Derivation of a minimal clinically important difference score for the WHODAS 2.0 in mild traumatic brain injury

Deborah L. Snell[a] and Noah D. Silverberg[b],*

[a]*Department of Orthopaedic Surgery and Musculoskeletal Medicine, University of Otago Christchurch, Christchurch, New Zealand*
[b]*Department of Psychology, University of British Columbia, Vancouver, BC, Canada*

Abstract.

BACKGROUND: Mild traumatic brain injury (mTBI) treatment research is hindered by lack of clinically meaningful and responsive outcome measures. One promising measure is the World Health Organisation Disability Assessment Schedule 2.0 (WHODAS 2.0), although minimal clinically important differences (MCID) for have not been established.

OBJECTIVE: To estimate MCID for the WHODAS 2.0 for mTBI.

METHODS: We analysed two prospectively collected mTBI datasets ($n = 225$) attending adult outpatient clinics in British Columbia, Canada. Participants completed the 12-item WHODAS 2.0, Patient Global Impression of Change scale, and Rivermead Post-Concussion Symptoms Questionnaire. We used anchor- and distribution-based methods to explore MCIDs in WHODAS 2.0 scores.

RESULTS: For Study 1 ($n = 131$), the anchor and distribution-based approaches produced minimal change estimates ranging from 1.3 to 2.8 interval scores. For Study 2 ($n = 94$), the anchor and distribution-based approaches produced minimal change estimates from 2.2 to 3.2 interval scores. For certain subgroups based on age, sex, and post-concussion severity, minimal change estimates were slightly higher.

CONCLUSION: An MCID of 3.5 interval WHODAS 2.0 points would conservatively capture meaningful change in adults of varying age, sex, and post-concussion symptom severity. Such a uniform metric will assist future mTBI intervention studies to improve standards of care and evaluation of outcomes.

Keywords: WHODAS 2.0, mTBI, mild traumatic brain injury, MCID

1. Introduction

Mild traumatic brain injury (mTBI) treatment research has likely been stalled by a lack of clinically meaningful and responsive outcome measures. Recent systematic reviews of treatments for mTBI continue to highlight diverse outcomes and outcome measures across studies, making it difficult to synthesize evidence in systematic reviews (Arbabi et al., 2020; Horton, Rhodes, & Wilson, 2018). The most commonly used measure of global functioning in traumatic brain injury research is the Glasgow Outcome Scale-Extended (GOSE: (Wilson, Pettigrew, & Teasdale, 1998)), a structured interview that outputs a single ordinal rating on an 8-point scale from dead to complete recovery. The GOSE has been criticized for being insensitive to subtle changes in functioning, making it suboptimal for mTBI clinical trials (Alali

*Address for correspondence: Noah D. Silverberg, PhD, The University of British Columbia, 3505-2136 West Mall, Vancouver V6T 1Z4, BC, Canada. E-mail: noah.silverberg@ubc.ca.

et al., 2015; Horton et al., 2018; Menon & Maas, 2015). Evaluation and validation of alternate outcome measures for this patient population are needed.

The Word Health Organisation Disability Assessment Schedule (WHODAS 2.0) (Federici, Bracalenti, Meloni, & Luciano, 2016; Üstün, Kastanjsek, Chatterji, & Rehm, 2010) is a generic measure of disability that may be suitable as a clinical endpoint for mTBI clinical trials. The 12-item version has encouraging psychometric properties in the mTBI population (Snell, Iverson, Panenka, & Silverberg, 2017; Snell, Siegert, & Silverberg, 2020), as well as in related health conditions such as chronic pain and anxiety.(Axelsson, Lindsater, Ljotsson, Andersson, & Hedman-Lagerlöf, 2017; Saltychev et al., 2016) The 12-item WHODAS 2.0 focuses on six domains of functioning derived from the International Classification of Functioning, Disability and Health (ICF), where disability has been defined as the interaction between a person's health condition and their context, resulting in activity limitations and participation restrictions (http://www.who.int/classifications/icf/whodasii/en/). While the WHODAS 2.0 has been widely used in both clinical practice and research contexts across a range of populations including mTBI, the interpretation of scores including responsiveness to change and minimally clinically important differences (MCID) have not yet been well defined (Katajapuu, Heinonen, & Saltychev, 2020).

The MCID is an indicator of the minimal change in scores that is considered clinically relevant and has been defined as the smallest difference in the domain of interest that is perceived as beneficial by patients and clinicians (Jaeshke, Singer, & Guyatt, 1989). Knowledge of the MCID helps to inform the target difference for a clinical trial and contextualizes treatment group differences. An intervention that is similar to current standard of care with regard to cost and side effect profile might be considered useful if the estimated between-group difference is greater than the MCID. An intervention that has a less favourable side effect profile or is more costly would need to show an effect larger than the MCID to be considered (Barrett, Brown, Mundt, & Brown, 2005).

There are a range of methods for estimating MCIDs and these can be broadly categorised into three conceptually different methods: opinion-based, distribution-based and anchor-based methods (Mouelhi, Jouve, Castelli, & Gentile, 2020). There are limitations to both anchor and distribution-based methods. Anchor-based approaches risk recall bias,

with current state potentially influencing retrospective ratings of a person's previous state (Shulman, Kasza, & Myles, 2020; Wijeysundera & Johnson, 2016). On the other hand, distribution-based methods lack external reference points (Revicki, Haysb, Cellac, & Sloan, 2008; Shulman et al., 2020). Using anchor-based methods to provide primary evidence for the estimation of the MCID, supported by distribution-based approaches, may optimally integrate these sources of information (Revicki et al., 2008; Shulman et al., 2020) and produce change estimates that are patient-centred and explicitly linked to self-reported experience (Wijeysundera & Johnson, 2016).

There are a few studies that have attempted to estimate the MCID for the 12-item WHODAS 2.0 in patient populations such as anxiety disorders (Axelsson et al., 2017), chronic musculoskeletal pain (Katajapuu et al., 2020), and peri- and post-surgical populations (Shulman et al., 2020). However no studies have yet considered MCID for the WHODAS 2.0 in mTBI samples. Some authors have expressed concerns regarding the multidimensionality of the 12-item WHODAS 2.0, which may limit the responsiveness of this scale (Katajapuu et al., 2020). We have previously shown that the 12-item WHODAS 2.0 met the requirements for unidimensionality using the Rasch model in an mTBI sample (Snell et al., 2020), whereas prior studies in other health conditions analysed raw (ordinal) total scores, using interval scores derived using the Rasch model (Snell et al., 2020), which offer an opportunity to estimate MCIDs for mTBI with greater precision.

1.1. Study objectives

The aim of this study was to explore to changes in WHODAS 2.0 scores and patient perceptions of the importance of these changes using a Rasch derived interval scoring guide from an earlier validation study (Snell et al., 2020).

2. Materials and methods

2.1. Study design

The present study is an exploratory secondary analysis of two prospectively collected datasets. Both original studies received ethical approval from the University of British Columbia Behavioral Research Ethics Board and operational approval from the

Table 1
Eligibility criteria for studies 1 and 2

Eligibility criterion	Study 1	Study 2
Age	18–60 years	18–70 years
Injury-recruitment timeframe	mTBI <3 months prior	mTBI 1-12 months prior
Injury diagnosis	Per referring clinician	WHO Neurotrauma Taskforce definition
Language	Fluent in English	Fluent in English
Symptoms	At least 3 RPQ symptoms	≥3 moderate-severe symptoms on RPQ

mTBI=mild traumatic brain injury. WHO (World Health Organisation) Neurotrauma Taskforce definition (Holm, Cassidy, Carroll, & Borg, 2005). RPQ = Rivermead Post Concussion Symptom Questionnaire (King et al., 1995).

Vancouver Coastal Health Research Institute and the Fraser Health Research Institute.

i) Study 1

Study 1 (Snell et al., 2020) involved 131 treatment-seeking adults, recruited from two outpatient clinics in Vancouver, Canada. Study 1 eligibility criteria are shown in Table 1. Litigation was not an exclusion criterion. Recruited participants were assessed on two occasions, on average 10 and 20 weeks following their mTBI. Recruitment methods are described in more detail elsewhere (Silverberg et al., 2020).

ii) Study 2:

Ninety-four new participants were recruited from the same two outpatient clinics in Vancouver, Canada (Silverberg et al., 2021). The Study 2 eligibility criteria are also shown in Table 1. Participants were assessed at on average of 18 and 32 weeks following mTBI. This dataset included participants who met additional eligibility criteria for an intervention study (clinicaltrials.gov NCT03972579) as well as those who did not, the latter enrolled in an observational arm. All participants in the dataset had unrestricted access to usual clinical care and litigation was not an exclusion criterion. Those enrolled in the intervention study received additional specialized treatment from occupational therapy and psychology providers.

2.2. Measures and variables

i) WHODAS 2.0 (12-item version) (Üstün et al., 2010)

Both studies used the 12-item WHODAS 2.0 (interview version). This measure evaluates disability across six ICF activity and participation domains including cognition, mobility, self-care, interpersonal functioning, life activities, and participation. The WHODAS 2.0 asks respondents how much difficulty they have had in the past 30 days in relation to their health problems for each of the 12 items, with options: 0 = none, 1 = mild, 2 = moderate, 3 = severe,

4 = extreme/cannot do (high scores mean greater disability). The 12-item WHODAS 2.0 has been validated for mTBI and demonstrated high internal consistency (Cronbach's alpha > 0.9) (Snell et al., 2017; Snell et al., 2020). Total scores were calculated and then converted to interval scores (range 0-48) using Rasch tables (Snell et al., 2020). Total interval scores could only be calculated for complete data sets and no imputation for missing data points was undertaken.

ii) Rivermead Post-Concussion Symptom Questionnaire (RPQ: (King, Crawford, Wenden, Moss, & Wade, 1995))

Both studies had RPQ data available. This is a 16-item self-report symptom questionnaire including common symptoms following mTBI. The RPQ queries somatic symptoms (headaches, feelings of dizziness, nausea and vomiting, noise sensitivity, sleep disturbance, light sensitivity and double vision); cognitive symptoms (forgetfulness/ poor memory, poor concentration and taking longer to think); and emotional symptoms (being irritable/easily angered, feeling depressed or tearful, feeling frustrated or impatient). Participants rate the presence and problem status of these symptoms on a scale of 0–4 (0 = not experienced at all; 1 = no more of a problem than before the injury; 2 = a mild problem; 3 = a moderate problem; 4 = a severe problem). This self-report symptom questionnaire was included to help characterize the sample and to examine responsiveness to change for the WHODAS 2.0. Total scores were calculated for RPQ items (excluding items rated as 1).

iii) Patient-reported Global Impression of Change (PGIC)

The PGIC is typically a Likert scale ranking improvement or decline in clinical status and includes a neutral state (Kamper, Maher, & Mackay, 2009). Patient-reported GIC variables were collected in both studies to evaluate perceptions of changes

Table 2
Anchor-based estimates of the minimal clinically important difference in WHODAS 2.0 interval scores anchored to patient reported global impression of change from first assessment to second assessment (Study 1; n = 111)

	N	Assessment 1 M (SD)	Assessment 2 M (SD)	Mean Difference[1] M (SD) 95% CI
Completely better	10	14.8 (4.9)	8.2 (6.4)	6.5 (4.1) 3.6, 9.5
Much improved	59	17.8 (6. 5)	13.0 (5.2)	4.8 (3.4) 3.4, 6.1
Some improvement	**35**	**21.0 (4.6)**	**19.7 (5.3)**	**1.3 (4.0) -0.1, 2.6**
No improvement	1	27.3	25.8	1.6
A little worse	1	18.9	15.8	3.1
Much worse	1	19.8	16.9	3.0

Anchoring question: How well do you feel you are recovering/have recovered from your injury?.

Table 3
Anchor-based estimates of the minimal clinically important difference in WHODAS 2.0 interval scores anchored to patient reported global impression of change from baseline assessment to outcome assessment (Study 2; n = 83)

	N	Baseline assessment M (SD)	Outcome assessment M (SD)	Mean Difference[1] M (SD) 95% CI
Very much improved	19	17.5 (4.2)	8.3 (5.7)	9.1 (5.7) 6.4, 11.9
Much improved	43	20.0 (3.7)	15.4 (4.6)	4.6 (3.9) 3.4, 5.8
Minimally improved	**16**	**20.5 (3.8)**	**17.3 (3.7)**	**3.2 (2.1) 2.1, 4.3**
No change	0			
Minimally worse	2	23.5 (2.0)	21.9 (2.9)	1.6 (1.0)
Much worse	0			
Very much worse	0			

Anchoring question: Compared to how you were before treatment, how are you doing overall?.

in symptoms between assessment timepoints. The two samples were not combined for PGIC analyses because of the differences in assessment timepoints, methods and indicators of change (see Tables 2 and 3).

iv) Demographic and clinical variables

Both studies collected a range of demographic and clinical variables. Demographic variables included age, sex, ethnicity, educational attainment and return to productivity. Clinical variables included self-reported loss of consciousness, mechanism of injury, prior psychiatric history (depression, anxiety) and past history of traumatic brain injury.

2.3. Data analyses

The data were analysed using SPSS version 27. Demographic and clinical variables were summarised using descriptive statistics and are presented as mean (SD), or frequency (n (%)), as specified. No a priori calculations for sample size or power were performed as the sample size was restricted to that of the combined studies.

Most authors recommend using multiple approaches to estimating MCIDs and triangu-

lating results (Mouelhi et al., 2020; Revicki et al., 2008; Wells et al., 2001). Accordingly we used a combination of anchor and distribution-based approaches and triangulated the findings, prioritising anchor-based results.

i) Anchor-based approach:

One anchor-based approach is to evaluate change in an outcome score anchored to change on a PGIC scale (Wells et al., 2001). The anchor-based perceptions of change in studies 1 and 2 were calculated by asking patients to rate their mTBI recovery from first assessment to second assessment. Response options varied slightly across the two studies and these are shown in Table 2 (Study 1) and Table 3 (Study 2). For each response category the average change in WHODAS scores was calculated along with a 95% confidence interval (CI). The mean difference between assessment timepoints for the first improvement category was considered to reflect a minimal change.

ii) Distribution-based methods

There is divergence in the literature on what constitutes the best way to approach distribution-based MCID estimation but most authors agree that more

than one method should be used (Wells et al., 2001). These methods are generally straight forward and use various measures of statistical distribution (Shulman et al., 2020). For this study we selected two broad distribution-based approaches recommended by previous authors (Shulman et al., 2020; Wells et al., 2001), standard deviation (SD) and the standard error of measurement (SEM).

For SD approaches we used an indicator of effect size, calculated as a multiple (typically 0.5 (Mouelhi et al., 2020)) of the standard deviation of the mean change in scores between the two assessment time-points. We calculated the SEM as the SD multiplied by the square root of 1-intraclass coefficient (Cronbach's alpha) as recommended by others (Mouelhi et al., 2020), using baseline and follow up assessment data.

Following Shulman et al. (2020), we calculated a range of MCID estimates across different demographic and clinical variables because we wished to determine if these estimates varied by key patient characteristics. In this study we compared change estimates between males and females, by age, and by post-concussion symptom severity (RPQ scores). For age we calculated a dichotomous variable based on the distribution in the samples (median split) and for the RPQ we calculated tertiles.

3. Results

3.1. Description of study samples

Demographic and clinical characteristics of participants from Study 1 ($n = 131$) and Study 2 ($n = 94$) are summarised in Table 4. The mean age of participants across both studies was 40.9 years, and 36.9% were male. A majority had sustained their mTBI in a motor vehicle accident (46.2%). Just over half of the participants in each study were receiving compensation at time of study participation (study 1 : 63%; study 2 : 51%). In study 2, 25% were in litigation. Litigation status was not collected from study 1 participants.

3.2. Anchor-based estimate

For anchor-based estimation using PGIC, $n = 111$ (Study 1) and $n = 83$ (Study 2) participants had completed both baseline and follow up assessments. For Study 1, participants who reported "some improvement" on the PGIC had a mean WHODAS change score of 1.3 (95% CI -01, 2.6). For Study 2, partici-

pants who endorsed feeling "minimally improved" on the PGIC for that study, the mean WHODAS change score was 3.2 (95% CI 2.1, 4.3) (Table 3).

3.3. Distribution-based estimates

Distribution-based estimates were calculated from baseline and follow up assessment data from each sample (Study 1: $n = 127$; Study 2: $n = 90$). The MCID estimates were broadly similar for the SD and SEM methods, with averaged estimates varying between 1.7 and 2.6 points for baseline and follow up assessments (Table 4). Considering an effect size approach, using 0.5 x SD of the mean change between baseline and follow up assessments, resulted in an MCID estimate of 3.5 (Study 1) and 2.3 (Study 2). Sensitivity analyses found subtle trends towards higher estimates for females, older participants, and those with fewer post-concussion symptoms, depending on the cohort and assessment timepoint (see Table 5).

3.4. Triangulation of findings by study sample

Figure 1 provides a visual overview of the findings across anchor and distribution-based methods and samples. For Study 1, the anchor-based method revealed an estimated minimal change perceived by participants of 1.3 interval points. Averaging distribution-based findings produced an estimated minimal change of 2.8. For Study 2, the anchor-based method revealed an estimated minimal change of 3.2. interval points. Averaging distribution-based findings produced an MCID estimate of 2.2 interval points. Variability was evident across the datasets in regard to timing for assessments and study methods, with Study 1 yielding a smaller anchor-based estimate of minimal change compared with Study 2, but greater variability in score distributions.

4. Discussion

Across two longitudinal studies involving 225 people with mTBI, 'at least some change' perceived by participants was reflected in WHODAS 2.0 interval score differences ranging between 1-3 for the Study 1 sample (corresponding to 2-5 ordinal points) and 2-4 interval points (corresponding to 3-6 ordinal points) for the Study 2 sample. There was general consistency across the range of methods used with minor variability associated with different patient characteristics such as age, sex, and post-concussion symptom sever-

Table 4
Demographic and clinical characteristics of the study samples ($n = 225$)

	Study 1 ($n = 131$) N (%)	Study 2 ($n = 94$) N (%)
Demographic characteristics		
Age [(y), mean (SD), range]	40.5 (12.1), 18-68	41.3 (11.9), 20-65
Sex [male, n (%)]	50 (33.3)	33 (35.1)
Ethnicity [n (%)]		
- Caucasian	97 (65.5)	72 (76.6)
- African American	3 (2.0)	1 (1.1)
- Asian/Asian Canadian	38 (25.3)	15 (16.0)
- Aboriginal/First Nations	6 (4.1)	1 (1.1)
- Hispanic	–	1 (1.1)
Education [post-secondary degree n (%)]	104 (69.3)	69 (74.2)
Return to productivity at second assessment [n (%)]		
- Full return to productivity	44 (29.3)	30 (36.1)
- Partial return to productivity	31 (20.7)	28 (33.7)
- Not working	32 (21.3)	25 (30.1)
Prior psychiatric history (depression or anxiety) [yes, *n* (%)][1]	51 (39.2)	63 (67.0)
Prior traumatic brain injury [yes, *n* (%)]	56 (37.3)	34 (37.4)
Clinical characteristics		
Loss of consciousness [*n* (%)]		
- Yes	23 (15.3)	11 (11.7)
- No	101 (67.3)	67 (71.3)
- Unclear or unknown	8 (5.3)	16 (17.0)
Injury Mechanism [n (%)]		
- Motor vehicle accident	71 (47.3)	33 (35.1)
- Fall	24 (16.0)	20 (21.3)
- Assault	2 (1.3)	1 (1.1)
- Sport	24 (16.0)	23 (24.5)
- Other	26 (17.3)	17 (18.1)
Receiving compensation at enrolment [yes, n (%)]	82 (62.6)	48 (51.1)
In litigation [yes, n (%)]	n/a	23 (24.5)
WHODAS 2.0 first assessment [M (SD), range][2]	19.0 (6.0), 0.0-39.4	19.9 (3.9), 10.1-27.9
WHODAS 2.0 second assessment [M (SD), range][3]	15.1 (7.0), 0.0-29.9	14.3 (5.7), 0.0-26.2
RPQ at first assessment [M (SD), range][2]	33.1 (13.5), 3-62	34.5 (14.4), 0-61
RPQ at second assessment [M (SD), range][3]	27.9 (13.0), 0-59	23.0 (14.8), 0-54

1. Positive prior psych history means a response of yes to "prior to your injury, were you ever diagnosed with or treated for depression or anxiety.". 2. Time from injury to first assessment (Study: 1 $m = 10.6$ weeks; Study 2: $m = 18.1$ weeks). 3. Time from injury to second assessment (Study 1: $m = 20.0$ weeks; Study 2: $m = 32.2$ weeks). RPQ = Rivermead Post Concussion Symptom Questionnaire (total scores response option 1 excluded); WHODAS 2.0 = 12-item World Health Organisation Disability Assessment Schedule 2.0 (interval (Rasch-converted) scores).

ity. There was however some variability between the two samples, possibly due to differences in the timing of the initial assessment (Study 1 participants were assessed sooner post-injury) and differences in treatment access during the test-retest period (the majority of Study 2 participants received more intensive treatment through the intervention study).

The findings from the present analyses are consistent with the small number of prior studies considering MCIDs for the 12-item WHODAS 2.0 in other populations. For example Katajapuu et al. (2020) reported MCIDs ranging between 3-5 ordinal points among people with chronic musculoskeletal pain ($n = 1988$) using distribution-based methods. This was a cross-sectional study without the opportunity to examine MCID against self-reported anchors

and time. In another study (Axelsson et al., 2017), MCID estimates ranged between 2-7 ordinal points for people with anxiety disorders ($n = 160$), based on an effect size approach and a clinician-based anchor. In peri- and post-surgical samples (Shulman et al., 2020), MCID estimates for the 12-item WHODAS 2.0 using a range of anchor and distribution methods were reported as at least 5% (2-3 ordinal points) across included datasets.

4.1. Strengths and limitations

Using a range of approaches across two different mTBI samples increases robustness of our results. It has been recommended that multiple independent anchors and methods, as well as confirming respon-

Table 5
Distribution-based estimates of differences in WHODAS 2.0 interval scores for baseline and follow up assessments

Study 1 Variables	N	0.5 SD	SEM	Average	Study 2 Variables	N	0.5 SD	SEM	Average
First assessment	**127**	**3.0**	**1.7**	**2.3**	**Baseline assessment**	**90**	**2.0**	**1.5**	**1.7**
Sex					Sex				
Male	41	2.7	1.5	2.1	Male	33	2.0	1.4	1.7
Female	85	3.2	1.8	2.5	Female	57	2.0	1.5	1.8
Age[1]					Age[1]				
≤41	63	2.4	1.4	1.9	≤40	44	1.9	1.4	1.7
≥42	62	3.5	2.0	2.7	≥41	46	2.0	1.5	1.8
RPQ (tertiles)[1]					RPQ (tertiles)[1]				
≤25	41	2.8	1.6	2.2	≤28	31	1.6	1.2	1.4
26-39	40	1.8	1.0	1.4	29-41	29	1.3	0.9	1.1
≥40	43	2.6	1.4	2.0	≥42	29	1.5	1.1	1.3
Second assessment	**111**	**3.5**	**1.7**	**2.6**	**Outcome assessment**	**83**	**2.9**	**2.0**	**2.4**
Sex					Sex				
Male	37	3.5	1.7	2.6	Male	30	2.2	1.5	1.9
Female	74	3.5	1.7	2.6	Female	53	3.1	2.2	2.7
Age[1]					Age[1]				
≤41	55	3.0	1.4	2.2	≤40	43	3.0	2.1	2.6
≥42	55	4.0	1.9	3.0	≥41	40	2.7	1.9	2.3
RPQ (tertiles)[1]					RPQ (tertiles)[1]				
≤21	38	3.5	1.5	2.5	≤14	28	2.6	1.8	2.2
22-32	36	3.2	0.9	2.0	15-29	26	1.6	1.1	1.4
≥33	33	2.3	1.1	1.7	≥30	28	1.7	1.7	1.4

1. Age and RPQ tertiles based on distribution in the sample; RPQ = Rivermead Post Concussion Symptom Questionnaire (total scores response option 1 excluded); WHODAS 2.0 = 12-item World Health Organisation Disability Assessment Schedule 2.0 (interval (Rasch-converted) scores).
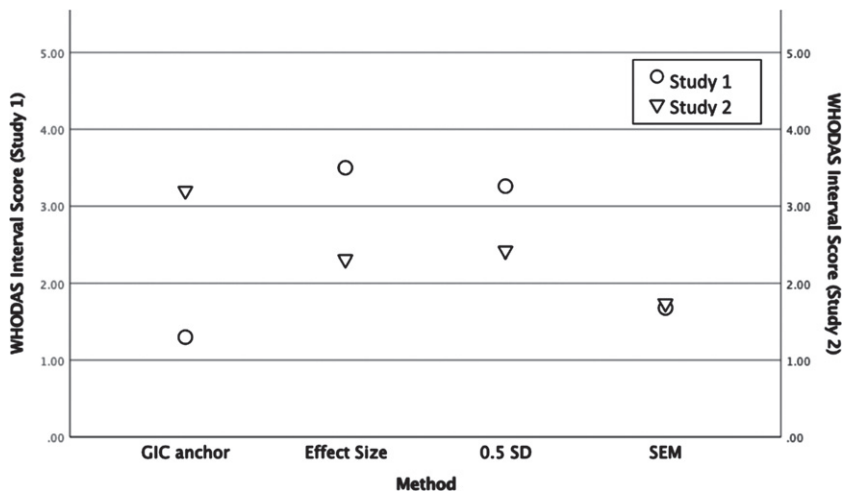

Fig. 1. Summary of anchor and distribution-based interval change scores for the 12-item WHODAS 2.0 in two mTBI samples. (WHODAS 2.0 = 12-item World Health Organisation Disability Assessment Schedule 2.0; GIC = Global Impression of Change; SD = standard deviation; SEM = standard error of measurement. If more than one value was available for a given type of estimate, the average was plotted.

siveness across multiple samples, is required when estimating MCID (Revicki et al., 2008). Our study was an exploratory and secondary evaluation of two datasets and accordingly, further validation of the MCID estimates is recommended before applying them to mTBI cohorts recruited from different settings or at different timepoints.

The PGIC scales we used to develop anchors of clinically important change on the 12-item WHODAS 2.0 varied slightly across the two studies. These scales have face validity and are a cornerstone for MCID (Myles, Myles, Chew, & Dennis, 2016; Shulman et al., 2020) but there is no universally agreed wording and it is unclear how subtle wording dif-

ferences affect MCID estimates. Future work should include alternative anchors such as patient acceptable symptom states (PASS) (Tubach et al., 2005). This means the minimal detectable change relevant for an individual might vary, for example in the study by Katajapuu et al. (2020) with chronic pain patients, the minimal detectable change exceeded the MCID derived by distribution-based methods. Some authors suggest that the PASS approach may be less susceptible to recall bias because it is less reliant on patients comparing their current functioning to a prior level of functioning (Tubach et al., 2006). PASS approaches have also been considered less influenced by age and gender (Kvien, Heiberg, & Hagen, 2007).

Other potential limitations include ceiling effects of certain items (e.g., getting dressed) from the 12-item WHODAS 2.0 in patients with mTBI.(Snell et al., 2017; Snell et al., 2020) These might result in skewed MCID estimates, especially using distribution-based methods (Kohn, Sidovar, Kaur, Zhu, & Coleman, 2014; Revicki et al., 2008). However, total scores were approximately normally distributed in both our study samples. In addition, the sample sizes for anchoring PGIC options were small because we could not merge the two samples together for these analyses. Finally, litigation is a known confound in the mTBI population and individuals with secondary gain may be more motivated to report more symptoms (Hanks et al., 2019). In study 2 approximately 25% of participants were in litigation at time of study participation. This information was not collected in study 1 however as the two studies recruited patients from the same clinics and in a similar manner, we presume the percentage in litigation was similar to that in Study 2. Litigation status should be considered when evaluating our findings.

5. Conclusion

Change in 12-item WHODAS 2.0 interval scores ranged between 1-4 interval points across two longitudinal mTBI datasets. An MCID of 3.5 interval points might conservatively capture meaningful change in adults of varying age, sex, and post-concussion symptom severity. This information provides a uniform metric that can be considered for designing and interpreting future mTBI clinical trials with WHODAS as the primary outcome. Evidence of minor variability between cohorts and previous research suggests that the change estimates derived

from the present study may not generalize to other clinical settings and contexts.

Conflict of interest

The authors report no declarations or conflicts of interest.

References

Alali, A., Vavrek, D., Barber, J., Dikmen, S., Nathans, A., & Temkin, N. (2015). Comparative Study of Outcome Measures and Analysis Methods for Traumatic Brain Injury Trials. *Journal of Neurotrauma*, *32*, 581-589.

Arbabi, M., Sheldon, R., Bahadoran, P., Smith, J., Poole, N., & Agrawal, N. (2020). Treatment outcomes in mild traumatic brain injury: a systematic review of randomized controlled trials. *Brain Injury*, *34*(9), 1139-1149.

Axelsson, E., Lindsater, E., Ljotsson, B., Andersson, E., & Hedman-Lagerlöf, E. (2017). The 12-item self-report World Health Organization Disability Assessment Schedule (WHODAS) 2.0 administered via the Internet to individuals with anxiety and stress disorders: a psychometric investigation based on data from two clinical trials. *JMIR Mental Health*, *4*, e58.

Barrett, B., Brown, D., Mundt, M., & Brown, R. (2005). Sufficiently Important Difference: Expanding the Framework of Clinical Significance. *Medical Decision Making*, *25*(3), 250-261.

Federici, S., Bracalenti, M., Meloni, F., & Luciano, J. (2016). World Health Organization disability assessment schedule 2.0: An international systematic review. *Disability and Rehabilitation*, DOI: 10.1080/09638288.2016.1223177. Available at: http://dx.doi.org/10.1080/09638288.2016.1223177.

Hanks, R., Rapport, l., Seagly, K., Millis, S., Scott, C., & Pearson, C. (2019). Outcomes after Concussion Recovery Education: Effects of Litigation and Disability Status on Maintenance of Symptoms. *Journal of Neurotrauma*, *36*(4), 554-558.

Holm, L., Cassidy, J., Carroll, L., & Borg, J. (2005). Summary of the WHO Collaborating Centre for Neurotrauma Task Force on Mild Traumatic Brain Injury. *Journal of Rehabilitation Medicine*, *37*(3), 137-141. doi:10.1080/16501970510027321

Horton, L., Rhodes, J., & Wilson, L. (2018). Randomized Controlled Trials in Adult Traumatic Brain Injury: A Systematic Review on the Use and Reporting of Clinical Outcome Assessments. *Journal of Neurotrauma*, *35*, 2005-2014.

Jaeshke, R., Singer, J., & Guyatt, G. (1989). Measurement of health status. Ascertaining the minimal clinically important difference. *Controlled Clinical Trials*, *10*, 407-415.

Kamper, S., Maher, C., & Mackay, G. (2009). Global rating of change scales: A review of strengths and weaknesses and considerations for design. *The Journal of Manual & Manipulative Therapy*, *17*(3), 163-170.

Katajapuu, N., Heinonen, A., & Saltychev, M. (2020). Minimal clinically important difference and minimal detectable change of the World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) amongst patients with chronic musculoskeletal pain. *Clinical Rehabilitation*, *34*(12), 1506-1511.

King, N., Crawford, S., Wenden, F., Moss, N., & Wade, D. (1995). The Rivermead Post Concussion Symptoms Questionnaire: a measure of symptoms commonly experienced after head injury and its reliability. *Journal of Neurology*, *242*, 587-592.

Kohn, C., Sidovar, M., Kaur, K., Zhu, Y., & Coleman, C. (2014). Estimating a minimal clinically important difference for the EuroQol 5-dimension health status index in persons with multiple sclerosis. *Health and Quality of Life Outcomes*, *12*, 66.

Kvien, T., Heiberg, T., & Hagen, K. (2007). Minimal clinically important improvement/difference (MCII/ MCID) and patient acceptable symptom state (PASS): what do these concepts mean? *Annals of Rheumatic Disease*, *66*(Suppl III):iii40–iii41.

Menon, D., & Maas, A. (2015). Progress, failures and new approaches for TBI research. *Nature Reviews Neurology*, *11*, 71-72.

Mouelhi, Y., Jouve, E., Castelli, C., & Gentile, S. (2020). How is the minimal clinically important difference established in health-related quality of life instruments? Review of anchors and methods. *Health and Quality of Life Outcomes*, *18*, 136-154.

Myles, P., Myles, D., Chew, C., & Dennis, A. (2016). Minimal Clinically Important Difference for Three Quality of Recovery Scales. *Anesthesiology*, *125*, 39-45.

Revicki, D., Haysb, R., Cellac, D., & Sloan, J. (2008). Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *Journal of Clinical Epidemiology*, *61*, 102-109.

Saltychev, M., Bärlund, E., Mattie, R., McCormick, Z., Paltamaa, J., & Laimi, K. (2016). A study of the psychometric properties of 12-item World Health Organization Disability Assessment Schedule 2.0 in a large population of people with chronic musculoskeletal pain. *Clinical Rehabilitation*, *31*(2), 262-272.

Shulman, M., Kasza, J., & Myles, P. (2020). Defining the Minimal Clinically Important Difference and Patient-acceptable Symptom State Score for Disability Assessment in Surgical Patients. *Anesthesiology*, *132*, 1362-1370.

Silverberg, N., Cairncross, M., Brasher, P., Vranceanu, A., Snell, D., Yeates, K.,... Burke, M. (2021). Feasibility of concussion rehabilitation approaches tailored to psychological coping styles: A randomized controlled trial. *Archive of Physical Medicine and Rehabilitation* (in press).

Silverberg, N., Panenka, W., Lizotte, P.-P., Bayley, M., Dance, D., & Li, L. (2020). Promoting early treatment for mild traumatic brain injury in primary care with a guideline implementation tool: a pilot cluster randomised trial. *BMJ Open*, *10*(e035527).

Snell, D., Iverson, G., Panenka, W., & Silverberg, N. (2017). Preliminary Validation of the World Health Organization Disability Assessment Schedule 2.0 for Mild Traumatic Brain Injury. *Journal of Neurotrauma*, *34*(23), 3256-3261.

Snell, D., Siegert, R., & Silverberg, N. (2020). Rasch analysis of the World Health Organization Disability Assessment Schedule 2.0 in a mild traumatic brain injury sample. *Brain Injury*, *34*(5), 610-618.

Tubach, F., Dougados, M., Falissard, B., Baron, G., Logeart, I., & Ravaud, P. (2006). Feeling good rather than feeling better matters more to patients. *Arthritis & Rheumatism (Arthritis Care & Research)*, *55*(4), 526-530.

Tubach, F., Ravaud, P., Baron, G., Falissard, B., Logeart, I., Bellamy, N.,... Dougados, M. (2005). Evaluation of clinically relevant states in patient- reported outcomes in knee and hip osteoarthritis: the patient acceptable symptom state. *Annals of the Rheumatic Diseases*, *64*(1), 34-37.

Üstün, T., Kastanjsek, N., Chatterji, S., & Rehm, J. (2010). *Measuring Health and Disability. Manual for WHO Disability Assessment Schedule. WHODAS 2.0*. Geneva: World Health Organization.

Wells, G., Beaton, D., Shea, B., Boers, M., Simon, L., Strand, V.,... Tugwell, P. (2001). Minimal Clinically Important Differences: Review of Methods. *Journal Rheumatology*, *28*, 406-412.

Wijeysundera, D., & Johnson, S. (2016). How Much Better Is Good Enough? Patient-reported Outcomes, Minimal Clinically Important Differences, and Patient Acceptable Symptom States in Perioperative Research. *Anesthesiology*, *125*(1), 7-10.

Wilson, J., Pettigrew, L., & Teasdale, G. (1998). Structured interviews for the Glasgow Outcome Scale and the extended Glasgow Outcome Scale: guidelines for their use. *Journal of Neurotrauma*, *15*, 573-585.