



Strong agreement between interview-obtained and self-administered Wexner and St. Mark's scores using a single questionnaire

Stig Norderval^{1,2} · Mona Birgitte Rydningen¹ · Ragnhild Sørum Falk³ · Arvid Stordahl⁴ · Hege Hølmo Johannessen⁵

Received: 15 December 2018 / Accepted: 28 March 2019 / Published online: 6 June 2019
© The International Urogynecological Association 2019

Abstract

Introduction and hypothesis St Mark's incontinence score (SMIS) and the Wexner score have been constructed and validated as interview-based scoring systems. We developed a single questionnaire from which a separate SMIS or Wexner score could be derived. This study aimed to demonstrate the level of agreement between self-administered (sSMIS and sWexner) and interview-based (iSMIS and iWexner) scores using this questionnaire.

Methods One hundred five consecutive patients (30 male) seen in the incontinence outpatient clinics at the Østfold Hospital Trust, Sarpsborg, and University Hospital of North Norway, Tromsø, completed the self-administered incontinence questionnaire prior to the appointment. Following clinical investigation, the patients were interviewed about their symptoms according to the SMIS and Wexner scores, with the interviewers blinded to the results from self-reported questionnaire. Agreement between total scores and between subscores of the various items were determined using interclass correlation coefficient (ICC) and kappa statistics, respectively.

Results The self-administered questionnaire was incomplete in six cases (5.7%) and the interview-based was incomplete in two cases. Agreement was almost perfect between the iSMIS and sSMIS and between the iWexner score and sWexner score (ICC 0.90 and 0.92, respectively). Agreement was substantial to almost perfect for all items in both scoring systems, with kappa values ranging from 0.64–0.94. Mean iSMIS was 9.48 versus 9.53 for sSMIS ($p = 0.90$) and 8.26 versus 8.44 for the iWexner and sWexner score, respectively ($p = 0.42$).

Conclusion The SMIS and Wexner scores can be completed by the patients using a single questionnaire, and the derived SMIS and Wexner scores are highly consistent with scores obtained by interview.

Keywords Fecal incontinence · Anal incontinence · Wexner score · St Mark's score · Scoring systems

Introduction

Fecal incontinence is a bothersome condition affecting 3–14% of the western population in general [1, 2], with higher prevalence rates in various subgroups. The wide range in reported prevalence can be explained by differences in the definition of fecal incontinence and in methods for data collection. In a large systematic review where these factors were adjusted for, Sharma and co-workers estimated the prevalence of fecal incontinence occurring at least once per month to 8.3–8.4% in face-to-face or telephone interviews and 11.2–11.4% in postal surveys [3]. The prevalence of anal incontinence, which also includes involuntary loss of gas, is even higher [3]. Following increased focus on the condition over the last 30 years, several

✉ Stig Norderval
norde@online.no; stig.norderval@unn.no

¹ Department of Gastrointestinal Surgery, University Hospital of North Norway, 9038 Tromsø, Norway

² Gastrosurgical Research Group, Department of Clinical Medicine, UiT the Arctic University of Norway, Tromsø, Norway

³ Oslo Centre of Biostatistics and Epidemiology, Oslo University Hospital, Oslo, Norway

⁴ Department of Surgery, Østfold Hospital Trust, Sarpsborg, Norway

⁵ Department of Physical Medicine and Rehabilitation, Østfold Hospital Trust, Sarpsborg, Norway

scoring systems have been developed to quantify the symptoms and to assess response to treatment. Two of the most commonly used scoring systems are the Wexner score [4] and the St. Mark's incontinence score (SMIS) [5]. The two scores are similar, but the SMIS is based on symptoms over the last 4 weeks only. In addition, the SMIS also records the use of a plug, the presence of urgency symptoms and the use of constipating medication.

Although frequently used as self-administered questionnaires [6–8], both scoring systems are constructed and validated as interview scores. However, there are several advantages to self-administered questionnaires: they enable a more effective workup as the patients can score their symptoms prior to the consultation, reduce the risk of bias introduced by the interviewer and enable a more efficient follow-up regimen. In a research setting, the use of self-administered questionnaires clearly facilitates an easier study setup than interviews. Four of the authors of the present study (SN, MBR, AS and HHJ) are members of the steering committee of the Norwegian National Quality Registry of Anal Incontinence. This registry has decided to use both the SMIS and Wexner score as instruments to assess the severity of incontinence prior to and after treatment of the included patients, preferably by using one questionnaire. Substantial agreement between interview-based and self-administered SMIS and Wexner scores has been reported [9, 10], but not within the frame of a single questionnaire. We aimed to assess the level of agreement between self-administered Wexner scores (sSMIS and sWexner) and scores obtained by interview (iSMIS and iWexner) using a single questionnaire. Our aim was not to establish a new scoring system, but to simplify information retrieval by using a single questionnaire from which a separate SMIS or Wexner score could be derived.

Materials and methods

All consecutive patients aged ≥ 18 years seen in the incontinence outpatient clinics at the Østfold Hospital Trust, Sarpsborg, and University Hospital of North Norway, Tromsø, were offered inclusion in the study. The patients were new referrals, follow-ups after primary obstetric sphincter repairs or planned controls after conservative or surgical treatment for anal incontinence. Included patients completed the self-administered incontinence questionnaire in a separate room next to the investigation room prior to the appointment. After the scheduled clinical investigation, the patients were interviewed by the attending colorectal surgeon, incontinence physiotherapist or incontinence nurse, using the same questionnaire but blinded to the results from the self-administered questionnaire. The interviewers made sure that the patients had understood the various questions by rephrasing them and giving examples if needed. Age and gender were also

recorded. The study was approved by the Data Protection Officer at the two hospitals, and informed consent was obtained in all cases. The study was classified as a quality improvement project; hence, approval from the Regional Committee for Medical Research Ethics was not required.

Recording of incontinence symptoms

The original interview versions of the Wexner score [4] and SMIS [5] stratify the frequency of incontinence to gas, liquid and solid stool in addition to the frequency of lifestyle alteration due to incontinence symptoms. Both scoring systems also record use of pads, but while the Wexner score registers the frequency of pad use, the SMIS only registers use or no use as a dichotomous variable. The SMIS furthermore dichotomously records the use of a plug, constipating medication and symptoms of urgency. The maximum symptom score is 20 in the Wexner score and 24 in the SMIS. There are also slight differences in how “never,” “rarely” and “sometimes” are defined in the two scoring systems. While the SMIS defines “never” as no symptoms during the last 4 weeks, no such restriction is made in the Wexner score. This implies that, for example, one incontinence episode for liquid stool 2 months prior to assessment is classified as “rarely” (< 1 episode/month) and generates one point in the Wexner score, but in the SMIS is classified as “never” and thereby adds zero points. Furthermore, one such incontinence episode during the last 4 weeks is classified as “sometimes,” generating two points in the Wexner score, but classified as “rarely” and adding one point to the SMIS. The questionnaire including both the SMIS and Wexner scoring systems therefore had to incorporate these differences. In contrast to the original interview forms, the layout of the questionnaire was set up in order to facilitate completion of the form by addressing one question at the time. The questionnaire and instructions how to calculate separate SMIS and Wexner scores are shown in Fig. 1. The primary outcome was the level of agreement between self-reported scores (sWexner and sSMIS) and scores obtained by interview (iWexner and iSMIS). Secondary outcomes were comparison of total scores and subscores and the degree of correlation between the sWexner and iWexner score and between the sSMIS and iSMIS.

Statistics

Data were analyzed using IBM SPSS Statistics for Mac, version 25 (IBM, Armonk, NY, USA). Level of agreement was assessed according to the COSMIN recommendations [11], using Cohen's kappa (κ) statistics for dichotomous variables and quadratic weighted kappa statistics for ordinal level variables, applying different weights depending on the magnitude of disagreement [12]. Agreement was rated as slight ($\kappa \leq 0.2$), fair ($0.2 < \kappa \leq 0.4$), moderate ($0.4 < \kappa \leq 0.6$), substantial ($0.6 < \kappa \leq 0.8$) and almost perfect ($0.8 < \kappa \leq 1$) [13].

Fig. 1 Questionnaire including the St. Mark's and Wexner incontinence scoring systems. To derive the Wexner score, only questions 1–5a are assessed: “never” = 0; “less than once a month” = 1; “once last 4 weeks” = 2; “2–3 times last 4 weeks” = 2; “once a week or more, but not daily” = 3; “daily” = 4. To obtain the total score, add the points from each item. Minimum score = 0; maximum score = 20. To derive the St. Mark's score, all questions are assessed. Questions 1–4: “never” = 0; “less than once a month” = 0; “once last 4 weeks” = 1; “2–3 times last 4 weeks” = 2; “once a week or more, but not daily” = 3; “daily” = 4. Question 5a: “never” = 0; “less than once a month” = 0; “once last 4 weeks” or more frequent = 2. Question 5b: “no” = 0; “yes” = 2, but 0 if points already obtained in question 5a. Question 6: “no” = 0; “yes” = 2. Question 7: “15 minutes or more” = 0; “less than 15 minutes” = 4. To obtain the total score, add the points from each item. Minimum score = 0; maximum score = 24

In this questionnaire you are being asked to answer questions regarding incontinence for gas or stool. Please tick the relevant boxes. Please remember to answer **all** questions.

1) How often do you experience incontinence for solid stool?

- Never
 Less than once a month
 Once last 4 weeks
 2-3 times last 4 weeks
 Once a week or more, but not daily
 Daily

2) How often do you experience incontinence for liquid stool?

- Never
 Less than once a month
 Once last 4 weeks
 2-3 times last 4 weeks
 Once a week or more, but not daily
 Daily

3) How often do you experience incontinence for gas?

- Never
 Less than once a month
 Once last 4 weeks
 2-3 times last 4 weeks
 Once a week or more, but not daily
 Daily

4) Have your incontinence symptoms caused you to alter your lifestyle?

- Never
 Less than once a month
 Once last 4 weeks
 2-3 times last 4 weeks
 Once a week or more, but not daily
 Daily

5a) Do you wear pads because of stool incontinence?

- Never
 Less than once a month
 Once last 4 weeks
 2-3 times last 4 weeks
 Once a week or more, but not daily
 Daily

5b) Are you using a plug because of stool incontinence

- No
 Yes

6) Are you taking constipating medication because of stool incontinence?

- No
 Yes

7) For how long are you usually able to defer defecation?

- 15 minutes or more
 Less than 15 minutes

Level of agreement and level of correlation between the total scores of the interview-based and self-administered SMIS (iSMIS and sSMIS) and between the iWexner and sWexner were assessed with the two-way single-measure intraclass correlation coefficient (ICC) and Spearman's rho, respectively. Paired *t*-test was used for comparison of means, with 95% confidence interval (CI) of the difference. Related

samples Wilcoxon signed-ranks test with interquartile (IQ) range was used to compare ordinal data and chi-square test for comparison of dichotomous data. Two-sided *p* values of < 0.05 were considered statistically significant. There is no standardized way to perform a power analysis in agreement studies, but based on the considerations of Liao [14] we aimed to include about 100 patients.

Results

In the period from 1 September 2017 to 31 December 2017, a total of 105 patients (75 females and 30 males) were included from the two incontinence outpatient clinics. Mean age was 55.2 years (95% CI 52.5–57.9; range: 24–85). The self-administered questionnaire was incomplete in six cases (5.7%) affecting the SMIS only in two cases. Furthermore, the interview-based questionnaire was incomplete in two cases in which the interviewer had not stated the presence or absence of urgency. Incomplete questionnaires were excluded from analysis. Comparable SMISs were obtainable in 97 of 105 (92%) patients and Wexner scores in 101 of 105 (96%) patients.

Agreement was almost perfect between the total scores for sSMIS and iSMIS (ICC = 0.90; 95% CI 0.85–0.93) as well as between the total sWexner and iWexner scores (ICC = 0.92; 95% CI 0.88–0.94). Agreement between the individual items was substantial except for ‘use of pad or plug’ and ‘incontinence for gas’ in the SMIS and ‘use of pad’ and ‘incontinence for gas’ in the Wexner score where agreement was almost perfect. The kappa values for the individual items in the SMIS and Wexner score are presented in Tables 1 and 2, respectively. Agreement between the total sWexner and iWexner score was 0.88 (95% CI 0.73–0.94) in 24 women interviewed by a nurse or a physiotherapist, compared with 0.92 (95% CI 0.88–0.95) in 77 women interviewed by a colorectal surgeon. Agreement between the total sSMIS and iSMIS score was 0.87 (95% CI 0.72–0.94) in 24 women interviewed by a nurse or a physiotherapist compared with 0.89 (95% CI 0.83–0.93) in 73 women interviewed by a colorectal surgeon.

Mean iSMIS and sSMIS were 9.48 and 9.53, respectively (95% CI of the difference: -0.67 – 0.58 , $p = 0.90$). Mean iWexner and sWexner score was 8.26 and 8.44, respectively (95% CI of the difference: -0.28 – 0.65 , $p = 0.42$). Incontinence for solid stool was rated as more severe in the

sSMIS and sWexner than in the iSMIS and iWexner, while alteration of lifestyle was rated as more frequent in the iSMIS and iWexner than in the sSMIS and sWexner (Table 3 and 4).

Spearman’s rho showed a strong correlation between the sSMIS and iSMIS ($r = 0.90$, $n = 97$, $p < 0.01$) and between the sWexner and iWexner scores ($r = 0.91$, $n = 101$, $p < 0.01$).

Discussion

To our knowledge, this is the first study, assessing the level of agreement between self-administered and interview-obtained data for both the Wexner score and the SMIS using a single questionnaire including the two incontinence scoring systems. Our results show that there is a strong level of agreement and correlation between the self-reported and interview-based SMIS and Wexner score within the frame of the questionnaire. Furthermore, the mean total SMIS and Wexner score did not differ between interview-based and self-administered assessments.

While several studies have explored various aspects of the Wexner score and SMIS when used as a self-administered questionnaire [7, 8, 15, 6, 16], we are aware of just two previous studies assessing the level of agreement between the self-administered SMIS and interview-based SMIS [9, 10], with one study also assessing the Wexner score [10]. In the study by Johannessen et al. [9] involving 147 women at follow-up after first delivery or after primary repair of obstetric anal sphincter tears, the total SMIS was somewhat higher when patients completed the SMIS themselves than the score obtained by standard interview (score 4.0 and 4.3, respectively), and the kappa values for the various items varied between 0.22 and 0.72 [9]. The patients also rated more severe incontinence to both solid and liquid stool in the self-administered SMIS compared with the interview-based one. In our study, a similar discrepancy was only revealed for solid stool incontinence (Tables 3 and 4). It has been shown that people report

Table 1 Agreement between self-administered and interview-based variables for St. Mark’s incontinence scoring system ($n = 97$)

Variables	Exact agreement (crude, %)	Agreement within ± 1 point (crude, %)	Kappa
Ordinal variables (score 0–4)			
Incontinence for solid stool	81	93	$\kappa = 0.75^*$
Incontinence for liquid stool	75	89	$\kappa = 0.73^*$
Incontinence for gas	72	91	$\kappa = 0.84^*$
Alteration in lifestyle	69	85	$\kappa = 0.72^*$
Dichotomous variables			
Need to wear a pad or plug	97	–	$\kappa = 0.94$
Taking constipating medicines	91	–	$\kappa = 0.70$
Lack of ability to defer defecation for 15 min	82	–	$\kappa = 0.64$

*Quadratic weighted kappa statistics

Table 2 Agreement between self-administered and interview-based variables for the Wexner incontinence scoring system ($n = 101$)

Variables	Exact agreement (crude, %)	Agreement within ± 1 point (crude, %)	Kappa*
Ordinal variables (score 0–4)			
Incontinence for solid stool	80	93	$\kappa = 0.78$
Incontinence for liquid stool	67	89	$\kappa = 0.74$
Incontinence for gas	68	95	$\kappa = 0.84$
Alteration in lifestyle	65	84	$\kappa = 0.77$
Need to wear pad	84	88	$\kappa = 0.86$

*Quadratic weighted kappa statistics

embarrassing symptoms more willingly in a questionnaire than to an interviewer [17]. In the systematic review by Sharma and colleagues, incontinence for solid and liquid stool as well as for gas was reported more frequently in studies using postal surveys than in face-to-face or telephone interviews [3]. This phenomenon could explain the higher frequency of incontinence for solid stool in the self-administered questionnaire than in the the interview-based questionnaire. However, we would expect incontinence for liquid stool and gas to be equally embarrassing [18], but the outcome in the two groups regarding these two items did not differ at all in our study. We have no good explanation for why this discrepancy was seen only for solid stool in the present study. It might be that patients who need to clean themselves repeatedly after bowel movements define the small amount of remaining stool as leakage, but we have no data that can shed further light upon this.

In the study by Hussain and co-workers [10] including 36 patients, the level of agreement between patient-administered and interview-obtained SMIS and Wexner scores was assessed at two time points 6 weeks apart. The first interview was undertaken by a physician, the second by a nurse. The SMIS and Wexner scores were almost identical between the patient’s self-reported rating and the rating done by the nurse, with an ICC of 0.94 and 0.95, respectively. Between patient and physician, however, median SMIS differed by 4.0 points and Wexner score by 3.5 as the physician rated the incontinence as less severe than the patients, and ICC was 0.79 for both scoring systems. No such difference was seen in the current study, reflecting a high degree of agreement between the health carers undertaking the interviews. It is important that health personnel who assess incontinent patients agree in the rating of symptoms, especially when determining response to

Table 3 Median values for the various subscores and frequencies of the dichotomous variables of self-administered and interview-based St Mark’s incontinence score ($n = 97$)

Variables	iSMIS	sSMIS	iSMIS > sSMIS (%)	sSMIS > iSMIS (%)	<i>p</i> value
Ordinal variables (score 0–4)					
	Median (IQ range)	Median (IQ range)			
Incontinence for solid stool	0 (0–1)	0 (0–2)	5	13	0.05*
Incontinence for liquid stool	0 (0–2)	0 (0–2)	12	12	0.63*
Incontinence for gas	3 (0–4)	3 (0–4)	14	13	0.81*
Alteration in lifestyle	1 (0–4)	0 (0–3.5)	22	9	0.02*
Dichotomous variables					
	Yes/no (%)	Yes/no (%)			
Need to wear a pad or plug	48/52	47/53	1	2	0.87**
Taking constipating medicines	18/82	21/79	3	6	0.58**
Lack of ability to defer defecation for 15 min	57/43	64/36	5	12	0.30**

sSMIS: self-administered St. Mark’s incontinence score

iSMIS: interview-based St. Mark’s incontinence score

IQ: interquartile

*Related samples Wilcoxon signed-ranks test

**Chi-square test

Table 4 Median values for the various subscores of the self-administered and interview-based Wexner incontinence score ($n = 101$)

Variables	iWexner Median (IQ range)	sWexner Median (IQ range)	iWexner>sWexner (%)	sWexner>iWexner (%)	P value*
Ordinal variables (score 0–4)					
Incontinence for solid stool	0 (0–2)	0 (0–2)	5	15	0.01
Incontinence for liquid stool	1 (0–2.5)	1 (0–2)	16	17	0.77
Incontinence for gas	3 (1–4)	3 (1–4)	13	19	0.32
Alteration in lifestyle	2 (0–4)	1 (0–4)	22	13	0.03
Need to wear pad	0 (0–4)	0 (0–4)	10	5	0.09

sWexner: self-administered Wexner incontinence score

iWexner: interview-based Wexner incontinence score

IQ: Interquartile

*Related samples Wilcoxon signed-ranks test

treatment. Using self-administered questionnaires where the patient rates her or his symptoms prior to and after treatment reduces this risk of bias.

The reported frequency of the item “alteration of lifestyle” was lower in the self-administered assessments than in the interviewed-based assessments (Tables 1 and 2). It is interesting that patients with a relatively high mean SMIS or Wexner score rate the impact on lifestyle as less severe when completing the questionnaire by themselves than in an interview setting. Four of the authors of this article (SN, MBR, AS and HHJ) have > 10 years of experience each in interviewing patients about anal incontinence using the SMIS and Wexner scores, and the impression is that quite a few patients are neglecting their lifestyle alterations unless being questioned in detail. This may explain why the interviewer reported this item to be more frequent than the self-administered questionnaire. However, as no such difference was found in the study by Johannesen and colleagues [9], our finding must be interpreted with caution.

The level of agreement between the self-administered and interview-based SMIS in the present study was higher than found in the study by Johannesen et al. [9]. There are several possible reasons for this. The symptom burden was more severe among the patients in our study, with mean SMIS being twice as high as in the study by Johannesen. If the vast majority of the patients report no symptoms, this will reduce the kappa value between two assessments even if the number of cases with exact agreement is high. This can explain why the level of agreement in the study by Johannesen et al. was lower than in the present study, despite a similar frequency of exact agreements between the various items in the sSMIS and sSMIS [9]. Furthermore, the patients in our study had sought medical help for their incontinence and consequently had discussed their problems with health personnel prior to completion of the questionnaire. In contrast, the patients in the study by

Johannesen and co-workers were assessed at follow-up after first delivery or after obstetric anal sphincter tears, and many of the patients had minor or no incontinence. It is possible that patients with mild and infrequent incontinence are less aware of their symptoms, resulting in a higher level of inconsistency between the two assessment methods the first time this subject is addressed.

The layout of the questionnaire may also be important. The present questionnaire deals with one symptom at a time, and the patients have to tick the box indicating the correct frequency for a specific symptom before proceeding to the next question. In previous studies where the Wexner score or SMIS was used as a questionnaire for self-assessment, the original setup of the two scoring systems was used [7, 19, 15, 9, 10]. This setup is basically a cross table, which could be somewhat difficult for the patients to comply with, thereby increasing the possibility of giving an incorrect answer unintentionally.

Using one questionnaire instead of two almost identical questionnaires has the advantage that patients, and also interviewers, have to spend less time completing the forms when both scoring systems are to be used. Attempts to include the Wexner and SIMS in one questionnaire have been made previously. Bols and colleagues designed a questionnaire from which the SMIS and Wexner scores were meant to be derived [19]. Unfortunately, the level of agreement between the recorded interview-based and self-administered items was not assessed, and the authors did not consider the difference between the two scoring systems regarding the frequency categories “rarely” and “sometimes.” Hence, only the SMIS could actually be recorded and reported correctly, thereby reducing the usefulness of this questionnaire. The present questionnaire incorporates these differences between the SMIS and Wexner score and enables correct scores as intended for both scoring systems. Nevertheless, questionnaires for self-completion will always carry the risk that some patients will fail to answer

all the questions, and in the present study this occurred in almost 6% despite written instruction to answer all questions. Incompletely filled-out questionnaires reduce the usefulness of this method of information retrieval. This problem is overcome by using electronic questionnaires where all questions have to be answered before submitting, and this method should be chosen in the future also for this questionnaire.

The present study has some limitations. First, the interviews were undertaken the same day as the patients completed the self-administered questionnaire. This introduces a possible bias, as patients may remember their initial answers and repeat them during the following interview, even if realizing that their former answer was not quite correct. However, as the study by Johannesen and co-workers was performed in a similar way, the higher level of agreement found in the present study is less likely explained by this factor. Second, the study was carried out on patients with moderate to severe incontinence, and it is not known whether similar results would have been obtained if the patients had less severe symptoms like in the study of Johannesen et al. This has, however, been investigated by Roos and co-workers, who found SMIS to be reliable regardless of the severity of incontinence symptoms [20]. Third, the study does not assess the questionnaire's test-retest stability, internal consistency and sensitivity to detect changes over time. However, as the questionnaire only consists of the various items of the SMIS and Wexner score, the above-mentioned assessments will equal these assessments for the two scoring systems separately. The test-retest stability and consistency are shown to be substantial to excellent for both the SMIS and Wexner score [5, 16, 21, 10], and both scoring systems are highly sensitive to detect changes after treatment [5, 21]. Hence, it is generally agreed that both the Wexner incontinence score and SMIS provide important information about the degree and extent of symptoms [5, 15, 19, 8, 20]. To comply with both the SMIS and Wexner score, the present questionnaire contains five answer options in each item 1 to 5a compared with four options in the SMIS and Wexner score, as previously explained. This new constellation could, however, induce changes in the response to less frequent symptoms compared with if the answers were given according to the original SMIS or Wexner questionnaire. Although we regard this potential risk as minimal, we have no data that can shed further light upon this.

In conclusion, our study shows that the presented questionnaire can be completed by the patients on their own. The questionnaire enables calculation of both the SMIS and Wexner scores, and the derived scores are highly consistent with the SMIS and Wexner scores obtained in an interview setting.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Conference presentation Poster presentation at the International Continence Society 48th annual meeting, Philadelphia, USA, 28–31 August 2018.

References

1. Ilnyckyj A. Prevalence of idiopathic fecal incontinence in a community-based sample. *Can J Gastroenterol*. 2010;24(4):251–4.
2. Menees SB, Almario CV, Spiegel BMR, Chey WD. Prevalence of and factors associated with fecal incontinence: results from a population-based survey. *Gastroenterology*. 2018;154(6):1672–1681 e1673. <https://doi.org/10.1053/j.gastro.2018.01.062>.
3. Sharma A, Yuan L, Marshall RJ, Merrie AE, Bissett IP. Systematic review of the prevalence of faecal incontinence. *Br J Surg*. 2016;103(12):1589–97. <https://doi.org/10.1002/bjs.10298>.
4. Jorge JM, Wexner SD. Etiology and management of fecal incontinence. *Dis Colon Rectum*. 1993;36(1):77–97.
5. Vaizey CJ, Carapeti E, Cahill JA, Kamm MA. Prospective comparison of faecal incontinence grading systems. *Gut*. 1999;44(1):77–80.
6. van Meegdenburg MM, Trzpis M, Broens PM. Fecal incontinence and parity in the Dutch population: A cross-sectional analysis. *United European Gastroenterol J*. 2018;6(5):781–90. <https://doi.org/10.1177/2050640618760386>.
7. Rygh AB, Komer H. The overlap technique versus end-to-end approximation technique for primary repair of obstetric anal sphincter rupture: a randomized controlled study. *Acta Obstet Gynecol Scand*. 2010;89(10):1256–62. <https://doi.org/10.3109/00016349.2010.512073>.
8. Bakx R, Sprangers MA, Oort FJ, van Tets WF, Bemelman WA, Slors JF, et al. Development and validation of a colorectal functional outcome questionnaire. *Int J Color Dis*. 2005;20(2):126–36. <https://doi.org/10.1007/s00384-004-0638-9>.
9. Johannesen HH, Norderval S, Stordahl A, Falk RS, Wibe A. Interview-based versus self-reported anal incontinence using St Mark's incontinence score. *Int Urogynecol J*. 2018;29(5):667–71. <https://doi.org/10.1007/s00192-017-3363-y>.
10. Hussain ZI, Lim M, Stojkovic S. The test-retest reliability of fecal incontinence severity and quality-of-life assessment tools. *Dis Colon Rectum*. 2014;57(5):638–44. <https://doi.org/10.1097/dcr.000000000000118>.
11. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res Int J Qual Life Asp Treat Care Rehab*. 2010;19(4):539–49. <https://doi.org/10.1007/s11136-010-9606-8>.
12. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull*. 1968;70(4):213–20.
13. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–74.
14. Liao JJ. Quantifying an agreement study. *Int J Biostat*. 2015;11(1):125–33. <https://doi.org/10.1515/ijb-2014-0030>.
15. Maeda Y, Pares D, Norton C, Vaizey CJ, Kamm MA. Does the St. Mark's incontinence score reflect patients' perceptions? A review of 390 patients. *Dis Colon Rectum*. 2008;51(4):436–42. <https://doi.org/10.1007/s10350-007-9157-4>.

16. Cam C, Selcuk S, Asoglu MR, Tug N, Akdemir Y, Ay P, et al. Validation of the Wexner scale in women with fecal incontinence in a Turkish population. *Int Urogynecol J*. 2011;22(11):1375–9. <https://doi.org/10.1007/s00192-011-1464-6>.
17. Tourangeau R, Yan T. Sensitive questions in surveys. *Psychol Bull*. 2007;133(5):859–83. <https://doi.org/10.1037/0033-2909.133.5.859>.
18. Johannessen HH, Morkved S, Stordahl A, Sandvik L, Wibe A. Anal incontinence and quality of life in late pregnancy: a cross-sectional study. *BJOG*. 2014;121(8):978–87. <https://doi.org/10.1111/1471-0528.12643>.
19. Bols EM, Hendriks HJ, Berghmans LC, Baeten CG, de Bie RA. Responsiveness and interpretability of incontinence severity scores and FIQL in patients with fecal incontinence: a secondary analysis from a randomized controlled trial. *Int Urogynecol J*. 2013;24(3):469–78. <https://doi.org/10.1007/s00192-012-1886-9>.
20. Roos AM, Sultan AH, Thakar R. St. Mark's incontinence score for assessment of anal incontinence following obstetric anal sphincter injuries (OASIS). *Int Urogynecol J Pelvic Floor Dysfunct*. 2009;20(4):407–10. <https://doi.org/10.1007/s00192-008-0784-7>.
21. Sansoni J, Hawthorne G, Fleming G, Marosszeky N. The revised faecal incontinence scale: a clinical validation of a new, short measure for assessment and outcomes evaluation. *Dis Colon Rectum*. 2013;56(5):652–9. <https://doi.org/10.1097/DCR.0b013e318279c2ac>.

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