widely in this service-user group as a research tool and, in Australia, as a routine clinical outcome tool for many years. The LSP provides ratings on five subdomains and a total score, with higher scores denoting greater social functioning. Comparisons between LSP ratings and the seven MHRS domains that appeared most relevant to social functioning (managing mental health, self-care, living skills, social networks, work, relationships and responsibilities) were made to test for convergent validity.

Analysis 4: convergence between staff-only and staff–service-user collaborative ratings

A subsample of service users rated in analysis 1 were randomly selected by a computer-generated list to participate in a collaborative discussion with one of the two staff members who had rated them previously. They then agreed their MHRS ratings together. Comparisons of ratings were made to assess the extent to which discussion with service users modified MHRS ratings completed by staff only. Staff and service users were asked their opinion about the MHRS's usability and usefulness. Socio-demographic details (age, gender, ethnicity), diagnosis and length

of contact with mental health services were also collected from service users and corroborated by staff and case-note data.

Analysis 5: convergent validity of service-user ratings

The same service users who participated in analysis 4 also rated their recovery using the Mental Health Recovery Measure (MHRM),[11] a standardised self-report measure that assesses the subjective experience of recovery from mental illness and that has been shown to have good convergent validity with other measures of empowerment and resilience. Comparisons between the MHRM rating and MHRS ratings were made to test for convergent validity.

Analysis 6: test–retest reliability of staff–service-user collaborative ratings

A further subsample of service users who completed a collaborative rating in analysis 4 were randomly selected, using a computer-generated list, to repeat a collaborative rating with the same staff member 1–2 weeks later to assess the stability of the rating.

Analysis

Data were entered into an IBM SPSS Statistics (version 19) database for analysis on Windows. All analyses were carried out by S.W. Test–retest reliability (analyses 1 and 6) and interrater reliability (analysis 2) were assessed using ICCs. We interpreted ICCs above 0.7 as indicating acceptable reliability. Convergent validity (analyses 3 and 5) was assessed by investigating the correlation between the MHRS domain ratings and the standardised measures of social functioning (LSP[10]) and recovery (MHRM[11]) using Pearson correlation coefficients (reported with 95% confidence intervals), with a coefficient of 0.7 and above suggesting acceptable convergence. Analysis 4 investigated the degree of change in staff-only MHRS ratings after discussion with service users. This was examined first using descriptive statistics and ICCs were calculated to assess the agreement between staff-only and collaborative ratings.

Results

In total 182 service users gave informed consent to participate in the study, although for 9, no ratings were received. There were few differences between characteristics of service-user participants in the different analyses (Table 1). There were 120 staff involved in rating service users at least once. Their characteristics are given in Table 2.

Analysis 1: test–retest reliability of staff-only MHRS ratings

Data on 34 service users could not be included (27 did not have ratings performed by the same staff member and 7 repeat ratings had been made more than 1 month apart). The mean time between ratings for the remaining 138 service users was 14 days (median 14, range 3–29). The ICCs for all ten MHRS domains were above 0.7, indicating good test–retest reliability (Table 3).

Analysis 2: convergent validity of staff-only MHRS ratings and a measure of social functioning

A total of 140 ratings were available for this analysis. Pearson correlation coefficients were calculated between the seven domains of the MHRS that appeared to be assessing social functioning and the five LSP subscales and LSP total score. Managing mental health, self-care and living skills had acceptable convergent validity with the total LSP score (ICC > 0.7); managing mental health and

Table 1 Service-user characteristics

	Total sample (n = 182)	Analysis 1 (n = 138)	Analysis 4 (n = 95)	Analysis 6 (n = 39)
Age, years				
Total n	180	136	94	39
Mean (s.d.)	42.4 (13.1)	43.5 (13.1)	41.3 (12.0)	42.2 (12.6)
Min–max	20–74	20–74	20–65	22–67
Gender, n (%)				
Total n	182	138	95	39
Males	100 (54.9)	71 (51.4)	54 (56.8)	19 (48.7)
Ethnicity, n (%)				
Total n	181	137	95	39
White	138 (76.2)	112 (81.8)	70 (73.7)	30 (76.9)
Black	21 (11.6)	11 (8.0)	13 (13.7)	6 (15.4)
Asian	4 (2.2)	2 (1.5)	3 (3.2)	0
Other	18 (9.9)	12 (8.8)	9 (9.5)	3 (7.7)
Diagnosis, n (%)				
Total n	175	135	92	38
Schizophrenia	71 (40.6)	48 (35.6)	40 (43.5)	19 (50.0)
Schizoaffective disorder	11 (6.3)	10 (7.4)	5 (5.4)	2 (5.3)
Other psychosis	8 (4.6)	5 (3.7)	8 (8.7)	3 (7.9)
Bipolar affective disorder	18 (10.3)	14 (10.4)	8 (8.7)	1 (2.6)
Depression	29 (16.6)	25 (18.5)	12 (13.0)	6 (15.8)
Anxiety/OCD/PTSD	7 (4.0)	6 (4.4)	4 (4.34)	–
Personality disorder	30 (17.1)	26 (19.3)	15 (16.3)	7 (18.4)
Autism spectrum disorder	1 (0.6)	1 (0.7)	–	–
Type of setting recruited from, n (%)				
Total n	182	138	95	39
Day service	76 (41.8)	66 (47.8)	38 (40.0)	11 (28.2)
In-patient: low secure	26 (14.3)	18 (13.0)	19 (20.0)	4 (10.3)
In-patient: medium secure	39 (21.4)	29 (21.0)	22 (23.2)	14 (35.9)
Community residential facility	41 (22.5)	25 (18.1)	16 (16.8)	10 (25.6)
Length of time receiving care from mental health services, months				
Total n	181	138	95	39
Mean (s.d.)	163 (128)	172 (131)	178 (130)	206 (128)
Min–max	2–540	2–540	2–456	8–468
Site of recruitment, n (%)				
Total n	182	138	95	39
St Andrews Healthcare	63 (34.6)	47 (34.1)	40 (42.1)	16 (41.0)
Camden and Islington NHS	61 (33.5)	43 (31.2)	35 (36.8)	12 (30.8)
Hampshire Partnership NHS	17 (9.3)	12 (8.7)	4 (4.2)	5 (12.8)
Northumberland, Tyne and Wear NHS	41 (22.5)	36 (26.1)	16 (16.8)	6 (15.4)

Min–max, minimum to maximum; NHS, National Health Service; OCD, obsessive–compulsive disorder; PTSD, post-traumatic stress disorder.

self-care had acceptable convergent validity with the LSP self-care subdomain; and social networks approached had acceptable convergent validity with the LSP social contacts subdomain (Table 4).

Analysis 3: staff-only MHRS interrater reliability

We calculated ICCs for the level of agreement between the ratings of two members of staff for each domain of the MHRS. The analysis was restricted to ratings completed within 1 month of each other. A total of 87 ratings could not be included (67 did not have two ratings completed by different staff and 20 had been completed more than 1 month apart). The mean time between the remaining 85 ratings was 8.5 days (median 7, range 0–31). The results are shown in Table 3. Only the MHRS work domain had acceptable interrater reliability.

Analysis 4: staff-only and staff–service-user collaborative rating convergence

Disparities in scores between the staff-only MHRS ratings and the staff–service-user collaborative ratings were summarised using descriptive statistics for each domain. A total of 95 ratings were

available for this analysis. Change scores were calculated as staff-only rating minus collaborative rating. Therefore a positive change score meant the staff-only ratings were higher (greater recovery) than the collaborative ratings; a negative score meant the collaborative ratings were higher than the staff-only ratings (Table 5). The mean change scores for all but two of the MHRS domains were negative (although the median change scores for most domains were zero), suggesting that staff scored service users slightly more negatively when completing the MHRS alone than when rating collaboratively with the service user. The ICCs for the staff–service-user collaborative ratings are also presented to allow comparison with other results of convergence between two ratings. These suggest that interrater reliability for staff-only and collaborative ratings was acceptable for only the work domain.

Analysis 5: convergent validity of service-user MHRS ratings

Pearson correlation coefficients were calculated between each of the ten domains of the MHRS and the seven MHRM subscales and total score. No MHRS domains had an acceptable level of convergence (online Table DS1).

Table 2 Staff characteristics

Staff characteristics	Sample (*n* = 120)
Age, years: mean (s.d.) Min–max	40.8 (11.1) 20–63
Gender, *n* (%)	
Total *n*	113
Male	34 (30.1)
Female	79 (69.9)
Ethnicity, *n* (%)	
Total *n*	110
White	82 (74.5)
Black	22 (20.0)
Asian	3 (2.7)
Other	3 (2.7)
Discipline, *n* (%)	
Total *n*	105
Nurse	62 (59.0)
Occupational therapist	8 (7.6)
Social worker	4 (3.8)
Doctor	2 (1.9)
Art therapist	2 (1.9)
Psychologist	1 (1.0)
Other	26 (24.8)
Length of time working in mental health services, months: mean (s.d.) Min–max	144.2 (115.5) 1.8–444
Employer, *n* (%)	
Total *n*	120
St Andrews Healthcare	53 (44.2)
Camden and Islington NHS	30 (25.0)
Hampshire Partnership NHS	20 (16.7)
Northumberland, Tyne and Wear NHS	17 (14.2)
Min–max, minimum to maximum; NHS, National Health Service.	

Analysis 6: Staff–service-user collaborative MHRS rating test–retest reliability

We calculated ICCs for each of the ten domains of the MHRS; 39 ratings were available for this analysis. The mean time between ratings was 12 days (median 11, range 4–28). Test–retest reliability was good, with all domains having an ICC greater than 0.7 (Table 3).

Acceptability of the MHRS

Of the 183 MHRS staff-only ratings, 125 (68%) staff reported that it took less than 30 min to complete, and 55 (30%) reported that it took 30–60 min. Overall, 120 (66%) staff felt it was easy/very easy to decide on a score for each domain and 152 (83%) felt that it was easy to use. A total of 168 (92%) felt it was useful/very useful for care planning and 106 (58%) felt it was useful/very useful as a clinical outcome measure, although 67 (37%) did not answer this last question.

Of the 92 collaborative ratings, 42 (46%) staff reported that it took less than 30 min to complete and 39 (42%) that it took 30–60 min. Overall, 54 (59%) staff felt it was easy/very easy to decide on a score for each domain, with 43 (47%) reporting that it was easier to score collaboratively than alone and 19 (21%) reporting this as more difficult. Seventy-five (82%) reported it as being easy to use, 78 (85%) that it was useful/very useful for care planning and 39 (42%) that it was useful/very useful as a clinical outcome measure (47 (51%) did not answer this last question).

Of the 92 (57%) service users who completed a collaborative rating, 52 reported that this took less than 30 min and 34 (37%)

Table 3 Test–retest and interrater reliability of staff-only and collaborative Mental Health Recovery Star (MHRS) ratings

MHRS domains	Intraclass correlation coefficient[a] (95% CI)		
	Staff-only MHRS test–retest reliability	Staff-only MHRS interrater reliability	Staff–service-user collaborative MHRS test–retest reliability
Mental health	0.83 (0.77–0.88)	0.69 (0.56–0.79)	0.75 (0.57–0.86)
Self-care	0.89 (0.84–0.92)	0.55 (0.38–0.69)	0.74 (0.57–0.86)
Living skills	0.83 (0.76–0.87)	0.67 (0.53–0.77)	0.77 (0.60–0.87)
Social networks	0.70 (0.61–0.78)	0.67 (0.53–0.77)	0.76 (0.59–0.87)
Work	0.86 (0.81–0.90)	0.77 (0.67–0.85)	0.82 (0.68–0.90)
Relationships	0.79 (0.71–0.84)	0.53 (0.36–0.67)	0.82 (0.69–0.90)
Addictive behaviour	0.85 (0.80–0.89)	0.46 (0.27–0.62)	0.79 (0.63–0.88)
Responsibilities	0.80 (0.73–0.85)	0.60 (0.44–0.72)	0.78 (0.62–0.88)
Identity and self-esteem	0.81 (0.74–0.86)	0.58 (0.44–0.72)	0.78 (0.62–0.88)
Trust and hope	0.78 (0.70–0.84)	0.62 (0.46–0.73)	0.71 (0.49–0.84)
a. >0.7 considered acceptable.			

Table 4 Convergent validity of staff-only Mental Health Recovery Star (MHRS) ratings with social functioning

Life Skills Profile (LSP) subscale	MHRS domains, coefficient[a] (95% CI)						
	Managing mental health	Self-care	Living skills	Social networks	Work	Relationships	Responsibilities
Self-care	0.7 (0.61–0.78)	0.71 (0.62–0.78)	0.66 (0.56–0.74)	0.47 (0.33–0.59)	0.33 (0.17–0.47)	0.23 (0.07–0.38)	0.51 (0.38–0.63)
Non-turbulence	0.43 (0.29–0.56)	0.46 (0.32–0.58)	0.49 (0.35–0.60)	0.41 (0.26–0.54)	0.34 (0.18–0.48)	0.27 (0.11–0.42)	0.57 (0.44–0.67)
Social contact	0.64 (0.53–0.73)	0.61 (0.50–0.71)	0.62 (0.51–0.72)	0.69 (0.59–0.77)	0.36 (0.20–0.49)	0.38 (0.23–0.51)	0.53 (0.40–0.64)
Communication	0.38 (0.23–0.51)	0.41 (0.26–0.54)	0.46 (0.32–0.58)	0.38 (0.23–0.52)	0.28 (0.12–0.43)	0.29 (0.13–0.44)	0.35 (0.19–0.48)
Responsibilities	0.53 (0.39–0.64)	0.54 (0.42–0.65)	0.55 (0.43–0.66)	0.44 (0.29–0.56)	0.30 (0.14–0.44)	0.29 (0.13–0.43)	0.52 (0.39–0.63)
Total LSP	0.7 (0.61–0.78)	0.7 (0.61–0.78)	0.71 (0.62–0.78)	0.6 (0.48–0.70)	0.41 (0.26–0.54)	0.36 (0.21–0.50)	0.64 (0.53–0.73)
a. Pearson correlation coefficient (>0.7 considered acceptable).							

Table 5 Change scores (staff-only minus collaborative rating) and intraclass correlation coefficient (ICC) of staff-only and collaborative Mental Health Recovery Star (MHRS) ratings

MHRS domains	Mean	Median	Lower quartile	Upper quartile	Minimum	Maximum	ICC (95% CI)
Managing mental health	−0.5	0	−1	0	−5	4	0.65 (0.51–0.76)
Self-care	−0.6	−1	−2	0	−8	4	0.63 (0.48–0.75)
Living skills	−0.5	0	−2	0.25	−6	4	0.64 (0.50–0.75)
Social networks	−0.3	0	−1	1	−7	6	0.63 (0.50–0.74)
Work	−0.2	0	−1	0	−8	6	0.70 (0.58–0.79)
Relationships	−0.7	0	−2	1	−5	5	0.50 (0.33–0.64)
Addiction	0.1	0	−1	1	−8	6	0.68 (0.55–0.78)
Responsibilities	−0.8	0	−2	0	−7	3	0.60 (0.42–0.74)
Identity and self-esteem	−0.9	−1	−2	0	−6	3	0.59 (0.35–0.74)
Trust	0.8	−1	−2	0	−7	4	0.59 (0.41–0.72)

reported that it took 30–60 min. Sixty-one (66%) service users reported that it was easy/very easy to decide on a score for each domain, with 65 (70%) reporting the MHRS as being easy to use. Seventy-nine (85%) service users felt the measure was useful/very useful in helping them and the staff understand how they were getting on and 79 (85%) felt it was useful/very useful for helping them and the staff plan the support they needed.

Discussion

Main findings

This study evaluated the acceptability, reliability and convergent validity of the MHRS in assessing people with severe and enduring mental health problems. We found the measure was acceptable to staff and service users, with few reporting it as difficult to complete and most reporting completion within 30–60 min. The majority of staff and service users felt it to be useful for care planning, but fewer staff reported it to be useful as a clinical outcome measure (although we note the lower response to this question, suggesting perhaps that some staff did not understand the question or did not feel able to give a view on this).

Due to the unusual, collaborative rating of the measure, it was not possible to assess interrater reliability *per se*. Instead, interrater reliability of staff-only ratings and the influence of collaboration between staff and service users on ratings were investigated. The measure had good test–retest reliability for staff-only and collaborative ratings. However, interrater reliability of staff-only ratings was inadequate. This is a serious problem and of particular relevance in mental health services where staff turnover and multidisciplinary working mean that different members of staff need to be able to assess service users reliably. Collaboration between staff and service users in rating the measure influenced the score, with staff tending to rate slightly lower when completing the rating alone. This is in keeping with previous findings that have reported that staff rate service users as having higher needs than service users rate themselves.[12] Clearly it is not possible to say whether either rating is accurate without another objective measure. Convergent validity with a routinely used social function measure was acceptable for three of the seven MHRS subscales assessed (and a further subscale almost met the ICC threshold). However, the MHRS had poor convergence with an existing service user-rated measure of subjective recovery. This suggests that the scale is more likely to be assessing social functioning than the personal experience of recovery described in the introductory paragraph of this paper.[7] It also highlights

the difficulty in defining the contemporary concept of recovery for measurement.

A recently published, detailed and systematic 'review of reviews'[13] identified 22 measures of subjective service user experience of recovery that have been developed since 1995. Only one measure had been developed in the UK (the MHRS), two in Australia and the rest in the USA. None met all four of the psychometric properties considered most important by the authors (internal consistency, concurrent validity, test–retest reliability and sensitivity to change), although four met the first two of these. Concurrent validity was assessed against a variety of other measures (recovery, empowerment, resilience, self-esteem, social support, well-being and hope). None of the measures had been tested for interrater reliability or sensitivity to change. Although two measures appeared promising (the Recovery Assessment Scale[14] and the Illness Management and Recovery Scale[15]), the authors concluded that no measures of subjective recovery had been adequately tested to be able to recommend their use as routine outcome measures.

Writers on the subject of 'values-based care' have eloquently described how the assessment of quality and outcomes are often conflated, with measures of process often being reported as measures of outcome, when, in fact, process and outcome are separate constructs in the assessment of service quality.[16] The MHRS was developed in the context of UK mental health policy that 'puts people who use services at the heart of everything we do'.[3] In this context, the service user experience has come to be considered as an outcome itself, when it is a actually a measure of process. The MHRS has been referred to as a PROM, but although the 'ladder of change' appears to relate to the subjective experience of recovery,[7] our findings suggest that the measure mixes this process with the specific outcome of social functioning. This mixed construct may explain some of the difficulties with its psychometrics. Such is the current enthusiasm in many services to include PROMs in routinely collected data that measures are adopted without adequate understanding of what exactly they are measuring, whether they are appropriate for the intended purpose and whether they have adequate psychometric credentials.

Limitations

The limitations of the study resources meant that we were unable to examine the MHRS's ability to assess service users' progress over time. Our sample size was adequate for our purposes but did not allow for more complex regression modelling to investigate staff and service user factors associated with change

in scores between staff-only and collaborative ratings or to investigate differences between service types.

Implications

The inadequate interrater reliability of this measure does not support its recommendation for use as a routine clinical outcome tool at present. Further refinement may improve this and testing of the tool's sensitivity to change would then also be required. However, the tool appears to assess social functioning more than recovery and other, reliable measures of social function already exist. Nevertheless, it is acceptable to staff and service users and its novel, collaborative rating and visual appeal may be useful in promoting service user involvement in care planning and help to focus the content of discussions between staff and service users.

Helen Killaspy, MBBS, PhD, FRCPsych, Mental Health Sciences Unit, University College London; **Sarah White**, PhD, Research Resource Unit, Section of Mental Health, Division of Population Health Sciences and Education, St Georges, University of London; **Tatiana L. Taylor**, BA(Hons), MSc, **Michael King**, MBChB, PhD, FRCGP, FRCP, FRCPsych, Mental Health Sciences Unit, University College London, UK

Correspondence: Helen Killaspy, MBBS, PhD, FRCPsych, Mental Health Sciences Unit, University College London, Charles Bell House, 67–73 Riding House Street, London W1W 7EJ, UK. Email: h.killaspy@ucl.ac.uk

First received 20 Dec 2011, final revision 6 Mar 2012, accepted 15 Mar 2012

References

1 Department of Health. *Liberating the NHS. Transparency in Outcomes – A Framework for the NHS*. Department of Health, 2011.

2 Appleby J. Patient reported outcome measures: how are we feeling today? *BMJ* 2012; **344**: d8191.

3 HM Government. *No Health without Mental Health: A Cross-Government Mental Health Outcomes Strategy for People of all Ages*. Department of Health, 2011.

4 Shepherd G, Boardman J, Slade M. *Making Recovery a Reality*. Sainsbury Centre for Mental Health, 2008.

5 MacKeith J, Burns S. *Mental Health Recovery Star*. Mental Health Providers Forum and Triangle Consulting, 2008.

6 Burgess P, Pirkis J, Coombs T, Rosen A. Assessing the value of existing recovery measures for routine use in Australian mental health services. *Aust NZ J Psychiatry* 2011; **45**: 267–80.

7 Anthony WA. Recovery from mental illness: the guiding vision of the mental health service system in the 1990s. *Psychol Rehab J* 1993; **16**: 11–24.

8 Mental Health Providers Forum. *Recovery Star Online Analysis – The Growing Picture Continued*. Mental Health Providers Forum, 2011.

9 Beazley PI. The Recovery Star: is it a valid tool? (letter). *Psychiatrist* 2011; **35**: 196–7.

10 Parker G, Rosen A, Emdur N, Hazipavlov D. The Life Skills Profile: psychometric properties of a measure assessing function and disability in schizophrenia. *Acta Psychiatr Scand* 1991; **83**: 145–52.

11 Young SL, Bullock WA. *The Mental Health Recovery Measure*. University of Toledo, 2003.

12 Najim H, McCrone P. The Camberwell Assessment of Need: comparison of assessments by staff and patients in an inner-city and a semi-rural community area. *Psychiatr Bull* 2005; **29**: 13–7.

13 Burgess P, Pirkis J, Coombs T, Rosen A. Assessing the value of existing recovery measures for routine use in Australian mental health services. *Aust N Z J Psychiatry* 2011; **45**: 267–80.

14 Flinn S. *Reliability and Validity of the Recovery Assessment Scale for Consumers with Severe Mental Illness Living in Group Home Settings*. Kent State University, 2005.

15 Hasson-Ohayon I, Roe D, Kravetz S. The psychometric properties of the Illness Management and Recovery Scale: client and clinician versions. *Psychiatr Res* 2007; **160**: 228–35.

16 Porter M. What is value in health care? *N Engl Med J* 2010; **363**: 26.

EXTRA CONTENT ONLINE