

# The psychometric properties and minimal clinically important difference for disability assessment using WHODAS 2.0 in critically ill patients

Alisa M Higgins, Ary Serpa Neto, Michael Bailey, Jonathan Barrett, Rinaldo Bellomo, D James Cooper, Belinda Gabbe, Natalie Linke, Paul S Myles, Michelle Paton, Steve Philpot, Mark Shulman, Meredith Young and Carol L Hodgson, on behalf of the PREDICT Study Investigators

Millions of patients survive critical illness each year.<sup>1</sup> As survival rates have improved,<sup>2,3</sup> the focus has shifted to improving the quality of recovery. Disability is common after critical illness,<sup>4-7</sup> and while the new physical, psychological and cognitive impairments that occur following critical illness (collectively termed post-intensive care syndrome<sup>8</sup>) are often assessed, a global measure of disability or impaired functioning may enable the overall burden of critical illness to be determined. Many current outcome assessment tools measure physical, psychological or cognitive impairments in isolation. As a result, core outcome sets often require a large number of measures to ensure different impairments are captured.<sup>9</sup> This increases the burden on patients, health care professionals, and researchers, and may increase withdrawal rates in research.

The World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) is a generic assessment instrument developed by WHO to provide a standardised method for measuring health and disability according to the International Classification of Functioning, Disability and Health (ICF) framework.<sup>10</sup> It provides a cross-cultural, standardised method for measuring limitations and restrictions on individuals' activities and participation in their society.<sup>11</sup> It captures the level of function across the six ICF activity and participation domains: cognition, mobility, self-care, interpersonal relationships, work and household roles, and participation in society. It has been used extensively and has been shown to be reliable and valid in numerous populations, including surgical patients, patients with chronic musculoskeletal pain, and patients with depression.<sup>12-14</sup> However, its psychometric properties have not been explored in a critically

## ABSTRACT

**Objectives:** The 12-item World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) provides a standardised method for measuring health and disability. This study aimed to determine its reliability, validity and responsiveness and to establish the minimum clinically important difference (MCID) in critically ill patients.

**Design:** Prospective, multicentre cohort study.

**Setting:** Intensive care units of six metropolitan hospitals.

**Participants:** Adults mechanically ventilated for > 24 hours.

**Main outcome measures:** Reliability was assessed by measuring internal consistency. Construct validity was assessed by comparing WHODAS 2.0 scores at 6 months with the EuroQoL visual analogue scale (EQ VAS) and Lawton Instrumental Activities of Daily Living (IADL) scale scores. Responsiveness was evaluated by assessing change over time, effect sizes, and percentage of patients showing no change. The MCID was calculated using both anchor and distribution-based methods with triangulation of results.

**Main results:** A baseline and 6-month WHODAS 2.0 score were available for 448 patients. The WHODAS 2.0 demonstrated good correlation between items with no evidence of item redundancy. Cronbach  $\alpha$  coefficient was 0.91 and average split-half coefficient was 0.91. There was a moderate correlation between the WHODAS 2.0 and the EQ VAS scores ( $r = -0.72$ ;  $P < 0.001$ ) and between the WHODAS 2.0 and the Lawton IADL scores ( $r = -0.66$ ;  $P < 0.001$ ) at 6 months. The effect sizes for change in the WHODAS 2.0 score from baseline to 3 months and from 3 to 6 months were low. Ceiling effects were not present and floor effects were present at baseline only. The final MCID estimate was 10%.

**Conclusion:** The 12-item WHODAS 2.0 is a reliable, valid and responsive measure of disability in critically ill patients. A change in the total WHODAS 2.0 score of 10% represents the MCID.

Crit Care Resusc 2021; 23 (1): 103-112

ill population. There are also limited data about a minimal clinically important difference (MCID), although this has recently been published for a surgical population.<sup>15</sup>

The aim of the current study was to assess the reliability, validity and responsiveness of the 12-item WHODAS 2.0 in patients who have survived critical illness, and to determine the MCID using data collected in a large multicentre observational study.

## Methods

### Study design

This was a multicentre, prospective cohort study conducted at six metropolitan intensive care units (ICUs) in the state of Victoria, Australia. Three public tertiary teaching hospitals, one public metropolitan hospital and two private hospitals were included (Online Appendix, eTable 1). Ethics committee approval, including a waiver of consent for data collection, was obtained at each site.

### Participants

Consecutive eligible patients were identified from the hospital clinical information system at each participating site. Patients were eligible if they had been admitted to the ICU and received more than 24 hours of mechanical ventilation. Patients who were aged less than 18 years, who had a proven or suspected acute primary brain process that was likely to result in global impairment of consciousness or cognition (eg, traumatic brain injury or stroke), or who did not speak English were excluded.

### Data collection

Demographic data were obtained from electronic health records for all eligible patients under a waiver of consent. Patients who survived the hospital admission were contacted initially by mail and invited to participate in telephone interviews at 3 and 6 months after ICU admission. Patients who were still in an acute hospital at 3 or 6 months after ICU admission were not called for the follow-up to reduce any burden on the patient while they were acutely ill.

Patient-reported long term outcomes were assessed at three pre-specified time points: baseline, 3 and 6 months. Baseline health and disability (defined as the status one month before ICU admission) were assessed retrospectively at the 3-month interview. All assessments were performed by telephone by trained central outcome assessors, who were blinded to the details of the patient's hospital admission. Data were entered into an electronic data capture system (REDCap, Vanderbilt University, USA).

### Outcome measures

Outcome measures included the 12-item WHODAS 2.0, the EQ-5D-5L health-related quality of life (HRQoL) questionnaire, the Lawton Instrumental Activities of Daily Living (IADL) scale, and a global disability rating scale where

patients rated their current disability on a 5-point scale with the options of 1, no disability; 2, mild disability; 3, moderate disability; 4, marked disability; or 5, severe disability. Each WHODAS 2.0 item has scores from 0 (no difficulty) to 4 (extreme difficulty) and a total WHODAS 2.0 score range from 0 to 48, with higher scores representing greater disability.<sup>11</sup> The total score is divided by 48 and multiplied by 100 to convert it to a percentage of maximum disability. As outlined in the WHODAS 2.0 manual, where a single WHODAS 2.0 item was missing, the mean value of the remaining 11 items was assigned to the missed item.<sup>10</sup> The WHODAS 2.0 score was not calculated when more than one item was missed.

The EQ-5D-5L is a generic preference-based health status instrument which is comprised of five dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. It also incorporates the EuroQoL visual analogue scale (EQ VAS), which is a quantitative measure of the patient's own judgement of their health status on a scale ranging from the worst imaginable health state (0) to the best imaginable health state (100).<sup>16</sup> The EQ-5D-5L is recommended as a core outcome measure for critically ill patients.<sup>9</sup>

The Lawton IADL scale is a valid and reliable functional assessment instrument that measures independent living skills across eight domains of function.<sup>17</sup> A summary score ranges between 0 and 8, with higher scores indicating greater levels of independence.

### Statistical analysis

Reliability was assessed by measuring internal consistency. An inter-item correlation matrix was visually inspected before measuring Cronbach  $\alpha$  and split-half reliability coefficients. Correlations were defined as excellent ( $> 0.75$ ), moderate (0.50–0.74), fair (0.25–0.49), and no meaningful correlation ( $< 0.25$ ).<sup>18</sup> With a sample size of 450 patients, this study has a 99% power (two-sided  $P = 0.05$ ) to detect a correlation coefficient of 0.2 or greater.

The construct validity of the WHODAS 2.0 at 6 months after ICU admission was evaluated by hypothesising that the WHODAS 2.0 would be more closely related to HRQoL (EQ VAS score) and functionality measures (Lawton IADL total score) and less correlated with weight. Spearman rank correlation coefficients were used to assess the correlation between the WHODAS 2.0 values and continuous variables.

Responsiveness was evaluated by assessing change over time, effect sizes, and percentage of patients showing no change. The WHODAS 2.0 values were compared to assess significance of change over time from baseline to 3 months and from 3 months to 6 months using a Wilcoxon signed rank test. The percentage of patients showing no change was calculated between each time point. The ability of WHODAS 2.0 to detect a meaningful change in the clinical

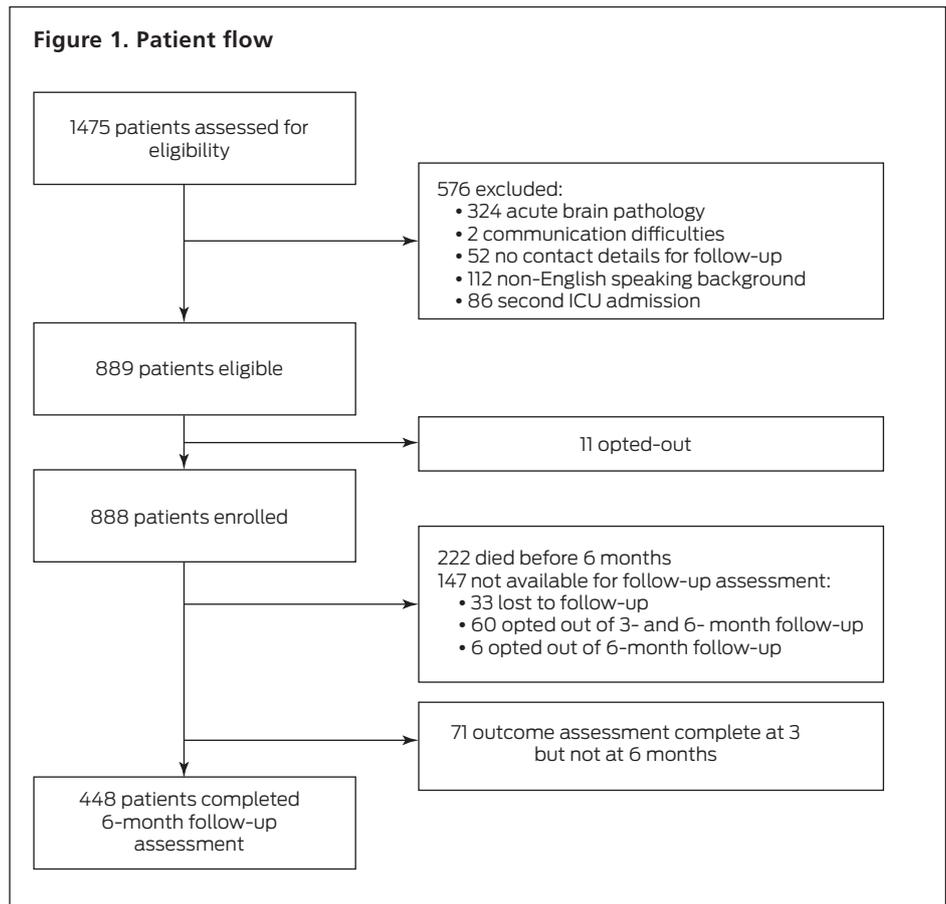
state of a patient was quantified using the Cohen effect size. This is the mean difference in scores between two time points, divided by the average of their standard deviations. Floor and ceiling effects were determined by assessing the number and percentage of participants scoring the lowest value (0) and the highest value (100) at baseline, 3 and 6 months. Floor and ceiling effects below 15% were considered acceptable.<sup>19</sup>

The minimal clinically important difference is the smallest change in score on a given scale that corresponds to a meaningful change in clinical state from the patient's perspective.<sup>20,21</sup> A combination of anchor and distribution-based methods with triangulation (averaging) of results was used to determine the MCID. Three distribution-based methods were used considering the 6-month WHODAS 2.0 scores: i) 0.3 times the standard deviation (SD); ii) the standard error of the measurement (SEM; calculated as the standard deviation multiplied by the square root of 1 minus the reliability estimate, calculated from the Cronbach  $\alpha$ ); and iii) 5% of the score range. These methods have been recommended and previously used to determine the WHODAS 2.0 MCID in surgical patients.<sup>15,22</sup>

Anchor-based methods estimate the MCID by relating a change in patient-centred outcome score to a change in clinical scenario or a change on a patient-reported global rating scale. The MCID was estimated anchored to the patient-reported global disability rating scale at 6 months, with a worsening or improving of one step (eg, no disability to mild disability, or mild disability to moderate disability) in the global disability rating at 6 months compared with baseline being considered consistent with a change equivalent to the MCID. The distribution and anchor-based estimates were then averaged to reach a final MCID estimate. As the MCID may be sensitive to different population groups and clinical scenarios, the MCID was further compared in the following subgroups:

- gender (women *v* men);
- age ( $\leq 65$  *v*  $> 65$  years old);
- type of admission (medical *v* surgical);
- presence of sepsis (yes *v* no);

**Figure 1. Patient flow**



- cardiac arrest (yes *v* no);
- acute respiratory failure (yes *v* no); and
- severity of illness (Acute Physiology and Chronic Health Evaluation [APACHE] III score  $\leq 64$  *v*  $> 64$ ).

All analyses were performed considering a two-sided hypothesis test, with a significance level of 0.05. Analyses were performed using the software R v.4.0.3 (R Core Team, 2016, Austria).

## Results

### Participants

A total of 1475 patients, admitted to the ICU between May 2017 and June 2018 were screened, of which 899 patients met the eligibility criteria. Eleven patients opted out of all data collection at the time of follow-up, resulting in 888 enrolled patients. Of these, 222 patients (25%) died by 6 months and 66 (7%) opted out of the 3- and/or 6-month telephone follow-up. The final sample for the current study includes 448 participants (67.3% of surviving participants at 6 months) who completed the baseline WHODAS 2.0 and a follow-up WHODAS 2.0 at 6 months (Figure 1). Baseline data for these patients are shown in Table 1. Characteristics of surviving patients who were included in

**Table 1. Patient baseline characteristics**

Characteristics	Overall
Total number of patients	448
Age (years), median (IQR)	58.2 (44.7–67.7)
Male sex	291 (65.0%)
Body mass index (kg/m <sup>2</sup> ), median (IQR)	27.2 (23.4–31.2)
APACHE III score, median (IQR)	57.0 (44.0–76.0)
Type of admission	
Medical	238 (53.8%)
Surgical	204 (46.2%)
Diagnosis category	
Cardiovascular	150 (33.5%)
Gastrointestinal	32 (7.1%)
Gynaecological	1 (0.2%)
Haematological	1 (0.2%)
Metabolic	24 (5.4%)
Musculoskeletal and skin	9 (2.0%)
Neurological	19 (4.2%)
Renal and genitourinary	5 (1.1%)
Respiratory	89 (19.9%)
Sepsis	52 (11.6%)
Trauma	66 (14.7%)
Co-existing disorders	
Chronic respiratory failure	
▶ Lung transplant	37 (8.4%)
▶ Other	13 (2.9%)
Chronic cardiovascular disease	25 (5.7%)
Chronic liver disease	9 (2.0%)
Chronic kidney disease	6 (1.4%)
Chronic immune disease	4 (0.9%)
Chronic immunosuppression	27 (6.1%)
Diabetes	158 (35.3%)
ICU source of admission	
Operating room	204 (46.3%)
Emergency room	132 (29.9%)
Ward	43 (9.8%)
Other ICU	37 (8.4%)
Other hospital	25 (5.7%)
Clinical frailty score at ICU admission	
Non-frail	284 (83.0%)
Mild-to-moderate frail	55 (16.1%)
Severely frail	3 (0.9%)

APACHE = Acute Physiology and Chronic Health Evaluation; ICU = intensive care unit. Percentages may not total 100 because of rounding. Denominators are shown when the overall sample size was not available.

the study and of the surviving patients who did not complete the follow-up assessment are shown in the Online Appendix, eTable 2. A higher proportion of non-responders were male, but there were no other baseline differences between responders and non-responders.

### Reliability

The WHODAS 2.0 demonstrated excellent reliability. The inter-item correlation matrix for the WHODAS 2.0 at 6 months (Online Appendix, eTable 3 and eFigure 1) demonstrated good correlation between items with no evidence of item redundancy, indicated by almost all inter-item correlations between 0.4 and 0.8. Cronbach  $\alpha$  coefficient was 0.91 and average split-half coefficient was 0.91 (minimum and maximum split-half coefficients were 0.77 and 0.96 respectively).

### Validity

Construct validity was demonstrated through a moderate correlation between the WHODAS 2.0 and the EQ VAS scores ( $r = -0.72$ ; 95% CI,  $-0.77$  to  $-0.67$ ;  $P < 0.001$ ), and between the WHODAS 2.0 and Lawton IADL scores ( $r = -0.66$ ; 95% CI,  $-0.71$  to  $-0.61$ ;  $P < 0.001$ ) at 6 months (Figure 2, A and B). In addition, no significant correlation was found between WHODAS 2.0 scores at 6 months and the unrelated variable patient body weight ( $r = 0.07$ ; 95% CI,  $-0.03$  to  $0.17$ ;  $P = 0.163$ ) (Figure 2, C).

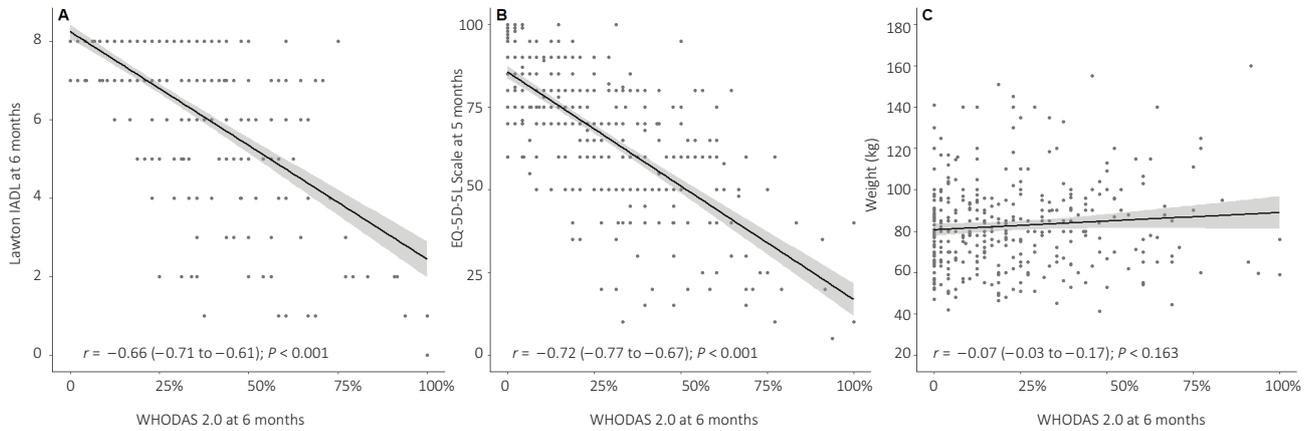
### Responsiveness

The effect sizes for change in the WHODAS 2.0 score from baseline (median, 12.5; interquartile range [IQR], 0.0–39.6) to 3 months (median, 20.8; IQR, 6.2–41.7) and to 6 months (median 16.7; IQR, 4.2–35.4) were low (Figure 3). The WHODAS 2.0 score changed from baseline to 3 months ( $P = 0.001$ ) and from 3 to 6 months ( $P < 0.001$ ). At 6 months, the WHODAS 2.0 had a very good scaling property: the 10th, 25th, 50th, 75th and 90th percentiles were 0.0, 4.2, 16.7, 35.4 and 52.7 respectively.

### Floor and ceiling effects

A floor effect was present only at baseline, with 29.1%, 12.1% and 15.8% of the patients at baseline, 3 and 6 months having a WHODAS 2.0 score of 0. A ceiling effect was low at all time points (0.4%, 0.6% and 0.4% at baseline, 3 and 6 months

**Figure 2. Assessment of convergent and divergent validity of the World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) at 6 months**



IADL = Instrumental Activities of Daily Living. Correlation between WHODAS 2.0 and Lawton Instrumental Activities of Daily Living (IADL) at 6 months (A), between WHODAS 2.0 and EQ-5D-5L at 6 months (B), and between WHODAS 2.0 at 6 months and weight (C). *r* is the Spearman rank correlation coefficient.

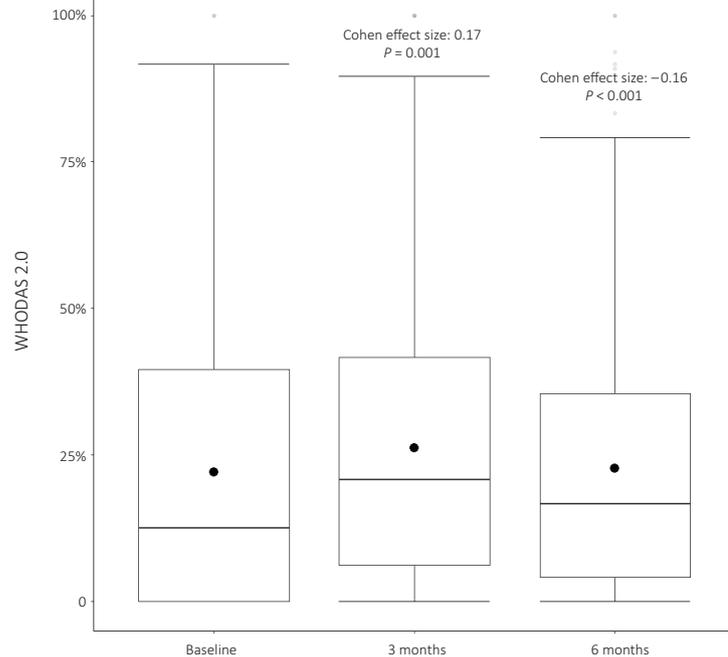
respectively) (Figure 4). The percentage of patients showing no change from baseline to 3 months was 14.2% and from 3 to 6 months it was 12.6%.

**Minimal clinically important difference**

The distribution-based estimates of the MCID were similar for SD, SEM and 5% instrument range methods (Table 2). Sensitivity analysis found that MCID estimates were consistent in all but in one subgroup (a lower MCID was found in patients with cardiac arrest). Anchor-based estimates of the MCID, based on Global Disability Score, were higher than the distribution-based estimated (Table 3). The average of the mean differences from the two methods of anchor-based estimation was 12%.

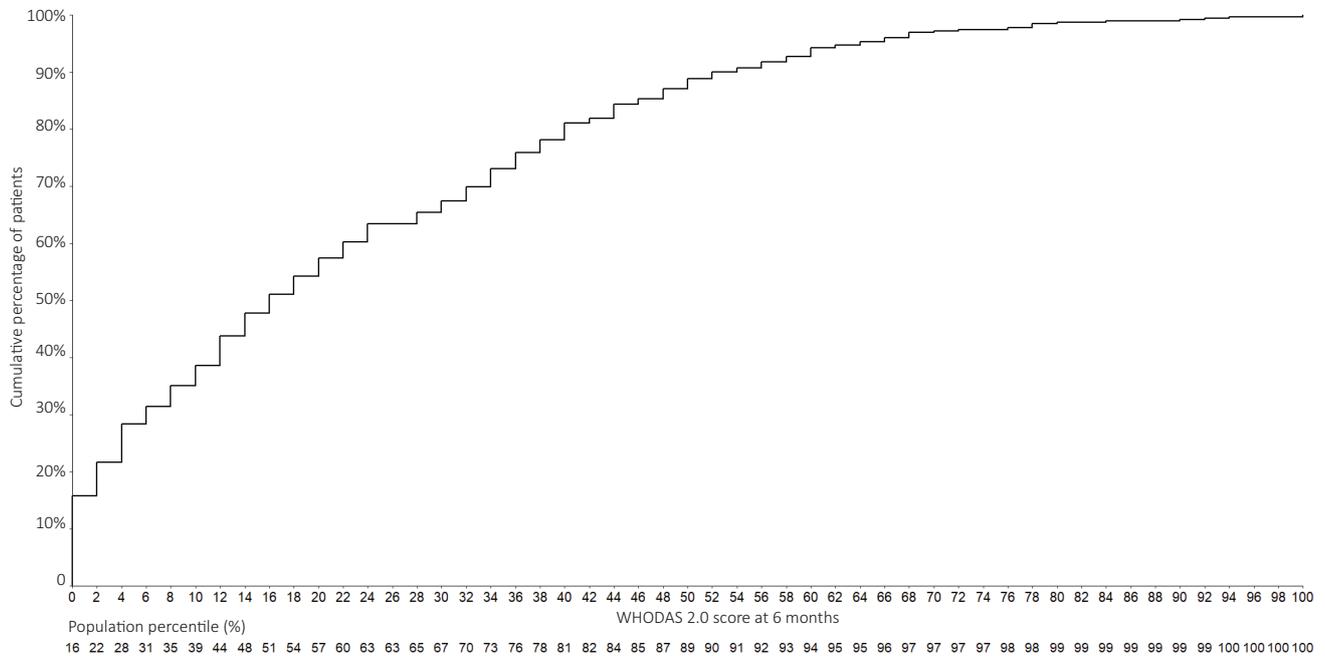
Triangulating the average distribution-based MCID estimate (6%) and two anchor-based MCID (12%), a final MCID estimate of 10% was reached. Using this MCID value to quantify the change in disability from baseline to 6 months, 148 patients (36.4%) had a clinically important increase in disability, 106 patients (26.1%) had a clinically important decrease in disability, and 152 patients (37.4%) had no change in disability.

**Figure 3. Assessment of responsiveness of the World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) score at 3 and 6 months**



Boxes represent median and interquartile range (IQR). Whiskers extend 1.5 times the IQR beyond the first and third quartiles per the conventional Tukey method. Transparent circles beyond the whiskers represent outliers. Filled circles represent mean values. Cohen effect size calculated as the mean difference in scores in two time points, divided by the average of their standard deviations. *P* values from Wilcoxon signed rank test.

**Figure 4. Cumulative percentage of the World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) scores at 6 months**



The curve is the empirical cumulative distribution function of WHODAS 2.0 scores at 6 months.

## Discussion

The 12-item WHODAS 2.0 provides a standardised method to assess health and disability according to the ICF framework. The current study has shown the WHODAS 2.0 has strong psychometric qualities in critically ill patients, including excellent reliability, validity and responsiveness. We estimate that the minimal clinically important difference for WHODAS 2.0 in critically ill patients is 10%, indicating that a change in WHODAS 2.0 score of 10% or more is consistent with a clinically meaningful increase or decrease in the level of disability.

The WHODAS 2.0 is being used increasingly as an outcome measure in clinical research, with a recent systematic review finding that, of 811 publications using the WHODAS 2.0, more than 50% were published in the most recent 3 years.<sup>23</sup> The WHODAS 2.0 has been used to measure disability in critical care studies,<sup>6,24</sup> and its use in ongoing registries and trials is increasing (eg, NCT03793257, NCT02735707, NCT04588311, NCT03671447). Despite this, we are unaware of any previous studies that have evaluated its psychometric properties in the critically ill. A systematic review of the psychometric properties of the 12-item WHODAS 2.0 in the general population and in people with non-acute physical causes of disability concluded

that the WHODAS 2.0 is an internally consistent and reliable scale demonstrating overall good correlation with other measures of disability.<sup>25</sup> A study in surgical patients also found that the WHODAS 2.0 had excellent reliability, criterion and convergent validity and responsiveness.<sup>14</sup> Our findings in a critically ill population are consistent with this research in other patient groups.

The correlation between the WHODAS 2.0 and existing, commonly used critical care outcome measures (EQ VAS and Lawton IADL) was used to determine the validity of the WHODAS 2.0.<sup>9,26</sup> The WHODAS 2.0 showed a moderate correlation with both the EQ VAS ( $r = -0.72$ ) and the Lawton IADL ( $r = -0.66$ ), which is expected as the scales measure slightly different constructs. Divergent validity was shown through the lack of correlation between the WHODAS 2.0 and body weight ( $r = 0.07$ ).

Consistent with recommendations, we used a number of alternative methods to estimate the MCID for the WHODAS 2.0, including distribution-based and anchor-based approaches.<sup>27</sup> Based on these methods, our results suggest that the MCID for the WHODAS 2.0 is 10%, which is slightly higher than the 8% MCID in non-surgical populations,<sup>10,28</sup> and higher than the 5% MCID recently reported for a surgical population.<sup>15</sup> The distribution of

**Table 2. Distribution-based estimates of the minimal clinically important difference in the World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) score measured at 6 months**

	Number of patients	6-month WHODAS 2.0 percentage score (mean $\pm$ SD)	0.3 SD	SEM	5% of range	Average
Total population	448	22.7 $\pm$ 21.4	6.4	6.4	5.0	6
Sex						
Male	291	20.9 $\pm$ 20.5	6.2	6.2	5.0	6
Female	157	25.9 $\pm$ 22.7	6.8	6.8	5.0	6
Age, years						
$\leq$ 65	305	23.6 $\pm$ 21.6	6.5	6.5	5.0	6
$>$ 65	143	20.7 $\pm$ 21.1	6.3	6.3	5.0	6
Type of admission						
Medical	238	25.7 $\pm$ 22.5	6.8	6.8	5.0	6
Surgical	204	19.2 $\pm$ 19.8	5.9	5.9	4.5	5
Sepsis						
Yes	118	26.1 $\pm$ 22.1	6.6	6.6	4.7	6
No	330	21.5 $\pm$ 21.1	6.3	6.3	5.0	6
Cardiac arrest						
Yes	42	15.6 $\pm$ 14.7	4.4	4.4	2.6	4
No	399	23.3 $\pm$ 22.0	6.6	6.6	5.0	6
Acute respiratory failure						
Yes	28	26.2 $\pm$ 21.8	6.5	6.5	3.9	6
No	404	22.2 $\pm$ 21.3	6.4	6.4	5.0	6
APACHE III						
$\leq$ 64	271	21.1 $\pm$ 21.4	6.4	6.4	5.0	6
$>$ 64	170	24.9 $\pm$ 21.5	6.5	6.5	5.0	6

APACHE = Acute Physiology and Chronic Health Evaluation; SD = standard deviation; SEM = standard error of the mean.

**Table 3. Anchor-based estimates of the minimal clinically important difference in World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) scores**

	WHODAS 2.0 score		
	Baseline score	Score at 6 months	Mean difference
Decline of one category in global disability rating scale	10.6 $\pm$ 16.1	23.2 $\pm$ 19.4	12.7
Improvement of one category in global disability rating scale	30.6 $\pm$ 24.0	19.4 $\pm$ 17.2	-11.2

the WHODAS 2.0 differs in the surgical population, where higher proportions of patients had lower WHODAS 2.0 scores (75% of scores  $<$  16.7%).<sup>15</sup> A change of 10% may be represented by larger changes across two items in the WHODAS 2.0 (eg, moving from no difficulty in walking for long distances and standing for long periods to severe difficulty walking for long distances and moderate difficulty standing for long periods) or by small changes across up to five items (eg, moving from no difficulty in washing,

getting dressed, concentrating for 10 minutes, learning a new task, and maintaining a friendship to mild difficulties in each of these tasks). The MCID can be used by clinicians and researchers to determine sample size for future studies and to determine the clinical significance of changes in WHODAS 2.0 scores both in research and in clinical practice.

Recovery following critical illness is complex and physical, cognitive and psychological impairments are common.<sup>4-6,29-35</sup> There is increasing recognition of the need

for systematic and frequent assessment for impairments associated with post-intensive care syndrome.<sup>36</sup> Identifying these impairments, including characterisation of patients at risk, is essential to enable development of appropriate preventive and treatment interventions. The core outcome set for evaluating post-discharge outcomes of acute respiratory failure survivors includes the EQ-5D (with the optional addition of SF-36 version 2), the Hospital Anxiety and Depression Scale, and the Impact of Events Scale-Revised.<sup>9</sup> Completion of multiple assessment measures may not be feasible on a large scale due to logistical and resource barriers. A simple measure that is able to identify patients with disability may be useful for both clinical practice and research. The 12-item WHODAS 2.0 may serve as a reliable and valid tool to screen patients for disability following critical illness, with subsequent referral of those identified as having disability to enable identification of specific impairments and development of management plans.

The strengths of this study include its prospective, multicentre, longitudinal design, the inclusion of a heterogeneous group of critically ill patients, the collection of a range of detailed clinical parameters (allowing comparisons to other populations), and central assessment of long term outcome measures. Our sample size of over 400 patients is large compared with other validation studies in critically ill patients.<sup>37-40</sup> Our study also has several limitations. We only included patients who were ventilated for over 24 hours, and the psychometric properties of the WHODAS 2.0 in non-ventilated critically ill patients are unknown. Of the 666 survivors at 6 months, only 448 (67.3%) were included in the current study. While the only difference in the baseline characteristics of patients who were included and patients who were not included was that a higher proportion of responders were male, there may be important unmeasured characteristics that could influence our results. A further limitation is that our baseline WHODAS 2.0 score and the baseline global disability rating scale were assessed retrospectively at 3 months after ICU admission. While it is not possible to obtain a baseline measure from sedated, mechanical ventilated patients, the retrospective assessment may have introduced recall bias.

## Conclusion

Our results demonstrate that the 12-item WHODAS 2.0 is a reliable, valid, and responsive measure of disability in critically ill patients, and that a change in the total WHODAS 2.0 score of 10% represents the minimal clinically important difference. The current study provides support to use the WHODAS 2.0 as a common, international, standardised instrument to measure disability following critical illness.

**Acknowledgements:** The PREDICT study was funded by the National Health and Medical Research Council (NHMRC) of Australia (GNT1132976). Carol Hodgson is supported by a Heart Foundation Fellowship and an NHMRC Investigator Grant (GNT1173271). Belinda Gabbe is supported by an Australian Research Council Future Fellowship (FT170100048).

## Competing interests

None declared.

## Author details

Alisa M Higgins<sup>1</sup>  
 Ary Serpa Neto<sup>1,2,3</sup>  
 Michael Bailey<sup>1,2</sup>  
 Jonathan Barrett<sup>4,5</sup>  
 Rinaldo Bellomo<sup>1,2,3</sup>  
 D James Cooper<sup>1,6</sup>  
 Belinda Gabbe<sup>7</sup>  
 Natalie Linke<sup>1</sup>  
 Paul S Myles<sup>7,8</sup>  
 Michelle Paton<sup>1,9</sup>  
 Steve Philpot<sup>10</sup>  
 Mark Shulman<sup>7,8</sup>  
 Meredith Young<sup>6</sup>  
 Carol L Hodgson<sup>1,6</sup>

On behalf of the PREDICT Study Investigators\*

\* The full list of the PREDICT Study Investigators is available in the Online Appendix.

- 1 Australian and New Zealand Intensive Care Research Centre, School of Public Health and Preventive Medicine, Monash University, Melbourne, VIC, Australia.
- 2 Department of Critical Care, University of Melbourne, Melbourne, VIC, Australia.
- 3 Department of Intensive Care, Austin Health, Melbourne, VIC, Australia.
- 4 Intensive Care Unit, Epworth Healthcare, Melbourne, VIC, Australia.
- 5 Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, VIC, Australia.
- 6 Department of Intensive Care and Hyperbaric Medicine, The Alfred, Melbourne, VIC, Australia.
- 7 School of Public Health and Preventive Medicine, Monash University, Melbourne, VIC, Australia.
- 8 Department of Anaesthesiology and Perioperative Medicine, The Alfred, Melbourne, VIC, Australia.
- 9 Department of Physiotherapy, Monash Health, Melbourne, VIC, Australia.
- 10 Intensive Care Unit, Cabrini Health, Melbourne, VIC, Australia.

**Correspondence:** carol.hodgson@monash.edu

## References

- 1 Wunsch H, Angus DC, Harrison DA, et al. Variation in critical care services across North America and Western Europe. *Crit Care Med* 2008; 36: 2787-93, e1-9.
- 2 Kaukonen KM, Bailey M, Suzuki S, et al. Mortality related to severe sepsis and septic shock among critically ill patients in ANZ, 2000-2012. *JAMA* 2014; 311: 1308-16.
- 3 Australian and New Zealand Intensive Care Society Centre for Outcome and Resource Evaluation. 2019 Report. Melbourne: ANZICS, 2020. <https://www.anzics.com.au/wp-content/uploads/2020/11/2019-CORE-Report.pdf> (viewed Dec 2020).
- 4 Herridge MS, Chu LM, Matte A, et al. The RECOVER Program: disability risk groups and 1-year outcome after 7 or more days of mechanical ventilation. *Am J Respir Crit Care Med* 2016; 194: 831-44.
- 5 Herridge MS, Tansey CM, Matte A, et al. Functional disability 5 years after acute respiratory distress syndrome. *N Engl J Med* 2011; 364: 1293-304.
- 6 Hodgson CL, Udy AA, Bailey M, et al. The impact of disability in survivors of critical illness. *Intensive Care Med* 2017; 43: 992-1001.
- 7 Iwashyna TJ, Cooke CR, Wunsch H, Kahn JM. Population burden of long-term survivorship after severe sepsis in older Americans. *J Am Geriatr Soc* 2012; 60: 1070-7.
- 8 Needham DM, Davidson J, Cohen H, et al. Improving long-term outcomes after discharge from intensive care unit: report from a stakeholders' conference. *Crit Care Med* 2012; 40: 502-9.
- 9 Needham DM, Sepulveda KA, Dinglas VD, et al. Core outcome measures for clinical research in acute respiratory failure survivors: an international modified Delphi consensus study. *Am J Respir Crit Care Med* 2017; 196: 1122-30.
- 10 Üstün TB, Kostanjsek N, Chatterji S, Rehm J; editors. Measuring health and disability: manual for WHO Disability Assessment Schedule (WHODAS 2.0). Geneva; 2010. <https://apps.who.int/iris/handle/10665/43974> (viewed Dec 2020).
- 11 Ustun TB, Chatterji S, Kostanjsek N, et al. Developing the World Health Organization Disability Assessment Schedule 2.0. *Bull World Health Organ* 2010; 88: 815-23.
- 12 Luciano JV, Ayuso-Mateos JL, Fernández A, et al. Psychometric properties of the twelve item World Health Organization Disability Assessment Schedule II (WHO-DAS II) in Spanish primary care patients with a first major depressive episode. *J Affect Disord* 2010; 121: 52-8.
- 13 Saltychev M, Barlund E, Mattie R, et al. 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. *Clin Rehabil* 2017; 31: 262-72.
- 14 Shulman MA, Myles PS, Chan MT, et al. Measurement of disability-free survival after surgery. *Anesthesiology* 2015; 122: 524-36.
- 15 Shulman MA, Kasza J, Myles PS. defining the minimal clinically important difference and patient-acceptable symptom state score for disability assessment in surgical patients. *Anesthesiology* 2020; 132: 1362-70.
- 16 Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011; 20: 1727-36.
- 17 Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist* 1969; 9: 179-86.
- 18 Fan E, Ciesla ND, Truong AD, et al. Inter-rater reliability of manual muscle strength testing in ICU survivors and simulated patients. *Intensive Care Med* 2010; 36: 1038-43.
- 19 Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007; 60: 34-42.
- 20 Revicki D, 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.
- 21 Wells G, Beaton D, Shea B, et al. Minimal clinically important differences: review of methods. *J Rheumatology* 2001; 28: 406-12.
- 22 Myles PS, Myles DB, Gallagher W, et al. Minimal clinically important difference for three quality of recovery scales. *Anesthesiology* 2016; 125: 39-45.
- 23 Federici S, Bracalenti M, Meloni F, Luciano JV. World Health Organization disability assessment schedule 2.0: an international systematic review. *Disabil Rehabil* 2017; 39: 2347-80.
- 24 Myles PS, Bellomo R, Corcoran T, et al. Restrictive versus liberal fluid therapy for major abdominal surgery. *N Engl J Med* 2018; 378: 2263-74.
- 25 Saltychev M, Katajapuu N, Barlund E, Laimi K. Psychometric properties of 12-item self-administered World Health Organization disability assessment schedule 2.0 (WHODAS 2.0) among general population and people with non-acute physical causes of disability — systematic review. *Disabil Rehabil* 2019; 1-6.
- 26 Hopkins RO, Suchyta MR, Kamdar BB, et al. Instrumental activities of daily living after critical illness: a systematic review. *Ann Am Thorac Soc* 2017; 14: 1332-43.
- 27 Yost KJ, Eton DT. Combining distribution- and anchor-based approaches to determine minimally important differences: the FACIT experience. *Eval Health Prof* 2005; 28: 172-91.
- 28 Andrews G, Kemp A, Sunderland M, et al. Normative data for the 12 item WHO Disability Assessment Schedule 2.0. *PLoS One* 2009; 4: e8343.
- 29 Brummel NE, Bell SP, Girard TD, et al. Frailty and subsequent disability and mortality among patients with critical illness. *Am J Respir Crit Care Med* 2017; 196: 64-72.
- 30 Davydow DS, Gifford JM, Desai SV, et al. Depression in general intensive care unit survivors: a systematic review. *Intensive Care Med* 2009; 35: 796-809.
- 31 Davydow DS, Gifford JM, Desai SV, et al. Posttraumatic stress

## ORIGINAL ARTICLES

- disorder in general intensive care unit survivors: a systematic review. *Gen Hosp Psychiatry* 2008; 30: 421-34.
- 32 Hopkins RO, Herridge MS. Quality of life, emotional abnormalities, and cognitive dysfunction in survivors of acute lung injury/acute respiratory distress syndrome. *Clin Chest Med* 2006; 27: 679-89.
- 33 Iwashyna TJ, Ely EW, Smith DM, Langa KM. Long-term cognitive impairment and functional disability among survivors of severe sepsis. *JAMA* 2010; 304: 1787-94.
- 34 Jackson JC, Pandharipande PP, Girard TD, et al. Depression, post-traumatic stress disorder, and functional disability in survivors of critical illness in the BRAIN-ICU study: a longitudinal cohort study. *Lancet Respir Med* 2014; 2: 369-79.
- 35 Pandharipande PP, Girard TD, Jackson JC, et al. Long-term cognitive impairment after critical illness. *New Engl J Med* 2013; 369: 1306-16.
- 36 Elliott D, Davidson JE, Harvey MA, et al. Exploring the scope of post-intensive care syndrome therapy and care: engagement of non-critical care providers and survivors in a second stakeholders meeting. *Crit Care Med* 2014; 42: 2518-26.
- 37 Denehy L, de Morton NA, Skinner EH, et al. A physical function test for use in the intensive care unit: validity, responsiveness, and predictive utility of the physical function ICU test (scored). *Phys Ther* 2013; 93: 1636-45.
- 38 Hodgson CL, Needham D, Haines K, et al. Feasibility and inter-rater reliability of the ICU Mobility Scale. *Heart Lung* 2014; 43: 19-24.
- 39 Parry SM, Granger CL, Berney S, et al. Assessment of impairment and activity limitations in the critically ill: a systematic review of measurement instruments and their clinimetric properties. *Intensive Care Med* 2015; 41: 744-62.
- 40 Tipping CJ, Bailey MJ, Bellomo R, et al. The ICU Mobility Scale has construct and predictive validity and is responsive. A multicenter observational study. *Ann Am Thorac Soc* 2016; 13: 887-93.