Reliability and Validity of a Modified Bristol Stool Form Scale for Children

Mariella M. Lane, PhD, Danita I. Czyzewski, PhD, Bruno P. Chumpitazi, MD, MPH, and Robert J. Shulman, MD

Objectives This study sought to: evaluate the ability of children to reliably use a modified Bristol Stool Form Scale for Children (mBSFS-C), evaluate criterion-related validity of the mBSFS-C, and identify the lower age limit for mBSFS-C use.

Study design The mBSFS-C comprises 5 stool form types described and depicted in drawings. Children 3 to 18 years old rated stool form for 10 stool photographs. Because of low reliability when stool form descriptors were not read aloud (n = 119), a subsequent sample of children (n = 191) rated photographs with descriptors read.

Results Intraclass correlation coefficients for descriptor-unread versus -read samples were 0.62 and 0.79, respectively. Children were increasingly reliable with age. Percentage of correct ratings varied with stool form type, but generally increased with age. With descriptors unread, children 8 years and older demonstrated acceptable interobserver reliability, with >78% of ratings correct. With descriptors read, children 6 years and older demonstrated acceptable reliability, with >80% of ratings correct.

Conclusions The mBSFS-C is reliable and valid for use by children, with age 6 years being the lower limit for scale use with descriptors read and age 8 years being the lower limit without descriptors read. We anticipate that the mBSFS-C can be effectively used in pediatric clinical and research settings. *(J Pediatr 2011;159:437-41).*

Alterations in stool form are associated with numerous gastrointestinal symptoms and conditions that affect children. Patient report of stool form changes are used clinically for diagnosis and treatment and are measured in research settings to assess clinical outcomes.[1,2]

Despite the importance of accurately assessing stool form changes, a stool form scale has not been validated for this use in pediatric patients. In adults, the Bristol Stool Form Scale (BSFS) is an increasingly used self-report instrument for assessing stool form, although it was validated as a measure of gastrointestinal transit time[3-8] without assessment of the ability of adults to reliably and accurately identify stool form type. The scale allows classification of stool form in 7 types, ranging from "separate hard lumps like nuts" (type 1) to "watery, no solid pieces" (type 7).[3] The BSFS has been used to evaluate stool form in a variety of clinical studies,[9-16] and the Rome Foundation has recommended its use to assess stool form in adults with functional gastrointestinal disorders.[17]

In children, gathering information about stool form may be challenging because caregivers may not directly observe all stools. As such, healthcare providers are often placed in a position in which a child's description of his/her stools is required for diagnostic and clinical decision-making. However, procuring accurate descriptions from children may be particularly challenging without an objective tool. Because no stool form scale has been validated for use in children, we sought to validate a modified Bristol Stool Form Scale for children (mBSFS-C) for this indication.

Thus, the aims of this study were to evaluate the ability of children to reliably use a modified BSFS, evaluate evidence for criterion-related validity of the modified scale by assessing children's ability to correctly identify stool form type, and identify the lower limit of age for which use of the mBSFS-C is appropriate. We expected to demonstrate evidence of inter-observer reliability and criterion-related validity of the mBSFS-C, hypothesizing both to increase with age and predicting that we could identify a minimum age for appropriate scale use.

Methods

As described previously, the original BSFS was adapted by decreasing the number of stool categories from 7 to 5.[18] This reduces the discriminations children are required to make, with the goal of maintaining scale usefulness while increasing accuracy of classification by young children who may have more difficulty

BSFS	Bristol Stool Form Scale
mBSFS-C	modified Bristol Stool Form Scale for Children

Table I. Percent correct ratings for each stool photograph (descriptors read)

		Age range (years)				
	Total sample	3-5 (n = 34)	6-7 (n = 36)	8-10 (n = 43)	11-13 (n = 36)	14-18 (n = 42)
Photograph D (type 1)	99.0	100	97.2	97.7	100	100
Photograph G (type 1)	92.7	91.2	94.4	90.7	91.7	95.2
Photograph F (type 2)	68.6	35.3	66.7	76.7	86.1	73.8
Photograph J (type 2)	74.9	44.1	77.8	74.4	86.1	88.1
Photograph A (type 3)	96.9	97.1	88.9	100	97.2	100
Photograph H (type 3)	89.5	76.5	91.7	93.0	94.4	90.5
Photograph B (type 4)	71.7	44.1	52.8	81.4	86.1	88.1
Photograph E (type 4)	61.8	35.3	61.1	62.8	77.8	69.0
Photograph C (type 5)	91.1	76.5	86.1	93.0	97.2	100
Photograph I (type 5)	92.1	91.2	86.1	95.3	91.7	95.2
Overall percent correct	83.8	69.1	80.3	86.5	90.8	90.0

Expert rater agreement for each photograph was: photograph A, 100%; photograph B, 100%; photograph C, 100%; photograph D, 100%; photograph E, 92.9%; photograph F, 78.6%; photograph G, 100%; photograph H, 100%; photograph I, 92.9%; photograph J, 92.9%.[18]

attending to and discriminating between the original 7 categories. Reducing response categories for self-report scale use by children is common practice.[19-21]

The 5 categories were chosen to adequately describe a range of stool form from hard through watery. Type 3 ("like a sausage or snake but with cracks on its surface") and type 5 ("soft blobs with clear cut edges") were eliminated as response options, with the remaining types 1, 2, 4, 6, and 7 comprising the modified scale. The final scale (Figure; available at www.jpeds.com) consists of 5 stool form descriptors accompanied by drawings similar to the Rome II and Rome III renderings of the BSFS.[17,22]

In an initial evaluation of our modification of the BSFS, 14 pediatric gastroenterologists used the scale to rate 32 color photographs of stool. Photographs were obtained from publicly accessible areas of the internet and depicted focused, close-up photographs of entire bowel movements.[18] The modified scale had a high degree of inter- and intra-rater reliability when used by the expert raters.[18]

Selection of Stool Photographs for Rating by Children

To select which photographs would be used as stimuli for child ratings, the distributions of stool form ratings by the expert raters were evaluated for the 32 stool photographs. Two photographs of each stool form type with high absolute agreement among the 14 expert raters were selected (Table I).

Recruitment and Data Collection

The Baylor College of Medicine institutional review board approved the study, and consent/assent was obtained. Participants included patients or siblings ages 3 to 18 years who were attending a scheduled outpatient clinic visit at Texas Children's Hospital. Data were collected in the waiting room or an examination room. A trained research assistant presented children with the 10 stool photographs in random order, with the mBSFS-C printed beneath each. Children were asked to use the mBSFS-C to assign a stool form category to each photograph by pointing to their selection on the scale.

In our initial sample, a research assistant presented the mBSFS-C without reading descriptors aloud. When these data indicated less than acceptable reliability for the total sample and particularly for the youngest children (see Results), we sought to evaluate whether reading descriptors aloud could improve children's ability to use the scale. In a second sample, stool form descriptors were read aloud for each photograph.

Statistical Analysis

Statistics were performed with SPSS software version 17.0 (SPSS Inc, Chicago, Illinois). Single measures intraclass correlation coefficients (two-way random effects model with absolute agreement) were used to assess inter-observer reliability. For both the descriptor-read and -unread samples, values are presented for the total sample and 5 age ranges.

Sources of variance in the ratings (ie, variance because of raters, photographs, or the interaction of rater and photograph) were examined for the descriptor-read total sample, in which a high degree of variance caused by photographs versus other sources provides strongest support for the scale. We also explored whether children's rating deviations from the expert rating (ie, rating errors) were random or systematic; difference scores were not normally distributed, so non-parametric one-sample Wilcoxon signed ranks tests were used. Kruskal-Wallis tests then explored whether rating errors were related to age group.

Criterion-related validity was evaluated by comparing the child stool form ratings to the correct classifications determined by the physician ratings in our earlier study. Percentages of children selecting the correct stool form classification for each photograph were calculated for the total descriptor-read sample and for each age range. To compare differences in criterion-related validity between the descriptor-read and the descriptor-unread administration methods, overall percent correct for the total unread sample and for each age range was also calculated.