Received: 30 September 2019 | Revised: 19 December 2019 | Accepted: 29 January 2020

DOI: 10.1002/mpr.1821

ORIGINAL ARTICLE

WILEY

Critical consideration of assessment methods for clinically significant changes of mental distress after psycho-oncological interventions

Lusine Vaganian[1] | Sonja Bussmann[1] | Alexander L. Gerlach[1] | Michael Kusch[2] | Hildegard Labouvie[2] | Jan C. Cwik[1]

[1]Department of Clinical Psychology and Psychotherapy, Faculty of Human Sciences, University of Cologne, Cologne, Germany

[2]Department of Internal Medicine, Section: Clinical Psycho-Oncology, Working Group Psycho-Oncological Health Services Research, University Hospital of Cologne, Cologne, Germany

Correspondence
Lusine Vaganian, Department of Clinical Psychology and Psychotherapy, University of Cologne, Pohligstr 1, 50969 Köln, Germany.
Email: lusine.vaganian@uni-koeln.de

Funding information
Innovation Fond of the Joint Federal Committee, Grant/Award Number: 01NVF17022-IsPo

Abstract

Objectives: Considering the heterogeneity of cancer entities and the associated disease progression, personalized care of patients is increasingly emphasized in psycho-oncology. This individualization makes the use of measurements of individual clinically significant change important when studying the efficacy and effectiveness of psycho-oncological care. Two conceptualizations for the measurement of clinical significance are critically contrasted in this study: the Reliable Change Index (RCI) and the Minimal Important Difference (MID) method.

Methods: In total, 2,121 cancer patients participated in the study and a subsample of 708 patients was reassessed about 4 months later. Psychological distress was measured using the Hospital Anxiety and Depression Scale. We evaluated two measures of clinical significance (RCI, MID) by comparing the respective numbers of improved, unimproved, and deteriorated patients.

Results: Individually significant changes were observed with both methods; however, determined rates of improvement differed substantially: MID (66.67%) and RCI (48.23%). Most importantly, according to MID, 17.93% of patients were identified as being improved, although their respective improvements were not statistically significant and thus unreliable.

Conclusions: The benefits of RCI outweigh MID, and therefore, the RCI is recommended as a measure to assess change.

KEYWORDS

clinical significance, hospital anxiety and depression scale, minimal important difference, psycho-oncology, reliable change index

1 | INTRODUCTION

Coping with the diagnosis of cancer, and mastering the associated tasks and changes, can be a significant challenge for patients. Physical and mental distress are often associated with cancer, and can deplete patients' quality of life, disease progression, and survival rates (Chan, Ahmad, Yusof, Ho, & Krupat, 2015; Karakas & Okanli, 2014; Linden, Vodermaier, MacKenzie, & Greig, 2012). The most common psychological consequences are anxiety and depression (Bussmann et al., 2018; Linden et al., 2012). Due to the heterogeneity of cancer entities

and the individual disease progression, the focus in recent years has increasingly been on patient-oriented medicine (Sinaiko et al., 2017). Consequently, it is required of the care system to provide the right patient at the right time with the right care at the right place (Kusch, Labouvie, Gerlach, Hellmich, & Hallek, 2016; Kusch, Labouvie, & Hein-Nau, 2013). In order to ensure evidence-based patient-centered care, care providers are also developing psycho-oncological programs that can be used to provide individualized quality-assured patient care, for example, continuous screenings for stress, psychoeducation, and stepped psycho-oncological treatments (Fann, Ell, & Sharpe, 2012; Forsythe et al., 2013).

Usually, statements of effectiveness are based on the analysis of its statistical significance. On a group level, statistical significance gives information about differences found in terms of a probability level lower than would be expected if occurring by chance (Page, 2014). In controlled and well-conducted group studies, significance testing provides meaningful and necessary evidence about the impact of specific interventions on a given population (Bothe & Richardson, 2011). However, the disadvantages of analyses on a group level are that very marginal differences can be statistically significant if the sample size is large enough (Hays, Brodsky, Johnston, Spritzer, & Hui, 2005; Hays, Spritzer, Sherbourne, Ryan, & Coulter, 2018). Furthermore, the results yield no information about individual change and, thus, cannot be used as an indicator of clinical significance (Bothe & Richardson, 2011; Lambert & Ogles, 2009). Similarly, when it comes to identifying responders to a particular treatment, the mere use of group-level significance can lead to misclassification of patients as a responder if they show no change on an individual level (Hays et al., 2018).

Therefore, in addition to statistical significance on a group level, the relevance of clinical significance and related concepts are increasingly being used to improve change measurement and clinical decision-making. These approaches are also increasingly used for the assessment and improvement of psycho-oncological care programs (Bedard et al., 2013; Guyatt, Osoba, Wu, Wyrwich, & Norman, 2002; Ogles, Lunnen, & Bonesteel, 2001). The concept of clinical significance represents the assessment of significant change on an individual level. The methods used to accomplish this are either distribution-based or anchor-based approaches (Ogles et al., 2001; Page, 2014; Wyrwich, Norquist, Lenderking, Acaster, & The Industry Advisory Committee of International Society for Quality of Life Research, 2013).

Many different concepts have been developed to assess clinical significance. One of the leading concepts is the Reliable Change Index (RCI) (Jacobson & Truax, 1991). The definition of clinical significance is that a patient has returned from a so-called dysfunctional (clinical) state to a so-called functional (healthy) state (Jacobson & Truax, 1991). To observe this, it is necessary to combine two criteria, a statistically significant change and a clinically significant change. Only based on both criteria, the individual change can be classified within defined categories. A patient is classified as "recovered," if the difference between the pre- and the posttest value is greater than the RCI (i.e., is statistically reliable), and if the posttest score has passed a predetermined cut-off point. Put another way, this classification may only take place if there is a statistically and clinically significant change. Accordingly, a patient is classified as "improved," if there is a statistically change, but the values did not pass the predetermined cut-off point. Thus, the patient's dysfunctional symptoms are still present subsequent to treatment. Furthermore, there is a category of patients who are classified as "unimproved" or "deteriorated." Unimproved means that patients revealed no statistical change, regardless of whether the cut-off point was crossed. Furthermore, patients who report a statistically significant worsening of symptoms are classified as "deteriorated."

However, besides the RCI, another concept of clinical significance, the minimal differences between two measurement points have been suggested, also known as Minimum/Minimal Important Difference (MID). Within this approach, the minimum significant difference or change for the patient should be represented as a score (Guyatt et al., 2002; Jaeschke, Singer, & Guyatt, 1989; Revicki et al., 2006; Wyrwich et al., 2013). The statistical significance is not a requirement for the calculation (in contrast to the RCI) (Page, 2014). First, the MID was defined as the "smallest difference in score in the domain of interest which patients perceive as beneficial and which would mandate, in the absence of troublesome side effects and excessive cost, a change in the patient's management" (Jaeschke et al., 1989, p. 408). The MID is calculated by the smallest significant difference between pre- and posttest value, which represents a "significant" change (Copay, Subach, Glassman, Polly, & Schuler, 2007; Revicki et al., 2006; Revicki, Hays, Cella, & Sloan, 2008). In order to define the significance of this change, the anchor-based approach compares the results with other measures using an external anchor or criterion, whereas distribution-based methods calculate the MID by using a measure of variability (Copay et al., 2007; Crosby, Kolotkin, & Williams, 2003). The advantage of the anchor-based methods is the comparison of results with an external anchor, while the advantage of distribution-based methods is that changes are presented free of random variations (Crosby et al., 2003). Each MID value for a given instrument may vary with regard to the studied population and the given context (Revicki et al., 2008). The aim of the MID is to provide feedback to the patient, as well as to the clinician, about the benefits and implications for further treatment. To balance the advantages and disadvantages of anchor-based and distributed-based approaches within the concept of MID, it is commonly recommended to calculate the MID using a combination of both methods (Bedard et al., 2013; Guyatt et al., 2002; Revicki et al., 2008). However, there is still no agreement which method or combination is the best (Guyatt et al., 2002). Because of the potential relevance of this decision for each individual patient, it is essential to make the right clinical decision. One study has already provided an overview about three different methods of clinical significance including standard error of measurement (*SEM*), standard error of prediction, and the RCI (Hays, Brodsky, et al., 2005). Note that the *SEM* is often used to calculate the MID (Ousmen et al., 2018). Whereas in the study of Hays, Brodsky, et al. (2005), the importance of examining the significant change at the individual level for improvement, consistency, or deterioration is emphasized, in many studies, significance on an individual level is often not ensured (Breitbart et al., 2015). Furthermore, no recommendation was

made with regard to which of the methods should be used. Because of the high relevance of clinical decision-making, this study aims to critically contrast the two measures of clinical significance (RCI, MID) based on the change in the symptoms of anxiety and depression by cancer patients due to psycho-oncological treatment.

2 | METHODS

2.1 | Participants and procedure

Data collection took part as part of a standardized program of the Clinic I of Internal Medicine (Clinical Psychology) in cancer patients of the Centre of Integrated Oncology Cologne and from the region. All participants provided written informed consent. The data were collected at two measurement time points. The first measurement time point (t1) was at the time of inpatient admission of the cancer patients and the second measurement time point (t2) was 4 months later. The questionnaire was handed out at t1 as part of the standardized care program (Kusch et al., 2014). At t2, the questionnaire was handed out again if patients stayed in the hospital or sent to patients by mail, if they were discharged from the inpatient unit before t2. In total, 2,121 cancer patients (1,643 women [77.5%]) with mean age of 53.02 (*SD* = 13.50) and mean Hospital Anxiety and Depression Scale total score (HADS-T) of 16.91 (*SD* = 8.56) participated in the study at t1 and 708 patients (582 women [82.2%]); mean age of 53.23 (*SD* = 13.00) and mean HADS-T of 13.67 (*SD* = 7.88) filled out the questionnaire a second time at t2. The cancer diagnoses among the participants are presented in Table 1. All procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. The work was approved by the Ethics Commission of Cologne University's Faculty of Medicine (reference number 15-048).

TABLE 1 Percentage of cancer diagnoses among participants at t1

Types of cancer	Percentage (%)
Breast	40.9
Lymphoid, hematopoietic, and related tissue	18.2
Female genital organs	9.1
Digestive organs	7.0
Respiratory and intrathoracic organs	3.9
Eye, brain and other parts of central nervous system	3.3
Male genital organs	1.7
Thyroid and other endocrine glands	1.6
Skin	1.5
Urinary tract	1.4
Mesothelial and soft tissue	1.3
Lip, oral cavity and pharynx	1.1
Ill-defined, secondary and unspecified sites	0.8
Bone and articular cartilage	0.4
Residual category (including different forms of cancer)	7.8

2.2 | Measures

Self-reported distress was measured by using the German version of the HADS (Herrmann-Lingen, Buss, & Snaith, 2011). The scale is an established tool for the assessment of anxiety and depression in cancer patients (Mitchell, Meader, & Symonds, 2010; Vodermaier & Millman, 2011). Furthermore, it is recommended as a screening tool for the measurement of psychological strain according to the S3-guidelines for the psycho-oncological management of adult cancer patients (Deutsche Krebsgesellschaft, Deutsche Krebshilfe, & AWMF, 2014). The HADS is a self-rating questionnaire and consists of 14 items with a total score ranging from 0 to 42. It is also possible to calculate subscales scores, but in psycho-oncological contexts, patients often show combined and fluctuating manifestations of anxious and depressive symptoms so that a global measure of the HADS-T can best represent the clinical situation (Herrmann-Lingen et al., 2011). Based on the HADS-T, cancer patients suffering from significant distress can be reliably distinguished from cancer patients without distress (Mitchell et al., 2010). Accordingly, the analyses of the present study will focus on the HADS-T score. There are different cut-off scores recommended for HADS-T. Specifically, in cancer patients, a sum score of HADS-T ≥ 15 can be used as the cut-off value to identify patients with an increased need for psycho-oncological care and especially for depression symptoms (Mitchell et al., 2010; Vodermaier & Millman, 2011). The psychometric properties of the HADS indicate a reliable and valid instrument (Bjelland, Dahl, Haug, & Neckelmann, 2002; Herrmann-Lingen et al., 2011). In the present study, Cronbach's alpha was excellent (HADS-T: $\alpha = 0.91$).

2.3 | Statistical analyses

Patients' characteristics are described by using means and standard deviations. Kolmogorov–Smirnov test results indicated that the data were not normally distributed. Nonetheless, the statistical measure of change on group level was examined by multivariate analysis of variance for repeated measures, since the *F*-test is robust in a large sample if the assumption of normal distribution is not met. Tests were calculated two-tailed with an assumed significance level of $p < .05$.

For the analysis of clinical significance, only cancer patients who had higher values at the first measurement time point than a critical threshold of HADS-T ≥ 15 were included. The rationale for this decision was that the present study is focused on the methodological comparison of two concepts of clinical change and whether these two concepts would influence clinical decision-making differentially. We presumed that individuals with relevant self-reported distress would more likely improve through psycho-oncological care allowing for the

best possible comparison of the two concepts. Note, however, that we do not want to imply or indicate that patients with values under the cut-off threshold of ≥15 could not potentially benefit from psycho-oncological care. The calculation formula for RCI used here is: $RCI = \frac{Y-X}{Std\,(Y-X)} = \frac{Y-X}{Std(X)\sqrt{2(1-Rel(X))}}$.

According to this formula, the RCI is calculated as the difference between the pre- (Y) and posttest (X) values divided by the standard measurement error of the difference (Jacobson & Truax, 1991). As a measure of reliability, α is recommended and will be used for the test instrument and the examined sample (Lambert & Ogles, 2009).

A RCI > 1.96 corresponds to a 95% confidence interval and indicates that the individual change is statistically significant. In the present case, the change between the pre- and posttest values had to be greater than or equal to five to indicate a statistically significant change for each individual patient. To assess whether the change is also clinically significant, the postvalue must pass a predetermined cut-off point, which separates the dysfunctional from the functional population (Jacobson & Truax, 1991). This cut-off point can be defined in three different ways (Jacobson & Truax, 1991; Lambert & Ogles, 2009). In the present study, the most conservative cut-off point was used. This cut-off point is based on information from functional and dysfunctional population and is recommended to determine the significant clinical change if functional standards are known. We used the calculation formula for this cut-off point: $c = \frac{(SDpatient\ Mnonpatient) + (SDnonpatient\ Mpatient)}{(SDpatient + SDnonpatient)}$ (Lambert & Ogles, 2009). Thus, data from a healthy population were used to define the cut-off point (HADS-T: $M = 9.45$; $SD = 6.80$) (Hinz & Brähler, 2011). In consequence according to the calculation formula, the cut-off point of HADS-T = 16.52 was used in addition to the RCI to determine significant clinical change in the present study.

In order to define clinically significant change employing MID, studies were searched in which an MID was established for the HADS-T. We found one study that had calculated an MID of 1.5 for the HADS-T in a sample of patients with chronic obstructive pulmonary disease (Puhan, Frey, Büchi, & Schünemann, 2008). The MID of 1.5 was calculated using an anchor-based approach employing a linear regression analysis. Two self-report instruments that assessed the burden of disease (symptomatology and affect) were used as anchors. In the same study, an effect size of 0.5 *SD* units of change score was additionally used to estimate a MID of 1.17 for the HADS-T. In this case, the MID was estimated based on a distribution-based method as an alternative to the anchor-based approach (Puhan et al., 2008). Investigators regularly consider an effect size of 0.5 *SD* units as an adequate estimate of clinically significant change (Walters & Brazier, 2003). Given that the HADS-T results only in positive integers, no actual difference results in using the anchor- or distribution-based method. More specifically, the MID between pre- and postmeasurement always had to be ≥2-point change on the HADS-T.

3 | RESULTS

At t1 1,251 of all 2,121 patients (59.0%) exceeded a HADS-T score ≥ 15, indicating relevant distress; the mean HADS-T score of these patients was 22.61 (*SD* = 5.85). From all 708 patients who took part at both measurement time points, 396 patients (55.9%), who exceeded a HADS-T ≥ 15, reported a mean HADS-T at t1 of 22.08 (*SD* = 5.59) and at t2 of 17.40 (*SD* = 7.16). Significant differences between the group who took part at only t1 (group 1; 1,413 patients) and the group who took part at both t1 and t2 (group 2; 708 patients) were found with respect to gender (a higher proportion of women in group 2; $\chi^2[1] = 13.68$, $p < .001$) and their respective HADS-T scores (higher scores in group 1; $t[2119] = 2.75$, $p = .006$) at measurement time point t1. There were no differences between the groups with regard to age ($t[2118] = -0.488$, $p = .626$). Overall, there was a statistically significant improvement on the group level (HADS-T: $F[1,707] = 95.35$, $p < .001$; $\eta_p^2 = .119$) from t1 to t2 in all patients of group 2. Furthermore, patients who exceeded the critical threshold of the HADS-T score of ≥15 at the time of pre-examination showed also a significant change at t2 ($F[1,395] = 176.75$, $p < .001$; $\eta_p^2 = .309$), illustrating an even stronger effect of treatment in this subgroup of patients of group 2.

For the analysis of statistical and clinical significance on an individual level, first the RCI was calculated. Additionally, an external MID value was used as alternative reference. In order to better illustrate the effects of using the two different types of clinically significant change,

TABLE 2 Frequency and percentage of different change based on RCI and MID for cancer patients for HADS-T from pre- to postexamination

	MID—deteriorated (pre–post difference ≤ −2)	MID—unimproved (2 > pre–post difference > −2)	MID—improved (pre–post difference ≥ 2)	Total
RCI—deteriorated (pre–post difference ≤ −5)	29 (7.32%)	0	0	29 (7.32%)
RCI—unimproved (5 > pre–post difference > −5)	39 (9.85%)	64 (16.16%)	71 (17.93%)	174 (43.94%)
RCI—improved (pre–post difference ≥ 5)	0	0	193 (48.74%)	193 (48.74%)
Total	68 (17.17%)	64 (16.16%)	264 (66.67%)	396 (100%)

in this analysis, only patients who reached or exceeded the critical threshold for psychological strain as measured by HADS-T scores ≥15 at the time of the pre-examination were included (i.e., the group of patients for whom psycho-oncological treatment is indicated).

According to RCI concept, 193 of 396 (48.74%) cancer patients exhibited a statistically reliable change of self-reported distress (improved; see Table 2). Statistically significant and clinically significant change in symptoms was seen in 191 cancer patients (48.23%) (recovered). In turn, 29 (7.32%) patients showed a reliable worsening of symptoms (deteriorated) and 174 (43.94%) patients showed neither statistical nor clinically significant changes (unimproved). The estimated change of ≥2 points on the HADS-T was used as a cut-off point to assess the clinical change using the MID concept. For HADS-T, 264 of the 396 analyzed patients (66.67%) achieved a change of at least ≥2 points in HADS-T between t1 and t2. However, among the 264 patients, 71 patients (17.93%) did not meet the criteria of a reliable statistical change and 73 patients (18.43%) of significant clinical change according to the concept of the RCI. Table 2 illustrates this pattern of results using a crosstab to give an overview on patients' changes with regard to depressive and anxiety symptoms as determined by using the RCI and MID concept.

4 | DISCUSSION

Before interpreting the data, it should be called to mind that the aim of this study is not the evaluation of a specific psycho-oncological treatment package, but rather a comparison of two commonly used methods of determining clinical significance. The efficacy of treatment is not emphasized in the present study, so the relatively high dropout rate and the missing control group are not exceptional given the naturalistic sample. Thus, this study aimed to critically contrast two concepts for the measurement of clinical significance (RCI and MID) within a sample of cancer patients with respect to the identifiable ratio of individual improvement and deterioration in symptoms between the methods.

Overall, the patients with signs of substantial psychological stress before treatment showed significant improvement in symptoms of anxiety and depression. These results are in line with other studies, which previously highlighted the potentially highly positive effects of integrative psycho-oncological treatment on anxiety and depression in cancer patients (Grassi, Spiegel, & Riba, 2017; Kost et al., 2009). To transfer these group-level results to measurements of clinical relevance for each individual patient, RCI and MID were calculated and critically compared. Both measurements of clinical change supported the claim of patient-oriented psycho-oncological care as being efficacious. Based on the RCI concept, 48.74% of patients exhibited a reliable statistical change (improvement) with regard to symptoms of anxiety and depression. The results of the analysis are in line with other studies focusing on individual case analyses (Grassi et al., 2017; Kost et al., 2009). In addition, 48.23% of patients reported a statistically and clinically significant change (recovered) based on their scores on the HADS-T. On the downside, 29 (7.32%) patients reported statistically increased levels of depression and anxiety.

The estimation of change and clinical significance was also performed employing the MID. According to the MID concept, 66.67% of patients showed a clinically meaningful improvement. Thus, compared with the RCI, the estimated number of patients with clinically significant improvements was clearly higher when using MID. However, 71 patients (17.93%), who supposedly had improved, did not meet the requirements for a statistically reliable change. Moreover, 73 patients (18.43%) did not meet the criteria for significant clinical change according to the RCI concept (i.e., rather belong to a healthy as compared to a psychologically stressed population). In addition, 39 patients (9.85%) were classified as deteriorated according to MID. Although again, this numerical change was not statistically reliable. Thus, on the one hand, the MID value as a cut-off point appears less stringent than the RCI's cut-off points, and on the other hand, it is not ensured that identified changes are indeed reliable changes and consequently there is a substantial risk of overestimating the results when studying the effects of an intervention. Interestingly, using MID did not result in a larger group of individuals being detected whose psychological stress had deteriorated. Thus, the lower threshold of the MID only resulted in a more liberal detection of improvement and not of deterioration.

An important advantage of RCIs compared with the MID is that statistical significance of each individual change is a requirement for determining clinical significance. Or with other words, clinical significance is always statistically safeguarded (Lambert & Ogles, 2009). Also, the calculation of an individual's clinical significance using the RCI concept is quick and efficient (Lambert & Ogles, 2009). A criticism is that the RCI is a rather conservative method and the cut-off points used are relatively strict criteria. Individuals with low initial symptom severity have little chance of undergoing a clinically relevant change as they have little room for improvement (Lambert & Ogles, 2009). However, we believe that being conservative when determining the benefit of an intervention is a merit rather than being problematic. Before ending treatment, it should be safely ensured that a person indeed has improved rather than discharging a patient who only ostensibly has improved. Another potential problem with the MID was not much of a problem in our sample due to the instrument we used. There are a number of different definitions for the calculation of the MID, which often have different aims (King, 2011). For the approach of the anchor-based methods as well as for the distribution-based methods, different authors suggest different types of calculations (King, 2011; Revicki et al., 2008). In a recent structured review, Ousmen et al. (2018) reported that the most commonly used distribution-based method was the 0.5 *SD*, followed by the *SEM*. Note that as the best approximation, a combination of both approaches is recommended (Bedard et al., 2013; Guyatt et al., 2002). However, this approach rarely used in practice (Ousmen et al., 2018). Furthermore, Hays, Farivar, and Liu (2005) highlighted that despite the fact that several measures (e.g., *SEM*) are related to the MID, these measures do not provide direct information about the MID (e.g., new information about the size of change) and that their use should therefore be discouraged. In our sample, given that only integer numbers are calculated as individual scores on the HADS-T, these different MID scores

resulted always in the same HADS-T change score of 2. Thus, in the present study, it seemed rather inconsequential, which method to use.

4.1 | Study limitations

Besides the strength of large sample size, examination on patients' reports, and the benefits for clinical and research practice, the study also has limitations. One limitation is that the difference from pre- to postexamination does not necessarily have to be related solely to psycho-oncological treatment. Frequently, patients respond to cancer diagnosis with signs of anxiety and depression, which often remit spontaneously without the need for psycho-oncological support (Cook, Salmon, Hayes, Byrne, & Fisher, 2018). To examine this aspect in a study with an additional randomized control group would be interesting and important. Additionally, the data showed a very high drop-out rate of approximately 66.6% in the period between pre- (2,121 cancer patients) and postexamination (708 cancer patients). This can primarily be explained by the fact that the data collection did not take place as part of a research project, but within routine care. It is possible that patients switched to outpatient care were no longer in treatment or even died. This was possibly especially likely because the data were collected in a highly specialized treatment facility, recruiting patients from a relative wide catchment area. Nonetheless, there are differences between the patients who dropped out and who took part at both measurement time points with regard to gender and the initial HADS-T score. To control these differences would be interesting and important in further studies. Note, however, that the differences on the HADS-T are significant but that the mean HADS-T values of the drop-out group (M = 17.27, SD = 8.69) and the nondrop-out group (M = 16.18, SD = 8.24) nonetheless had higher values than the critical threshold of HADS-T ≥ 15. Arguably, the self-reported psychological distress was clearly relevant in both groups. Moreover, this group difference had no direct influence on the analysis of possible differences between the individual change measurements.

4.2 | Clinical implications

Concepts to determine the clinical significance of treatments are increasingly being used for adequate change measurement and clinical decision-making. These approaches are also increasingly used in the assessment and improvement of psycho-oncological care. Therefore, it is important to critically assess which method can be used for a clinical decision. Although the MID ostensibly shows a higher percentage of improvements, it is statistically unreliable and as a consequence of its use, an overestimation of the effects of this form of intervention is possible. In addition, errors in clinical decision-making may result if patients' treatments are ended prematurely. Due to these weaknesses of the MID and the substantive advantages of the RCI, the RCI is recommended as a measure of change in the care research of cancer patients.

ACKNOWLEDGMENTS

The preparation of this manuscript was partially funded by the Innovation Fond of the Joint Federal Committee in the form of the project "isPO—Integrated, cross-sectoral psycho-oncology" (grant number: 01NVF17022-IsPo).

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

ORCID

Lusine Vaganian https://orcid.org/0000-0001-9718-4106
Alexander L. Gerlach https://orcid.org/0000-0001-6794-5349
Jan C. Cwik https://orcid.org/0000-0002-2290-353X

REFERENCES

Bedard, G., Zeng, L., Zhang, L., Lauzon, N., Holden, L., Tsao, M., … Chow, E. (2013). Minimal clinically important differences in the Edmonton symptom assessment system in patients with advanced cancer. *Journal of Pain and Symptom Management*, *46*(2), 192–200. https://doi.org/10.1016/j.jpainsymman.2012.07.022

Bjelland, I., Dahl, A. A., Haug, T. T., & Neckelmann, D. (2002). The validity of the hospital anxiety and depression scale: An updated literature review. *Journal of Psychosomatic Research*, *52*(2), 69–77. https://doi.org/10.1016/S0022-3999(01)00296-3

Bothe, A. K., & Richardson, J. D. (2011). Statistical, practical, clinical, and personal significance: Definitions and applications in speech-language pathology. *American Journal of Speech-Language Pathology*, *20*(3), 233–242. https://doi.org/10.1044/1058-0360(2011/10-0034)

Breitbart, W., Rosenfeld, B., Pessin, H., Applebaum, A., Kulikowski, J., & Lichtenthal, W. G. (2015). Meaning-centered group psychotherapy: An effective intervention for improving psychological well-being in patients with advanced cancer. *Journal of Clinical Oncology*, *33*(7), 749–754. https://doi.org/10.1200/JCO.2014.57.2198

Bussmann, S., Vaganian, L., Kusch, M., Labouvie, H., Gerlach, A. L., & Cwik, J. C. (2018). Angst und Emotionsregulation bei Krebspatienten. *Psychotherapeut*, *63*(3), 194–203. https://doi.org/10.1007/s00278-018-0283-3

Chan, C. M. H., Ahmad, W. A. W., Yusof, M. M., Ho, G. F., & Krupat, E. (2015). Effects of depression and anxiety on mortality in a mixed cancer group: A longitudinal approach using standardised diagnostic interviews. *Psycho-Oncology*, *24*(6), 718–725. https://doi.org/10.1002/pon.3714

Cook, S. A., Salmon, P., Hayes, G., Byrne, A., & Fisher, P. L. (2018). Predictors of emotional distress a year or more after diagnosis of cancer: A systematic review of the literature. *Psycho-Oncology*, *27*(3), 791–801. https://doi.org/10.1002/pon.4601

Copay, A. G., Subach, B. R., Glassman, S. D., Polly, D. W., & Schuler, T. C. (2007). Understanding the minimum clinically important difference: A review of concepts and methods. *The Spine Journal*, *7*(5), 541–546. https://doi.org/10.1016/j.spinee.2007.01.008

Crosby, R. D., Kolotkin, R. L., & Williams, G. R. (2003). Defining clinically meaningful change in health-related quality of life. *Journal of Clinical Epidemiology*, *56*, 395–407. https://doi.org/10.1016/s0895-4356(03)00044-1

Deutsche Krebsgesellschaft, Deutsche Krebshilfe, & AWMF. (2014). *S3-Leitlinie Psychoonkologische Diagnostik Beratung und Behandlung von erwachsenen Krebspatienten*. Berlin, Germany: Deutsche Krebsgesellschaft, Deutsche Krebshilfe, AWMF.

Fann, J. R., Ell, K., & Sharpe, M. (2012). Integrating psychosocial care into cancer services. *Journal of Clinical Oncology*, *30*(11), 1178–1186. https://doi.org/10.1200/jco.2011.39.7398

Forsythe, L. P., Rowland, J. H., Padgett, L., Blaseg, K., Siegel, S. D., Dingman, C. M., & Gillis, T. A. (2013). The cancer psychosocial care

matrix: A community-derived evaluative tool for designing quality psychosocial cancer care delivery. *Psycho-Oncology*, *22*(9), 1953–1962. https://doi.org/10.1002/pon.3254

Grassi, L., Spiegel, D., & Riba, M. (2017). Advancing psychosocial care in cancer patients. *F1000Research*, *6*, 2083. https://doi.org/10.12688/f1000research.11902.1

Guyatt, G., Osoba, D., Wu, A. W., Wyrwich, K. W., & Norman, G. R. (2002). Methods to explain the clinical significance of health status measures. *Mayo Clinic Proceedings*, *77*, 371–383. https://doi.org/10.4065/77.4.371

Hays, R. D., Brodsky, M., Johnston, M. F., Spritzer, K. L., & Hui, K.-K. (2005). Evaluating the statistical significance of health-related quality-of-life change in individual patients. *Evaluation & the Health Professions*, *28*(2), 160–171. https://doi.org/10.1177/0163278705275339

Hays, R. D., Farivar, S. S., & Liu, H. (2005). Approaches and recommendations for estimating minimally important differences for health-related quality of life measures. *COPD: Journal of Chronic Obstructive Pulmonary Disease*, *2*(1), 63–67. https://doi.org/10.1081/COPD-200050663

Hays, R. D., Spritzer, K. L., Sherbourne, C. D., Ryan, G. W., & Coulter, I. D. (2018). Group and individual-level change on health-related quality of life in chiropractic patients with chronic low Back or neck pain. *Spine*, *44*(9), 647–651. https://doi.org/10.1097/brs.0000000000002902

Herrmann-Lingen, C., Buss, U., & Snaith, P. R. (2011). *Hospital Anxiety and Depression Scale-Deutsche Version (HADS-D)*. Bern: Huber.

Hinz, A., & Brähler, E. (2011). Normative values for the hospital anxiety and depression scale (HADS) in the general German population. *Journal of Psychosomatic Research*, *71*, 74–78. https://doi.org/10.1016/j.jpsychores.2011.01.005

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*(1), 12–19. https://doi.org/10.1037/0022-006X.59.1.12

Jaeschke, R., Singer, J., & Guyatt, G. H. (1989). Measurement of health status: Ascertaining the minimal clinically important difference. *Controlled Clinical Trials*, *10*(4), 407–415. https://doi.org/10.1016/0197-2456(89)90005-6

Karakas, S. A., & Okanli, A. (2014). The relationship between meaning of illness, anxiety, depression, and quality of life for cancer patients. *Collegium Antropologicum*, *38*(3), 939–944.

King, M. T. (2011). A point of minimal important difference (MID): A critique of terminology and methods. *Expert Review of Pharmacoeconomics & Outcomes Research*, *11*(2), 171–184. https://doi.org/10.1586/erp.11.9

Kost, C., Labouvie, H., Kosfelder, J., Höhl, U., Herrmann-Lingen, C., & Kusch, M. (2009). Veränderung der psychischen Belastung von Krebspatienten im Rahmen einer strukturierten psychoonkologischen Versorgung. [change in psychological distress of cancer in-patients attending a structured psycho-oncological care program]. *Psychotherapie Psychosomatik Medizinische Psychologie*, *59*(12), 432–439. https://doi.org/10.1055/s-2008-1067575

Kusch, M., Labouvie, H., Gerlach, A. L., Hellmich, M., & Hallek, M. (2016). Kommt die personalisierte Psychoonkologie? *Tumor Diagnostik Und Therapie*, *37*(10), 560–564. https://doi.org/10.1055/s-0042-118685

Kusch, M., Labouvie, H., & Hein-Nau, B. (2013). *Klinische Psychoonkologie*. Berlin: Springer-Verlag.

Kusch, M., Labouvie, H., Hein-Nau, B., Schwarzkamp, U., Wolf, J., & Hallek, M. (2014). Integrierte Psychoonkologie: Implementierung eines psychoonkologischen Versorgungsprogramms am centrum für Integrierte Onkologie Köln Bonn, Standort Köln. [integrated psychooncology: Implementation of psychooncological health care at the Center for Integrated Oncology Cologne—Bonn]. *Deutsche Medizinische Wochenschrift*, *139*(46), 2357–2360. https://doi.org/10.1055/s-0034-1387384

Lambert, M. J., & Ogles, B. M. (2009). Using clinical significance in psychotherapy outcome research: The need for a common procedure and validity data. *Psychotherapy Research*, *19*(4–5), 493–501. https://doi.org/10.1080/10503300902849483

Linden, W., Vodermaier, A., MacKenzie, R., & Greig, D. (2012). Anxiety and depression after cancer diagnosis: Prevalence rates by cancer type, gender, and age. *Journal of Affective Disorders*, *141*(2), 343–351. https://doi.org/10.1016/j.jad.2012.03.025

Mitchell, A. J., Meader, N., & Symonds, P. (2010). Diagnostic validity of the hospital anxiety and depression scale (HADS) in cancer and palliative settings: A meta-analysis. *Journal of Affective Disorders*, *126*(3), 335–348. https://doi.org/10.1016/j.jad.2010.01.067

Ogles, B. M., Lunnen, K. M., & Bonesteel, K. (2001). Clinical significance: History, application, and current practice. *Clinical Psychology Review*, *21*(3), 421–446. https://doi.org/10.1016/S0272-7358(99)00058-6

Ousmen, A., Touraine, C., Deliu, N., Cottone, F., Bonnetain, F., Efficace, F., … Anota, A. (2018). Distribution- and anchor-based methods to determine the minimally important difference on patient-reported outcome questionnaires in oncology: A structured review. *Health and Quality of Life Outcomes*, *16*(1), 228. https://doi.org/10.1186/s12955-018-1055-z

Page, P. (2014). Beyond statistical significance: Clinical interpretation of rehabilitation research literature. *International Journal of Sports Physical Therapy*, *9*(5), 726–736.

Puhan, M. A., Frey, M., Büchi, S., & Schünemann, H. J. (2008). The minimal important difference of the hospital anxiety and depression scale in patients with chronic obstructive pulmonary disease. *Health and Quality of Life Outcomes*, *6*, 46–51. https://doi.org/10.1186/1477-7525-6-46

Revicki, D., Hays, R. D., Cella, D., & Sloan, J. (2008). Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *Journal of Clinical Epidemiology*, *61*(2), 102–109. https://doi.org/10.1016/j.jclinepi.2007.03.012

Revicki, D. A., Cella, D., Hays, R. D., Sloan, J. A., Lenderking, W. R., & Aaronson, N. K. (2006). Responsiveness and minimal important differences for patient reported outcomes. *Health and Quality of Life Outcomes*, *4*(1), 70–74. https://doi.org/10.1186/1477-7525-4-70

Sinaiko, A. D., Landrum, M. B., Meyers, D. J., Alidina, S., Maeng, D. D., Friedberg, M. W., … Rosenthal, M. B. (2017). Synthesis of research on patient-centered medical homes brings systematic differences into relief. *Health Affairs (Project Hope)*, *36*(3), 500–508. https://doi.org/10.1377/hlthaff.2016.1235

Vodermaier, A., & Millman, R. D. (2011). Accuracy of the hospital anxiety and depression scale as a screening tool in cancer patients: A systematic review and meta-analysis. *Supportive Care in Cancer*, *19*(12), 1899–1908. https://doi.org/10.1007/s00520-011-1251-4

Walters, S. J., & Brazier, J. E. (2003). What is the relationship between the minimally important difference and health state utility values? The case of the SF-6D. *Health and Quality of Life Outcomes*, *1*, 4. https://doi.org/10.1186/1477-7525-1-4

Wyrwich, K. W., Norquist, J. M., Lenderking, W. R., Acaster, S., & the industry advisory Committee of International Society for quality of life research (ISOQOL). (2013). Methods for interpreting change over time in patient-reported outcome measures. *Quality of Life Research*, *22*(3), 475–483. https://doi.org/10.1007/s11136-012-0175-x

How to cite this article: Vaganian L, Bussmann S, Gerlach AL, Kusch M, Labouvie H, Cwik JC. Critical consideration of assessment methods for clinically significant changes of mental distress after psycho-oncological interventions. *Int J Methods Psychiatr Res*. 2020;29:e1821. https://doi.org/10.1002/mpr.1821