

RESEARCH ARTICLE

# The Qualification of Outcome after Cervical Spine Surgery by Patients Compared to the Neck Disability Index

Roland Donk<sup>1</sup>, Andre Verbeek<sup>2</sup>, Wim Verhagen<sup>3</sup>, Hans Groenewoud<sup>2</sup>, Allard Hosman<sup>4</sup>, Ronald Bartels<sup>5,6\*</sup>

**1** Department of Orthopedic Surgery, Via Sana Clinics, Hoogveldseweg 1, 5451 AA, Mill, the Netherlands, **2** Department for Health Evidence, Radboud university medical center, Geert Groote Plein-zuid 10, 6525 GA, Nijmegen, the Netherlands, **3** Department of Neurology, Canisius Wilhelmina Hospital, Weg door Jonkerbos 100, 6532 SZ, Nijmegen, the Netherlands, **4** Department for Orthopedic Surgery, Radboud university medical center, Geert Groote Plein-zuid 10, 6525 GA, Nijmegen, the Netherlands, **5** Department of Neurosurgery, Radboud university medical center, Geert Groote Plein-zuid 10, 6525 GA, Nijmegen, the Netherlands, **6** Canisius Wilhelmina Hospital, Department of Neurosurgery, Weg door Jonkerbos 100, 6532 SZ, Nijmegen, the Netherlands

\* [ronald.bartels@radboudumc.nl](mailto:ronald.bartels@radboudumc.nl)



## Abstract

### Objective

The Neck Disability Index (NDI) is a patient self-assessed outcome measurement tool to assess disability, and that is frequently used to evaluate the effects of the treatment of neck-related problems. In individualized medicine it is mandatory that patients can interpret data in order to choose a treatment. A change of NDI or an absolute NDI is generally meaningless to a patient. Therefore, a correlation between the qualification of the clinical situation rated by the patient and the NDI score was evaluated.

### Methods

Patients who completed an NDI after anterior surgery because of symptomatic single level degenerative cervical disc disease were asked one month after completion of the NDI to qualify their clinical situation of a 5-item Likert scale varying from excellent to bad. Since a clear distinction between the categories was not possible based on the total NDI score, a ROC-curve was built, and the AUC computed in order to estimate best dichotomization in qualification of the clinical situation. The best corresponding cut-off point for the NDI total score was found by studying sensitivity and specificity for all possible cut-off points.

### Results

102 patients were included. The highest AUC was obtained by dichotomizing the qualification into a group with good outcome and less-good outcome. The highest sensitivity and specificity for the dichotomized qualification as good outcome corresponded to a NDI  $\leq 7$ . Sensitivity was 81.08% and specificity was 78.57%.

## OPEN ACCESS

**Citation:** 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(8): e0161593. doi:10.1371/journal.pone.0161593

**Editor:** Giovanni Grasso, Università degli Studi di Palermo, ITALY

**Received:** May 23, 2016

**Accepted:** August 8, 2016

**Published:** August 23, 2016

**Copyright:** © 2016 Donk et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** The authors received no specific funding for this work.

**Competing Interests:** The authors have declared that no competing interests exist.

## Conclusion

This is the first study that correlated the qualification of the situation by the patients themselves and NDI. An  $NDI \leq 7$  corresponded to a good outcome according to the patients. This is valuable information to inform patients in their decision for any treatment.

## Introduction

The Neck Disability Index (NDI) is a frequently used, well known, and in multiple languages validated outcome measurement tool to assess self-rated disability in patients with neck pain. It can be categorized as a patient reported outcome measurement tool (PROM). The NDI is frequently used in clinical practice, but also for research purposes [1–3]. The main purpose is the quantification of the difference in pre- and post-treatment condition according to the patients suffering from disabling neck pathology. The NDI addresses pain and functional items related to neck problems. It has been validated in both neck pain and, especially, whiplash patients [2, 4].

Informing the patient is crucial before installing any treatment. In modern times information can be gained through many resources, but the treating physician is still very important. It has been shown that fulfillment of preoperative expectations is related to the highest post-operative satisfaction. A mismatch of disease understanding and expectation between treating physician and patient might result in a less than favorable outcome according to the patient [5–7].

The information provided by PROMs as the NDI obtained from studies can contribute in sketching expectations while informing the patient before any treatment. The most useful tools in the process of gaining information or providing it are clear clinical outcomes: mortality, infection rate etc. However, PROMs including the NDI are not reporting on a clearly defined outcome but on a combination of surrogate outcomes.

An adequate interpretation of a PROM is difficult, especially since it has been demonstrated that the language used in the questionnaires is very difficult to understand for patients. As El-Daly, I. et al. stated in their conclusion: “the majority of PROMs analyzed are written at a level that is incomprehensible to the average UK adult” [8]. The usefulness of the results of PROMs with low readability is debatable.

However, the NDI was also incorporated in the earlier mentioned study. For a correct understanding of the NDI an education level of 13–15 year-old subject was needed indicating a readability level of standard English [8]. Since the translation of the NDI into Dutch has been validated [9] we feel confident that most of the patients did understand the questions and completed their questionnaires without difficulty.

Although it seemed related, the readability of a PROM is different than interpreting the result. For example and specifically for the NDI, what information is provided to the patient if he reads or hears that a mean total score of the NDI of 8 is achieved in a group of 100 patients after a certain treatment. Information should be presented in a way that is acceptable and useful for a patient [10]. In a survey among patients with scoliosis and their carers, it was advised that the information should be user friendly and in plain language [11]. For NDI grades of disability have been defined, although these also differ and are based on clinical information and not the patients qualification [2]. A grade of disability like “none to mild disability” is not very illustrative to a patient.

Therefore, we would like to correlate the total score of the NDI with a qualitative rating by the patient themselves in a way that everyone can understand. This will contribute to understanding and decision making for patients in the future.

## Methods

The STROBE statement was followed ([S1 File](#))[12]. 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 [13].

Patients who participated in the Procon trial Current (Controlled Trials ISRCTN41681847) [14], a comparison of different anterior cervical surgery techniques for symptomatic single level degenerative disc herniation without spinal cord involvement, and who completed a NDI were included. 142 patients participated of whom 140 completed and returned the NDI. One patient died unrelated to the trial, the other refused to return the NDI questionnaire. So, 140 patients were eligible. The mean time after surgery was  $9.1 \pm 1.9$  years (5.6–12.2 years).

Within two months after completion of the NDI, a questionnaire was sent to the patients about the qualification of their situation regarding the neck and its related problems at that moment. Although little is known about the bias introduced by sending reminders [15], we did not send reminders or contacted the non-responders.

A five-item Likert scale was used. We did not predefine the criteria, since we were interested in the qualitative judgment of the patients themselves without any bias introduced by the researcher. The possible qualifications of their situation were: excellent, very good, good, moderate, and bad (see [S2](#) and [S3](#) Files).

For statistical analyses SAS version 9.2 (SAS Institute Inc. Cary NC, USA) was used. Continuous variables are depicted as value  $\pm$  standard deviation (minimum-maximum). For data analysis the Student-t test was used. Dichotomization of the patient qualifications seemed to be appropriate. To estimate which qualifications could be best combined for each possible dichotomized set, a ROC curve was build and the area under the curve (AUC) was calculated. The combination with the highest AUC was chosen. To estimate the value of NDI that corresponded best with the dichotomized outcome, the cut-off value of the total NDI with the highest sensitivity and specificity was chosen. A P value  $< 0.05$  was assumed to be statistically significant.

## Results

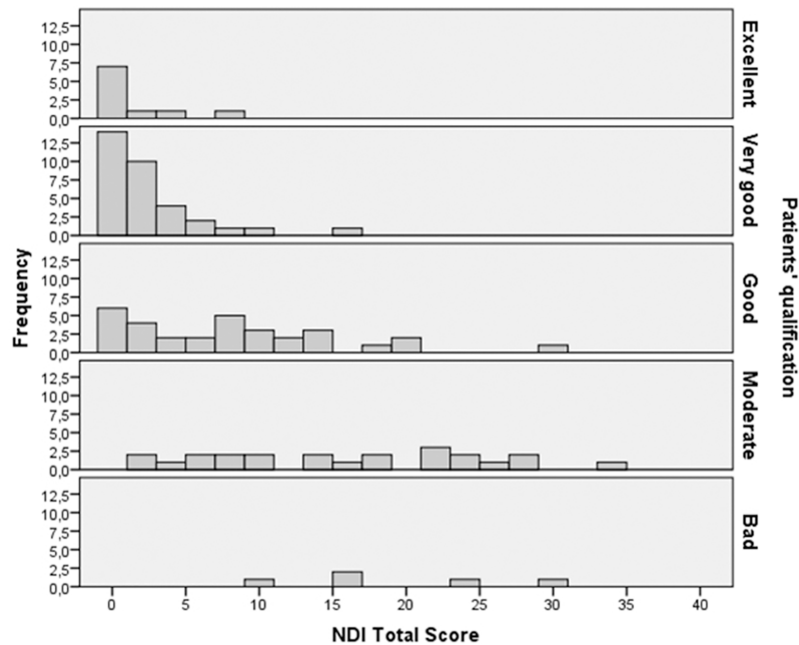
Of the 140 eligible patients, 102 consecutive patients completed the questionnaires (response rate: 72.9%). Mean NDI was  $7.5 \pm 8.6$  (0–34) for the responders and  $6.7 \pm 8.3$  for non-responders. The difference in NDI did not reach statistical significance ( $P = 0.6$ ). Ten patients rated their situation excellent, 33 very good, 32 good, 23 moderate, and 5 qualified their situation as bad. 73.5% of the patients rated their situation as good or better. In [Fig 1](#) NDI is represented in relation to the Likert qualification. It was not possible to distinct the qualifications clearly based on a total NDI score. Therefore, we decided to dichotomize qualification by the patient.

The biggest AUC was obtained by dichotomizing the qualifications in the group excellent, very good and good versus the combination of moderate and bad (AUC = 0.874). The first group consisted of those patients with a good outcome; the patients belonging to the latter will be regarded as having a less-good outcome.

Then a ROC was constructed ([Fig 2](#)). The highest sensitivity and highest specificity for a good outcome is obtained when NDI is seven or less: sensitivity was 81.08% and specificity was 78.57%. The distribution of patients after dichotomization in relation to the NDI is shown in [Table 1](#).

## Discussion

Currently, information about any treatment is very easy accessible to patients. However, interpretation of the data is very difficult or even impossible for most patients due to lack of

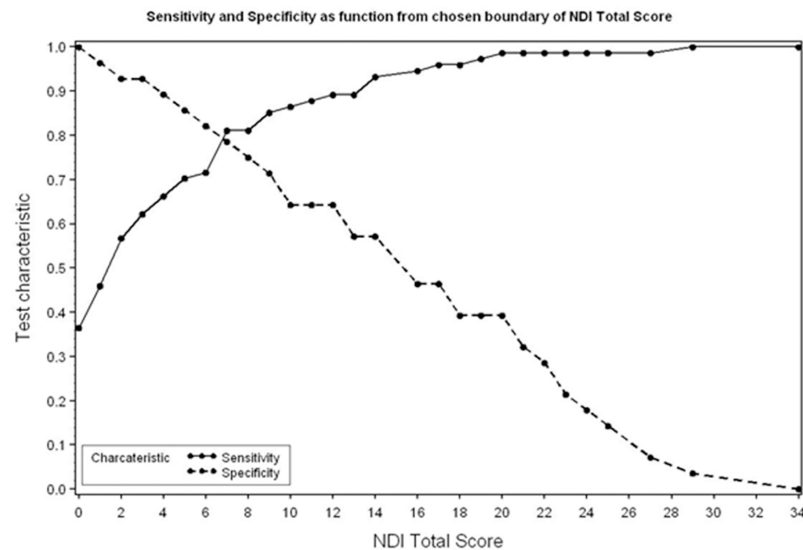


**Fig 1. distribution of total ND score in relation to patients' qualification.**

doi:10.1371/journal.pone.0161593.g001

adequate knowledge. Surrogate outcomes are provided that are valuable for scientific purposes, but are not easily transposed to lay terms.

The NDI is a questionnaire assessed by the patient self. The NID consists of ten questions, and for each question six answers are possible. The answers are ordered starting from no disability to maximal disability. The answers are graded from zero to six, and therefore the total NDI score can vary between zero and fifty. The best outcome will be a total score of zero.



**Fig 2. Figure depicting the cut-off value of the total NDI with the highest sensitivity and specificity.**

doi:10.1371/journal.pone.0161593.g002

**Table 1. Distribution of patients based on outcome defined as good or less-good in relation to NDI.**

	Good	Less-Good	Total
NDI $\leq$ 7	60	14	74
NDI $>$ 7	6	22	28
Total	66	46	102

doi:10.1371/journal.pone.0161593.t001

The NDI has not been uniformly divided in grades of disability [16–18]. A major concern is furthermore that the investigators predefined the qualifications of each grade. They correlated it to existing questionnaires or findings at physical examinations.

From an investigators point of view a total score of the NDI of zero would correspond with an excellent outcome. We have shown that only a proportion of the patients that rated their situation as excellent had a total NDI score of zero, whereas some patients that rated their situation as good or very good, also had a total NDI score of zero. Other (probably psychological) factors that are not taken into account in the NDI, might explain this.

Transforming a total NDI score into an expression that can easily be understood by patients will help them in making a decision about their eventual treatment, and is a contribution to individualized medicine. This is achieved not only by calculating a cut off value for the total NDI score (NDI  $\leq$  7 versus NDI  $>$  7), but also by dichotomizing the patients' qualification in good and less good.

Not actively motivating patients to respond might be considered a flaw of the study. However, comparison of the NDI between the group of responders and non-responders convinced us that the sample is representative. Especially when the response rate of more than 70%, that can be considered as good [19], is taking into account.

Another limitation of the study could be the lack of a pre-inquiry definition of the qualifications as rated by the patients. Therefore, the distribution of the NDI for any qualification is much wider than when the qualifications were defined prior to asking the patients. However, this would have been again the interpretation of the researcher, whereas at this moment we are convinced that the qualifications really represented the perspective of the patient.

Finally, determining the cut off value of the NDI to consider a good or less good outcome can be subject of debate. We have chosen for a conservative approach by requesting the highest sensitivity in combination with the highest specificity. Increasing the NDI score would increase sensitivity and decrease specificity, and decreasing the NDI would induce a reverse effect creating, in our opinion, a less reliable definition of good and less outcome.

Although we did not investigated whether the patients have a better understanding of the expression of a good outcome compared to mild disability, we are convinced that the first is more appealing. From a patients perspective a total NDI score or a difference in NDI score, that is however important for scientific evaluation, is meaningless. It will not help him/her in decision-making about any treatment for neck-related problems.

In conclusion, to help the patient in the decision-making for any treatment of neck-related pathology it seems obvious that expressions should be used that are understandable. Therefore, we propose that a NDI of seven or less is qualified as a good outcome.

## Supporting Information

**S1 File. STROBE 2007 (v4) Statement—Checklist of items that should be included in reports of cohort studies.**

(DOCX)

**S2 File. NDI original data.**  
(XLSX)

**S3 File. Questionnaire translated into English.**  
(DOCX)

## Author Contributions

**Conceptualization:** RD RB WV AH.

**Formal analysis:** HG.

**Investigation:** RD RB.

**Methodology:** RD RB AV HG.

**Supervision:** RB.

**Visualization:** HG.

**Writing – original draft:** RD RB WV AH.

**Writing – review & editing:** RD AV WV HG AH RB.

## References

1. Godil SS, Parker SL, Zuckerman SL, Mendenhall SK, McGirt MJ. Accurately measuring the quality and effectiveness of cervical spine surgery in registry efforts: determining the most valid and responsive instruments. *Spine J.* 2015; 15(6):1203–9. doi: [10.1016/j.spinee.2013.07.444](https://doi.org/10.1016/j.spinee.2013.07.444) PMID: [24076442](https://pubmed.ncbi.nlm.nih.gov/24076442/).
2. MacDermid JC, Walton DM, Avery S, Blanchard A, Etruw E, McAlpine C, et al. Measurement properties of the neck disability index: a systematic review. *J Orthop Sports Phys Ther.* 2009; 39(5):400–17. doi: [10.2519/jospt.2009.2930](https://doi.org/10.2519/jospt.2009.2930) PMID: [19521015](https://pubmed.ncbi.nlm.nih.gov/19521015/).
3. Vernon H. The Neck Disability Index: state-of-the-art, 1991–2008. *J Manipulative Physiol Ther.* 2008; 31(7):491–502. doi: [10.1016/j.jmpt.2008.08.006](https://doi.org/10.1016/j.jmpt.2008.08.006) PMID: [18803999](https://pubmed.ncbi.nlm.nih.gov/18803999/).
4. Howell ER. The association between neck pain, the Neck Disability Index and cervical ranges of motion: a narrative review. *J Can Chiropr Assoc.* 2011; 55(3):211–21. PMID: [21886283](https://pubmed.ncbi.nlm.nih.gov/21886283/); PubMed Central PMCID: [PMCPMC3154067](https://pubmed.ncbi.nlm.nih.gov/PMC3154067/).
5. Mannion AF, Junge A, Elfering A, Dvorak J, Porchet F, Grob D. Great expectations: really the novel predictor of outcome after spinal surgery? *Spine (Phila Pa 1976).* 2009; 34(15):1590–9. PMID: [19521272](https://pubmed.ncbi.nlm.nih.gov/19521272/).
6. McGregor AH, Hughes SP. The evaluation of the surgical management of nerve root compression in patients with low back pain: Part 2: patient expectations and satisfaction. *Spine (Phila Pa 1976).* 2002; 27(13):1471–6; discussion 6–7. PMID: [12131749](https://pubmed.ncbi.nlm.nih.gov/12131749/).
7. Soroceanu A, Ching A, Abdu W, McGuire K. Relationship between preoperative expectations, satisfaction, and functional outcomes in patients undergoing lumbar and cervical spine surgery: a multicenter study. *Spine (Phila Pa 1976).* 2012; 37(2):E103–8. PMID: [21629159](https://pubmed.ncbi.nlm.nih.gov/21629159/).
8. El-Daly I, Ibraheim H, Rajakulendran K, Culpan P, Bates P. Are patient-reported outcome measures in orthopaedics easily read by patients? *Clin Orthop Relat Res.* 2016; 474(1):246–55. doi: [10.1007/s11999-015-4595-0](https://doi.org/10.1007/s11999-015-4595-0) PMID: [26472587](https://pubmed.ncbi.nlm.nih.gov/26472587/); PubMed Central PMCID: [PMCPMC4686523](https://pubmed.ncbi.nlm.nih.gov/PMC4686523/).
9. Jorritsma W, de Vries GE, Dijkstra PU, Geertzen JH, Reneman MF. Neck Pain and Disability Scale and Neck Disability Index: validity of Dutch language versions. *Eur Spine J.* 2012; 21(1):93–100. doi: [10.1007/s00586-011-1920-5](https://doi.org/10.1007/s00586-011-1920-5) PMID: [21814745](https://pubmed.ncbi.nlm.nih.gov/21814745/); PubMed Central PMCID: [PMCPMC3252449](https://pubmed.ncbi.nlm.nih.gov/PMC3252449/).
10. Pellise F, Sell P, EuroSpine Patient Line Task F. Patient information and education with modern media: the Spine Society of Europe Patient Line. *Eur Spine J.* 2009; 18 Suppl 3:395–401. doi: [10.1007/s00586-009-0973-1](https://doi.org/10.1007/s00586-009-0973-1) PMID: [19381695](https://pubmed.ncbi.nlm.nih.gov/19381695/); PubMed Central PMCID: [PMCPMC2899323](https://pubmed.ncbi.nlm.nih.gov/PMC2899323/).
11. Wellburn S, Bettany-Saltikov J, van Schaik P. An evaluation of web sites recommended by UK NHS consultants to patients with adolescent idiopathic scoliosis at the first point of diagnosis. *Spine (Phila Pa 1976).* 2013; 38(18):1590–4. PMID: [23649217](https://pubmed.ncbi.nlm.nih.gov/23649217/).
12. von Elm E, Altman DG, Egger M, Pocock SJ, Gotsche PC, Vandenbroucke JP, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting

- observational studies. *BMJ*. 2007; 335(7624):806–8. doi: [10.1136/bmj.39335.541782.AD](https://doi.org/10.1136/bmj.39335.541782.AD) PMID: [17947786](https://pubmed.ncbi.nlm.nih.gov/17947786/); PubMed Central PMCID: PMCPMC2034723.
13. Fuson RL, Sherman M, Van Vleet J, Wendt T. The conduct of orthopaedic clinical trials. *J Bone Joint Surg Am*. 1997; 79(7):1089–98. PMID: [9234889](https://pubmed.ncbi.nlm.nih.gov/9234889/).
  14. Bartels RH, Donk R, van der Wilt GJ, Grotenhuis JA, Venderink D. 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*. 2006; 7:85. doi: [10.1186/1471-2474-7-85](https://doi.org/10.1186/1471-2474-7-85) PMID: [17096851](https://pubmed.ncbi.nlm.nih.gov/17096851/); PubMed Central PMCID: PMCPMC1637105.
  15. Tam CC, Higgins CD, Rodrigues LC. Effect of reminders on mitigating participation bias in a case-control study. *BMC Med Res Methodol*. 2011; 11:33. doi: [10.1186/1471-2288-11-33](https://doi.org/10.1186/1471-2288-11-33) PMID: [21453477](https://pubmed.ncbi.nlm.nih.gov/21453477/); PubMed Central PMCID: PMCPMC3079699.
  16. Miettinen T, Leino E, Airaksinen O, Lindgren KA. The possibility to use simple validated questionnaires to predict long-term health problems after whiplash injury. *Spine (Phila Pa 1976)*. 2004; 29(3):E47–51. PMID: [14752363](https://pubmed.ncbi.nlm.nih.gov/14752363/).
  17. Vernon HT. Assessment of Self-Rated Disability, Impairment, and Sincerity of Effort in Whiplash-Associated Disorder. *J Musculoskel Pain*. 2000; 8:155–67.
  18. Sterling M, Jull G, Vicenzino B, Kenardy J. Characterization of acute whiplash-associated disorders. *Spine (Phila Pa 1976)*. 2004; 29(2):182–8. PMID: [14722412](https://pubmed.ncbi.nlm.nih.gov/14722412/).
  19. Fincham JE. Response rates and responsiveness for surveys, standards, and the Journal. *Am J Pharm Educ*. 2008; 72(2):43. PMID: [18483608](https://pubmed.ncbi.nlm.nih.gov/18483608/); PubMed Central PMCID: PMCPMC2384218.