

Minimal Important Difference in Patient Reported Outcome Measures in Persons with Severe Mental Illness, a Post-hoc Pre-Post Analysis

Titus Beentjes (✉ titus.beentjes@radboudumc.nl)

Radboudumc IQ healthcare <https://orcid.org/0000-0002-5645-1068>

Steven Teerenstra

Radboudumc IQ healthcare

Hester Vermeulen

Radboudumc IQ healthcare

Maria W.G. Nijhuis-van der Sanden

Radboudumc IQ healthcare

Betsie G.I. van Gaal

Radboudumc IQ healthcare

Peter J.J. Goossens

Dimence

Research

Keywords: Self-management, Recovery, Minimal important difference, Severe mental illness, Illness management and recovery

Posted Date: February 11th, 2020

DOI: <https://doi.org/10.21203/rs.2.23224/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: Complementary interventions for persons with severe mental illness (SMI) provide broad strategies for recovery and illness self-management. It is not known which outcome measure can be considered to be relevant for persons with SMI. This knowledge can motivate a professional to offer and stimulate a person to participate in that intervention. This paper aimed to identify the outcome measures that determine the most relevant and meaningful change and capture the benefits of a complementary intervention.

Methods: By using anchor-based and distribution-based methods, we estimated the minimal important difference (MID) to determine which outcome measure persons improved in beyond the MID to reflect a relevant change in pre-post effect of a complementary intervention, in casu the Illness Management and Recovery programme (IMR).

Results: The anchor MID was based on the results of the measure Rand General Health Perception (Rand-GHP). On all MIDs, the Mental Health Recovery Measure (MHRM) had the highest score on the effect compared to its MIDs, and also on all MIDs the MHRM had the highest percentages of participants that scored above the MID.

Conclusion: The Rand-GHP is considered to be an excellent measure for investigating the MID as a result of an intervention. The results of our study can be used in shared decision-making processes to determine which intervention is suitable for a person with SMI. A person who desires a recovery outcome, as measured by the MHRM, can be recommended to do the IMR programme.

Introduction

In the last decades, the focus of treatment for people with severe mental illness (SMI) changed from decreasing burden of symptoms towards living a meaningful life [1]. In the 1980s, the concept of recovery was introduced, defined as a deeply personal, unique process of changing one's attitudes, values, feelings, goals, skills, and/or roles [2]. The illness became less important in favour of quality of life. In the 1990s, as a result of better general health care, life expectancy grew and illnesses became chronic. The challenge to manage chronic illnesses and their consequences increased. The term self-management was introduced and defined as a dynamic and continuous process of self-regulation that refers to an individual's ability to manage the symptoms, treatment, physical and psychosocial consequences, and lifestyle changes inherent in living with a chronic condition [3]. This development has affected the perspectives on health and changed health from "not being ill" towards health as an ability to adapt and to self-manage [4].

In the field of people with SMI, self-management and symptom reduction represent the clinical or medical orientation, and recovery is used as an orientation on personal issues as defined by Anthony [2]. In this field, a dismissive attitude towards the word illness can be heard. Interventions with a focus on recovery are more in favour than medically focused interventions. Several interventions with a single focus on recovery have been developed to help persons with SMI to choose, get and keep valued roles. Complementary interventions also use the focus on illness self-management, providing broad strategies for facilitating recovery [5]. Examples of complementary interventions are Wellness Recovery Action Plan (WRAP) [6] and the Illness Management and Recovery programme (IMR) [7]. The discussion on medical versus recovery orientation made us curious about which outcome measures most likely capture participants' potential benefits from an intervention. The WRAP showed effects on symptom domains and recovery [6]. In different kinds of trials, the IMR showed effects on recovery [8–10], symptom reduction [8, 11], and illness self-management [12–14]. We wondered whether a complementary intervention, like the IMR, is able to realise effects in outcome measures from both orientations and in measures on quality of life and general health. Besides, McGuire et al. [15] recommend exploration of the effects of the IMR programme on recovery and symptoms severity outcomes.

Outcome measures are considered to be relevant when pre-post effects are meaningful to patients. An appropriate benchmark to assess this is the concept of minimal important difference (MID) [16–18]. Guyatt et al. [19] defined the MID as the smallest difference in score in the domain of interest that patients perceive as important, either beneficial or harmful, and that help shared decision making in considering a change in treatment. For instance, knowing that an intervention can enhance an important difference in a desired outcome may motivate a patient to use the intervention. The concept of MID has become a standard approach in determining clinical relevance of changes in patient reported outcomes (PROs). A PRO is defined as any report coming directly from patients about how they function or feel in relation to a health condition and its therapy [20]. No scientific literature on MID for patient reported outcome measures concerning people with SMI is available.

Considering the discourse of a medical versus recovery orientation in the field of people with SMI, this paper aimed to identify the PRO that had the most relevant and meaningful change and capture the benefits of a complementary intervention. To this end, we first estimated MIDs and then determined which PRO persons improved in beyond the MID to reflect a relevant change.

Methods And Materials

Trial design and settings

For the identification of the MIDs for PROs, we used the data from all participants in a registered cluster-randomised controlled trial which examined the effect of the e-health version of the Illness Management and Recovery (e-IMR) intervention compared to the standard IMR [21,22]. The trial was performed in mental health institutions that were members of the Dutch IMR network. As no relevant differences between the e-IMR and IMR was found, we pooled the data from the control and experimental groups and performed pre-post tests on different PROs to examine which PRO captured IMR's potential benefit the most.

Data collection and outcome measures

In this paper, we used the time points before starting the IMR programme (baseline) and after a year when finishing the IMR programme (endpoint). To describe the study population, we used the participant characteristics collected at trial baseline, i.e. age, gender, psychiatric diagnoses, psychiatric and somatic comorbidities, treatment history, cultural background, housing, social economic status, education level, and diagnosis conforming the Diagnostic and Statistical Manual of Mental Disorders, 4th edition (See Table 1).

In order to respond to the aim, we used validated PROs that measured illness management, recovery, self-management, symptom severity, quality of life, and general health.

With the complementary **Illness Management and Recovery Scale** (IMRS), we measured illness management, consisting of 15 items. The response levels, on a five-point scale (1–5), vary depending on the item. The IMRS' test-retest coefficient (r_{xx}) varies between 0.79 and 0.84 [23–26].

Recovery was measured with the **Mental Health Recovery Measure** (MHRM), consisting of 30 items referring to self-empowerment, learning and new potentials, and spirituality [27]. The response levels, on a five-point scale, vary from 'strongly disagree' (0) to 'strongly agree' (4), with 'neutral' in between (2). The MHRM's r_{xx} is 0.92 [28].

Self-management was measured with the **Patient Activation Measure** (PAM-13), consisting of 13 items referring to the individual's knowledge, skill, and confidence for managing his/her own health and health care [29]. The response levels, on a four-point scale, vary from 'strongly disagree' to 'strongly agree', and the fifth option is 'not applicable'. The PAM-13's r_{xx} is 0.76 [30].

Symptom severity was measured with the **Brief Symptom Inventory** (BSI), consisting of 53 items referring to the burden of physical and psychological symptoms during the past month. The response levels, on a five-point scale, vary from 'not at all' (0) to 'extremely' (4) [31]. The BSI's r_{xx} is 0.90 [32].

Quality of life was measured with the **Manchester Short Assessment of Quality of Life** (MANSA) [33], consisting of 12 items rating satisfaction with their life as a whole and with 11 other social, physical, and mental health domains, on a seven-point scale, varying from 'couldn't be worse' (1) to 'couldn't be better' (7). The MANSA's r_{xx} is 0.82 [34].

The participants' general health status was measured with the **Rand 36-item Health Survey**, consisting of 36 items assembled into nine concepts. In this study, we only used the concepts of general health perception (Rand-GHP) and health change (Rand-HC) [35]. The Rand-GHP estimates the participant's current perception of their general health (bio-psycho-social) by scoring five items. On a five-point scale, participants score 1) how many times their health status hindered them in social activities; and whether they agree with the statements that 2) they become ill more easily than other people, 3) their health status is just like other people they know, 4) the expectation that their health status will decline, and 5) that their health status is excellent. The Rand-GHP's r_{xx} is 0.80. With the Rand-HC, participants estimate their health compared to a year ago on a five-point scale varying from much or somewhat better, the same, and somewhat or much worse. The r_{xx} of the Rand-HC is 0.40 [36].

Statistical methods

Analyses were conducted using Statistical Package for the Social Sciences® (SPSS) 23 [37]. Mixed model multilevel regression analyses were used to examine the pre-post change in the outcome measures, taking into account clustering of participants. This method automatically uses the 'missing at random' assumption to handle missing data. Random effects on cluster, trainer, and individual participants nested within the cluster and fixed main effects for time trend were included in the model. The analyses were executed according to the intention-to-treat principle. Participants who did not complete the IMR sessions were included in the analyses. Non-completers are participants who attended fewer than 50% of the IMR sessions.

Methods for investigating minimal important differences

To assess on which PRO participants improved to a meaningful degree, we estimated the minimal difference that would likely be important for the participants. Four methods to estimate the MID are recommended [17,38–40]. Two are based on statistical distribution: using the effect size (ES) and the standard error of measurement (SEM), and two are anchor-based methods: using a global transition question and a clinical criterion. It is recommended to estimate the MID primarily by anchor-based methods [17,41] and to use distribution-based methods as supportive information [17]. We considered that the PRO with the highest effect/MID rate represents the most relevant change and is capable of capturing the potential benefits of the IMR programme.

Anchor-based method

For the 'global transition question' anchor-based method, we used the Rand-HC. In our study, the 'year ago' was the start of the IMR programme. The other anchor-based method uses a criterion, which is a measure health professionals are familiar with and is widely used in assessing patients' health status [39], such as clinical endpoints or person-based global improvement in PROs [17]. Since there is no such widely used criterion in mental health, we searched for a criterion in our own data that captures the richness and variation of the construct of a Quality of Life measure (QoL) [17]. Besides the Rand-HC, we examined a number of QoL anchor candidates using the change scores of 1) the first item of the MANSA estimating satisfaction with their life as a whole (MANSA-1), 2) the total MANSA, and 3) the Rand-GHP. The strength of the association between the anchor and the PRO needs to be determined because low or no correlation can provide misleading information [17,42]. A correlation of at least 0.30–0.35 is recommended [17]. Therefore, correlations between the anchor candidates and the PROs were analysed. Outliers should not drive a correlation to a significant level. In SPSS, scores that are 2.58 times the standard deviation are assigned as a probable outlier [43]. Probable outliers were assessed on their appropriateness and impact on the correlation, and a decision was made about removing or recoding to a reasonable level [44,45]. The anchor candidate with the highest correlation with the change scores in most PROs was considered to be the right anchor. Estimation of the MID based on an anchor proceeds as follows: the scores on the anchor were used to form five groups of participants

reflecting large negative, small negative, no, small positive, and large positive change. The mean of the four differences between the effects in the PROs in two succeeding change groups is the PRO's MID-anchor [39].

Distribution-based methods

To support the anchor-based MID method, we examined the two statistical-distribution-based MID methods: the ES and the SEM [17]. The ES of change scores on the PROs estimates the effect of the intervention related to the standard deviation of the change scores (SD_c), which is the endpoint scores minus the baseline scores of the participants. This relates to between-patient variation in change scores. To estimate the ES (effect/ SD_c), we used the estimated effects from the mixed model analyses. The one-half ES ($\frac{1}{2}$ -ES) is standard to estimate the PRO's MID based on the ES (MID-ES) [17,38–40]. The SEM is the measurement error of the outcome. The SEM is computed with the standard deviation (SD) and the test-retest coefficient index ($SD \cdot \sqrt{(1 - r_{xx})}$) [39,44,46,47]. To estimate the PRO's SEM, we used the SD and r_{xx} that were reported in psychometric studies of the PROs in populations comparable to ours as much as possible (See Table 2). A change smaller than the SEM is likely a result of the measure's unreliability rather than a true observed change. Therefore, one times the SEM (1-SEM) is the PRO's MID based on the SEM (MID-SEM).

After these calculations, in all PROs we estimated the percentages of participants that had improved above MID-anchor, MID-ES, and MID-SEM.

Results

Participant flow

From seven clusters, a total of ten IMR groups entered the trial, including 60 participants. Baseline characteristics of the participants are presented in Table 2. Out of the 60 participants, eighteen (30%) were identified as non-completers, fourteen of whom completed the trial. Out of the 60, fifteen (25%) were lost to follow-up, because of being too burdened by the interview. In total, 45 participants completed the trial, of whom change scores could be calculated (See Figure 1).

Pre-post effects of the IMR programme

Since the random effect of cluster was zero in nearly all the analyses, this factor was excluded from the mixed model analyses. The pre-post effects were significant for the IMRS, MHRM, the PAM-13, the BSI, the MANSA (all $p < 0.01$), and the Rand-GHP ($p < 0.05$) (See Table 2).

MID analyses

The Rand-HC showed no correlations to the other PROs and was therefore considered not useful to determine the MID-anchor. Compared to the MANSA-1 and the MANSA, the Rand-GHP showed the most frequent, highest, and significant ($p < 0.01$) correlations to the other PROs (See Figure 2). Outliers drove the correlation to the BSI. Examination identified the outliers as true scores, and to assess their impact on correlation, they were recoded into twice the SD_c ($= 1.02$) after which the correlation remained significant ($p < 0.01$) (See Table 3 and Figure. 2). Therefore, Rand-GHP was selected as anchor. Five change-groups were shaped with a score difference of 15 on the Rand-GHP change scores (see Table 4), resulting in the groups: large negative change ($n = 3$), small negative change ($n = 7$), no change ($n = 18$), small positive change ($n = 10$), and large positive change ($n = 7$). To be able to see how the mean group scores on the PROs relate, we estimated the ES in the groups and made a graph in Figure 3.

The pre-post effects in the PROs were (See Table 2):

≥ 1 -MID-anchor in the MHRM (1.29), the IMRS (1.14), and the PAM-13 (1.00);

≥ 1 -MID-ES in the MHRM (1.33) and the IMRS (1.02);

≥ 1 -MID-SEM in the MHRM (1.40) and the IMRS (1.07).

The PRO with the highest percentage of participants that improved ≥ 1 in all the MIDs was the MHRM: 51.1% ≥ 1 -MID-anchor, 55.6% ≥ 1 -MID-ES, and 55.6% ≥ 1 -MID-SEM (see Table 2).

Discussion

The aim of this paper was to identify the PRO that determines the most relevant and meaningful change for persons with SMI and captures the benefits of a complementary intervention, in casu the IMR programme. On all three MIDs, the MHRM had the highest score on the effect compared to its MIDs; also, on all three MIDs, the MHRM had the highest percentages of participants that scored above the MID. Therefore, we conclude that the recovery measure MHRM can be considered the most relevant PRO capable of capturing the benefit of the IMR programme (with or without the e-IMR).

Participants improved significantly on all the measures. The improvement on self-management, shown by the PAM-13 and the IMRS, might have enhanced their recovery more than the increase of burden of symptoms as measured by the BSI. This matches with Slade's statement that self-management is related to recovery as it can be a vital resource for supporting recovery [1]. Based on our finding, we conclude that the IMR programme is capable of enhancing recovery as in other research that claimed the recovery orientation of the IMR programme [15, 48, 49].

Comparing the three calculated MIDs in the PROs, we conclude that they do not differ much. This would be expected when in a SEM calculation the reliability index r_{xx} is 0.75, as then the MID-ES and the MID-SEM are equal [47]. The r_{xx} in the main PROs in our study ranged between 0.76 and 0.92. When an r_{xx} is

higher than 0.75, the MID-SEM is expected to be lower than the MID-ES. In our study, this is the case in the MHRM and BSI. In the MANSA, the opposite is present, which is due to difference in the SD between the population in the reference study [34] and the SD_c in our study. Nevertheless, we conclude that in our study the results in the three MIDs are reasonably consistent. And therefore, we conclude that the MID-ES and MID-SEM support our findings on the MID-anchor.

Considering the concept of MID, Revicki et al. [42] states that the $\frac{1}{2}$ -ES magnitude of change is certainly clinically significant but may not be the smallest non-ignorable difference. The $\frac{1}{2}$ ES in a PRO might be too large to be considered minimally important [42]. On the other hand, even the MID-SEM calculations might be questionable. The SEM accounts for the measurement error in a single measurement at one time point. However, change relates to two time points, which motivates to multiply the MID-SEM by $\sqrt{2}$ to reflect the error in measuring a change score, just like den Ouden et al. [38] did in their study. The $\sqrt{2}$ enlarges the MID-SEM and the distance to the smallest non-ignorable difference. In our study, the MID-anchor and the MID-ES were more in line with the MID-SEM without application of the $\sqrt{2}$.

The anchor-based method is our preferred method just as is recommended by Revecki et al. and Johnstone et al. [17, 41]. This effect/MID-anchor index is the closest to what the participants have reported about their health and what they might state as important. The Rand-HC 'global transition question' anchor-based method appeared to be inaccurate, which is in line with other studies that declares the inaccuracy with response shifts and recall bias [39, 50, 51]. An answer to such a global transition question might only be accurate when a person is confronted with detailed information about his or her own health statements from a year ago to which the person can compare their current health state. Nevertheless, this global transition question is still recommended in investigating the MID.

In our study, we found the Rand-GHP to be the best anchor. Although this choice was data driven, we also considered content validity. The five Rand-GHP questions contain important issues in estimating one's health status: global ratings of whether their health status hinders them in social activities and whether they estimate their health status to be excellent; social comparison in the two questions of comparing one's health to persons they know; and expectancy of deterioration in the statement that one expects their health status to decline. Social comparison is a ubiquitous social phenomenon. There is considerable evidence that evaluating oneself favourably in comparison with others is associated with having fewer health problems [52]. Social comparison also is an important behavioural change technique [53]. Because of the group-wise deliverance of the IMR programme in our study, participants got acquainted with peers. Comparing oneself to peers might be more realistic than the comparison to healthy persons. Perceiving one's health status as deteriorating is associated with a higher need for support with self-management tasks [54]. We conclude that the Rand-GHP captures the richness and variation of a construct of health perception and is an excellent measure for investigating the minimal important change as a result of an intervention.

Because of the variation in population and context, King [39] warns that there is no universal MID for a PRO. MIDs need to be investigated in different contexts, for instance the context of the country. The incidence of SMI in the United States of America (USA) is threefold the Netherlands' incidence, respectively 4.5% [55] and approximately 1.5% [56], which is due to the definition of illness duration. In the Netherlands, this is 'at least a couple of years', and in the USA this is 'sufficient to meet diagnostic criteria' [55], which can be much shorter than a couple of years. Therefore, our investigation needs to be repeated in other contexts.

This issue of context-dependency illustrates the dilemma of scientific research versus practical usability. The practical use of an MID is to contribute to shared decision-making processes [39]. A healthcare professional wants information that is as close as possible to the context in which an intervention needs to be chosen. Scientific research may have been conducted in different contexts, and/or interventions and may not be applicable to the context of interest. The MID-ES and MID-SEM might provide more generalisable information, but even these depend on context. Even if they are less context-dependent, they may overestimate the MID, i.e. indicate relevant changes that are not minimal. Therefore, these two MIDs might be more useful in interpretation guidelines becoming part of decision-making processes on a higher level than one patient-professional relationship. King [39] foresees a future in which MIDs are consolidated in guidelines. But, before a person with SMI can benefit from an intervention like the IMR programme, an institute must be able to provide it. Multidisciplinary guidelines can motivate institutes to train professionals in providing the intervention.

In the Dutch context of people with SMI, our findings can individually be used in shared decision-making processes. Knowing that an intervention can make an important difference on a desired domain can motivate a professional to offer and stimulate a person with SMI to participate in that intervention. In the case of IMR: a person who desires a recovery outcome, as measured by the MHRM, can be recommended to do the IMR programme.

Strength and limitations of the study

A number of limitations of this study should be noted. At first, we had to deal with a low number of participants and therefore lower level of certainty about conclusions. However, our sample size (more than 40) was large enough to detect correlation coefficients of 0.50 or higher with a power of 96% [44]. Besides this small sample, we were able to follow a number of participants with a low attendance rate, which makes our intention to treat analysis more realistic.

Conclusions

We have estimated MIDs for PROs in the context of people with SMI. Using these, we found that improvement in the recovery measure MHRM determined the most relevant change and captured the benefit of the IMR programme, a complementary intervention for people with SMI.

Implications for further research

More research needs to be done to get a more solid grounding for MIDs in the context of people with SMI. Future research should investigate the MIDs related to other recovery-oriented interventions as well.

Implications for further practise

The results of our investigation can be used in shared decision-making processes to determine which intervention is suitable for a person with SMI.

Declarations

Ethical Approval and Consent to participate

The ethical approval for conducting the e-IMR trial was provided by the Committee on Research Involving Human Subjects, Arnhem-Nijmegen (NL49693.091.14). All participants declared informed consent to participate in this research. The trial was registered in the Dutch Trial Register (NTR4772). During the research process, an independent researcher monitored the research procedure and administration.

Consent for publication

The authors declare consent for publication

Availability of supporting data

Data and materials will be made available after a request to the corresponding author.

Competing interest

The authors declare that they have no competing interests.

Funding

This study was funded by ZonMW (the Netherlands Organization for Health Care Research and Development) program, 'Tussen Weten en Doen' (Grant 520001001). The funder had no influence on the study design, the data collection, analysis and, the writing and the submitting the article.

Author's contributions

All authors contributed to the conception and design of the study. Titus Beentjes contributed to the data collection. All authors contributed to the analysis of the data, provided drafting and critical revision of the article and final approval of the article.

Acknowledgements

Not applicable

Authors' information

¹Radboud university medical center, Radboud Institute for Health Sciences, IQ healthcare, Nijmegen, the Netherlands. ²Center for Nursing Research, Saxion University of Applied Science, Deventer/Enschede, the Netherlands. ³Dimence Group Mental Health Care Centre, Deventer, the Netherlands. ⁴Radboud university medical center, Radboud Institute for Health Sciences, department for Health Evidence, group Biostatistics, Nijmegen, the Netherlands. ⁵HAN University of Applied Sciences, Faculty of Health and Social Studies, Nijmegen, the Netherlands. ⁶University Centre for Nursing and Midwifery, Department of Public Health and Primary Care, Faculty of Medicine and Health Sciences, Ghent University, Ghent, Belgium.

Abbreviations

½-ES: one-half ES

BSI: Brief Symptom Inventory

e-IMR: e-health version of the Illness Management and Recovery

ES: Effect Size

IMR: Illness Management and Recovery programme

IMRS: Illness Management and Recovery Scale

MANSA: Manchester Short Assessment of Quality of Life

MHRM: Mental Health Recovery Measure

MID: minimal important difference

PAM-13: Patient Activation Measure

PRO: patient reported outcomes

QoL: Quality of Life

Rand-GHP: Rand General Health Perception

RAND-HC: Rand Health Change

r_{xx} : test-retest coefficient

SD: Standard Deviation

SD_c : Standard Deviation of the Change Scores

SEM: Standard Error of Measurement

SMI: severe mental illness

SPSS: Statistical Package for the Social Sciences

USA: United States of America

WRAP: Wellness Recovery Action Plan

References

1. Slade M. *Personal Recovery and Mental Illness, A Guide for Mental Health Professionals*. Cambridge: Cambridge University Press; 2009.
2. Anthony WA. Recovery from mental illness: The guiding vision of the mental health service system in the 1990s. *Psychosoc. Rehabil. J.* [Internet]. 1993;16:11–23. Available from: <http://doi.apa.org/getdoi.cfm?doi=10.1037/h0095655>
3. Barlow J, Wright C, Sheasby J, Turner A, Hainsworth J. Self-management approaches for people with chronic conditions: a review. *Patient Educ. Couns.* 2002;48:177–87.
4. Huber M, Knottnerus JA, Green L, Horst H v. d., Jadad AR, Kromhout D, et al. How should we define health? *BMJ* [Internet]. 2011;343:d4163. Available from: <http://www.bmj.com/cgi/doi/10.1136/bmj.d4163>
5. Farkas M, Anthony WA. Psychiatric rehabilitation interventions: A review. *Int. Rev. psychiatry.* 2010;22:114–29.
6. Cook JA, Copeland ME, Floyd CB, Jonikas JA, Hamilton MM, Razzano L, et al. A Randomized Controlled Trial of Effects of Wellness Recovery Action Planning on Depression, Anxiety, and Recovery. *Psychiatr. Serv.* 2012;63:541.
7. Gingerich S, Mueser KT. *Illness Management & Recovery, Personalized Skills and Strategies for Those with Mental Illness, Implementation Guide*. 3rd ed. Dartmouth: Hazelden; 2011.
8. Mueser KT, Meyer PS, Penn DL, Clancy R, Clancy DM, Salyers MP. The Illness Management and Recovery program: rationale, development, and preliminary findings. *Schizophr. Bull.* 2006;32 Suppl 1:S32-43.
9. Salyers MP, Rollins AL, McGuire AB, Gearhart T. Barriers and facilitators in implementing illness management and recovery for consumers with severe mental illness: trainee perspectives. *Adm. Policy Ment. Health* [Internet]. 2009 [cited 2014 Oct 19];36:102–11. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19096924>
10. Salyers MP, Godfrey JL, McGuire AB, Gearhart T, Rollins AL, Boyle C. Implementing the Illness Management and Recovery program for consumers with severe mental illness. *Psychiatr. Serv.* 2009;60:483–90.
11. Fujita E, Kato D, Kuno E, Suzuki Y, Uchiyama S, Watanabe A, et al. Implementing the illness management and recovery program in Japan. *Psychiatr. Serv.* 2010;61:1157–61.
12. Färdig R, Lewander T, Melin L, Folke F, Fredriksson A. A randomized controlled trial of the illness management and recovery program for persons with schizophrenia. *Psychiatr. Serv.* [Internet]. 2011;62:606–12. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21632728>
13. Hasson-Ohayon I, Roe D, Kravetz S. A randomized controlled trial of the effectiveness of the illness management and recovery program. *Psychiatr. Serv.* [Internet]. 2007;58:1461–6. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17978257>
14. Levitt AJ, Mueser KT, Degenova J, Lorenzo J, Bradford-Watt D, Barbosa A, et al. Randomized controlled trial of illness management and recovery in multiple-unit supportive housing. *Psychiatr. Serv.* [Internet]. 2009;60:1629–36. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19952153>
15. McGuire AB, Kukla M, Green A, Gilbride D, Mueser KT, Salyers MP. Illness management and recovery: a review of the literature. *Psychiatr. Serv.* 2014;65:171–9.
16. Guyatt GH, Walter S, Norhfm G. Measuring change over time: assessing the usefulness of evaluative instruments. *J. Chronical Dis.* 1987;40:171–8.
17. Revicki DA, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J. Clin. Epidemiol.* 2008;61:102–9.
18. Angst F, Aeschlimann A, Angst J. The minimal clinically important difference raised the significance of outcome effects above the statistical level, with methodological implications for future studies. *J. Clin. Epidemiol.* 2017;82:128–36.
19. Guyatt GH, Osoba D, Wu AW, Wyrwich KW, Norman GR, Aaronson N, et al. Methods to explain the clinical significance of health status measures. *Mayo Clin. Proc.* 2002;77:371–83.

20. Jayadevappa R, Cook R, Chhatre S. Minimal important difference to infer changes in health-related quality of life—a systematic review. *J. Clin. Epidemiol.* [Internet]. Elsevier Inc; 2017;89:188–98. Available from: <https://doi.org/10.1016/j.jclinepi.2017.06.009>
21. Beentjes TAA, Goossens PJJ, Vermeulen H, Teerenstra S, Nijhuis-van der Sanden MWG, van Gaal BGI. E-IMR: e-health added to face-to-face delivery of Illness Management and Recovery programme for people with severe mental illness, an exploratory clustered randomized controlled trial. *BMC Health Serv. Res.* [Internet]. BMC Health Services Research; 2018;18:962. Available from: <https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-018-3767-5>
22. Beentjes TAA, van Gaal BGI, Goossens PJJ, Schoonhoven L. Development of an e-supported illness management and recovery programme for consumers with severe mental illness using intervention mapping, and design of an early cluster randomized controlled trial. *BMC Health Serv. Res.* [Internet]. 2016;16:20. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26782620>
23. Färdig R, Lewander T, Fredriksson A, Melin L. Evaluation of the Illness Management and Recovery Scale in schizophrenia and schizoaffective disorder. *Schizophr. Res.* 2011;132:157–64.
24. Goossens PJJ, Beentjes TAA, Knol S, Salyers MP, de Vries SJ. Investigating the reliability and validity of the Dutch versions of the illness management and recovery scales among clients with mental disorders. *J. Ment. Health* [Internet]. 2017;26:489–95. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26754026>
25. Hasson-Ohayon I, Roe D, Kravetz S. The psychometric properties of the Illness Management and Recovery scale: client and clinician versions. *Psychiatry Res.* 2008;160:228–35.
26. Salyers MP, Godfrey JL, Mueser KT, Labriola S. Measuring illness management outcomes: a psychometric study of clinician and consumer rating scales for illness self management and recovery. *Community Ment. Health J.* 2007;43:459–80.
27. van Nieuwenhuizen C, Wilrycx G, Moradi M, Brouwers E. Psychometric evaluation of the Dutch version of the mental health recovery measure (MHRM). *Int. J. Soc. Psychiatry.* 2014;60:162–8.
28. Bullock W. The Mental Health Recovery Measure (MHRM): Updated Normative Data and Psychometric Properties Overview [Internet]. Updat. Norm. Data Psychom. Prop. Toledo, OH; 2009. Available from: https://www.utoledo.edu/al/psychology/pdfs/MHRM_12-09.pdf
29. Hibbard JH, Mahoney ER, Stockard J, Tusler M. Development and testing of a short form of the patient activation measure. *Health Serv. Res.* [Internet]. 2005;40:1918–30. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16336556>
30. Moljord IEO, Lara-Cabrera ML, Perestelo-Pérez L, Rivero-Santana A, Eriksen L, Linaker OM. Psychometric properties of the Patient Activation Measure-13 among out-patients waiting for mental health treatment: A validation study in Norway. *Patient Educ. Couns.* 2015;98:1410–7.
31. Derogatis LR, Melisaratos N. The Brief Symptom Inventory—an introductory report. *Psychol. Med.* 1983;13:595–605.
32. Beurs E de, Zitman F. De Brief Symptom Inventory (BSI). De betrouwbaarheid en validiteit van een handzaam alternatief voor de SCL-90. [The Brief Symptom Inventory, reliability and validity of a convenient alternative]. *Maandbl. Geestelijke Volksgezond.* 2006;61:120–41.
33. Priebe S, Huxley P, Knight S, Evans S. Application and Results of the Manchester Short Assessment of Quality of Life (MANSA). *Int. J. Soc. Psychiatry.* 1999;45:7–12.
34. Birgisson ÓG, Sigurðsson BH. Psychometric Properties of the Icelandic Manchester Short Assessment of Quality of Life (MANSA) and its Possible Utility in Iceland. *Vísindagur meistaránema í klínískri sálfræði við Háskólann í.* 2016;
35. Hays RD, Sherbourne CD, Mazel RM. The RAND 36-Item Health Survey 1.0. *Health Econ.* 1993;2:217–27.
36. van der Zee KI, Sanderma R, Heyink JW, de Haes H. Psychometric Qualities of the RAND 36-Item Health Survey 1.0: A of Multidimensional Measure of General Health Status. *Int. J. Behav. Med.* 1996;3:104–22.
37. IBM C. IBM SPSS Statistics for Macintosh, Version 23. Armonk, NY: IBM Corp; 2015.
38. den Ouden BL, Zijlstra WP, de Vries J. The minimal clinical important difference in the World Health Organization Quality of Life instrument-100. *Support. Care Cancer* [Internet]. 2013;21:1295–301. Available from: <http://link.springer.com/10.1007/s00520-012-1664-8>
39. King MT. A point of minimal important difference (MID): A critique of terminology and methods. *Expert Rev. Pharmacoeconomics Outcomes Res.* 2011;11:171–84.
40. Turner D, Schünemann HJ, Griffith LE, Beaton DE, Griffiths AM, Critch JN, et al. The minimal detectable change cannot reliably replace the minimal important difference. *J. Clin. Epidemiol.* 2010;63:28–36.
41. Johnston BC, Ebrahim S, Carrasco-Labra A, Furukawa TA, Patrick DL, Crawford MW, et al. Minimally important difference estimates and methods: A protocol. *BMJ Open.* 2015;5:1–7.
42. Revicki DA, Cella D, Hays RD, Sloan JA, Lenderking WR, Aaronson NK. Responsiveness and minimal important differences for patient reported outcomes. *Health Qual. Life Outcomes.* 2006;4:1–5.
43. Field A. *Discovering Statistics Using IBM SPSS Statistics.* 4th ed. London: SAGE Publications Ltd; 2013.
44. Portney L, Watkins M. *Foundation of Clinical Research, application to practice.* Third edit. Upper Saddle River, New Jersey: Pearson Education, Inc.; 2015.
45. Polit DF, Beck CT. *Nursing Research: Generating and Assessing Evidence for Nursing Practice.* 10th ed. New York: Lippincot Williams & Wilkins; 2016.
46. Streiner DL, Norman GR, Cairney J. *Health Measurement Scales, a practical guide to their development and use.* 4th ed. Oxford: Oxford University Press; 2015.
47. Wyrwich KW, Bullinger M, Aaronson N, Hays RD, Patrick DL, Symonds T. Estimating clinically significant differences in quality of life outcomes. *Qual. Life Res.* 2005;14:285–95.

48. McGuire AB, Luther L, White D, White LM, McGrew J, Salyers MP. The "Critical" Elements of Illness Management and Recovery: Comparing Methodological Approaches. *Adm. Policy Ment. Health* [Internet]. 2014; Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25425013>
49. van Langen WJM, Beentjes TAA, van Gaal BGI, Nijhuis-van der Sanden MWG, Goossens PJJ. How the Illness Management and Recovery Program Enhanced Recovery of Persons With Schizophrenia and Other Psychotic Disorders: A Qualitative Study. *Arch. Psychiatr. Nurs.* [Internet]. 2016;30:552–7. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0883941716300115>
50. Banaji MR, Hardin C. Affect and Memory in Retrospective Reports. In: Schwarz N, Sudman S, editors. *Autobiographical Mem. Validity Retrospect. Reports*. New York: Springer-Verlag; 1994. p. 71–86.
51. Coughlin SS. Recall bias in epidemiologic studies. *J. Clin. Epidemiol.* 1990;43:87–91.
52. Buunk AP, Gibbons FX. Social comparison: The end of a theory and the emergence of a field. *Organ. Behav. Hum. Decis. Process.* 2007;102:3–21.
53. Bruin M De, Kok G, Schaalma H. Coding manual for behavioral change techniques [Internet]. 2007. Available from: <http://www.biomedcentral.com/content/supplementary/1748-5908-7-92-S2.pdf>
54. van Houtum L, Rijken M, Heijmans M, Groenewegen P. Self-management support needs of patients with chronic illness: Do needs for support differ according to the course of illness? *Patient Educ. Couns.* [Internet]. Elsevier Ireland Ltd; 2013 [cited 2013 Dec 15];93:626–32. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24029582>
55. NIMH. Serious Mental Illness (SMI) Among U.S. Adults [Internet]. 2013 [cited 2015 Sep 10]. Available from: <http://www.nimh.nih.gov/health/statistics/prevalence/serious-mental-illness-smi-among-us-adults.shtml>
56. Delespaul PH, Dutch consensus group SMI. [Consensus regarding the definition of persons with severe mental illness and the number of such persons in the Netherlands]. *Tijdschr. Psychiatr.* [Internet]. 2013;55:427–38. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23864410>

Tables

Table 1. Baseline participants' demographic and clinical characteristics

Variable	n(%)	Mean (SD)
Participants	60(100)	
Female	36(60)	
Age		45(11.6)
Diagnoses		
Psychotic disorders	20(33)	
Mood/anxiety disorders	25(42)	
Other disorders	15(25)	
Global Assessment of Functioning	50.5(8.7)	
Having a somatic comorbidity	30(50)	
Having a psychiatric comorbidity	38(63)	
Treatment history		
Years ago since first treatment		16.9(11.5)
Number of admissions		4(3.7)
Never admitted	9(15)	
Housing		
Independent living	38(63)	
Assisted living	22(37)	
Netto income		
≤ Minimal income	47(78)	
> Minimal income	13(22)	
Highest graduated education		
≤ Middle school	38(63)	
≥ High school	22(37)	

n: number; SD: standard deviation

Table 2. Analyses of Minimal Important Differences in patient reported outcome measures

Measures [references]	Min- Max scores	From references ¹		Effects		MID-Anchor			MID-ES			MID-SEM			
		r _{xx}	SD	Calculated SEM	Pre- post effect (mixed model)	SD _c	MID- anchor	Effect/MID- anchor	% Improvement > MID-anchor	MID- ES	Effect/MID- ES	% Improvement > MID-ES	MID- SEM	Effect/MID- SEM	% Improver >MID-SEI
IMRS [24]	15-75	.80	6.99	3.15	3.36**	6.62	2.9	1.14 ^b	48.9	3.3	1.02 ^b	44.4	3.2	1.07 ^b	44.4
PAM-13 [30]	0-120	.92	20.00	5.66	7.93**	11.90	6.1	1.29 ^b	51.1 [#]	6.0	1.33 ^b	55.6 [#]	5.7	1.40 ^b	55.6
MHRM [28]	0-100	.76	14.21	6.96	5.70**	12.32	5.7	1.00 ^b	44.4	6.2	.92	44.4	7.0	.82	42.2
MANSA [34]	0-4	.90	.72	0.23	-1.17**	.51	-.23 [¶]	.74 [¶]	33.3 [¶]	.26	.66	28.9	0.23	.74	33.3
BSI [32]	1-7	.82	1.080	0.46	.22**	.58	.25	.89	48.9	.29	.75	42.2	0.46	.47	037.2
Rand- GHP [36]	0-100	.80	22.70	10.15	4.73*	17.43	15.5	.31	17.8	8.7	.54	37.8	10.2	0.47	31.1

BSI: Brief Symptom Inventory; ES: Effect-size; MANSA: Manchester Short Assessment of Quality of Life; MHRM: Mental Health Recovery Measure; MID: Minimal Important Difference; PAM: Patient Activation Measure; r_{xx}: test-retest coefficient; Rand-GHP: Rand General Health Perception; SD_c: standard deviation of the change scores; SD_T: standard deviation of the baseline scores; SEM: Standard Error of Measurement (=SD*√(1- r_{xx})); **p<0.01; *p<0.05; ^b>MID; [#] highest % improvement > MID; [§]>15% Flc Ceiling effects; [¶] including outliers

Table 3 Correlations between change scores and anchor candidates

Measures	MANSA-1	MANSA	Rand-HC	Rand-GHP
IMRS		.034	.548**	.103
PAM-13		.028	.190	.493**
MHRM		.206	.282	.491**
MANSA		.383**	.049	.495**
BSI		-.247	-.344*	-.450**
BSI [¶]		-.256	-.361*	-.405**

BSI: Brief Symptom Inventory; MANSA: Manchester Short Assessment of Quality of Life; MHRM: Mental Health Recovery Measure; PAM: Patient Activation Measure; Rand-GHP: Rand General Health Perception; Rand-HC: Rand Health Change; [¶] Based on outliers recoded to 2-SD_c; **p<.01; *p<.05

Table 4. Participants' scores on Rand-GHP divided in change groups

Change groups	Endpoint minus baseline scores	Frequency	Cumulative
Large negative change	-35	1	
	-30	0	
	-25	2	3
Small negative change	-20	2	
	-15	1	
	-10	4	7
No change	-5	6	
	0	7	
	5	5	18
Small positive change	10	4	
	15	5	
	20	1	10
Large positive change	25	1	
	30	1	
	35	4	
	40	1	7
Total		45	45

Rand-GHP: Rand General Health Perception

Figures

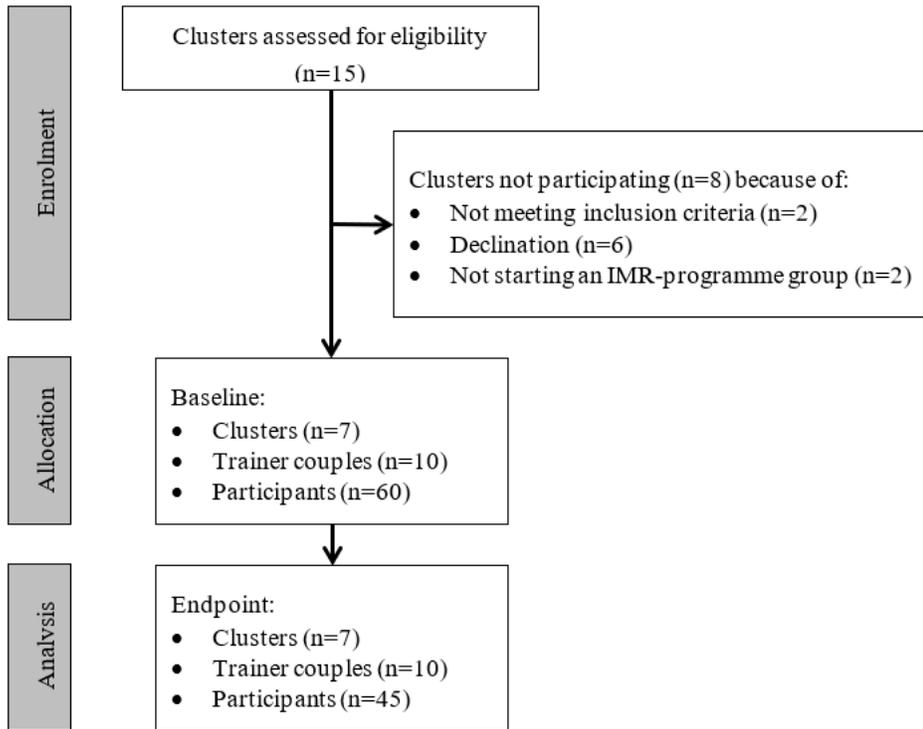


Fig. 1 Participants flow diagram

Figure 1

Participants flow diagram

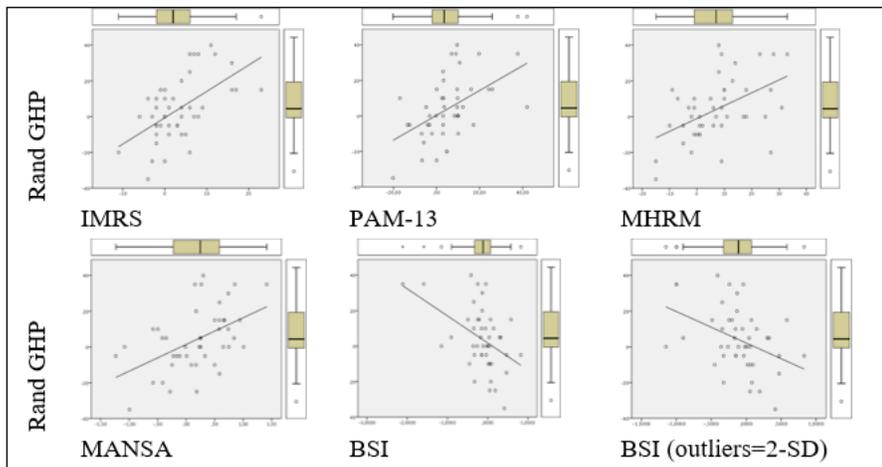


Figure 2

Box- and scatterplots from correlation calculations

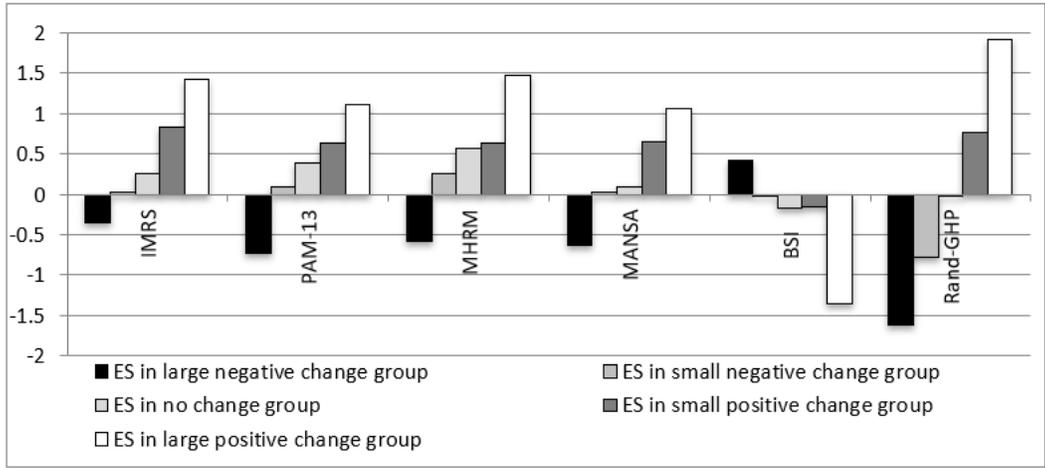


Figure 3

Effect sizes in the change-groups