

# Estimating the minimal clinically important difference for the Physical Component Summary of the Short Form 36 for patients with stroke

Vivian Fu<sup>1</sup>, Mark Weatherall<sup>2</sup> and Harry McNaughton<sup>1</sup> 

## Abstract

**Objective:** To determine the Physical Component Summary (PCS) score's minimal clinically important difference (MCID) on the Short Form 36 (SF-36) for people with stroke.

**Methods:** We conducted secondary analysis of data from a large randomized controlled trial (N = 400) in the post-hospital discharge phase of stroke rehabilitation with outcome measurement 6 and 12 months following stroke. Three methods were used for estimating the MCID: two anchor and one distribution. Method 1 compared SF-36 PCS scores at 12 months for responses to the SF-36's Perceived Health Change (PHC) question. Method 2 compared the change in PCS score between 6 and 12 months for responses to the PHC question. Method 3 used Cohen's method to estimate the MCID from the PCS score distribution.

**Results:** Method 1: the mean PCS score increased by 3.0 units (95% confidence interval [CI] 2.2–3.9) for each unit change in the PHC question. Method 2: the mean change in PCS score increased by 2.1 units (95% CI 1.4–2.8) for each unit change in the PHC question. Method 3: the MCID was estimated to be 1.8 units.

**Conclusions:** Our estimate of the MCID for the PCS in patients with stroke was 1.8 to 3.0 units.

## Keywords

Stroke rehabilitation, minimal clinically important difference, Short Form 36, stroke, outcome measure, physical component summary

Date received: 9 September 2021; accepted: 2 December 2021

## Corresponding author:

Harry McNaughton, Stroke/Rehabilitation Research, Medical Research Institute of New Zealand, Level 7, CSB Building, Wellington Hospital, Riddiford St., Newtown, Wellington 6021, New Zealand.  
Email: [harry.mcnaughton@mrinz.ac.nz](mailto:harry.mcnaughton@mrinz.ac.nz)

<sup>1</sup>Stroke/Rehabilitation Research, Medical Research Institute of New Zealand, Wellington, New Zealand

<sup>2</sup>Department of Medicine, University of Otago, Wellington, New Zealand



## Introduction

Progress in stroke rehabilitation research has been hampered by the absence of generally agreed outcome measurement tools.<sup>1,2</sup> Researchers, grant allocation committees and journal editors require tools that are valid, responsive to change, appropriate across the full range of stroke severity, but also meaningful to a non-stroke-specific audience. Ideally, these tools would also contribute to meaningful estimates of the cost-effectiveness of any intervention being tested. The Physical Component Summary (PCS) score of the Short Form 36 (SF-36) meets many of these conditions. The SF-36 is a 36-question self-report instrument that was developed to measure health-related quality of life. The PCS is derived from a selection of responses to yield a single score between 0 and 100, with a mean of 50 and standard deviation (SD) of 10 in population studies, which reflects “physical health.” The PCS has robust psychometric properties, has been shown to be responsive to changes in stroke, and has been used in a large number of trials in many different conditions. The questions from the SF-36 can be used to calculate tariffs for cost-effectiveness analysis.<sup>3-5</sup>

For patients with stroke, however, the minimal clinically important difference (MCID) for the PCS of the SF-36 has not been adequately defined. The MCID was originally defined as “the smallest difference in score in the domain of interest which patients perceive as beneficial and which would mandate, in the absence of troublesome side effects and excessive cost, a change in the patient’s management.”<sup>6</sup> A more concise definition is “the smallest benefit of value to patients.”<sup>7</sup> Tools with an established MCID for the condition under study allow for adequate powering of studies, a critical element in planning a trial and a universal requirement

for grant funding applications. Such tools also allow for appropriate interpretation of the clinical relevance of statistically significant study results. To define the MCID for a condition, the score difference must be related to a meaningful and beneficial change in health status perceived by the patient.<sup>8,9</sup> “Norman’s Law” proposes a standard MCID of 5 units (half the SD) for the PCS for all chronic conditions with any degree of severity, with no rigorous testing of this hypothesis for patients with stroke.<sup>10</sup> Others have argued that an MCID of 2.5 units is generally accepted for the PCS with rheumatological conditions.<sup>11</sup> Researchers have sought to establish MCIDs for the PCS in different conditions, with a range of 2 to 6 points described, clustering around 2 to 3 points.<sup>12-14</sup> As part of the Taking Charge after Stroke (TaCAS) study among 400 individuals after acute stroke, in this study, we aimed to describe the MCID of the PCS for patients following stroke.

## Methods

This was a secondary analysis of data from the TaCAS study,<sup>15</sup> a randomized controlled trial of a novel intervention (the “Take Charge” intervention) compared with controls for people discharged to community living following stroke. The full methods and main results of that study have been previously reported, according to CONSORT (Consolidated Standards of Reporting Trials) guidelines.<sup>15</sup> In the current analysis, we treated study participants as a single cohort, reported in line with the STROBE (STrengthening the Reporting of OBServational studies in Epidemiology) guidelines for an observational cohort study.<sup>16</sup> In the TaCAS study, the SF-36 was measured 12 months after the index stroke, representing the final follow-up. In addition, the Short Form 12 (SF-12, described below) was measured after

6 months. PCS scores were calculated from the raw data using proprietary software supplied under license from QualityMetric ([www.qualitymetric.com](http://www.qualitymetric.com)), which is a standard, validated approach to scoring.

### *Instrument and MCID estimation*

The SF-36 is a generic 36-item, patient-reported survey of health, widely used in clinical studies across many conditions.<sup>17</sup> The survey contains eight domains: vitality, physical functioning, bodily pain, general health perceptions, physical functioning, emotional functioning, social functioning, and mental health. There are no disease-specific questions. Possible scores range from 0 to 100, with higher scores representing better health status. These scores are Z-transformed and weighted to calculate PCS and Mental Component Summary (MCS) scores, which are norm based with a mean of 50 and an SD of 10.<sup>18</sup> The PCS score can be interpreted as describing the physical health of the person. The PCS has excellent psychometric properties, including for people with stroke.<sup>3,4</sup> The SF-12 includes 12 of the 36 items found on the SF-36. The PCS for the SF-12 is calculated in a similar fashion to that of the SF-36. The comparability of the PCS SF-12 and PCS SF-36 has been cross-validated using data from general population surveys in multiple countries but has not been assessed specifically for people with stroke.<sup>19</sup> Pearson correlation coefficients between PCS SF-36 and PCS SF-12 scores in those studies were very high, ranging from 0.94 to 0.96. Norm-based scoring and a fixed population-based mean of 50 (SD = 10) allows for treatment of the PCS SF-36 and PCS SF-12 as being on the same scale.

We estimated the MCID using three methods: two anchor based and one distribution based.<sup>8,20</sup> For anchor-based methods, a modified version of the Perceived

Health Change (PHC) question from the SF-36 after 12 months was used as the anchor. In line with our definition of MCID, we assumed that patients would generally agree that a one-level difference in this score would represent a meaningful change. Importantly, this question does not form part of the calculation of the PCS. We modified this question, which reads “Compared with 1 year ago, how would you rate your health in general now: much better, somewhat better, about the same, somewhat worse, or much worse?” to read “Compared with 6 months ago, how would you rate your health: much better, somewhat better, about the same, somewhat worse, much worse?” The reasoning for this was that if participants were asked this question at 12 months following stroke, there may be confusion about whether this question was a comparison with pre-stroke function rather than with the previous assessment at 6 months. Data summaries were calculated for the PCS score at 12 months for each level of the PHC question. The second anchor-based estimation for the MCID was the difference in PCS score at 12 months (based on the SF-36) and the PCS score at 6 months (based on the SF-12), for each level on the PHC question. Linear regression was used to estimate the difference in PCS scores (method 1) and PCS 12-month versus PCS 6-month difference scores (method 2) per 1-unit increase in PHC, with the resultant regression coefficient nominated as the MCID.

For the distribution-based method of MCID estimation, we used the Cohen effect-size benchmark.<sup>21</sup> According to the Cohen score, differences of 0.2 SD units correspond to small but important changes in treatment trials,<sup>20</sup> which is widely supported.<sup>20,22,23</sup> The root mean square error (RMSE) from the analysis of variance (ANOVA) for the main comparison was used to estimate the SD.

### Ethical considerations and statistical analysis

The study protocol for the TaCAS trial was approved by the Health and Disability Ethics Committee. Written informed consent was obtained from all participants. The trial is registered with the Australia New Zealand Clinical Trials Registry ACTRN12615001163594.

We used SAS version 9.4 for the analyses (SAS Institute, Cary, NC, USA).

### Results

A total of 400 adults were randomized in the TaCAS study. The mean (SD) age of participants was 72.0 (12.5) years and 234 (58.5%) were men. The mean (SD) days from stroke in the 6-month and 12-month assessments was 190 (31) and 354 (32) days, respectively. Of the original 400 participants, 369 (92.3%) were assessed at 6 months, and 351 of these (95%) completed the SF-12. Of 388 participants (97% of the total) contacted at 12 months, 381 (98.2%) completed the SF-36. Mean (SD) scores for the PCS at 6 months and 12 months for all participants combined were 43.5 (8.9) and 45.4 (9.2), respectively.

#### MCID assessment

Method 1: Perceived Health Change (PHC) question as anchor, PCS at 12 months

The summary data for the PCS score for each level of PHC are shown in Table 1. Mean PCS scores for responses of “same” and “somewhat better” for the PHC question were virtually identical (46.4 and 46.2, respectively). Otherwise, mean scores were distributed over a large range, from 22.2 for “much worse” to 48.3 for “much better.” The mean (95% confidence interval [CI]) PCS score increased by 3.0 units (2.2–3.9) for each unit change in the PHC question ( $p < 0.001$ ).

**Table 1.** Mean SF-36 PCS scores at 12 months after stroke for each level of the Perceived Health Change question.

Perceived health change at 12 months*	SF-36 PCS mean (SD)
1 = Much better, n = 110	48.3 (8.2)
2 = Somewhat better, n = 106	46.2 (8.3)
3 = Same, n = 107	46.4 (8.8)
4 = Somewhat worse, n = 55	37.4 (8.4)
5 = Much worse, n = 3	22.2 (4.3)

\* Question 2 from the Short Form 36 v2 (New Zealand version), with modification (see text).

SF-36 PCS, Physical Component Summary score of the Short Form 36; SD, standard deviation.

Method 2: Perceived Health Change (PHC) question as anchor, change in PCS score between 6 and 12 months.

The summary data for the difference in PCS scores for each level of PHC are shown in Table 2. Scores for the mean change in PCS were appropriately distributed, from –8.9 units for “much worse” to +4.2 units for “much better.” The mean (95% CI) change in PCS scores increased by 2.1 units (1.4–2.8) for each unit change in the PHC question ( $p < 0.001$ ).

Method 3: Distribution method

The RMSE from the ANOVA for the PCS SF-36 was 9.2 (Table 3). Based on the RMSE of 9.2 and 0.2 as a small effect size, a small effect would be  $0.2 \times 9.2 = 1.8$  units.

### Discussion

Using a combination of anchor-based and distribution-based methods, we estimated the MCID of the SF-36 PCS to be between 1.8 and 3.0 units. This was broadly similar to MCIDs of the PCS in other chronic conditions,<sup>11–14</sup> although smaller than the “Norman’s Law” value of 5 units.<sup>10</sup>

This was the first time that the MCID for the PCS has been defined in a patient population with stroke. This was a large

**Table 2.** Change in mean PCS over 6 months for each level of the Perceived Health Change question.

Perceived health change at 12 months*	SF-36 PCS at 12 months minus SF-12 PCS at 6 months mean (SD)
1 = Much better, n = 101	4.2 (7.1)
2 = Somewhat better, n = 94	2.6 (6.3)
3 = Same, n = 99	1.2 (6.9)
4 = Somewhat worse, n = 50	-2.4 (6.5)
5 = Much worse, n = 3	-8.9 (11.1)

\*Question 2 from the Short Form 36 v2 (New Zealand version), with modification (see text).

SF-36 PCS, Physical Component Summary score of the Short Form 36; SF-12 PCS, Physical Component Summary of the Short Form 12; SD, standard deviation.

**Table 3.** Analysis of variance for Physical Component Summary of the Short Form 36 versus three randomization arms in the Taking Charge after Stroke trial, 12 months after stroke.

Source	DF	Sum of squares	Mean square	F value	P
Model	2	973.4	486.7	5.75	0.0035
Error	378	31968.5	84.6 <sup>1</sup>		
Corrected total	380	32941.8			

<sup>1</sup>Root mean square error: 9.2.  
DF, degrees of freedom.

study, and both follow-up rates and PCS completion rates were excellent. The three methods used to estimate the MCID were appropriate and provided broadly similar estimates. It is not possible to determine which of these is the “best” estimate of the MCID for stroke. As others have described, the distribution method provided the smallest estimate for the MCID.<sup>13</sup> We suggest that in future intervention trials for stroke rehabilitation, powering the study to detect a difference of 2.5 units in the PCS, midway between our estimates from the two anchor-based methods, would be a reasonable compromise between potentially missing a small but important difference and maintaining manageable study numbers.

The PCS has much to offer as an outcome measure in stroke rehabilitation trials, and the addition of a stroke-specific MCID may encourage more researchers to consider its use. Scores are self-rated, aligning with the move toward greater use of

patient-reported outcomes in rehabilitation trials.<sup>24</sup> The modified Rankin Scale (mRS),<sup>25</sup> an uneven ordinal scale spanning multiple domains of function, remains the benchmark outcome measure for acute stroke trials despite its several weaknesses, with familiarity, speed, and zero cost rated as important by adherents.<sup>26</sup> The mRS is much less useful in stroke rehabilitation trials where few participants are likely to score in the lowest (0, 1) categories or to be included in the most severe (5) category. Thus, it may be difficult for clinically useful rehabilitation interventions to make any statistical impact on outcome across the three remaining categories with very large clinical thresholds for movement between categories.<sup>26</sup> Use of the PCS as a primary outcome measure allows for recruitment of participants across the full range of stroke severity and is responsive to small but clinically meaningful changes in outcome.<sup>15</sup> The SF-12 takes less than 10 minutes to

complete. The need for a license and the cost remain impediments to its use, but in the context of most stroke trial budgets, the outlay is modest.

## Conclusions

Our estimate for the MCID of the PCS in a population with stroke was 1.8 to 3.0 units. We recommend the use of 2.5 units for planning of clinical trials.

## Data sharing

Original data are available from the corresponding author on reasonable request with an appropriate hypothesis and analysis plan.

## Author information

The Taking Charge after Stroke (TaCAS) study group is as follows.

Wellington/Lower Hutt: Harry McNaughton, Vivian Fu, William Taylor, Tom Thomson, Mark Weatherall, Judith Riley, Kathryn Fernando, Anna Hunt, Gayle Williams, Claire Houghton, James Berry, Tony Mallon, Tanya Baker, Joanna Read, Allie Eathorne, Nick Shortt.

Palmerston North: Anna Ranta, Rebekah Higgs.

Christchurch: Carl Hanger, Deborah Allen, Haley Evans.

Counties-Manukau: Geoff Green, Amanda Retter.

Hawke's Bay: John Gommans, Eryn Kyle-Foulds.

Auckland: Anna McRae, Nicole Nancarrow, Lauren Lucas.

## Declaration of conflicting interest

The authors declare that there is no conflict of interest.

## Funding

The authors disclosed receipt of the following financial support for the research, authorship, and publication of this article: The study was funded by a grant from the Health Research Council of New Zealand (15/297). The study

fundors had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

## ORCID iD

Harry McNaughton  <https://orcid.org/0000-0003-3030-6062>

## References

1. Duncan PW, Jorgensen HS and Wade DT. Outcome measures in acute stroke trials: a systematic review and some recommendations to improve practice. *Stroke* 2000; 31: 1429–1438.
2. Salter KL, Teasell RW, Foley NC, et al. Outcome assessment in randomized controlled trials of stroke rehabilitation. *Am J Phys Med Rehabil* 2007; 86: 1007–1012.
3. Anderson C, Laubscher S and Burns R. Validation of the Short Form 36 (SF-36) health survey questionnaire among stroke patients. *Stroke* 1996; 27: 1812–1816.
4. McHorney CA, Ware JE Jr and Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Med Care* 1993; 31: 247–263.
5. Salaffi F, Carotti M, Ciapetti A, et al. A comparison of utility measurement using EQ-5D and SF-6D preference-based generic instruments in patients with rheumatoid arthritis. *Clin Exp Rheumatol* 2011; 29: 661–671.
6. Jaeschke R, Singer J and Guyatt GH. Ascertaining the minimal clinically important difference. *Cont Clin Trials* 1989; 10: 407–415.
7. McGlothlin AE and Lewis RJ. Minimal Clinically Important Difference: Defining What Really Matters to Patients. *JAMA* 2014; 312: 1342–1343.
8. Revicki D, Hays RD, Cella D, et al. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008; 61: 102–109.

9. Liang MH, Lew RA, Stucki G, et al. Measuring clinically important changes with patient-oriented questionnaires. *Med Care* 2002; 40: II45–II51.
10. Norman GR, Sloan JA and Wyrwich KW. Interpretation of Changes in Health-related Quality of Life: The Remarkable Universality of Half a Standard Deviation. *Med Care* 2003; 41: 582–592.
11. Strand V and Singh JA Improved health-related quality of life with effective disease-modifying antirheumatic drugs: evidence from randomized controlled trials. *Am J Manag Care* 2008; 14: 234–254.
12. Parker SL, Adogwa O, Mendenhall SK, et al. Determination of minimum clinically important difference (MCID) in pain, disability, and quality of life after revision fusion for symptomatic pseudoarthrosis. *Spine J* 2012; 12: 1122–1128.
13. Copay AG, Glassman SD, Subach BR, et al. Minimum clinically important difference in lumbar spine surgery patients: a choice of methods using the Oswestry Disability Index, Medical Outcomes Study questionnaire Short Form 36, and pain scales. *Spine J* 2008; 6: 968–974.
14. Carreon LY, Glassman SD, Campbell MJ, et al. Neck Disability Index, short form-36 physical component summary, and pain scales for neck and arm pain: the minimum clinically important difference and substantial clinical benefit after cervical spine fusion. *Spine J* 2010; 6: 469–474.
15. Fu V, Weatherall M, McPherson K, et al. Taking Charge after Stroke: A randomized controlled trial of a person-centered, self-directed rehabilitation intervention. *Int J Stroke* 2020; 15: 954–964.
16. Von Elm E, Altman DG, Egger M, et al. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies *BMJ* 2007; 335: 806.
17. Ware JE, Snow KK, Kosinski M, et al. *SF-36 Health Survey Manual and Interpretation Guide*. Boston, MA: The Health Institute, New England Medical Center, 1993.
18. Ware JE, Kosinski M., Keller SD. *SF-36 Physical and Mental Health Summary Scales: A User's Manual*. Boston, MA: The Health Institute, New England Medical Center, 1994.
19. Gandek B, Ware JE and Aaronson NK. Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: results from the IQOLA Project. *J Clin Epidemiol* 1998; 51: 1171–1178.
20. Samsa G, Edelman D, Rothman ML, et al. Determining clinically important differences in health status measures: a general approach with illustration to the Health Utilities Index Mark II. *Pharmacoeconomics* 1999; 15: 141–155.
21. Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale: Lawrence Erlbaum Assoc, 1988.
22. Eton DT, Cella D, Yost KJ, et al. A combination of distribution- and anchor-based approaches determined minimally important differences (MIDs) for four endpoints in a breast cancer scale. *J Clin Epidemiol* 2004; 57: 898–910.
23. Hawkes WG, Williams GR, Zimmerman S, et al. A clinically meaningful difference was generated for a performance measure of recovery from hip fracture. *J Clin Epidemiol* 2004; 57: 1019–1024.
24. Salinas J, Sprinkhuizen SM, Ackerson T, et al. An International Standard Set of Patient-Centered Outcome Measures After Stroke. *Stroke* 2016; 47: 180–186.
25. Van Swieten JC, Koudstaal PJ, Visser MC, et al. Interobserver agreement for the assessment of handicap in stroke patients. *Stroke* 1988; 19: 604–607.
26. Lees KR, Bath PM, Schellinger PD, et al. Contemporary outcome measures in acute stroke research: choice of primary outcome measure. *Stroke* 2012; 43: 1163–1170.