Reviewer ID: Marzena Zhou, Risa Fox

Type of Outcome Measure: Neurogenic Bowel Dysfunction Score (NBD) | **Total articles:** 3

Author ID Year	Study Design	Setting	Population (sample size, age) and Group
Krogh et al. 2006	Development/validation study	University Hospital of Aarhus, Denmark.	**SCI N:** 589 (424/72%responded) **Level:** cervical 174, thoracic 155, 79 lumbar; 254 sensory complete, 166 incomplete in 166 **Etiology:** trauma (75%), spinal surgery (8%), myelomeningocele (4%), infection (4%), spinal thrombosis or hemorrhage (3%), or other causes (6%). **Age:** Mean 41 years, Range8-88 years **Duration:** Mean 14 years, Range 0 to 59 years **% Female:** 29%
Erdem et al. 2017	Psychometrics study.	Dokuz Eylül University Faculty of Medicine, Turkey.	**SCI N:** 42 **Level:** 12 cervical, 24 thoracic 24,6 lumbar; AIS 23 A, 4 B, 11 C, 4 D **Etiology:** traffic accident 21, 4 firearm injury, 11 falling down, 3 crushing underweight , 1 diving into shallow water, 2 Others **Age:** Mean 39 years, SD 16 years **Duration:** ⩽1 year - 15 2–5 year- 17 6–9 year- 6 ⩾10 year- 4 **% Female:** 19%
Kelly et al. 2017	Development, Reliability and Validation Strudy	University of California Irvine, Urology Center, CHOC Children's Hospital, Orange, California	**Spina Bifida N:** 34 children **Control N**: 18 patients **Level:** 4 thoracic (12%), 16 lumbar (46%), 12 sacral (35%), 2 unknown (6%) **Ambulatory status**: 16 fully ambulatiory (47%), 18 wheelchair or walker (53%) **Etiology:** spina bifida **Age:** Mean 5.3yo **Ethnicity:** Caucasian 14 (27%) Hispanic 31 (60%) Asian 7 (13%) **% Female:** 50%

1. RELIABILITY

Author ID	Internal Consistency	Test-retest, Inter-rater, Intra-rater
Erdem et al. 2017	NBD score is not a Likert type scale, but the additivity of the total score was tested with Tukey's nonadditivity test and the additivity characteristic was demonstrated. Due to the presence of additivity characteristic of the test and ordinal characteristic of the answers, Cronbach alpha coefficient was calculated to determine the reliability of internal	Test-retest answers of each question r=1.000, P<0.001 Consistency of frequency distribution r=1.000, P<0.001

	consistency. Cronbach's alpha coefficient for internal consistency was 0.547	
Kelly et al. 2017	No data available	**Test-Retest Reliability** B/w each answer on the two questionnaires: 53% was found (k>0.8) 32% had k b/w 0.6 and 0.8 Difference b/w mean score of 1st and 2nd questionnaire not statistically significant. **Inter-Rater Reliability** B/w each answer on the two questionnaires: 79% with k>0.8 18% had k b/w 0.6 and 0.8

2. VALIDITY

Author ID	Validity
Krogh et al. 2006	**Reproducibility & validity of questions describing colorectal problems in patients with spinal cord lesions** **Item & Validity (k coefficient):** Frequency of bowel movements k = 0.92 Desire for defaecation (any) k = 0.89 Normal desire for defaecation k = 1.00 Abdominal discomfort k = 0.83 Perspiration, headache, or general discomfort during defecation k = 0.82 Oral laxatives k = 1.00 Enemas k = 0.77 Average time for defecation k = 0.79 Use of Clysma k = 0.90 Frequency of digital stimulation k = 0.77 Need helpfrom others for defecation k = 0.73 How much does disturbed defecation restrict social activities? k = 0.63 How much does disturbed defecation restrict quality of life? k = 0.83 Frequency of faecal incontinence k = 0.79 Flatus incontinence k = 0.53 Medication against faecal incontinence k = 1.00 Perianal skin problems k = 0.60 How much does faecal incontinence restrict social activities? k = 0.62 How much does faecal incontinence restrict quality of life? k = 0.52 **OR, level of significance and points in the NBD score** **for items significantly associated with impact on quality of life** Frequency of bowel movements OR = 6.1, p < 0.0001, points in NBD score = 6 Headache, perspiration or dyscomfort before or at defecation OR = 2.4, p < 0.01, points in NBD score = 2 Tablets against constipation OR = 1.9, p < 0.001, points in NBD score = 2 Drops against constipation OR = 2.3, p < 0.0001, points in NBD score = 2 Time used for defecation OR = 6.8, p < 0.0001, points in NBD score = 7 Digital stimulation or evacuation OR = 5.0 p < 0.01, points in NBD score = 6 Frequency of faecal incontinence OR = 13.1, p < 0.0001, points in NBD score = 13 Medication against faecal incontinence OR = 3.6, p < 0.01, points in NBD score = 4 Flatus incontinence OR = 1.8, p < 0.05, points in NBD score = 2 Perianal skin problems OR = 2.6, p < 0.01, points in NBD score = 3
Erdem et al. 2017	Correlations between total NBD score & SF-36 subscales, a statistically signicant negative correlation was detected between:

	• bodily pain (r=− 0.382, P=0.013) • general health (r=− 0.560, P<0.001) • vitality (r=− 0.626, P<0.001) • social role functioning (SF) (r=− 0.741, P<0.001) • emotional role functioning (r=− 0.604, P<0.001) and • mental health (r=− 0.687, P<0.001) subscales, No signicant correlation was found with the following subscales of SF-36: • physical functioning (PF) (r=− 0.233, P=0.138) and • physical role functioning (RP) (r=0.067, P=0.674) The highest correlation among these was found in the social function subscale. • Signicant negative correlation with the mental component summary score (MCS) (r=− 0.872, P< 0.001) • No signicant correlation with the physical component summary score (PCS) (r=− 0.187, P=0.235). • Signicant positive correlation between NBD total score and PGA (r=0.91, P<0.001). • Significant positive correlation was found between NBD total score and patients' assessment of impact of NBD on QoL (r=0.92, P<0.001). The patients were divided into 2 groups according to their NBD scores: • NBD scores between 0 and 9 were classied as mild and • NBD scores >10 were classied as signicant NBD. According to this, a statistically signicant difference was detected in all SF-36 subscales and MCS score (P<0.05), except PF and RP subscales and PCS score in both groups (P>0.05). When total NBD score before and after treatment was compared with Wilcoxon test, it was demonstrated that improvement in NBD score at the end of 2 months was significant (P=0.011). Also, there was a statistically significant positive correlation between Global Rating of Change scale and the change in total NBD score at the end of 2 months (r=0.821, P=0.007)
Kelly et al. 2017	**Construct Validity** A Spearman's rank-order correlation (r_s): r_s = 0.943 (P<0.0005)
3. RESPONSIVENESS – no data available	
4. FLOOR/CEILING EFFECT – no data available	
5. INTERPRETABILITY	
Author ID	**Interpretability**
Kelly et al. 2017	**Mean Scores:** Bowel subjects: 15.18 (SD ± 5.77) Control subjects: 4.68 (SD ± 2.98)