

ARTICLE



Neurogenic bowel dysfunction score in spinal cord-injured patients: translation and validation of the Dutch-language NBD score

Tess van Doorn ¹✉, Ilse M. Groenendijk¹, Jeroen R. Scheepe¹ and Bertil F. M. Blok ¹

© The Author(s), under exclusive licence to International Spinal Cord Society 2021

STUDY DESIGN: This is a prospective validation study.

OBJECTIVES: The neurogenic bowel dysfunction (NBD) score is a widely used symptom-based questionnaire evaluating bowel dysfunction and its impact on quality of life (QoL) in spinal cord-injured patients. This study aimed to translate and validate a Dutch-language NBD score in patients with SCI.

SETTING: Patients with SCI visiting the urology department or general practitioner (GP) in Rotterdam, the Netherlands.

METHODS: Standardized guidelines were followed for the translation and validation process of the NBD score. Adult patients with SCI visiting our urology department were asked to participate by filling in a set of questionnaires: the NBD score, the Fecal Incontinence Quality of Life scale (FIQL), the Fecal Incontinence Severity Index (FISI), and the European Quality of life 5-Dimension 3-Level questionnaire (EQ-5D-3L) at baseline and 1–2 weeks afterward. A control group recruited at a GP office completed the questionnaires once. The following measurement properties were evaluated: content validity, internal consistency, reproducibility, criterion-, and construct validity.

RESULTS: Fifty-eight patients and 50 references were included. Content validity was adequate, internal consistency was moderate (Cronbach's alpha 0.56 and 0.30) and reproducibility was adequate (ICC 0.87). Criterion validity was confirmed; NBD score correlated significantly with the FIQL, FISI, and EQ-5D-3L. NBD scores in the patient group were significantly higher than in references, demonstrating good construct validity.

CONCLUSIONS: The Dutch-language version of the NBD score showed moderate to good measurement properties, and therefore is a reliable tool to measure bowel dysfunction in patients with SCI. We recommend standardized usage of this questionnaire for clinical evaluation and research purposes.

Spinal Cord (2022) 60:223–227; <https://doi.org/10.1038/s41393-021-00668-8>

INTRODUCTION

Patients suffering from spinal cord injury (SCI) experience numerous symptoms affecting their quality of life (QoL). One of these symptoms is neurogenic bowel dysfunction (NBD). The most frequent manifestations of NBD are fecal incontinence (FI) and constipation. Imaginable, this bowel dysfunction can impact a patient's life in several areas. For example, it can lead to a reduced participation in social activities [1].

The pathophysiology of NBD is quite well studied in patients with SCI. It is known that the internal anal sphincter has an exaggerated smooth muscle response to rectal distention, which induces large rectal contractions. These contractions are associated with deep anal reflex and will most likely result in defecation without any obvious increase in intra-abdominal pressure. Because of the unpredictable nature of the spinal reflex, this mechanism makes FI a difficult problem for patients with SCI [2]. Furthermore, constipation in patients with SCI is most likely due to prolonged colonic transit time, and side effects of medication (e.g., anticholinergics) and immobilization [3]. Other

contributing factors are, e.g., loss of sensory function at the level of the rectum and perineum, incapability of active contraction of the pelvic floor muscles, variable loss of abdominal muscle contraction, and hence creation of intra-abdominal pressure. Most of patients experience both constipation and FI [4].

The NBD score is a questionnaire-based symptom score/patient-reported outcome (PRO) measure that was developed for clinical assessment of colorectal and anal dysfunction in patients with SCI [4].

Since bowel dysfunction can be a sensitive subject to discuss for patients, physicians should be aware of such dysfunction, and start the conversation about this subject and possible symptoms.

At this moment, there is no validated Dutch instrument available to measure NBD that can also be used as a tool to assess the impact on QoL in patients with SCI. The validation of the NBD score in the Dutch language will help to evaluate NBD objectively in the spinal cord-injured patients in the Netherlands. This PRO measure will make follow-up after therapy and clinical research possible and can function as an instrument to open up

¹Department of urology, Erasmus Medical Center, Rotterdam, The Netherlands. ✉email: t.vandoorn.1@erasmusmc.nl

Received: 4 February 2021 Revised: 25 June 2021 Accepted: 25 June 2021
Published online: 4 August 2021

the conversation about NBD between patients and physicians. Implementation of the NBD score in guidelines for the care of patients with SCI can help standardize these conversations. The aim of this study is to validate the Dutch-language NBD score in spinal cord-injured patients, so bowel dysfunction can be measured and followed over time or after treatment. Furthermore, such validated NBD score could be an important addition to value-based healthcare for neurogenic patients, in which all aspects of pelvic health are measured and evaluated with PRO's.

METHODS

Study design and study population

We conducted this prospective cohort validation study at the urology department of the Erasmus University Medical Center Rotterdam and a general practitioner (GP) office. The institutional Medical Ethics Review Board reviewed the research proposal and gave ethical approval (MEC 2018-1050).

Between August 2018 and February 2020, adult patients with SCI visiting our urology department were invited to participate in this study. Patients were not eligible for this validation study if they had difficulty reading and/or understanding the Dutch language, suffered from cognitive impairment or inflammatory bowel disease, had a recent gastroenterological malignancy or a bowel stoma. After signing the informed consent form, participants were asked to fill in a set of questionnaires at baseline. A second set of questionnaires was handed out on paper and returned by post after 1–2 weeks. Questionnaires were excluded if they were returned ≥ 35 days after the first set was completed. Patients' clinical characteristics were retrieved from their medical files.

For this validation process, we used data of a control group that was collected earlier for a similar trial in patients with multiple sclerosis (MS). This reference group was recruited at the participating GP office in Rotterdam, the Netherlands. Adult patients visiting the GP in September/October 2018 were asked to participate, while waiting for their appointment. Patients were found eligible if they had no difficulty reading and/or understanding the Dutch language. Exclusion criteria were similar to the patient group. All participants signed informed consent and references filled in the same set of questionnaires once.

Questionnaires

The NBD score is a questionnaire developed for patients with SCI and consists of symptom-based questions covering both constipation and FI [4]. The 10 multiple-choice questions have weighted answer options. Higher total scores are representing more severe bowel dysfunction (0–6 very minor, 7–9 minor, 10–13 moderate, and 14 or more severe). Some minor linguistic adjustments have been made after the original publication, and a scale question was added on general satisfaction of the current bowel management [5]. The latter is supposed to give the physician an insight of possible needs of the patient. Since this question was not included in the original questionnaire, and therefore not validated, we decided to omit this question in the current validation study. See attachment 1 for the Dutch-language version.

Three additional questionnaires were included in the set for the validation process. The first questionnaire is the Fecal Incontinence Quality of Life scale (FIQL), which measures the severity of incontinence for gas, mucus, liquid, and solid stool by scoring its frequency [6, 7]. This questionnaire has two different rating options, one for specialists and one for patients. The latter rating option was used for this study. The second questionnaire is the Fecal Incontinence Severity Index (FISI), which scores the severity of FI in patients by asking about the degree of occurrence of four aspects of incontinence: gas, mucus, liquid stool, and solid stool [7, 8]. Similar to the FIQL, this questionnaire has two rating options, a specialist specific and a patient-specific one. Again, for this study, patient-specific ratings were used. The final questionnaire is the European Quality of life 5-Dimension 3-Level questionnaire (EQ-5D-3L), a widely used instrument to measure health-related quality of life (HRQOL). The HRQOL is measured through five dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. This questionnaire also includes the EuroQol Visual Analog Scale, where patient can rate his or her health state from "best imaginable health state" to "worst imaginable health state".

Linguistic validation

Standardized guidelines on linguistic validation were followed for the cross-cultural adaptation of the original English-language NBD score into

the Dutch language [9]. During this process, three professional native Dutch-speaking translators performed the forward translation individually. During a meeting hosted by the researchers, consensus on the Dutch version of the questionnaire was reached by all translators and clinicians. After minor adjustments not affecting the context of the questionnaire, the translated version of the questionnaire was approved by two medical consultants with clinical experience in the treatment of patients with SCI. To finalize the translation procedure, an English native speaker performed the backward translation.

Measurement properties

Content validity. During the process of cross-cultural adaptation, content validity was assessed during face-to-face interviews with 17 patients at the outpatient's clinic. These patients were asked to give their opinion on clarity of the questions and easiness to fill in. The researchers also performed an assessment focusing on the questionnaire items and the correlation to the known clinical symptoms, which were found to be adequate.

Internal consistency. Internal consistency determines if items in a questionnaire are correlated, and if these questions measure the same underlying concept. Cronbach's alpha was measured for the total score of the NBD score, since no subscales are available. Values between 0.70 and 0.95 were considered sufficient, confirming adequate internal consistency [10].

Reproducibility. Reproducibility of a questionnaire is the degree of similarity of answers measured at different time points in a clinical stable situation. The test–retest period of 1–2 weeks was chosen following the recommendations of Terwee et al. [10], to determine the reproducibility of the NBD score. This time period is thought to be long enough to prevent recall bias, but short enough to prevent clinical imbalance. The interclass correlation coefficient (ICC) is used to determine the agreement of repeated test and retest measurements of the total NBD score. An outcome > 0.70 is considered adequate. The limits of agreement (LOA) are calculated as the mean change in scores of repeated measures \pm standard deviations (SD) [11, 12].

Criterion validity. Criterion validity is determined by measuring the correlation between the NBD score and a gold standard. Such true gold standard does not exist for the NBD score, so we chose the FIQL, FISI, and the EQ-5D-3L as a suitable combined gold standard. When linear associations are seen, the Pearson's correlation coefficient (range $-1, 1$) is used to measure the correlation of the NBD score with the chosen questionnaires. If no linear association is seen, the Spearman correlation coefficient (range $-1, 1$) will be used to measure such correlation.

Construct validity. Construct validity was determined by testing the following predefined hypotheses and was considered adequate when $\geq 75\%$ is confirmed [10].

Scores of references on the NBD score will be lower than scores of patients.

Patients who score lower on the FIQL will score higher on the NBD score.

Patients who score higher on the FISI will score higher on the NBD score.

Patients who score lower on the EQ-5D will score higher on the NBD score.

Floor and ceiling effects. When the lowest or highest score possible on a questionnaire is reached by $\geq 15\%$ of all respondents, floor and ceiling effects are present [10]. These effects were assessed for the total score of the NBD score at baseline, for both the patients and the reference group.

Statistical methods

A sample size of 50 patients and 50 controls was chosen based on guidelines on validation of questionnaires that were followed during this validation study [10]. All statistical analyses were performed using SPSS version 25 (IBM Corp, Armonk, NY). Mean and SD are reported in case of continuous data, and numbers or percentages for categorical data. Differences between patients and references were tested with Chi-

square tests for categorical variables and Student's *t* test for continuous variables. *P* values of <0.05 were considered statistically significant.

RESULTS

Between August 2018 and February 2020, patients were asked to participate in this trial. Fifty-eight patients agreed after being informed and the signed informed consent form. Two patients were excluded because the period between the first and the second sets was longer than 35 days. One patient was excluded because only the informed consent form was signed and further participation was declined. Eight patients were not included in the retest analysis, because they did not return the second set of questionnaires.

In September or October 2018, 50 control patients at the GP office agreed to participate in this trial and filled in the set of questionnaires, while waiting for their appointment. These references were significantly younger than the patients and there were significantly more men in the patient group, compared to the reference group. Significant differences of baseline total scores of all questionnaires between patients and references were seen. Characteristics of the patient and reference group can be found in Table 1.

Content validity

After professional medical translators translated the questionnaire into Dutch, 17 patients were interviewed to assess the validity of content. The questions were found to be relevant, clear, and easy to fill in; content validity was confirmed without the necessity of adjustments.

Internal consistency

Since the NBD score consists of only one total score measurement, if one or more questions are left open, this could not be calculated. For this reason total scores of six patients are missing. Five of these six patients left the same questions open at test and retest, namely, the questions on "digital evacuation of stool" and "frequency of FI". Regarding the last question, patients stated to miss an answer option "never" and left the question open because no suitable answer option was present. Internal consistency was found to be moderate to low for the total scores of the NBD score, Cronbach's alpha was measured to be 0.56 for the test and 0.30 for the retest.

Reproducibility

The mean time between completing the first and second sets of questionnaires was 17.4 (SD ± 7.4) days. The mean change between total scores of the NBD score of the test and retest was -0.5 ± 3.46 . For the total score of the questionnaire, the ICC was 0.87, indicating adequate reliability. The LOA ranges of the total scores were -7.28 to 6.78.

Criterion validity

Regarding criterion validity, significant correlations were found between total scores of the NBD score and total scores of the FIQL, FISl, and EQ-5D-3L. These correlations were found to be moderate. (Table 2).

Construct validity

Good construct validity was found, all predefined hypotheses were confirmed:

- References had significant lower scores in the NBD score than patients (Table 1).
- Patients who scored lower in the FIQL had significantly higher scores in the NBD score (Table 2).
- Patients who scores lower on the EQ-5D-3L had significantly higher scores in the NBD score (Table 2).

Table 1. Characteristics and baseline scores, presented as mean ± standard deviation or numbers (%).

	Patients <i>n</i> = 55 ^a	References <i>n</i> = 50 ^a	<i>p</i> Value
Age (years)	54 ± 15.8	38.4 ± 14.4	<0.001 ^b
Gender			
Male	36 (65.5)	11 (22)	<0.001 ^c
Female	19 (34.5)	39 (78)	
Duration SCI (years)	13.6 ± 8.4		
Level of SCI			
Cervical	14 (25.5)		
Thoracic	29 (52.7)		
Lumbar	9 (16.4)		
Unknown	3 (5.4)		
Mobility			
Fully ambulatory	5 (9.1)		
Limited walking	11 (20)		
Wheelchair bound	34 (61.8)		
Unknown	5 (9.1)		
ASIA Impairment Scale			
A	15 (27.3)		
B	7 (12.7)		
C	10 (18.2)		
D	9 (16.4)		
E	1 (1.8)		
Unkown	13 (23.6)		
Baseline scores			
NBD score	11.87 ± 6.36 (<i>n</i> = 54)	1.81 ± 3.67 (<i>n</i> = 47)	<0.001 ^b
NBD general satisfaction	5.89 ± 2.08 (<i>n</i> = 54)	7.61 ± 1.62 (<i>n</i> = 49)	<0.001 ^b
FIQL total	3.20 ± 0.81	3.99 ± 0.16	<0.001 ^b
Domain 1: lifestyle	3.25 ± 0.91	3.92 ± 0.23	<0.001 ^b
Domain 2: coping/behavior	3.14 ± 0.87	3.94 ± 0.19	<0.001 ^b
Domain 3: depression/self-perception	3.42 ± 0.88	4.12 ± 0.20	<0.001 ^b
Domain 4: embarrassment	2.98 ± 0.96	3.98 ± 0.10	<0.001 ^b
FISl total score	27.74 ± 13.21 (<i>n</i> = 54)	18.50 ± 13.59	0.001 ^b
EQ-5D-3L index score	0.50 ± 0.32	0.92 ± 0.15	<0.001 ^b
EQ-5D-3L VAS score	65.79 ± 17.19 (<i>n</i> = 47)	75.92 ± 15.11 (<i>n</i> = 38)	0.006 ^b

^aUnless stated otherwise.

^bStudent's *t* test.

^cChi-square test.

- Patients who scored higher in the FISl had significantly higher scores in the NBD score (Table 2).

Floor and ceiling effects

There were no floor and ceilings effects present in the patient group; one patient (1.8%) had the lowest score possible, no patient had the highest score possible. Floor effects were seen in the reference group, 52% of the participants had the lowest score (0) possible, indicating they experience no neurogenic bowel

Table 2. Criterion validity test.

	FIQL	FISI	EQ-5D-3L index score	EQ-5D-3L VAS score
Test, NBD total score	-0.648 ^a ($p < 0.001$)	0.367 ^a ($p < 0.001$)	-0.589 ^a ($p < 0.001$)	-0.428 ^a ($p < 0.001$)
Retest, NBD total score	-0.491 ^a ($p < 0.001$)	0.301 ^b ($p = 0.042$)	-0.394 ^a ($p = 0.007$)	-0.499 ^a ($p = 0.001$)

^aCriterion validity tested by Spearman correlation coefficient.

^bCriterion validity tested by Pearson's correlation coefficient.

symptoms. Ceiling effects were not present in the reference group; no one scored the highest score possible.

DISCUSSION

In the present study, the NBD score was translated into the Dutch language and this version was validated in patients with SCI. We demonstrated that the questionnaire is reliable and valid tool to measure NBD, and the impact of such dysfunction on a patient QoL. This enables physicians to evaluate NBD in a quick, easy, and objective manner.

The total scores of the NBD score of patients differed significantly from the references, showing that the NBD score can discriminate between those experiencing problems and those who are not. The reference group was significantly younger than the patient group and there were significantly more men in the patients group (Table 1). This discrepancy has had probably no major influence on the total scores of the NBD score, since the scores significantly differed, and therefore had no influence on the validation process.

Content validity was assessed during face-to-face interviews with 17 patients. In contrast to our other validation study in MS patients, no patients needed clarification on any of the questions [13]. A possible explanation for this is that the questionnaire is developed for SCI patients who are generally familiar with the used terminology like for example "digital evacuation of stool". In addition, bowel dysfunction in patients with SCI is more consistent and in MS these symptoms are dependent on the course of the disease [3].

Internal consistency was measured for the total scores of the NBD score with Cronbach's alpha and showed to be moderate to low (0.56 for the test and 0.30 for the retest). These measures are in line with the validation study of the Dutch-language NBD score in MS patients and a previous validation study by Erdem et al. [13, 14]. A low internal consistency alone is not problematic for a validation process according to guidelines on measurements in medicine [15]. The validity of a questionnaire is established by more than internal consistency, if the construct measured is evident, the questionnaire is still valid and reliable. A possible explanation for the lower internal consistency is that both FI and constipation are questioned in limited questions, with weighted answers and no subscales are available.

Criterion validity was measured using reliable questionnaires. Since no gold standard questionnaires for the validation of the NBD score exist, we choose the FISI, the FIQL, and the EQ-5D-3L as proxy gold standards, although the first two questionnaires solely measure FI and not constipation. Since all measurements significantly correlated, as could be seen in Table 2, we state that these questionnaires are suitable for this validation study.

Construct validity was determined by measuring the correlation of the total scores of the NBD score and the FIQL, FISI and the EQ-5D-3L. All hypotheses that were predefined were confirmed, showing good validity of the NBD score.

No floor and ceiling effects were seen in the patient group, no patient scored the highest or the lowest score possible. As expected, floor effects were seen in the reference group. A total

of 52% of the references had the lowest score possible, indicating no bowel problems.

One of the strengths of this study is that the translation and validation process was performed, according to standardized and widely used guidelines. The adequate sample size was reached for both the patient as the reference group. Another strength is the homogeneous nature of the patient group with SCI, showing good validity and adequate reliability in this specific patient group.

A limitation of this study is the dropout rate for the retest analysis. Eight patients were excluded for this analysis because the second sets of questionnaires were not returned. This retest only consisted of measuring reproducibility of the NBD score. Patients' characteristics of the excluded patients did not differ from the included patients, except for age. The excluded patients were significantly older, so clear instructions are recommended for the older patient population with SCI.

In conclusion, the NBD score is a valid and reliable tool, which enables physicians to open up the conversation on bowel dysfunction in this specific patients population and evaluate the impact on QoL. In addition, it will make it possible to evaluate different treatment options and it can also be used in future clinical research.

Furthermore, our group has previously validated the SF-Qualiveen on bladder symptoms [16] and the Multiple Sclerosis Intimacy and Sexuality Questionnaire on sexual dysfunction [17]. Together with these PRO's, the NBD score forms a complete set evaluating pelvic health in neurogenic patients.

DATA AVAILABILITY

The generated dataset from this trial is available from the corresponding author on reasonable request.

REFERENCES

- Krogh K, Christensen P, Laurberg S. Colorectal symptoms in patients with neurological diseases. *Acta Neurol Scand.* 2001;103:335-43.
- Sun WM, MacDonagh R, Forster D, Thomas DG, Smallwood R, Read NW. Anorectal function in patients with complete spinal transection before and after sacral posterior rhizotomy. *Gastroenterology.* 1995;108:990-8.
- Krogh K, Christensen P. Neurogenic colorectal and pelvic floor dysfunction. *Best Pr Res Clin Gastroenterol.* 2009;23:531-43.
- Krogh K, Christensen P, Sabroe S, Laurberg S. Neurogenic bowel dysfunction score. *Spinal Cord.* 2006;44:625-31.
- Coloplast. Coloplast - test your bowel symptoms: coloplast. 2020. <https://www.coloplast.us/Global/US/Continence/Peristeen/NBD%20Score.pdf>.
- Rockwood TH, Church JM, Fleshman JW, Kane RL, Mavrantonis C, Thorson AG, et al. Fecal Incontinence Quality of Life Scale: quality of life instrument for patients with fecal incontinence. *Dis Colon Rectum.* 2000;43:9-16. discussion -7.
- t Hoen LA, Utomo E, Schouten WR, Blok BF, Korfaage IJ. The fecal incontinence quality of life scale (FIQL) and fecal incontinence severity index (FISI): validation of the Dutch versions. *Neurourol Urodyn.* 2017;36:710-5.
- Rockwood TH, Church JM, Fleshman JW, Kane RL, Mavrantonis C, Thorson AG, et al. Patient and surgeon ranking of the severity of symptoms associated with fecal incontinence: the fecal incontinence severity index. *Dis Colon Rectum.* 1999;42:1525-32.
- Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol.* 1993;46:1417-32.

10. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60:34–42.
11. de Vet HC, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. *J Clin Epidemiol.* 2006;59:1033–9.
12. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986;1:307–10.
13. van Doorn T, Groenendijk IM, Scheepe JR, Blok BFM. The neurogenic bowel dysfunction score in patients with multiple sclerosis: validation of the Dutch-language version. (Under Review).
14. Erdem D, Hava D, Keskinoglu P, Bircan C, Peker O, Krogh K, et al. Reliability, validity and sensitivity to change of neurogenic bowel dysfunction score in patients with spinal cord injury. *Spinal Cord.* 2017;55:1084–7.
15. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in medicine: a practical guide.* Cambridge University Press; 2011.
16. Reuvers SHM, Korfage IJ, Scheepe JR, Blok BFM. The urinary-specific quality of life of multiple sclerosis patients: Dutch translation and validation of the SF-Qualiveen. *Neurourol Urodyn.* 2017;36:1629–35.
17. Noordhoff TC, Scheepe JR, t Hoen LA, Sluis TAR, Blok BFM. The Multiple Sclerosis Intimacy and Sexuality Questionnaire (MSISQ-15): validation of the Dutch version in patients with multiple sclerosis and spinal cord injury. *Neurourol Urodyn.* 2018;37:2867–74.

ACKNOWLEDGEMENTS

First, we would like to thank Dr. Frank Poelstra (GP in Rotterdam) for facilitating the reference group. We would also like to thank Prof. Klaus Krogh for his clarifications on the NBD score.

AUTHOR CONTRIBUTIONS

TvD performed the data collection. IMG, BFMB, and JRS were involved in the translation and validation process and supervised the consensus meetings. TvD prepared the manuscript, together with IMG, JRS, and BFMB.

COMPETING INTERESTS

The authors declare no competing interests.

STATEMENT OF ETHICS

The institutional Medical Ethics Review Board reviewed the research proposal and gave ethical approval (MEC 2018-1050) We certify that all applicable institutional and national regulations concerning human volunteers were followed during the course of this research.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41393-021-00668-8>.

Correspondence and requests for materials should be addressed to T.v.D.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.