

# Measurement of Disability-free Survival after Surgery

Mark A. Shulman, M.B., B.S., M.P.H., F.A.N.Z.C.A.,  
 Paul S. Myles, M.B., B.S., M.P.H., M.D., F.A.N.Z.C.A., F.R.C.A.,  
 Matthew T. V. Chan, M.B., B.S., F.A.N.Z.C.A., David R. McLlroy, M.B., B.S., M.Clin.Epi, F.A.N.Z.C.A.,  
 Sophie Wallace, M.P.H., Jennie Ponsford, B.A.(Hons), M.A.(Clin Neuropsych), Ph.D.

## ABSTRACT

**Background:** Survival and freedom from disability are arguably the most important patient-centered outcomes after surgery, but it is unclear how postoperative disability should be measured. The authors thus evaluated the World Health Organization Disability Assessment Schedule 2.0 in a surgical population.

**Methods:** The authors examined the psychometric properties of World Health Organization Disability Assessment Schedule 2.0 in a diverse cohort of 510 surgical patients. The authors assessed clinical acceptability, validity, reliability, and responsiveness up to 12 months after surgery.

**Results:** Criterion and convergent validity of World Health Organization Disability Assessment Schedule 2.0 were supported by good correlation with the 40-item quality of recovery scale at 30 days after surgery ( $r = -0.70$ ) and at 3, 6, and 12 months after surgery with physical functioning (The Katz index of independence in Activities of Daily Living;  $r = -0.70$ ,  $r = -0.60$ , and  $\rho = -0.47$ ); quality of life (EQ-5D;  $r = -0.57$ ,  $-0.60$ , and  $-0.52$ ); and pain interference scores (modified Brief Pain Inventory Short Form;  $r = 0.72$ ,  $0.74$ , and  $0.81$ ) (all  $P < 0.0005$ ). Construct validity was supported by increased hospital stay ( $6.9$  vs.  $5.3$  days,  $P = 0.008$ ) and increased day 30 complications ( $20\%$  vs.  $11\%$ ,  $P = 0.042$ ) in patients with new disability. There was excellent internal consistency with Cronbach's  $\alpha$  and split-half coefficients greater than 0.90 at all time points (all  $P < 0.0005$ ). Responsiveness was excellent with effect sizes of 3.4, 3.0, and 1.0 at 3, 6, and 12 months after surgery, respectively.

**Conclusions:** World Health Organization Disability Assessment Schedule 2.0 is a clinically acceptable, valid, reliable, and responsive instrument for measuring postoperative disability in a diverse surgical population. Its use as an endpoint in future perioperative studies can provide outcome data that are meaningful to clinicians and patients alike. (ANESTHESIOLOGY 2015; 122:524-36)

*"To cure sometimes, to relieve often, to comfort always"*

—Hippocrates

THE main aims of surgery are to cure or to at least relieve distressing symptoms from many conditions. Anesthetic and other perioperative research outcome measures have traditionally centered on surrogate endpoints and recovery times<sup>1</sup> and, far less often, major complications and death.<sup>2,3</sup> As exemplified in the PeriOperative ISchemia Evaluation (POISE) trial,<sup>4</sup> it can be difficult to ascribe a relative weight or harm to outcomes such as myocardial infarction or stroke, particularly when the long-term consequences of these outcomes vary substantially. In addition, such endpoints may not reflect the patient's perception of their subsequent health status after surgery.<sup>5,6</sup>

Previous research suggests that a return or maintenance of health, functional capacity, and emotional well-being are highly valued patient goals following surgery.<sup>2,7,8</sup> Accordingly,

### What We Already Know about This Topic

- Although survival is commonly measured after surgery, survival without disability is rarely measured, and it is unclear whether disability measures used in medical populations are appropriate to define disability after surgery

### What This Article Tells Us That Is New

- In a multicenter, multinational study of over 500 patients, the World Health Organization Disability Assessment Schedule 2.0 was shown to be a clinically acceptable, valid, reliable, and responsive instrument for measuring postoperative disability in a diverse surgical population

contemporary anesthetic and other perioperative research sometimes includes patient-centered outcome measures such as quality of recovery<sup>7,9,10</sup> and quality of life<sup>8,11,12</sup> after surgery. But new or residual disability after surgery is of particular concern to patients and clinicians alike.

This article is featured in "This Month in Anesthesiology," page 1A. Corresponding article on page 481. Supplemental Digital Content is available for this article. Direct URL citations appear in the printed text and are available in both the HTML and PDF versions of this article. Links to the digital files are provided in the HTML text of this article on the Journal's Web site ([www.anesthesiology.org](http://www.anesthesiology.org)).

Submitted for publication August 20, 2014. Accepted for publication November 11, 2014. From the Department of Anaesthesia and Perioperative Medicine, Alfred Hospital and Monash University, Melbourne, Victoria, Australia (M.A.S., P.S.M., D.R.M., S.W.); Department of Anaesthesia and Intensive Care, The Chinese University of Hong Kong, Prince of Wales Hospital, Hong Kong Special Administrative Region (M.T.V.C.); and Monash-Epworth Rehabilitation Research Centre, Epworth Hospital and Monash University, Richmond, Victoria, Australia (J.P.).

Copyright © 2015, the American Society of Anesthesiologists, Inc. Wolters Kluwer Health, Inc. All Rights Reserved. Anesthesiology 2015; 122:524-36

Current definitions of disability make distinction between the physical or mental impairment caused by a health condition and the impact that impairment has on the person's ability to work, care for themselves, and interact with society.<sup>13,14</sup> The World Health Organization (WHO) International Classification of Functioning, Disability and Health classifies disability as "difficulties in any area of functioning as they relate to environmental and personal factors."<sup>15</sup>

An instrument used to measure postoperative disability should include, but not be limited to, an assessment of physical functioning or quality of life. Furthermore, rather than focusing specifically on the presence (or even extent) of symptoms, it should assess the *impact* of these symptoms on the patient's life in the dimensions of psychological well-being, social involvement, life role activities, and cognitive well-being.<sup>16</sup>

Although it is tempting to use quality of life measures as a proxy for measuring postoperative disability, this approach is scientifically unsound, and there is currently no validated generic measure of long-term postoperative disability that accords with the WHO classification. The ideal instrument should be easy to administer, reliable, responsive to change, and be specifically validated in a surgical population.

The WHO Disability Assessment Schedule 2.0 (WHODAS) was developed to measure disability cross-culturally, in the aged, and for disease-related states.<sup>17</sup> It asks about limitations over the last 30 days in six major life domains: cognition, mobility, self-care, interpersonal relationships, work and household roles, and participation in society. WHODAS has excellent psychometric properties, is easy to use and score, and is available on the public domain in self-report, proxy, and telephone-based versions that can be administered in around 5 min.<sup>18</sup> WHODAS has been used to assess disability following trauma,<sup>19,20</sup> stroke,<sup>21,22</sup> spinal cord injury,<sup>23</sup> and in those with numerous and varied chronic diseases.<sup>24</sup> It has not, however, been specifically evaluated in a surgical setting.

The aim of this study was to evaluate WHODAS in a diverse surgical cohort with varying degrees of comorbid medical disease, disability, and health. A secondary aim was to characterize disability-free survival after surgery.

## Materials and Methods

This multicenter prospective observational cohort study was conducted in five hospitals in Australia and Hong Kong, and institutional review board approval was sought and obtained at each site (see table 1, Supplemental Digital Content 1, <http://links.lww.com/ALN/B131>). We specifically aimed to recruit a diverse low- to high-risk surgical population in order to properly evaluate diagnostic utility.<sup>25</sup> Patients were included in the study if aged 18 yr or over, able to provide informed consent, and were scheduled to have ambulatory, intermediate, major noncardiac, cardiac, or nonelective surgery. Patients were excluded if they were not expected to be available for

follow-up over the following year, had poor language comprehension, known or suspected cognitive impairment, current psychiatric disease, or substance abuse. While patients having nonelective surgery were included, patients having time-critical surgery (*e.g.*, requiring urgent transfer to the operating theater) were excluded due to insufficient time or patient ability to complete baseline testing and consent. Patients were excluded from analysis if they did not have surgery or if they were consented but no further data were collected.

In an effort to maintain unbiased sampling, a broad range of patients undergoing different types of surgery of varying extent were selected in consecutive order from operating theater booking lists. To increase study power, we planned to recruit a greater proportion of patients from the major noncardiac, cardiac, and nonelective surgery groups as they were more likely to have a complicated recovery after surgery, with a broader range of outcomes (including disability).

After providing informed consent, patients were given instructions in completing the predetermined standardized questionnaires, which they then completed without prompting from research staff. Efforts were made to follow-up all patients so that data were not lost from sicker or older patients.

We began our study with the intention to create a novel postoperative disability scale, but further literature review identified the WHODAS as being a likely valid measure that had as yet not been formally evaluated in a surgical population. We thus included this scale in our suite of perioperative measurements after commencement of the study and defined this as the revised primary aim of our study.

Following enrollment, patient medical and demographic data were collected and patients were provided with instructions in completing each of the health status questionnaires being used for validity testing:

1. The 12-item WHODAS.<sup>17</sup>
2. The 40-item quality of recovery (QoR-40) score,<sup>26</sup> as a global, patient-centered measure of health status at 30 days after surgery.
3. The EQ-5D Scale,<sup>27,28</sup> measuring quality of life.
4. The Katz index of independence in Activities of Daily Living (Katz ADL) scale,<sup>29</sup> measuring physical functioning.
5. The modified Brief Pain Inventory Short Form (mBPI-sf),<sup>30</sup> measuring daily pain.

The 12-item WHODAS (fig. 1) was scored as previously described.<sup>18,31</sup> Numerical values were attributed to each item on a 5-point Likert scale: none = 0; mild = 1; moderate = 2; severe = 3; and extreme = 4. The total score, between 0 and 48, is then divided by 48 and multiplied by 100 to convert it to a percentage of the maximum disability score. One site (Hong Kong) scored WHODAS from 1 to 5 as originally described in the WHODAS user manual. This was resolved by subtracting 1 point from each WHODAS item score at this site. Missing data were handled according to guidelines in the WHODAS manual,<sup>18</sup> whereby if a single item was



## WHODAS 2.0

WORLD HEALTH ORGANIZATION  
DISABILITY ASSESSMENT SCHEDULE 2.0

### 12-item version, self-administered

This questionnaire asks about difficulties due to health conditions. Health conditions include diseases or illnesses, other health problems that may be short or long lasting, injuries, mental or emotional problems, and problems with alcohol or drugs.

Think back over the past 30 days and answer these questions, thinking about how much difficulty you had doing the following activities. For each question, please circle only one response.

In the past 30 days, how much difficulty did you have in:						
S1	<u>Standing for long periods</u> such as <u>30 minutes</u> ?	None	Mild	Moderate	Severe	Extreme or cannot do
S2	Taking care of your <u>household responsibilities</u> ?	None	Mild	Moderate	Severe	Extreme or cannot do
S3	<u>Learning a new task</u> , for example, learning how to get to a new place?	None	Mild	Moderate	Severe	Extreme or cannot do
S4	How much of a problem did you have <u>joining in community activities</u> (for example, festivities, religious or other activities) in the same way as anyone else can?	None	Mild	Moderate	Severe	Extreme or cannot do
S5	How much have <u>you</u> been <u>emotionally affected</u> by your health problems?	None	Mild	Moderate	Severe	Extreme or cannot do

In the past 30 days, how much difficulty did you have in:						
S6	<u>Concentrating</u> on doing something for <u>ten minutes</u> ?	None	Mild	Moderate	Severe	Extreme or cannot do
S7	<u>Walking a long distance</u> such as a <u>kilometre</u> [or equivalent]?	None	Mild	Moderate	Severe	Extreme or cannot do
S8	<u>Washing your whole body</u> ?	None	Mild	Moderate	Severe	Extreme or cannot do
S9	Getting <u>dressed</u> ?	None	Mild	Moderate	Severe	Extreme or cannot do
S10	<u>Dealing with people you do not know</u> ?	None	Mild	Moderate	Severe	Extreme or cannot do
S11	<u>Maintaining a friendship</u> ?	None	Mild	Moderate	Severe	Extreme or cannot do
S12	Your day-to-day <u>work</u> ?	None	Mild	Moderate	Severe	Extreme or cannot do

H1	Overall, in the past 30 days, <u>how many days</u> were these difficulties present?	Record number of days ____				
H2	In the past 30 days, for how many days were you <u>totally unable</u> to carry out your usual activities or work because of any health condition?	Record number of days ____				
H3	In the past 30 days, not counting the days that you were totally unable, for how many days did you <u>cut back</u> or <u>reduce</u> your usual activities or work because of any health condition?	Record number of days ____				

**Fig. 1.** The 12-item World Health Organization Disability Assessment Schedule.<sup>18</sup> Reproduced, with permission of WHO, from *Measuring Health and Disability: Manual for WHO Disability Assessment Schedule*. Geneva, World Health Organization, 2010 (WHODAS 2.0 12-item version self-administered [www.who.int/classifications/icf/whodasii/en](http://www.who.int/classifications/icf/whodasii/en)).

missed, the mean value of the remaining items was assigned to the missed item. The WHODAS score was not calculated when more than one item was missed.

We considered a disability score of greater than or equal to 25% to indicate “disability,” based on the WHODAS and WHO International Classification of Functioning,

Disability and Health: none (0 to 4%); mild (5 to 24%); moderate (25 to 49%); severe (50 to 95%); and complete (96 to 100%) disability.<sup>18</sup> New disability was defined if a patient had an increase in the WHODAS score of greater than or equal to 8% from their preoperative assessment.<sup>31</sup>

A preoperative WHODAS was not included in the study procedures until 3 months after the commencement of the study, when we became aware of its potential utility in the perioperative setting. Up until that time, WHODAS had not featured in any surgical or anesthetic literature. As a result, a number of patients (n = 81) did not complete a preoperative WHODAS questionnaire. For those participants, we adopted strict criteria (preoperative scores of Katz ADL = 12, EQ-5D 100-point scale  $\geq$  80, and QoR-40  $\geq$  180) to classify them as being free of baseline disability for some secondary evaluations. We tested these criteria in the complete cohort and only 16 out of 151 (10.6%) of participants with a preoperative WHODAS score of less than 10 were misclassified as having preoperative disability.

The QoR-40 is a validated 40-item questionnaire measuring quality of recovery following anesthesia and surgery.<sup>26,32</sup> It consists of five dimensions: (1) physical comfort (12 items), (2) emotional state (nine items), (3) physical independence (five items), (4) psychological support (seven items), and (5) pain (seven items). The QoR-40 has a possible score of 40 (extremely poor quality of recovery) to 200 (excellent quality of recovery). Missing data were imputed by assigning the mean value of other items within that domain to the missing item. The EQ-5D has five dimensions, each ranked on a three-level scale as well as a 100-point scale where participants can rate their health from 0 ("worst imaginable") to 100 ("best imaginable"). The Katz ADL scale contains six domains of physical functioning, each scored between 0 ("little or no difficulty") and 2 ("unable"). The mBPI-sf has two parts. The first assesses "worst," "least," and "average" pain over the previous 24 h as well as pain "right now" from 0 ("no pain") to 10 ("pain as bad as you can imagine"). The second part assesses the degree to which pain interferes with seven life domains and is again scored from 0 ("does not interfere") to 10 ("completely interferes"). The mean interference score can be calculated if four or more of the seven items have been completed on a given administration.

Baseline patient surgical risk and health status were assessed by classifying patients according to the Portsmouth Physiological and Operative Severity Score for the Enumeration of Mortality and Morbidity (P-POSSUM) score,<sup>33,34</sup> the American Society of Anesthesiologists' physical status (ASA) score, and the Canadian Study of Health and Aging Clinical Frailty scale.<sup>35</sup>

The P-POSSUM score estimates the risk of postoperative morbidity and in-hospital mortality using defined physiological and operative variables and was calculated as previously described.<sup>33,36</sup> Physiological and operative scores were summed and applied to formula:  $\ln R/1 - R = -9.065 + (0.1692 \times \text{physiological score}) + (0.1550 \times \text{operative severity$

score). As no radiological data were collected, a history of congestive cardiac failure was substituted for cardiomegaly in the P-POSSUM scoring system. Values for other missing data: electrocardiogram (n = 53); hemoglobin (n = 37); urea (n = 35); potassium (n = 37), sodium (n = 37), and heart rate (n = 31) were assumed to be normal as these were most likely not measured for healthy participants having minor procedures.

The Canadian Study of Health and Aging Clinical Frailty scale is a subjective measurement of patient frailty based on their appearance and history. Both the attending anesthesiologist and an investigator independently determined the patient's level of frailty, and the average score was used to quantify the clinical level of frailty: no frailty, 1.0 to 3.9; vulnerable, 4.0 to 4.9; mild, 5 to 5.9; moderate, 6 to 6.9; and severe,  $\geq$ 7.0. Where the anesthesiologist did not complete the clinical frailty score (n = 13), the investigator-determined score was used.

Intraoperative data were recorded by the anesthesiologist and included the type, extent, duration, and urgency of surgery and estimated blood loss. The extent of surgery was classified according to the P-POSSUM system (minor, intermediate, major, and major+).<sup>34</sup> Nursing staff collected postoperative data, including temperature on arrival and length of stay in the postanesthesia care unit. For patients going directly to the intensive care unit, the duration of tracheal intubation was recorded from intensive care unit charts or discharge summary.

Discharge data, including the occurrence of postoperative complications, duration of hospital stay, and discharge destination, were collected from the patient's medical record and the hospital electronic discharge system. Duration of hospital stay was calculated as the number of full days spent in hospital as an inpatient.

Patients were followed up with self-assessment questionnaires and by telephone at 30 days, 3, 6, and 12 months after surgery. At 30 days, we ascertained whether they had experienced any postoperative complications (including readmission to hospital, readmission to the intensive care unit, myocardial infarction, respiratory complications, stroke, wound infection, or intraabdominal collection). At subsequent follow-up times, we recorded patients' current living situation (home with or without nursing assistance, rehabilitation, nursing home, or hospital). At each telephone interview, patients were asked to rate how worthwhile they felt their surgery was and the effect the surgery had on their lives, using 5-point Likert scales. The batch of questionnaires (WHODAS, QoR-40 [30 days only], EQ-5D, Katz ADL, and mBPI-sf) was sent to patients with a stamped self-addressed envelope for return postage.

### Psychometric Evaluation of WHODAS

Psychometric evaluation of a health status instrument should occur in the population and setting of interest and include assessment of the clinical acceptability, validity, reliability, and responsiveness of the scale.<sup>16,37</sup> The WHODAS

has previously undergone extensive psychometric evaluation,<sup>17,18,24,31,38</sup> but not in a surgical population.

*Clinical acceptability* was assessed by measuring WHO-DAS completion rates over time and the comparative completion rates of WHODAS and the other instruments at 12 months after surgery. Analysis was limited to 12-month completion rates for pragmatic reasons and because 12 months was considered a relevant time point for measuring long-term disability after surgery. The denominator for completion rate included all living patients remaining in the study that had not actively withdrawn or been lost to follow-up (*i.e.*, participants answering calls but not returning surveys were included).

The *validity*, or accuracy, of WHODAS was assessed according to guidelines<sup>16,25</sup>:

1. Content validity: WHODAS was developed and has been extensively validated as a responsive measure of health-related disability as defined by the International Classification of Functioning, Disability and Health.<sup>18</sup>
2. Concurrent (criterion) validity: WHODAS was compared to the QoR-40, Katz ADL, and mBPI-sf scales. Sensitivity analyses were done to explore whether correlations were modified by patient age or extent of surgery.
3. Convergent validity: WHODAS was compared to the EQ-5D 100-point quality of life health scale.
4. Construct validity:
  - a. Discriminative validity (construct validation by extreme groups): A good and poor quality of recovery at 30 days and good and poor quality of life at 3, 6, and 12 months were identified by using the upper and lower quartiles of the day 30 QoR-40 and EQ-5D 100-point scales, respectively. WHO-DAS scores were then compared between good and poor groups.
  - b. We measured the relationship between WHODAS and clinical variables likely to be associated with higher rates of disability after surgery: duration of hospital stay, complications, and unplanned readmission within 30 days after surgery.

*Reliability* was assessed by measuring internal consistency. An interitem correlation matrix was visually inspected before measuring Cronbach's  $\alpha$  and split-half reliability coefficients. In other words, we assessed the degree to which different items in the WHODAS scale agree with each other and with the overall measure of disability.

The repeatability of WHODAS has been evaluated extensively in previous studies<sup>17</sup> and was not assessed in this study as it would not be expected to be different in this population.

The *responsiveness*, or the ability of WHODAS to detect a meaningful change in the clinical state of a patient, was quantified using the Cohen effect size.<sup>39</sup> This is the mean difference in scores from baseline to the time point of interest, divided by the SD at baseline. The subgroup of patients with a baseline WHODAS score of less than or equal to 4% was used to define a group of patients with little to no

preoperative disability. An effect size of greater than 0.8 was considered to provide strong evidence that the score is responsive to change in health status.

Following psychometric evaluation, disability-free survival was calculated as the percentage of participants who were both alive and had a WHODAS score of less than 25% at each time point after surgery. Further exploratory analyses of the surgical population were undertaken to examine the relationship between disability-free survival and patient age, medical comorbidity, and surgical type and extent.

### Statistical Analysis

Our sample size calculation was based primarily on data from our previous quality of recovery studies,<sup>26,40</sup> using MedCalc version 12.3.0 (Ostend, Belgium). To have a probability of greater than or equal to 80% to detect a relationship between two variables at a two-sided 0.05 significance level, looking for a greater than or equal to 15% change in the dependent variable, with an assumption that the SD of the independent and dependent variables is 4 (on a 10-point scale), required 350 patients. To account for possible ineligible or incomplete questionnaires, and to support subgroup exploratory analysis, we increased the sample size to at least 500 patients.

Data are presented as mean  $\pm$  SD, median [interquartile range], number (%), or 95% confidence intervals. All percentages of 10 or more are rounded to the nearest integer. Associations were measured using Pearson correlation coefficients ( $r$ ) or Spearman rank correlation ( $\rho$ ) for nonnormal data. When comparing scales with reverse direction of scores indicating improved health, resultant correlations will be negative. Associations for ordinal data were measured using chi-square for trend. Internal consistency was measured using split-half reliability and Cronbach's  $\alpha$ .<sup>41</sup> Changes in numerical data from baseline were compared using the paired  $t$  test. Interrater agreement was measured using Cohen's  $\kappa$  coefficient.<sup>42</sup> The null hypothesis was rejected if the two-tailed  $P$  was less than 0.05. All analyses were performed using SPSS for Windows v22.0 (SPSS Inc., Chicago, IL).

### Results

Patient demographics (table 1) indicate a typical population of patients presenting to a university hospital for a broad range of surgical procedures. The mean age of patients was 56 yr (range, 18 to 90 yr), and 42% had an ASA score of III or IV. There was a high level of agreement between an investigator- and attending anesthesiologist-determined clinical frailty assessments,  $\kappa$  0.60 ( $P < 0.0005$ ). Additional results are provided in Supplemental Digital Content 1, <http://links.lww.com/ALN/B131>, tables 1–13.

Despite only 4.5% of patients having mild or moderate frailty before surgery (table 1), there was a high level of preoperative disability, with 115 patients (27%) having a WHODAS score greater than or equal to 25%. The mean P-POSSUM predicted in-hospital mortality was 2.0%. There was also a broad range of surgical type and

**Table 1.** Baseline Demographics (n = 510)

Patient Characteristics	
Age (yr)	55.9 ± 14.9; range 18–90
Gender	
Male/female (% male)	298/212 (58)
Body mass index (kg/m <sup>2</sup> )	27.1 ± 6.7
Preexisting medical condition	
Smoker	96 (19)
Hypertension	207 (41)
Ischemic heart disease	67 (13)
Previous myocardial infarction	39 (7.6)
Cardiac failure	25 (4.9)
Stroke or transient ischemic attack	33 (6.5)
Asthma or chronic obstructive pulmonary disease	81 (16)
Diabetes	70 (14)
Cancer	181 (36)
Chronic pain requiring daily medication	95 (19)
Plasma albumin (g/l)	33.8 ± 5.4
Hypoalbuminemia (plasma albumin <30 g/l)	24 (4.7)
Plasma hemoglobin (g/l)	133 ± 18.7
Plasma sodium (mmol/l)	140 ± 2.8
Plasma creatinine (μmol/l)	72 [62–85]
Medications	
Aspirin	102 (20)
Warfarin	19 (3.7)
Nitrate	26 (5.1)
Statin	119 (23)
ACE inhibitor/angiotensin-II blocker	127 (25)
β blocker	87 (17)
Insulin	14 (2.7)
Opioid	69 (14)
Chemotherapy	14 (2.7)
ASA physical status	
I	86 (17)
II	207 (41)
III	188 (37)
IV	26 (5)
CSHA Clinical Frailty Scale*	
No frailty (1.0–3.9)	330 (81)
Vulnerable (4.0–4.9)	59 (14)
Mild (5–5.9)	16 (4.0)
Moderate (6–6.9)	2 (0.5)
Severe (≥7.0)	0
P-POSSUM predicted mortality*	2.02 ± 3.7
Lives alone	86 (17)
Preoperative disability† (n = 426)	115 (27)

Values are presented as mean ± SD, number (%), or median [interquartile range].

\*See text for full description of CSHA Clinical Frailty Scale and P-POSSUM score. †Clinically significant disability defined as a WHODAS score ≥ 25%.

ACE = angiotensin-converting enzyme; ASA = American Society of Anesthesiologists; CSHA = Canadian Study of Health and Aging; P-POSSUM = Portsmouth Physiological and Operative Severity Score for the Enumeration of Mortality and Morbidity; WHODAS = World Health Organization Disability Assessment Schedule 2.0.

extent (table 2), including 42 patients (8.3%) undergoing nonelective surgery. The median length of stay was 5 days (interquartile range, 2 to 8), and the majority of patients

**Table 2.** Perioperative Characteristics (n = 510)

Patient Characteristics	
Extent of surgery*	
Minor	40 (7.8)
Moderate	158 (31)
Major	192 (38)
Major+	120 (24)
Nonelective surgery	42 (8.3)
Type of surgery	
General	175 (34)
Orthopedic	93 (18)
Arthroplasty	26 (5.1)
Thoracic	60 (12)
Cardiac	50 (10)
Neurosurgery	50 (10)
Urology	33 (6.5)
Vascular	20 (3.9)
Duration of surgery (min)	140 [69–210]
Duration of recovery room stay (min)	85 [60–125]
Intensive care unit admission	84 (17)
Time until extubation (h)	11 [7.0–17]
Duration of hospital stay (d)	4.9 [2.1–8.0]
Discharge destination (n = 499)	
Home	447 (90)
Home with nursing assistance	7 (1.4)
Rehabilitation	29 (5.8)
Nursing home	0
Other hospital	16 (3.2)

Values are presented as number (%) or median [interquartile range].

\*Surgery of moderate severity includes appendectomy, cholecystectomy, mastectomy, transurethral resection of prostate; major surgery includes any laparotomy, bowel resection, cholecystectomy with choledochotomy, peripheral vascular procedure, or major amputation; major+ surgery includes any aortic procedure, abdominopineal resection, pancreatic or liver resection, or esophagogastrectomy.<sup>36</sup>

(90%) were discharged home or to a rehabilitation facility (5.8%). By day 30, 15% of patients had at least one postoperative complication, 5 patients (1%) had died, and 35 patients (7.2%) had an unplanned readmission to hospital (table 3).

Of the 510 patients enrolled in the study, 68 (13%) had either withdrawn or been lost to follow-up at 6 months and 72 (14%) by 12 months. When comparing the baseline characteristics of patients with complete data at 6 and 12 months to patients with incomplete data (withdrawn or lost to follow-up), those with incomplete data were more likely to be female, having more minor surgery, with lower P-POSSUM scores but higher rates of preoperative disability (see tables 2 and 3, Supplemental Digital Content 1, <http://links.lww.com/ALN/B131>, which are tables describing baseline demographics for patients with complete and incomplete data at 6 and 12 months).

WHODAS demonstrated good clinical acceptability with completion and postal return rates of greater than or equal to 88% at all time points. At 12 months after surgery, WHODAS, Katz ADL, and EQ-5D all had 92% completion and postal return rates. Five patients had a missing value for WHODAS, allowing valid imputation.

**Table 3.** Postoperative Morbidity, Mortality, and Disability-free Survival

Time	Postoperative Complication (n = 491)	Complication Rate (%)	Unplanned Readmission to Hospital*	Cumulative Mortality%	Disability†	New Disability‡
Day 30	Wound infection	36 (7.3)	35 (7.2) (n = 487)	5 (1) (n = 495)		
	Respiratory complication	17 (3.5)				
	Myocardial infarction	12 (2.4)				
	Unplanned ICU admission	2 (0.4)				
	Stroke	2 (0.4)				
Any complication	73 (15)					
3 months			29 (6.4) (n = 454)	6 (1.3) (n = 471)	91 (22)	65 (18)
6 months			32 (7.2) (n = 442)	7 (1.5) (n = 458)	74 (18)	52 (14)
12 months			30 (7.3) (n = 413)	22 (5.0) (n = 438)	60 (16)	46 (13)

Values are presented as number (%).

\*Unplanned readmission to hospital as a consequence of index surgery. †Moderate or severe disability defined as a WHODAS score  $\geq$  25%. ‡New disability compared to preoperative state, defined as a decrement in WHODAS score  $\geq$  8%.

ICU = Intensive Care Unit; WHODAS = World Health Organization Disability Assessment Schedule 2.0.

The correlation between preoperative WHODAS, QoR-40, and Katz ADL was tested to explore the relationship between baseline disability (WHODAS has been extensively validated in nonsurgical patients) and the instruments proposed for subsequent postoperative criterion validity testing. As expected, there was moderate correlation with the Katz ADL scale ( $r = -0.56$ ,  $P < 0.0005$ ) and QoR-40 score ( $r = -0.60$ ,  $P < 0.0005$ ). Preoperative WHODAS had modest correlation with the Canadian Study of Health and Aging Clinical Frailty scale ( $\rho = 0.28$ ,  $P < 0.0005$ ) and ASA score ( $\rho = 0.22$ ,  $P < 0.0005$ ). There was no correlation ( $r = 0.01$ ) between patient age and preoperative disability (see table 4, Supplemental Digital Content 1, <http://links.lww.com/ALN/B131>, which is a table describing the preoperative correlations between WHODAS and other health assessment scales).

Concurrent validity of WHODAS in the early postoperative period was tested by measuring its correlation with the QoR-40 score at day 30. There was moderate to strong correlation globally, with  $r = -0.70$  ( $P < 0.0005$ ), and with each dimension of the QoR-40 scale, although as could be expected disability was less related to perceived patient support (table 4). The correlation between WHODAS and QoR-40 was maintained in subgroups stratified by extent of surgery and patient age (see table 5, Supplemental Digital Content 1, <http://links.lww.com/ALN/B131>, which is a table describing the day 30 correlation between WHODAS and QoR-40).

Concurrent validity was further assessed by the correlation between the Katz ADL and WHODAS at 3, 6, and 12 months, with  $r = -0.61$ ,  $r = -0.60$ , and  $\rho = -0.47$ , respectively (all  $P < 0.0005$ ). Again correlation was maintained when stratified for extent of surgery and patient age (see table 7, Supplemental Digital Content 1, <http://links.lww.com/ALN/B131>, which is a table describing correlations between WHODAS and Katz ADL).

There was a good correlation between WHODAS and the mBPI-sf pain scores (table 5), and strong correlation with mean pain interference scores, which increased over time at day 30 ( $r = 0.69$ ), 3 months ( $r = 0.72$ ), 6 months ( $r = 0.74$ ), and 12 months ( $r = 0.81$ ) after surgery (all  $P < 0.0005$ ).

Convergent validity was tested by the correlation between WHODAS and EQ-5D 100-point scale over time at day 30 ( $r = -0.55$ ), 3 months ( $r = -0.57$ ), 6 months ( $r = -0.60$ ), and 12 months ( $r = -0.52$ ) (all  $P < 0.0005$ ) (table 4).

Discriminative validity was excellent. WHODAS was able to discriminate between those with a good and poor quality of recovery after surgery at day 30 and quality of life at 3, 6, and 12 months (table 6).

Construct validity was further assessed by comparing length of stay and complications in patients with and without new

**Table 4.** The Association between WHODAS and the Five Dimensions of the QoR-40 Scale at 30 Days after Surgery (n = 298) and the EuroQOL EQ-5D Scale at 3, 6, and 12 Months after Surgery

Time	Health Instrument	Pearson Correlation with WHODAS 2.0
Day 30	QoR-40 dimension	
	Comfort	-0.58
	Emotions	-0.60
	Physical independence	-0.59
	Patient support	-0.40
	Pain	-0.50
	Total QoR-40 score	-0.70
EQ-5D 100-point scale	EQ-5D 100-point scale	-0.55
	EQ-5D 100-point scale	-0.57
6 months	EQ-5D 100-point scale	-0.60
12 months	EQ-5D 100-point scale	-0.52

All  $P < 0.0005$ .

EQ-5D = the EuroQOL EQ-5D Scale; QoR-40 = the 40-item quality of recovery score; WHODAS = World Health Organization Disability Assessment Schedule 2.0.

**Table 5.** Correlations between WHODAS and the Modified Brief Pain Inventory at Day 30 and at 3, 6, and 12 Months after Surgery

Time Point	Correlation* with WHODAS
Day 30	
Worst pain	0.54
Average pain	0.48
Mean interference	0.69
3 months	
Worst pain	0.54
Average pain	0.47
Mean interference	0.72
6 months	
Worst pain	0.52
Average pain	0.53
Mean interference	0.74
12 months	
Worst pain	0.64
Average pain	0.67
Mean interference	0.81

All  $P < 0.0005$ .

\*Pearson correlation coefficient used.

WHODAS = World Health Organization Disability Assessment Schedule 2.0.

**Table 6.** A Comparison of WHODAS Scores for Those with a Poor or Good Recovery and Quality of Life, Both Defined by the Upper and Lower Quartiles for QoR-40 and EQ-5D 100-Point Scale, Respectively

	Lowest and Highest Quartiles	n	WHODAS Scores	P Value
QoR-40				
30 days	≤169	65	38 [17–53]	0.0005
	≥191	83	0 [0–6]	
EQ-5D 100-point scale				
3 months	≤70	139	21 [6–42]	0.0005
	≥90	140	0 [0–6]	
6 months	≤70	129	17 [4–34]	<0.0005
	≥90	163	0 [0–4]	
12 months	≤70	111	17 [0–40]	<0.0005
	≥90	141	0 [0–4]	

Values are presented as median [interquartile range].

EQ-5D = the EuroQOL EQ-5D; QoR40 = the 40-item quality of recovery score; WHODAS = World Health Organization Disability Assessment Schedule 2.0.

disability at day 30. Those with new disability had a longer hospital stay (median 6.89 *vs.* 5.34 days,  $P = 0.008$ ) and were more likely to have a complication (20% *vs.* 11%,  $P = 0.042$ ). There was a nonsignificant increase in unplanned hospital readmission (9.8% *vs.* 4%,  $P = 0.06$ ) in patients with new disability. The direction and magnitude of change in WHO-DAS score at day 30 also varied with the type of complication, with a mean decrease in disability score of 16% (95% CI, -28 to -3.8,  $P = 0.01$ ) in patients with postoperative myocardial infarction, and a mean increase of 68% (95% CI, 33 to 100,  $P < 0.0005$ ) in patients with a postoperative stroke (see table 10,

Supplemental Digital Content 1, <http://links.lww.com/ALN/B131>, which is a table describing the association between day 30 complications and change in day 30 WHODAS score from baseline). Of the patients that had a myocardial infarction by day 30, 9 of 12 had undergone cardiac surgery. As such, the observed decrease in WHODAS may reflect improved early postoperative function in the cardiac surgery cohort. As expected, older patients were more likely to develop disability postoperatively, with  $r = 0.19, 0.21, \text{ and } 0.22$  at 3, 6, and 12 months, respectively (all  $P < 0.0005$ ).

WHODAS demonstrated excellent reliability. The interitem correlation matrix for WHODAS at 6 months is shown in table 7 and demonstrated good correlation between items with no evidence of item redundancy, indicated by almost all interitem correlations between 0.4 and 0.8. Similar results were obtained for the interitem matrices at day 30 and 3 and 12 months after surgery (results not shown). Cronbach's and split-half coefficients greater were than 0.90 at all time points (table 8). The Cohen effect size was very high at all times demonstrating excellent responsiveness (table 8).

WHODAS had very good scaling properties. The 10th, 25th, 50th, 75th, and 90th centiles were 0, 0, 2.1, 17, and 33, respectively. A floor effect was present,<sup>43</sup> with more than 40% of patients having little or no disability at 6 months, but otherwise there was very good spread of data. The scaling properties are demonstrated in figure 2, with 40% of patients having a score of zero and 85% of patients having a WHODAS score of less than 25%.

Disability-free survival at day 30 and at 3, 6, and 12 months after surgery was 72% and 74%, 80%, and 76%, respectively. When analyzing disability-free survival, disability had a greater contribution to disability-free survival rates than patient mortality (table 9). The pattern of recovery after surgery varied according to the patient's ASA physical status (fig. 3), with higher rates of disability-free survival for patients with lower ASA scores at all times ( $P$  for trend  $< 0.0005$ ). Compared to their preoperative state, ASA I and II patients tended to have less disability by day 30 and continued to improve out to 6 months. By contrast, ASA III and IV patients tended to have a more delayed recovery, with decreased disability-free survival at day 30 and significant recovery not occurring until 3 months. In general, all ASA groups plateaued by 6 months, with only slight decreases in disability-free survival afterward.

Disability-free survival and new disability also varied according to the type of surgery, with the lowest rates of disability-free survival at 6 months being in patients having orthopedic (67%) or neurosurgery (58%), and the highest rates of new disability occurring in patients having thoracic surgery (see table 12, Supplemental Digital Content 1, <http://links.lww.com/ALN/B131>, which is a table comparing the rate of disability-free survival and new disability at 6 months according to the type of surgery).

**Table 7.** Interitem Correlation Matrix for the 12-item WHODAS at 6 Months

WHODAS Item	1	2	3	4	5	6	7	8	9	10	11	12
1	—											
2	0.76	—										
3	0.55	0.65	—									
4	0.53	0.66	0.76	—								
5	0.53	0.61	0.44	0.57	—							
6	0.55	0.61	0.71	0.65	0.52	—						
7	0.78	0.73	0.52	0.51	0.58	0.53	—					
8	0.56	0.72	0.66	0.62	0.45	0.58	0.54	—				
9	0.58	0.68	0.67	0.60	0.38	0.63	0.51	0.87	—			
10	0.52	0.60	0.70	0.75	0.55	0.64	0.48	0.58	0.61	—		
11	0.49	0.55	0.63	0.69	0.47	0.60	0.44	0.59	0.60	0.78	—	
12	0.63	0.73	0.64	0.71	0.63	0.59	0.61	0.62	0.60	0.67	0.62	—

1 = standing for long periods; 2 = household responsibilities; 3 = learning a new task; 4 = joining in community activities; 5 = emotionally affected by your health problems; 6 = concentrating for 10 min; 7 = walking a long distance; 8 = washing your whole body; 9 = getting dressed; 10 = dealing with people you do not know; 11 = maintaining a friendship; 12 = day-to-day work.

WHODAS = World Health Organization Disability Assessment Schedule 2.0.

**Table 8.** Reliability and Responsiveness\* Testing of WHODAS

Time	n	Cronbach's $\alpha$	Split-half Reliability	Responsiveness (Cohen Effect Size)
3 months	301	0.94	0.95	3.4
6 months	311	0.94	0.96	3.0
12 months	298	0.94	0.95	1.1

All  $P < 0.0005$ .

\*See Materials and Methods section for details.

WHODAS = World Health Organization Disability Assessment Schedule 2.0.

## Discussion

This study was able to confirm that WHODAS retains its excellent psychometric properties found in community and medical populations when measuring disability in an adult surgical population. The broad range of patient demographics, medical comorbidities, surgical type and extent, and consistent psychometric indices in selected strata offer strong support for the generalizability of our findings to other surgical settings.

Overall, we had a very good participant retention rate at 12 months after surgery (85%) and excellent clinical acceptability as reflected by WHODAS completion rates between 88% and 92% *via* postal survey. It is likely that patient acceptability would be further improved if WHODAS was completed as the sole telephone survey instrument instead of being one of several postal surveys.

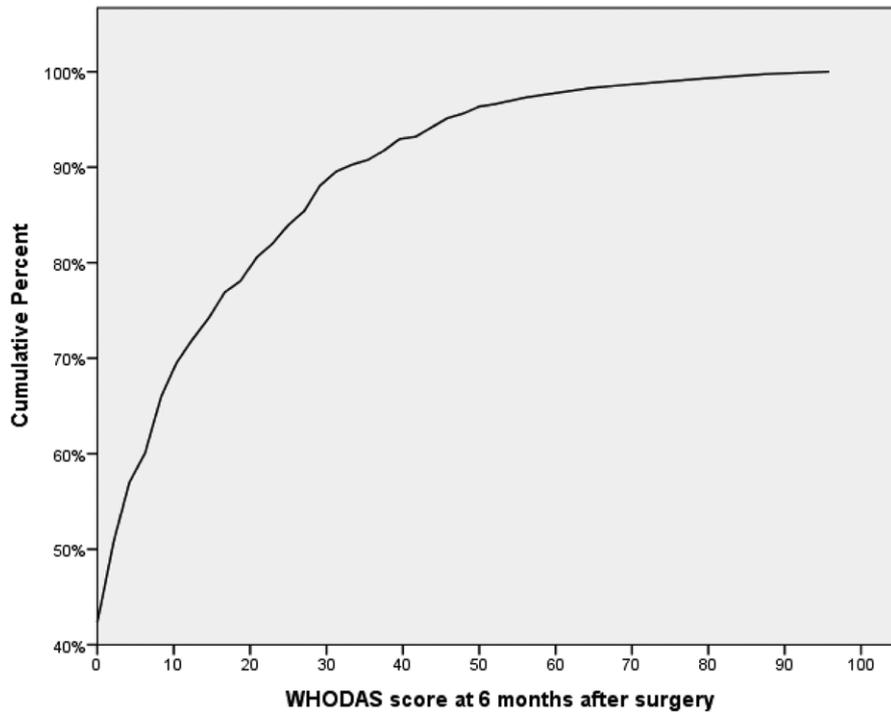
In the absence of a “definitive standard” patient-centered long-term outcome measure after surgery, the correlation of WHODAS with existing well-validated health status instruments (QoR-40, Katz ADL, EQ-5D, and mBPI-sf) that measure related but different constructs was used to assess validity. As expected, there was moderate but not high correlation ( $r = 0.5$  to  $0.7$ ) between scores, supporting a conclusion that these scales do not assess the same construct;

if so, WHODAS would seem redundant. There was strong correlation between the WHODAS score and the mean pain interference score of the mBPI-sf. This correlation increased with time after surgery and may be indicative of the influence chronic postsurgical pain has on persistent postoperative disability.

Construct validity testing revealed a number of interesting relationships between disability and patient characteristics in a surgical population. WHODAS demonstrated good discriminative validity, being able to clearly distinguish between patients with good or poor recovery at 30 days, and self-rated quality of life at 3, 6, and 12 months. Although there was modest correlation between preoperative disability and ASA physical status, there was no correlation between age and preoperative disability. By contrast, it is known that disability tends to increase with age in the community setting.<sup>31</sup> The lack of correlation in our study is almost certainly a true finding rather than a problem with performance of WHODAS in a surgical population, as these scores were established before surgery. By contrast, there was modest correlation ( $r = 0.17$  to  $0.21$ ) between patient age and postoperative disability. This finding should be consistent with most clinicians' experience, knowing that older patients having surgery are more likely to develop difficulties after surgery.<sup>3,44,45</sup>

Orthopedic patients had a low rate (67%) of disability-free survival at 6 months after surgery. This may reflect our cohort in that three of the recruiting hospitals have a trauma focus. However, 18 of the 75 orthopedic patients had elective hip or knee arthroplasty, and these patients had even lower rates (53%) of disability-free survival at 6 months. Orthopedic patients seem to have poorer rates of disability-free survival than anticipated, most probably due to persistent postsurgical pain in this group.<sup>46,47</sup>

There was modest correlation between preoperative frailty and disability measures. There are several reasons why



WHODAS score	N	Valid %	Cumulative %
0	170	41.8	41.8
2	34	8.4	50.1
4	26	6.4	56.5
6	13	3.2	59.7
8	24	5.9	65.6
10	15	3.7	69.3
13	10	2.5	71.7
15	9	2.2	74.0
17	11	2.7	76.7
19	5	1.2	77.9
21	10	2.5	80.3
23	6	1.5	81.8
25	8	2.0	83.8
27	6	1.5	85.3
29	11	2.7	88.0
31	6	1.5	89.4
33	3	0.7	90.2
35	2	0.5	90.7
38	4	1.0	91.6
40	5	1.2	92.9
42	1	0.2	93.1
44	4	1.0	94.1
46	4	1.0	95.1
48	2	0.5	95.6
50	3	0.7	96.3
52	1	0.2	96.6
56	3	0.7	97.3
58	1	0.2	97.5
63	2	0.5	98.0
65	1	0.2	98.3
79	4	1.0	99.3
88	2	0.5	99.8
96	1	0.2	100

**Fig. 2.** The cumulative percentage of World Health Organization Disability Assessment Schedule (WHODAS) scores at 6 months after surgery, depicting its scaling properties.

this correlation is lower than one might intuitively expect. Disability and frailty are different constructs and while most frail people are likely to have at least some disability, the reverse may not be true. In addition, poor correlation may exemplify the difference between subjective clinician- or investigator-rated scales and patient-rated assessments. Finally, the modest correlation may reflect the fact that there was a low rate of frailty in our cohort, thus limiting study power for this evaluation. As an aside, we were able to demonstrate a high level of agreement between anesthesiologist- and investigator-determined clinical frailty.

Although the study was not powered to determine the discriminant validity of WHODAS to reflect the impact of

postoperative complications, patients with a new disability after surgery were more likely to have had one or more postoperative complications and also had a longer hospital stay. There was also a trend toward increased unplanned readmission in patients with new disability. The relationship between the type of postoperative complication and subsequent change in WHODAS score highlights a major benefit of using a patient-centered outcome measure over traditional unweighted cardiovascular endpoints. The two patients that suffered a postoperative stroke had substantially increased disability, whereas the reverse was true for patients deemed to have had a postoperative myocardial infarction. Although most myocardial infarctions occurred in patients after cardiac surgery, this finding still demonstrates a potential problem when using traditional outcome measures, in that from a patient's perspective, stroke is likely to be a much more serious and disabling complication than myocardial infarction.

As in previous studies in other settings, WHODAS was found to be highly reliable and very responsive to change.<sup>17</sup> In our study, an expected floor effect was demonstrated, with 40% of patients having a WHODAS score of 0% (*i.e.*, no disability) at 6 months after surgery. However, this is similar to WHODAS scaling properties in the general population<sup>31</sup> and probably reflects a true incidence of people with no measurable disability rather than a problem with the lower end of the scale. Indeed, there was a good distribution of scores across the remainder of the scale.

While health-related quality of life is an important outcome measure in its own right,<sup>8,48</sup> such measures cannot be

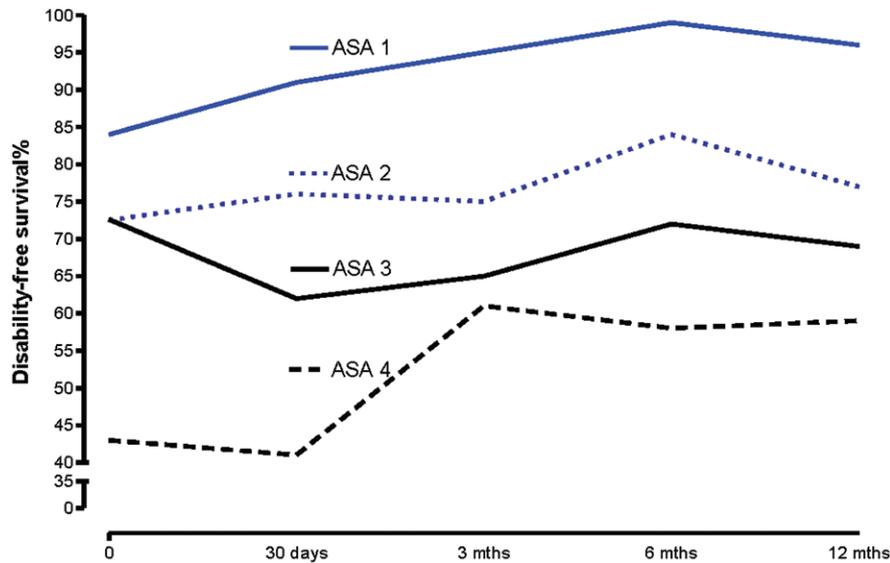
**Table 9.** Disability-free Survival According to ASA Physical Status at 6 Months

ASA	Survival at 6 Months			Disability-free Survival* at 6 Months		P for Trend
	n	Frequency	P for Trend	n	Frequency	
I	80	80 (100)	0.019	72	71 (99)	<0.0005
II	189	187 (99)		171	144 (84)	
III	162	159 (98)		144	103 (72)	
IV	24	22 (92)		24	14 (58)	

Values are presented as number (%).

\*Disability-free survival was calculated as the percentage of participants who were both alive and had a WHODAS score of  $\leq 25\%$  (indicating they had at least moderate disability) at a particular time point.

ASA = American Society of Anesthesiologists; WHODAS = World Health Organization Disability Assessment Schedule 2.0.



**Fig. 3.** Disability-free survival after surgery according to American Society of Anesthesiologists (ASA) physical status score.

simply dichotomized, are not designed for repeat testing or to be responsive to change, are heavily influenced by social and economic circumstances, and may overlook important aspects of functional independence. Of greater relevance, however, is that the general aims of surgery and other interventional procedures are to cure or relieve symptoms of a disease state. Survival and freedom from disability, therefore, should be measured after surgery.<sup>49</sup> Around 20% of elderly patients have one or more serious complications after surgery.<sup>3</sup> Many more never fully recover after their surgery and seem to have accelerated disability in the months and years that follow.<sup>50–52</sup>

Postoperative disability, as measured by WHODAS, is a valid and reliable clinical endpoint that is well suited for future anesthetic and surgical research. The WHODAS is simple to use and interpret and meaningful to clinicians and patients alike. The high rates of clinically significant preoperative (27%) and postoperative (16 to 22%) disability mean that comparative studies using disability as an endpoint would require modest sample sizes to obtain adequate statistical power. Bearing in mind that the current study contained patients with a mixed risk profile, disability rates can be expected to be higher in clinical trials enrolling high-risk surgical patients.

At present, investigators designing randomized trials have tended toward combining several complications or outcomes into one composite endpoint to increase the event rate and thereby decrease the sample size required to demonstrate a clinically important effect. This can be problematic.<sup>53</sup> Composite endpoints can be misleading when one of the outcomes in the composite has a higher incidence than the others or otherwise carries significantly less patient burden. Adopting disability-free survival as a primary endpoint in clinical studies should circumvent this problem.

Disability-free survival is an ideal study endpoint as it reflects the primary goal for most patients undergoing major surgery and can aid shared decision-making in surgical care.<sup>54</sup> It can be used as a single primary endpoint, and when using survival analysis, it has enhanced statistical power. It is particularly suitable for clinical trials in which groups have comparable baseline risk. It may be more difficult, however, to observe a clear disability signal when observing a group of patients with a mixed risk profile, having surgery of varying extent and type due to likely confounding effects. In this situation, it may be more useful to measure either rates of new disability or a significant change in WHODAS score. In addition, WHODAS would be an ideal measure for ongoing audit and clinical quality improvement processes.

#### Limitations of the Study

This study may be subject to nonresponder bias as patients who withdrew or were lost to follow-up were also more likely to have clinically significant preoperative disability. While some of these patients may have improved postoperatively, it is possible that missed patients may have developed worse disability and that the rates of disability were underestimated. This potential bias was minimized by low overall withdrawal and loss to follow-up rates. We acknowledge that not all postoperative disability may be directly attributable to the index surgery. This is particularly true at 6 and 12 months after surgery when intervening unrelated life events may result in overestimation of surgery-induced postoperative disability. On the other hand, the stress of surgery may precipitate a series of unrelated morbid events because of the patient's vulnerable status—the so-called post-hospital syndrome.<sup>50</sup> Of note, 54 (11%) patients in our study had further planned or unrelated procedures in the 12 months following their index surgery.

Based on previous literature,<sup>18</sup> we used a WHODAS score of greater than or equal to 25% to define clinically significant disability and the change in WHODAS score of greater than or equal to 8% to define a minimal clinically important difference.<sup>31</sup> These cut points require further verification to ensure they correspond to clinically meaningful endpoints in surgical populations.

In conclusion, WHODAS is a clinically acceptable, valid, reliable, and responsive instrument for measuring disability in a surgical population. Freedom from disability after surgery is a meaningful outcome for clinicians and patients alike. We recommend disability-free survival as an important endpoint in clinical trials.

### Acknowledgments

The authors thank Ed O'Loughlin, M.B., B.S., F.A.N.Z.C.A., M.Clin.Research (Fremantle Hospital and University of Western Australia), Kate Turnahan, B.Sc.(nurs), M.N. (Fremantle Hospital, Western Australia), Daniel Myles, B.A., G.Dip.Psych. (Monash University, Melbourne, Australia), and Marie Backstrom, B.Sc.(nurs) (Monash Medical Centre, Clayton, Australia), for their assistance with data collection.

This study was supported by The Australian and New Zealand College of Anaesthetists (Melbourne, Victoria, Australia) and a Direct Grant for Research (4054079). The Chinese University of Hong Kong (Shatin, New Territories, Hong Kong) provided funding for this work.

### Competing Interests

The authors declare no competing interests.

### Correspondence

Address correspondence to Dr. Myles: Department of Anaesthesia and Perioperative Medicine, Alfred Hospital, Commercial Road, Melbourne, Victoria, 3004, Australia. p.myles@alfred.org.au. This article may be accessed for personal use at no charge through the Journal Web site, www.anesthesiology.org.

### References

1. Fisher DM: Surrogate outcomes: Meaningful not! *ANESTHESIOLOGY* 1999; 90:355–6
2. Lee A, Lum ME: Measuring anaesthetic outcomes. *Anaesth Intensive Care* 1996; 24:685–93
3. Story DA, Leslie K, Myles PS, Fink M, Poustie SJ, Forbes A, Yap S, Beavis V, Kerridge R; REASON Investigators, Australian and New Zealand College of Anaesthetists Trials Group: Complications and mortality in older surgical patients in Australia and New Zealand (the REASON study): A multicentre, prospective, observational study. *Anaesthesia* 2010; 65:1022–30
4. Devereaux PJ, Yang H, Yusuf S, Guyatt G, Leslie K, Villar JC, Xavier D, Chrolavicius S, Greenspan L, Pogue J, Pais P, Liu L, Xu S, Malaga G, Avezum A, Chan M, Montori VM, Jacka M, Choi P: Effects of extended-release metoprolol succinate in patients undergoing non-cardiac surgery (POISE trial): A randomised controlled trial. *Lancet* 2008; 371:1839–47
5. Myles PS, Williams DL, Hendrata M, Anderson H, Weeks AM: Patient satisfaction after anaesthesia and surgery: Results of a prospective survey of 10,811 patients. *Br J Anaesth* 2000; 84:6–10
6. Hopkins PM: A shame about the patient? *Br J Anaesth* 2000; 84:1–2
7. Myles PS, Hunt JO, Nightingale CE, Fletcher H, Beh T, Tanil D, Nagy A, Rubinstein A, Ponsford JL: Development and psychometric testing of a quality of recovery score after general anaesthesia and surgery in adults. *Anesth Analg* 1999; 88:83–90
8. Mangione CM, Goldman L, Orav EJ, Marcantonio ER, Pedan A, Ludwig LE, Donaldson MC, Sugarbaker DJ, Poss R, Lee TH: Health-related quality of life after elective surgery: Measurement of longitudinal changes. *J Gen Intern Med* 1997; 12:686–97
9. Royse CF, Newman S, Chung F, Stygall J, McKay RE, Boldt J, Servin FS, Hurtado I, Hannallah R, Yu B, Wilkinson DJ: Development and feasibility of a scale to assess postoperative recovery: The post-operative quality recovery scale. *ANESTHESIOLOGY* 2010; 113:892–905
10. Hogue SL, Reese PR, Colopy M, Fleisher LA, Tuman KJ, Twersky RS, Warner DS, Jamerson B: Assessing a tool to measure patient functional ability after outpatient surgery. *Anesth Analg* 2000; 91:97–106
11. Myles PS, Viira D, Hunt JO: Quality of life at three years after cardiac surgery: Relationship with preoperative status and quality of recovery. *Anaesth Intensive Care* 2006; 34:176–83
12. Crawford RS, Pedraza JD, Chung TK, Corey M, Conrad MF, Cambria RP: Functional outcome after thoracoabdominal aneurysm repair. *J Vasc Surg* 2008; 48:828–35
13. World Health Organization: World Report on Disability 2011. Geneva, World Health Organization, 2011, pp 1–90
14. Leonardi M, Bickenbach J, Ustun TB, Kostanjsek N, Chatterji S; MHADIE Consortium: The definition of disability: What is in a name? *Lancet* 2006; 368:1219–21
15. World Health Organization: International Classification of Functioning, Disability and Health. Geneva, World Health Organization, 2001, pp 1–299
16. McDowell I: Measuring Health: A Guide to Rating Scales and Questionnaires, 3rd edition. New York, Oxford University Press, 2006, pp 3–149
17. Ustun TB, Chatterji S, Kostanjsek N, Rehm J, Kennedy C, Epping-Jordan J, Saxena S, von Korff M, Pull C; WHO/NIH Joint Project: Developing the World Health Organization Disability Assessment Schedule 2.0. *Bull World Health Organ* 2010; 88:815–23
18. Ustun TB, Kostanjsek N, Chatterji S, Rehm J: Measuring Health and Disability: Manual for WHO Disability Assessment Schedule (WHODAS 2.0). Geneva, World Health Organization, 2010, pp 1–88
19. Soberg HL, Finset A, Roise O, Bautz-Holter E: The trajectory of physical and mental health from injury to 5 years after multiple trauma: A prospective, longitudinal cohort study. *Arch Phys Med Rehabil* 2012; 93:765–74
20. Derrett S, Samaranyaka A, Wilson S, Langley J, Ameratunga S, Cameron ID, Lilley R, Wyeth E, Davie G: Prevalence and predictors of sub-acute phase disability after injury among hospitalised and non-hospitalised groups: A longitudinal cohort study. *PLoS One* 2012; 7:e44909
21. Küçükdeveci AA, Kutlay Ş, Yıldızlar D, Öztuna D, Elhan AH, Tennant A: The reliability and validity of the World Health Organization Disability Assessment Schedule (WHODAS-II) in stroke. *Disabil Rehabil* 2013; 35:214–20
22. Schlote A, Richter M, Wunderlich MT, Poppendick U, Möller C, Schwelm K, Wallesch CW: WHODAS II with people after stroke and their relatives. *Disabil Rehabil* 2009; 31:855–64
23. Wolf AC, Tate RL, Lannin NA, Middleton J, Lane-Brown A, Cameron ID: The World Health Organization Disability Assessment Scale, WHODAS II: Reliability and validity in the measurement of activity and participation in a spinal cord injury population. *J Rehabil Med* 2012; 44:747–55
24. Garin O, Ayuso-Mateos JL, Almansa J, Nieto M, Chatterji S, Vilagut G, Alonso J, Cieza A, Svetszkova O, Burger H, Racca V, Francescutti C, Vieta E, Kostanjsek N, Raggi A, Leonardi

- M, Ferrer M; MHADIE consortium: Validation of the "World Health Organization Disability Assessment Schedule, WHODAS-2" in patients with chronic diseases. *Health Qual Life Outcomes* 2010; 8:51
25. Streiner DL, Norman GR: *Health Measurement Scales: A Practical Guide to Their Development and Use*, 4th edition. Ontario, Oxford University Press, 2008, pp 1–423
  26. Myles PS, Weitkamp B, Jones K, Melick J, Hensen S: Validity and reliability of a postoperative quality of recovery score: The QoR-40. *Br J Anaesth* 2000; 84:11–5
  27. The EuroQol Group: EuroQol—A new facility for the measurement of health-related quality of life. *Health Policy* 1990; 16:199–208
  28. Dyer MT, Goldsmith KA, Sharples LS, Buxton MJ: A review of health utilities using the EQ-5D in studies of cardiovascular disease. *Health Qual Life Outcomes* 2010; 8:13
  29. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW: Studies of illness in the aged. The index of ADL: A standardized measure of biological and psychosocial function. *JAMA* 1963; 185:914–9
  30. Mendoza TR, Chen C, Brugger A, Hubbard R, Snabes M, Palmer SN, Zhang Q, Cleeland CS: The utility and validity of the modified brief pain inventory in a multiple-dose postoperative analgesic trial. *Clin J Pain* 2004; 20:357–62
  31. Andrews G, Kemp A, Sunderland M, Von Korff M, Ustun TB: Normative data for the 12 item WHO Disability Assessment Schedule 2.0. *PLoS One* 2009; 4:e8343
  32. Gornall BF, Myles PS, Smith CL, Burke JA, Leslie K, Pereira MJ, Bost JE, Kluijvers KB, Nilsson UG, Tanaka Y, Forbes A: Measurement of quality of recovery using the QoR-40: A quantitative systematic review. *Br J Anaesth* 2013; 111:161–9
  33. Prytherch DR, Whiteley MS, Higgins B, Weaver PC, Prout WG, Powell SJ: POSSUM and Portsmouth POSSUM for predicting mortality. Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity. *Br J Surg* 1998; 85:1217–20
  34. Brooks MJ, Sutton R, Sarin S: Comparison of Surgical Risk Score, POSSUM and p-POSSUM in higher-risk surgical patients. *Br J Surg* 2005; 92:1288–92
  35. Rockwood K, Song X, MacKnight C, Bergman H, Hogan DB, McDowell I, Mitnitski A: A global clinical measure of fitness and frailty in elderly people. *CMAJ* 2005; 173:489–95
  36. Copeland GP, Jones D, Walters M: POSSUM: A scoring system for surgical audit. *Br J Surg* 1991; 78:355–60
  37. Fitzpatrick R, Davey C, Buxton MJ, Jones DR: Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess* 1998; 2:i–iv, 1–74
  38. Sousa RM, Dewey ME, Acosta D, Jotheeswaran AT, Castro-Costa E, Ferri CP, Guerra M, Huang Y, Jacob KS, Rodriguez Pichardo JG, Garcia Ramirez N, Llibre Rodriguez J, Calvo Rodriguez M, Salas A, Sosa AL, Williams J, Prince MJ: Measuring disability across cultures—The psychometric properties of the WHODAS II in older people from seven low- and middle-income countries. The 10/66 Dementia Research Group population-based survey. *Int J Methods Psychiatr Res* 2010; 19:1–17
  39. Kazis LE, Anderson JJ, Meenan RF: Effect sizes for interpreting changes in health status. *Med Care* 1989; 27(3 Suppl):S178–89
  40. Gower ST, Quigg CA, Hunt JO, Wallace SK, Myles PS: A comparison of patient self-administered and investigator-administered measurement of quality of recovery using the QoR-40. *Anaesth Intensive Care* 2006; 34:634–8
  41. Cronbach LJ: Coefficient alpha and the internal structure of tests. *Psychometrika* 1951; 6:297–334
  42. Cohen J: A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960; 20:37–46
  43. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC: Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007; 60:34–42
  44. Khuri SF, Henderson WG, DePalma RG, Mosca C, Healey NA, Kumbhani DJ; Participants in the VA National Surgical Quality Improvement Program: Determinants of long-term survival after major surgery and the adverse effect of postoperative complications. *Ann Surg* 2005; 242:326–41; discussion 341–3
  45. Silbert BS, Evered LA, Scott DA: Cognitive decline after surgery and illness. *ANESTHESIOLOGY* 2010; 112:1282; author reply 1283–5
  46. Fuzier R, Serres I, Bourrel R, Palmaro A, Montastruc JL, Lapeyre-Mestre M: Analgesic drug consumption increases after knee arthroplasty: A pharmacoepidemiological study investigating postoperative pain. *Pain* 2014; 155:1339–45
  47. Wyld V, Hewlett S, Learmonth ID, Dieppe P: Persistent pain after joint replacement: Prevalence, sensory qualities, and postoperative determinants. *Pain* 2011; 152:566–72
  48. Lawrence VA, Hazuda HP, Cornell JE, Pederson T, Bradshaw PT, Mulrow CD, Page CP: Functional independence after major abdominal surgery in the elderly. *J Am Coll Surg* 2004; 199:762–72
  49. Chalmers I, Clarke M: Outcomes that matter to patients in tombstone trials. *Lancet* 2001; 358:1649
  50. Krumholz HM: Post-hospital syndrome—An acquired, transient condition of generalized risk. *N Engl J Med* 2013; 368:100–2
  51. Amemiya T, Oda K, Ando M, Kawamura T, Kitagawa Y, Okawa Y, Yasui A, Ike H, Shimada H, Kuroiwa K, Nimura Y, Fukata S: Activities of daily living and quality of life of elderly patients after elective surgery for gastric and colorectal cancers. *Ann Surg* 2007; 246:222–8
  52. Ferrucci L, Guralnik JM, Studenski S, Fried LP, Cutler GB Jr, Walston JD; Interventions on Frailty Working Group: Designing randomized, controlled trials aimed at preventing or delaying functional decline and disability in frail, older persons: A consensus report. *J Am Geriatr Soc* 2004; 52:625–34
  53. Myles PS, Devereaux PJ: Pros and cons of composite endpoints in anesthesia trials. *ANESTHESIOLOGY* 2010; 113:776–8
  54. Glance LG, Osler TM, Neuman MD: Redesigning surgical decision making for high-risk patients. *N Engl J Med* 2014; 370:1379–81