

# Substantial clinical benefit for Neck Disability Index revisited: establishing the goal for treatment?

Roland D. Donk<sup>1</sup> · Andre L. M. Verbeek<sup>2</sup> · Wim I. M. Verhagen<sup>3</sup> · Hans Groenewoud<sup>2</sup> · Allard T. F. Hosman<sup>4</sup> · Ronald H. M. A. Bartels<sup>5,6</sup>

Received: 22 April 2017/Revised: 11 July 2017/Accepted: 7 August 2017/Published online: 12 August 2017  
© The Author(s) 2017. This article is an open access publication

## Abstract

**Purpose** It is difficult for clinicians to inform patients about the success rate of a treatment as a cervical anterior discectomy procedure. Ideally, a proportion of good outcome as rated by patients is known. Patient-reported outcome measurements are helpful. The purpose is to relate the difference in Neck Disability Index (NDI) after a cervical anterior discectomy procedure for single level cervical degenerative disc disease with the patients' rating of their actual clinical situation after long-term follow-up to define the substantial clinical benefit (SCB).

**Methods** After completion of the NDI, patients who were surgically treated for cervical single level degenerative disc disease were asked to complete a five-item Likert scale to rate their clinical situation. After dichotomisation of the

outcome in good versus less than good, a cut-off value was defined by determining the value of the difference of NDI with the highest specificity and sensitivity. Funding was not obtained.

**Results** SCB for NDI after surgery for cervical single level degenerative disc disease should be set at ten with area under the curve (AUC) of 0.71 for sensitivity as well specificity.

**Conclusions** The goal for each treatment is a good outcome. While comparing treatments for cervical degenerative disc disease only those with an SCB of ten will be relevant for the patient, as patients who achieved this difference in NDI rated their actual situation at long-term follow-up as good.

**Keywords** Anterior cervical discectomy · NDI · PROM · Substantial clinical benefit · SCB · Outcome

✉ Ronald H. M. A. Bartels  
ronald.bartels@radboudumc.nl

<sup>1</sup> Department of Orthopaedic Surgery, Via Sana Clinics, Hoogveldseweg 1, 5451 AA Mill, The Netherlands

<sup>2</sup> Department of Health Evidence, Radboud University Medical Center, Geert Groote Plein-zuid 10, 6525 GA Nijmegen, The Netherlands

<sup>3</sup> Department of Neurology, Canisius Wilhelmina Hospital, Weg door Jonkerbos 100, 6532 SZ Nijmegen, The Netherlands

<sup>4</sup> Department of Orthopaedic Surgery, Radboud University Medical Center, Geert Groote Plein-zuid 10, 6525 GA Nijmegen, The Netherlands

<sup>5</sup> Department of Neurosurgery, Radboud University Medical Center, Geert Groote Plein-zuid 10, 6525 GA Nijmegen, The Netherlands

<sup>6</sup> Department of Neurosurgery, Canisius Wilhelmina Hospital, Weg door Jonkerbos 100, 6532 SZ Nijmegen, The Netherlands

## Introduction

Minimal clinically important difference (MCID) has been introduced to incorporate clinical relevance in the interpretation of the results of clinical trials. Originally, it was defined as “the smallest difference in score in the domain of interest which patients perceive as beneficial and would mandate, in the absence of troublesome side effects and excessive cost, a change in the patient's management” [1].

Although different methods to estimate MCID have been described [2], they are all constructed from a researcher's point of view. If patients are involved, they rate their outcome according to predefined terms and relative to their pre-treatment situation. Statements such as much better, somewhat better, about the same, somewhat worse and much worse are used for the comparison of patient's ratings and clinical outcome to estimate MCID.

MCID is associated with the clinical outcome related to the statement somewhat better than the pre-treatment situation.

A newer concept is the substantial clinical benefit (SCB). It was originally described for patients who were surgically treated for lumbar degenerative disease and defined as clinical improvement that represented a substantial clinical benefit after treatment [3]. The method was similar to some of those used for constructing MCID [2]. The patients rated their situation compared to the situation 1 year before [3].

The Neck Disability Index (NDI) is a frequently used outcome measurement to evaluate the treatment of neck-related disorders such as whiplash-associated disorder or degenerative disc disease. The MCID for NDI has been determined as 7.5 on a 0–50 scale by distinguishing “much better” from “somewhat better” patients [4]. SCB was set at 9.5 [4].

However, all parties involved want to have a good outcome after a treatment, also after long-term follow-up. Patients with severe neck and/or arm pain primarily opted for treatment not to achieve some improvement of their pre-treatment clinical situation, but to obtain a good clinical result at the end. In support of this hypothesis, a recent study on lumbar arthroplasty also focused on clinically relevant and long-lasting reduction of pain [5]. We were interested in the clinical outcome after a surgical treatment for degenerative cervical disc disease defined by the difference of the pre- and postoperative NDI ( $\Delta$ NDI) and the patients' ratings of their clinical situation after long-term follow-up. Ratings were not based on a comparison with their pre-treatment situation.

## Methods

The STROBE statement was adhered to [6]. The ethical board CMO Arnhem-Nijmegen approved the study. The study has been carried out in accordance with the World Medical Association Declaration of Helsinki [7].

Patients who participated in the Procon trial (Current Controlled Trials ISRCTN41681847) [8], a comparison of different anterior cervical surgery techniques for symptomatic single level degenerative disc herniation without spinal cord involvement, and who completed an NDI were eligible. Since treatments were not compared, the patients should be considered as part of a cohort.

The design of the trial from which the sample of the current was taken, was a prospective, double-blind, single-center randomized study with a three-arm parallel group design. Adult patients suffering from radicular signs and symptoms due to single level degenerative disc disease were included after informed consent. The experimental group was anterior cervical discectomy with arthroplasty,

whereas anterior cervical discectomy with fusion by cage standalone and anterior cervical discectomy were control groups. The primary outcome was NDI. The last follow-up after the index surgery was at the end of 2015. At that moment, the patients were asked to complete the NDI.

Within 2 months after completion of the NDI, a questionnaire was sent to the patients about the qualification of their clinical situation regarding the neck pain and arm pain at that moment.

For this questionnaire, a five-item Likert scale was used. We did not predefine the criteria, since we were interested in the qualitative judgment of the patients without any bias introduced by the researcher. The possible qualifications of their current clinical situation were: excellent, very good, good, moderate, and bad.

SAS version 9.2 (SAS Institute Inc. Cary NC, USA) was used for the statistical analysis. Continuous variables were depicted as mean value  $\pm$  standard deviation (minimum–maximum). For data analysis the Student's *t* test was used. At dichotomisation of outcome, two groups of patients could be defined: those with a good to excellent outcome, and those with a less than good outcome [9]. To estimate the value of  $\Delta$ NDI that corresponded best with the dichotomised outcome, the cut-off values with the highest sensitivity and specificity were chosen separately. A *P* value  $<0.05$  was assumed to be statistically significant.

## Results

Of the 140 patients in the trial, 80 patients (57.1%) were eligible. Women outnumbered men (61 versus 53); the mean age was  $45.3 \pm 6.8$  (29–49). Follow-up after the index surgery was  $9.1 \pm 1.9$  years (5.6–12.2 years). At that time patients completed the NDI. The mean preoperative NDI was  $18.6 \pm 7.1$  (1–36), whereas the mean postoperative NDI at the last follow-up was  $7.17 \pm 8.3$  (0–29).

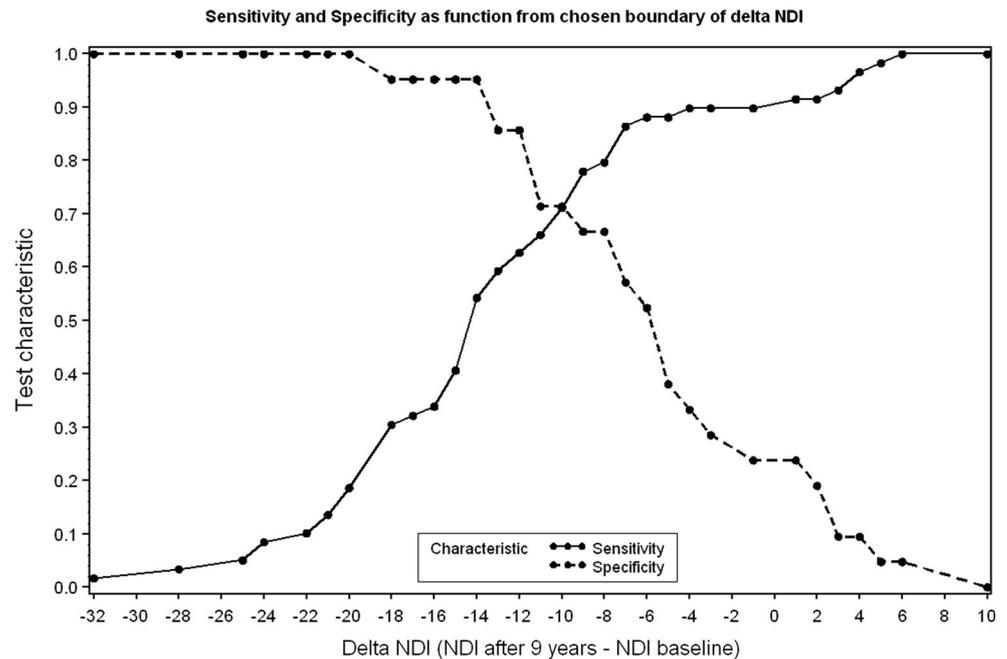
Eight patients rated their outcome as excellent, 23 as very good, 28 as good, 17 as moderate and 4 as bad.

To detect the optimal cut-off value for the  $\Delta$ NDI, a graph was constructed showing sensitivity and specificity as a function of all possible cut-off values (Fig. 1). The highest sensitivity and highest specificity (both 71.0%) for a good outcome were observed if  $\Delta$ NDI = 10.

## Discussion

Minimal clinically important difference (MCID) has been introduced in medical literature to define a threshold of improvement that is clinically important for relativizing the statistically significant differences. However, patients are

**Fig. 1** Sensitivity and specificity as function from chosen boundary of delta NDI



not expecting a minimal improvement, but are awaiting an optimal result.

Therefore, MCID is an important demarcation, but it could be considered a floor value rather than a goal in defining clinical success.

Although estimates for MCID and SCB for NDI have been made, the major drawback is the comparison of the patient's clinical situation with an earlier situation. Of course patients are generally interested in the improvement of their pre-treatment clinical state, but they are even more interested in the actual clinical situation at longer-term follow-up. A questionnaire sent to participants to prepare for a conference on MCID in 2001 revealed that 84% agreed that MCID should be validated for long-term outcome [10].

For two reasons, this study is unique. First it correlates the patients' ratings of their actual clinical situation and not in comparison to their pre-treatment situation. A good outcome is the preoperative goal for patients as well as treating physicians. Secondly, the outcome was rated a long time after the surgery. As far as we know, this has never been done before.

Therefore, SCB for NDI after long-term follow-up for cervical anterior discectomy is set at 10. This resembles the value of 9.5 reported by Carreon who compared the judgement of the actual situation compared to the preoperative situation [4]. Follow-up was shorter.

As Glassman has already pointed out, MCID is a floor value, and nobody goes for the minimal result. Nowadays, the implementation of treatments that only provide a minimal clinically important difference cannot be justified. A good outcome should be the goal.

Therefore, the value of MCID and SCB should be reconsidered. Although MCID was essential for the awareness that statistical results should also be translated into clinical relevance, nowadays only good outcomes are acceptable even though a good outcome for 100% of patients is impossible. We suggest focusing only on SCB to compare the two treatments.

However, the optimal relative difference in SCB between groups should still be defined. To evaluate a new treatment, the clinical result is compared to a known treatment, which serves as the gold standard. If the outcome of the new treatment compared to the pre-treatment situation is equal to the MCID and is similar to that of the gold standard, then an equal effect might be assumed. The same holds true for SCB. If a treatment is better than the old one, it cannot be expected that the clinical difference in treatment is equal to at least the absolute value of MCID or SCB. In our opinion, defining a minimum difference in the percentage of patients reaching SCB is a reasonable for making a comparison of the superiority or similarity of a treatment's efficacy. This value is best established by a worldwide board of physicians, representatives of insurance companies, governments and, most importantly, patients as consumers of healthcare.

The terms used to qualify the actual situation were not predefined. This can be interpreted as a flaw. However, we were reluctant to specify criteria due to the risk of a researcher-guided system and the subjective character of the rating being restricted so that it did not represent the patient's perspective. We feel confident that the chosen strategy contributed to a qualification that best resembled

the patient's subjective interpretation of the situation at that moment.

The different techniques for cervical anterior discectomy might be viewed as a drawback. However, the patients were randomized after the NDI was established. Therefore, the patient's rating was not influenced by knowing whether an implant was used and, if so, which type.

Retrieving information from less than 100% of the patients included in the trial from which the sample for this study was taken could be interpreted as a flaw. However, since the goal of the study was to establish the threshold for SCB for NDI, the outcome of interest should be related to the qualification of patients. For this purpose only a cohort of patients is needed, and therefore, we do not think that the chosen policy will alter the conclusion.

We did not take into account mental distress. Recently, it has been shown that patients with a high preoperative level of anxiety or depression showed at baseline but also at 2 years follow-up a worse outcome in NDI [11]. Therefore, SCB was not examined in cases without mental distress. Although SCB in this group will probably be lower, we think that the calculated SCB for NDI from our study will be more representative in general practice. Not every individual is currently screened for anxiety level. Furthermore, for the individual, it will not be possible to exactly determine SCB based on a score from an individual test for anxiety level.

Finally, determining the cut-off values of  $\Delta$ NDI in relation to a dichotomised outcome as good or less good is questionable. We chose a conservative approach by requesting the highest sensitivity in combination with the highest specificity. Increasing the values of  $\Delta$ NDI would increase sensitivity and decrease specificity, and decreasing the values of  $\Delta$ NDI would induce a reversal effect. This would result, in our opinion, in a less distinct cut-off value for  $\Delta$ NDI in relation to a good and a less good outcome.

In conclusion, a difference of ten between preoperative and postoperative NDI after a cervical anterior discectomy procedure for single level cervical degenerative disease corresponds to a clinical situation that patients rate as good at long-term follow-up. Reporting the proportion of patients with a good outcome will help to choose a treatment.

**Acknowledgements** We express our gratitude to Dave Thomas PhD from Radboud in'to Languages for editing English grammar and style.

**Compliance with ethical standards**

**Conflict of interest** The authors declares that they have no conflict of interest.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

1. Jaeschke R, Singer J, Guyatt GH (1989) Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials* 10:407–415
2. Wells G, Beaton D, Shea B, Boers M, Simon L, Strand V, Brooks P, Tugwell P (2001) Minimal clinically important differences: review of methods. *J Rheumatol* 28:406–412
3. Glassman SD, Copay AG, Berven SH, Polly DW, Subach BR, Carreon LY (2008) Defining substantial clinical benefit following lumbar spine arthrodesis. *J Bone Jt Surg Am* 90:1839–1847. doi:10.2106/JBJS.G.01095
4. Carreon LY, Glassman SD, Campbell MJ, Anderson PA (2010) 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* 10:469–474. doi:10.1016/j.spinee.2010.02.007
5. Plais N, Thevenot X, Cogniet A, Rigal J, Le Huec JC (2017) Maverick total disc arthroplasty performs well at 10 years follow-up: a prospective study with HRQL and balance analysis. *Eur Spine J*. doi:10.1007/s00586-017-5065-z
6. von Elm E, Altman DG, Egger M, Pocock SJ, Gotsche PC, Vandenbroucke JP, Initiative S (2007) The strengthening of reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med* 147:573–577
7. Fuson RL, Sherman M, Van Vleet J, Wendt T (1997) The conduct of orthopaedic clinical trials. *J Bone Jt Surg Am* 79:1089–1098
8. Bartels RH, Donk R, van der Wilt GJ, Grotenhuis JA, Venderink D (2006) Design of the PROCON trial: a prospective, randomized multi-center study comparing cervical anterior discectomy without fusion, with fusion or with arthroplasty. *BMC Musculoskelet Disord* 7:85. doi:10.1186/1471-2474-7-85
9. Donk R, Verbeek A, Verhagen W, Groenewoud H, Hosman A, Bartels R (2016) The qualification of outcome after cervical spine surgery by patients compared to the Neck Disability Index. *PLoS One* 11:e0161593. doi:10.1371/journal.pone.0161593
10. Wells G, Anderson J, Beaton D, Bellamy N, Boers M, Bombardier C, Breedveld F, Carr A, Cranney A, Dougados M, Felson D, Kirwan J, Schiff M, Shea B, Simon L, Smolen J, Strand V, Tugwell P, van Riel P, Welch VA (2001) Minimal clinically important difference module: summary, recommendations, and research agenda. *J Rheumatol* 28:452–454
11. Skeppholm M, Fransson R, Hammar M, Olerud C (2017) The association between preoperative mental distress and patient-reported outcome measures in patients treated surgically for cervical radiculopathy. *Spine J* 17:790–798. doi:10.1016/j.spinee.2016.02.037