

ORIGINAL ARTICLE

Comparative analysis of summary scoring systems in measuring fecal incontinence

Moo-Kyung Seong, Sung-Il Jung, Tae-won Kim, Hee-Kyung Joh¹

Departments of Surgery, ¹Family Medicine, Konkuk University School of Medicine, Seoul, Korea

Purpose: For measuring symptoms of fecal incontinence, summary scoring systems are widely used, but rigorous psychometric validation or assessment of such systems in terms of patients' subjective perception has rarely been done to date. This study was designed to assess the correlation between each severity measure and patients' subjective perception or clinicians' clinical assessment. We attempted to compare summary scoring systems of severity measures and searched for which of them showed higher validity among them. **Methods:** Consecutive patients who visited our clinic with fecal incontinence were prospectively evaluated. A total of 43 patients were included. Four summary scoring systems were chosen for comparison: the Rothenberger, Wexner, Vaizey and Fecal Incontinence Severity Index systems. They are correlated with subjective perception scores by patients, and also with clinical assessment scores by investigators. **Results:** There was no significant difference between clinical scores of two investigators (paired t-test, $P = 0.988$). Inter-observer reliability was 0.95 (Intra-class correlation coefficient, 95% confidence interval 0.91 to 0.98). Significant correlations were proved between patients' subjective perception scores and all the summary scoring systems, and also between the mean clinical scores and all the summary scoring systems. The highest was with the Wexner scale ($r = 0.66$, $P < 0.001$) ($r = 0.70$, $P < 0.001$), and the lowest was with the Rothenberger scale ($r = 0.58$, $P < 0.001$) ($r = 0.61$, $P < 0.001$) in both correlations. **Conclusion:** The Wexner scale correlates the most closely with subjective perception of severity of symptoms by patients, and also with clinical assessment by investigators. We recommend the Wexner scale among summary scoring systems as a tool for measuring fecal incontinence.

Key Words: Fecal incontinence, Measurement, Scores

INTRODUCTION

Fecal incontinence is defined as the inability to control feces and to expel it at a proper place and a proper time [1]. Its severity ranges from mild difficulty with gas control to complete loss of control over liquid and formed stool. While it is not clinically lethal, fecal incontinence is some-

times socially debilitating, and some patients inevitably change their lifestyle according to with their disease, depending on their personal character. In this context, it is a kind of disorder which needs a symptom-based approach rather than a traditional disease-based approach [2,3].

For symptom-based evaluation, fecal incontinence should be measured by subjective assessment. Objective

Received June 2, 2011, Revised July 30, 2011, Accepted August 29, 2011

Correspondence to: Moo-Kyung Seong
Department of Surgery, Konkuk University Medical Center, Konkuk University School of Medicine, 4-12 Hwayang-dong, Gwangjin-gu, Seoul 143-729, Korea
Tel: +82-2-2030-7582, Fax: +82-2-2030-7949, E-mail: recto@kuh.ac.kr

© Journal of the Korean Surgical Society is an Open Access Journal. All articles are distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

studies including anorectal manometry, electrophysiologic study, and endoanal ultrasonography are clinically useful for determining causes and treatment policy of fecal incontinence, but they do not measure incontinence, their findings can not always be matched with clinical symptoms, and so they have limitations in grading severity or evaluating treatment outcomes. For measuring symptoms of fecal incontinence, many systems of assessment have been developed. They can be broadly classified into descriptive measures, impact measures, and severity measures. Among them, severity measures are more commonly used than others in clinical practice [4].

Usual severity measures are summary scoring systems that assign values for certain categories of incontinence and produce summary scores based on the addition of values for each category [5-10]. Although summary scoring systems are widely used and some of them also accommodate impact components, rigorous psychometric validation or assessment in terms of patients' subjective perception has rarely been done to date [9]. This study was de-

signed to assess the correlation between each severity measure and patients' subjective perception or clinicians' clinical assessment. We attempted to compare existing summary scoring systems of severity measures and searched for which showed higher validity and utility.

METHODS

Participants in the present study were consecutive patients who visited our clinic with fecal incontinence between March 2009 and January 2010. They were prospectively evaluated. A total of 43 patients were included. There were 25 women with a mean age of 65.8 ± 11.2 years and 18 men with a mean age of 59.8 ± 16.2 years. The minimum sample size was calculated to detect a correlation of 0.55 or better at 5% significance level with 80% power was 19.

Four summary scoring systems, which are more common in use and their items of measurement seem repre-

Table 1. Summary of summary scoring systems in measuring fecal incontinence: frequency, type, and impact components

Variable	Frequency			Type				Impact components	
	Category number	Lowest ^{a)}	Highest	Gas	Liquid	Solid	Others	Pad usage	Lifestyle alteration
Rothenberger	4	Less than 1/mo	More than 1/wk	+	+	+	-	-	+
Wexner	5	Less than 1/mo	More than 1/day	+	+	+	-	+	+
Vaizey	5	1/mo	Every day	+	+	+	Urgency	+	+
FISI	6	1-3/mo	More than 2/day	+	+	+	Mucus	-	-

FISI, Fecal Incontinence Severity Index.

^{a)}Other than "never".

Table 2. Summary of summary scoring systems in measuring fecal incontinence; score

Variable	Score	
	Range	Weight
Rothenberger	0-30	Scores for gas incontinence are 1-3; liquid incontinence 4-6; solid incontinence 7-9; lifestyle alteration 10-12.
Wexner	0-20	All categories are given equal weights.
Vaizey	0-24	Gas, liquid, solid incontinence and lifestyle alteration are given equal weights. Others are scored variably.
FISI	0-61	Weights were already determined by patients.

FISI, Fecal Incontinence Severity Index.

representative of changing trends, were chosen for comparison: the Rothenberger [6] (modified Miller [5,6]), Wexner (Cleveland Clinic Florida) [8], Vaizey (St. Mark) [9], and Fecal Incontinence Severity Index (FISI) [10] systems. With FISI system, weightings by patients were used (Tables 1, 2).

Two investigators took the histories from each patient separately, one at OPD office during his or her first visit and the other at laboratory during physiologic study, usually two or three days after first visit. After all initial evaluations, patients' medical records were reviewed, including results of anal manometry, nerve conduction study, and endoanal ultrasonography. With information obtained, each of the investigators gave clinical scores on a scale of 0 to 10 (0 = perfect control, 10 = no control). They were blinded to each other's results. Apart from this process, patients were asked to complete all the written scoring systems mentioned above and check, additionally, their subjective perception scores. Frequency was determined upon episodes during the past four weeks. Another coordinating nurse, who was blind to patients' history, assisted the patients and explained how to fill in the paper work. Subjective perception scores by patients were scaled identically with clinical assessment scores by investigators.

Statistical analysis was done by SPSS ver. 16.0 (SPSS Inc., Chicago, IL, USA). Shapiro-Wilk test was used to confirm normality of data distribution. Paired t-test compared clinical assessment scores of the two investigators.

Table 3. Summary on the scores of summary scoring systems, patients' subjective perception scores, and investigators' clinical assessment scores

Variable	Mean	Percentage ^{a)}	SD	Range
Rothenberger (0-30)	9.0	63.3	9.1	3-30
Wexner (0-20)	9.1	45.5	5.6	2-20
Vaizey (0-24)	12.3	51.3	6.4	2-22
FISI (0-61)	28.0	45.9	16.8	4-61
Patients' perception (0-10)	6.4	64	2.4	2-10
Investigator's assessment, A (0-10)	6.3	63	2.2	1-9
Investigator's assessment, B (0-10)	6.2	62	2.1	1-9

FISI, Fecal Incontinence Severity Index.

^{a)}Percentage of the mean to maximum score.

Intra-class correlation coefficient (ICC) was used to determine inter-observer reliability. With Pearson correlation, subjective perception scores by patients and the mean of two clinical assessment scores by investigators were correlated with each of the summary scoring systems. The continuous variables were expressed as mean \pm SD. Statistical significance was set at $P < 0.05$.

RESULTS

Table 3 details all the variables. There was no significant difference between clinical assessment scores by the two investigators (paired t-test, $P = 0.988$). Inter-observer reliability was 0.95 (ICC, 95% confidence interval 0.91 to 0.98). Between-subject variation was 4.45, and within-subject variation was 0.26.

Significant correlations were proved between subjective perception scores by patients and all the summary scoring systems, and also between the mean clinical assessment scores by investigators and all the summary scoring systems (Table 4). The highest was with the Wexner scale, and the lowest was with the Rothenberger scale, in both the former and the latter correlations.

DISCUSSION

Quantifying a variety of symptoms into objective scale is mandatory for determining severity of incontinence or evaluating treatment outcomes. Various measures of in-

Table 4. Correlations of summary scoring systems with patients' subjective perception scores and investigators' clinical assessment scores

Variable	Patients' perception score		Clinical assessment score ^{a)}	
	Correlation coefficient	P-value	Correlation coefficient	P-value
Rothenberger	0.58	<0.001	0.61	<0.001
Wexner	0.66	<0.001	0.70	<0.001
Vaizey	0.64	<0.001	0.63	<0.001
FISI	0.63	<0.001	0.65	<0.001

FISI, Fecal Incontinence Severity Index.

^{a)}Mean of clinical assessment scores by two investigators.

continence symptoms are available to date, but which is better is still an ongoing problem for clinicians. They can be categorized into descriptive measures, impact measures, and severity measures.

Descriptive measures evaluate various aspects of fecal incontinence with numerous items of questions [11-13]. Each item is analyzed separately without giving any score. Impact measures focus on the impact of incontinence on individual quality of life. Generic impact measures [14] and fecal incontinence-specific impact measures [15,16] coexist.

Severity measures stratify incontinence directly by giving values to some aspects of incontinence. Two kinds are available, grading systems [17-20] and summary scoring systems [4-10]. Grading systems give a value to specific types of incontinence in an ordinal fashion. Summary scoring systems assign values for certain categories of incontinence and produce summary scores based on the addition of values for each category. These systems take into account the fact that various aspects of incontinence, mainly type and frequency, contribute to severity. Values for each type of incontinence are assigned according to the frequency of incontinent episodes. Summary scores are calculated based on the addition of values for each category, and thus they are much more likely to enable differentiation between groups and detection of clinically important change. Because of this benefit and relative convenience in clinical application, summary scoring systems of severity measures have been more commonly used than other measures [4].

Ten or more summary scoring systems are available to date. The Rothenberger, Wexner, Vaizey, and FISI are, however, chosen for comparison in this study, because they are relatively more common in use and their items of measurement seemed representative of changing trends.

Rothenberger scale [6], also known as modified Miller scale [5,6], gives variable weights to the same frequencies of different types of incontinence. Incontinence to liquid stool gets twice or more the value of incontinence to gas at the same frequency. Similarly incontinence to solid stool gets three times or more the value of incontinence to gas at the same frequency. But such distribution of weights is not based on patient perspective, and it may not reflect the

subjective experience of patients. Moreover, disproportionate weighting has patients inclined to high scores [5,6]. It was also shown in this study by a higher percentage of the mean (Table 3). This may cause a drop in discriminating power.

Wexner scale [8] is simple and easy to understand. It gives value to all types of incontinence equally, and therefore, the same frequencies of incontinence of gas and incontinence of solid stool contribute equally to the severity score. Actually, this equality adds a difficulty in stratifying the degree of sphincter impairment. This scale takes into account usage of pads and lifestyle alteration, which are associated with impact of incontinence. The inclusion of those items could cause an error in measuring severity, because items such as wearing of a pad might reflect the degree of individual fastidiousness, and it often relates to the presence of coexistent urinary leakage or vaginal discharge. Male patients tend not to use a pad. Scores for pad usage could be misleading for male incontinence [3].

Vaizey scale [9] additionally takes account of fecal urgency and constipating medicines. Fecal incontinence may be urgent or passive. Urgency means the inability to defer defecation. Passiveness means occurrence of episode without the patient's awareness. Although urgency can be present without overt fecal incontinence, it may be as serious to an affected individual as an overt one. Constipating, anti-diarrheal agents are sometimes used for incontinence treatment, and it may give a false impression of lower severity [3]. Thus, such addition of items could make the scale more meaningful. However, scores of this scale for patients with passive incontinence may not adequately reflect their symptom severity, because passiveness of incontinence is not considered for evaluation. Four points given for urgency could be too high [3].

FISI, which was developed by Rockwood et al. [10] gives variable weights to various frequencies and types of incontinence on the basis of subjective ratings of severity. Such ratings were developed by using both patient and surgeon input. Considering the subjective nature of incontinence, incorporation of patient values into severity measurement could be notable. But it has no aspect of impact such as alteration of life style, which the other three systems have. It has four types of incontinence including

mucus, in addition to usual gas, liquid and solid stool. Mucus is sometimes misleading to false sensation of liquid stool, and on occasion, patients may record a falsely high score.

Although there are some differences as mentioned above among these systems, similarities outweigh the differences. All the systems have some limitations in common. They regard frequency of incontinence as a major category of measurement, but patients often alter their lifestyle enough to avoid events of incontinence. The actual severity of symptoms measured by type and frequency of incontinence might not correlate with the subjective perception as some patients are depressed by only minor leakage, whereas others with major incontinence manage the symptoms by protective measures. Another point of limitations in common is the amount of leakage, which is missing in measurement. Hence, two patients similar in frequency but very different in amount may record the same score [3]. Above all, there are only limited data about the comparative validity of these systems [9], and thus, it is not easy to recommend the use of one over any other.

Searching for higher validity among summary scoring systems, the 0 to 10 (11 data points) scale was used as a comparison standard in this study. Actually, given the lack of objective measures, there is no exact criterion standard for comparison among the systems. The 0 to 10 scale is a well-known scale. It came from Visual Analog Scale, which had been well studied in the context of pain and known to allow patients to express the full spectrum of their problem in a simple scale [21]. Thus we used it as an additional comparison standard.

In conclusion, this study has shown that the Wexner scale correlates the most closely with subjective perception by patients, and also with clinical assessment by investigators about severity of symptoms, and the Rothenberger scale correlates the least closely. We recommend the Wexner scale among summary scoring systems as a tool for measuring fecal incontinence, although it has its own limitations as mentioned above, until a more precise and comprehensive tool is devised in the near future.

CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

ACKNOWLEDGEMENT

This study was supported by Konkuk University, 2009.

REFERENCES

1. Baeten CG, Kuijpers HC. Incontinence. In: Wolff BG, Beck DE, Church JM, Fleshman JW, Garcia-Aguilar J, Pemberton JH, et al. editors. The ASCRS textbook of colon and rectal surgery. New York: Springer; 2007. p.653-64.
2. Drossman DA, Dumitrascu DL. Rome III: New standard for functional gastrointestinal disorders. *J Gastrointest Liver Dis* 2006;15:237-41.
3. Maeda Y, Parés 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;51:436-42.
4. Baxter NN, Rothenberger DA, Lowry AC. Measuring fecal incontinence. *Dis Colon Rectum* 2003;46:1591-605.
5. Miller R, Bartolo DC, Locke-Edmunds JC, Mortensen NJ. Prospective study of conservative and operative treatment for faecal incontinence. *Br J Surg* 1988;75:101-5.
6. Rothenberger DA. Anal incontinence. In: Cameron JL, editor. *Current surgical therapy-3*. Philadelphia: BC Decker; 1989. p.186-94.
7. Pescatori M, Anastasio G, Bottini C, Mentasti A. New grading and scoring for anal incontinence. Evaluation of 335 patients. *Dis Colon Rectum* 1992;35:482-7.
8. Jorge JM, Wexner SD. Etiology and management of fecal incontinence. *Dis Colon Rectum* 1993;36:77-97.
9. Vaizey CJ, Carapeti E, Cahill JA, Kamm MA. Prospective comparison of faecal incontinence grading systems. *Gut* 1999;44:77-80.
10. 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.
11. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30:473-83.
12. Reilly WT, Talley NJ, Pemberton JH, Zinsmeister AR. Validation of a questionnaire to assess fecal incontinence and associated risk factors: fecal incontinence questionnaire. *Dis Colon Rectum* 2000;43:146-53.

13. Osterberg A, Graf W, Karlbom U, Pählman L. Evaluation of a questionnaire in the assessment of patients with faecal incontinence and constipation. *Scand J Gastroenterol* 1996;31:575-80.
14. Malouf AJ, Norton CS, Engel AF, Nicholls RJ, Kamm MA. Long-term results of overlapping anterior anal-sphincter repair for obstetric trauma. *Lancet* 2000;355:260-5.
15. 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.
16. Bug GJ, Kiff ES, Hosker G. A new condition-specific health-related quality of life questionnaire for the assessment of women with anal incontinence. *BJOG* 2001;108:1057-67.
17. Parks AG. Royal Society of Medicine, Section of Proctology; Meeting 27 November 1974. President's Address. Anorectal incontinence. *Proc R Soc Med* 1975;68:681-90.
18. Keighley MR, Fielding JW. Management of faecal incontinence and results of surgical treatment. *Br J Surg* 1983;70:463-8.
19. Williams NS, Patel J, George BD, Hallan RI, Watkins ES. Development of an electrically stimulated neoanal sphincter. *Lancet* 1991;338:1166-9.
20. Womack NR, Morrison JF, Williams NS. Prospective study of the effects of postanal repair in neurogenic faecal incontinence. *Br J Surg* 1988;75:48-52.
21. Lukacz ES, Lawrence JM, Burchette RJ, Lubner KM, Nager CW, Buckwalter JG. The use of Visual Analog Scale in urogynecologic research: a psychometric evaluation. *Am J Obstet Gynecol* 2004;191:165-70.