Published in final edited form as:

Female Pelvic Med Reconstr Surg. 2013 ; 19(5): 293–297. doi:10.1097/SPV.0b013e31828ab3e2.

Readability of Common Health-Related Quality-of-Life Instruments in Female Pelvic Medicine

Alexandria N. Alas, MD*, Jonathan Bergman, MD†, Gena C. Dunivan, MD‡, Rezoana Rashid, BS§, Shelby N. Morrisroe, MD†, Rebecca G. Rogers, MD‡, and Jennifer T. Anger, MD, MPH§

*Department of Obstetrics and Gynecology, Cedars-Sinai Medical Center, Beverly Hills

†Department of Urology, University of California Los Angeles, Los Angeles, CA

‡Division of Urogynecology, University of New Mexico, Albuquerque, NM

§Department of Surgery, Division of Urology, Cedars-Sinai Medical Center, Beverly Hills, CA

Abstract

Objectives—The average American adult reads below the eighth-grade level. To determine whether self-reported health-related quality-of-life questionnaires used for pelvic floor disorders are appropriate for American women, we measured reading levels of questionnaires for urinary incontinence(UI), pelvic organ prolapse(POP), and fecal incontinence (FI).

Methods—An online literature search identified questionnaires addressing UI, POP, and FI. Readability was assessed using Flesch-Kincaid reading level and ease formulas. Flesch-Kincaid grade level indicates the average grade one is expected to completely and lucidly comprehend the written text. Flesch-Kincaid reading ease score, from 0 to 100, indicates how easy the written text can be read.

Results—Questionnaires were categorized by UI, POP, FI, and combined pelvic floor symptoms. The median Flesch-Kincaid reading level was 7.2, 10.1, 7.6, and 9.7, for UI, POP, FI, and combined pelvic floor symptoms, respectively. Reading levels varied greatly between questionnaires, with only 54% of questionnaires written below the eighth-grade level.

Conclusions—We identified significant variation in reading levels among the questionnaires and found the 2 most commonly used questionnaires per survey in 2008 at Society of Urodynamics and Female Pelvic Medicine and Urogenital Reconstruction were above the recommended eighth-grade reading level. As specialty societies focus on standardizing questionnaires for research, reading levels should be considered so they are generalizable to larger populations of women with pelvic floor disorders.

Keywords

HRQOL; urinary incontinence; fecal incontinence; pelvic organ prolapse; pelvic floor disorder

Reprints: Jennifer T. Anger, MD, MPH, Department of Surgery, Division of Urology, Cedars-Sinai Medical Center, 99 La Cienega Blvd, #307, Beverly Hills, CA 90211. angerj@cshs.org.

The authors have declared they have no conflicts of interest.

The World Health Organization has suggested that health is not merely the absence of disease but also one's state of physical, social, and mental well-being.¹ Pelvic floor disorders (PFDs), including urinary incontinence (UI), fecal incontinence (FI), and pelvic organ prolapse (POP), are unique in that they adversely impact quality of life, often without affecting general health.² Disease-specific health-related quality-of-life (HRQOL) instruments, rather than general HRQOL instruments, have been shown to best quantify the severity and impact of PFDs on women.³ Pelvic floor disorders are treated based on history and physical examination, and administration of HRQOL instruments may not be standard for all practitioners. However, the data collected from these questionnaires can play an integral role in improving health by assisting in clinical decision-making, and providing accurate outcome measures of treatment that can be used in research to further our understanding of these diseases and improve treatments.^{4–9}

On the basis of a nationwide government survey of more than 19,000 participants from all 50 states in 2003, the National Assessment of Adult Literacy recommended that health information be written below an eighth-grade level. It was found that 50% of African Americans and Hispanics and 40% of seniors read below the fifth-grade level, illustrating that appropriate readability is even more necessary for disadvantaged minorities who may struggle with daily reading comprehension. In addition, health illiteracy has been shown to exacerbate health disparities based on a systematic review in 2004. It was shown that those with lower literacy levels were 1.5 to 3 times more likely to have an adverse event affecting their health condition or hospital stay.¹⁰

It has been suggested by the National Center for Education Statistics that most of the disease-specific HRQOL instruments, as well as patient-targeted pamphlets, Web sites, and questionnaires are written in a higher reading level than recommended.^{11,12} The purpose of this study was to assess the readability of the HRQOL instruments used for UI, POP, and FI.

MATERIALS AND METHODS

We performed a systematic search of HRQOL instruments available for PFDs in our descriptive study evaluating literacy levels using the Flesch-Kincaid readability statistics.

Literature Review and Instrument Characteristics

We used Web page searches of American Urogynecologic Society, Society for Urodynamics and Female Urology (SUFU), International Continence Society, and American Medical Association as well as searches on PubMed, Cochrane Database of Systematic Reviews, Patient Reported Outcome and Quality of Life Instruments Database (MapiResearch Institute, Lyon, France), Google, and Google Scholar to identify studies reported from 1999 to the November 2011 related to HRQOL, UI, POP, and FI. Initial search was performed by the same individual and repeated by a second individual for verification. We included all questionnaires in the study that pertained to PFDs. We excluded those that did not pertain to females. The main text-word strategies included “health-related quality of life,” “quality of life,” “questionnaire,” “instrument,” “urinary incontinence,” “fecal incontinence,” “pelvic organ prolapse,” and “pelvic floor disorder.” Text-word phrases were combined in all possible permutations. Manuscripts were then obtained and categorized by the HRQOL

instruments used. After the online questionnaire search, all authors reviewed the findings to ensure no further questionnaires needed to be included.

Readability Analysis

Before calculating readability for all HRQOL instruments, survey questions were converted from fragment sentences to complete sentences by adding necessary punctuation, as described by Bergman et al.¹³ For example, if the survey question was “I usually experience frequent urination,” we completed the sentence by adding a period. Proper formatting was required to ensure correct readability analysis. In addition, we combined root questions with corresponding stem answer choices to form complete sentences. For example, if a question asked, “I feel pain,” and the answer choices were “all the time,” or “some of the time,” we would connect the sentence fragments together such that the question would read, “I feel pain all the time.” All single word responses, such as “always” and “never” were excluded from the readability analysis to avoid skewing results. We also recorded the number of question items in each questionnaire.

Readability analysis was performed using Microsoft Word, 2003 version, Microsoft Institute, Seattle, Wash, by using the Flesch-Kincaid readability statistics, which were included in the software. The Flesch-Kincaid reading ease formula has been validated and was used by the United States Navy in 1975 to create a modified reading ease formula known as the Flesch-Kincaid reading grade level which has become one of the gold standards for calculating reading levels.^{11,14–16}

Our analysis was conducted by copying each modified HRQOL text into separate documents. Document settings were adjusted to display read ability statistics. Finally, the document text was selected and a spelling and grammar check was completed. Through this process, the Flesch-Kincaid readability statistics were displayed for the highlighted text as well. We analyzed the individual survey items as well as the entire survey. The Flesch-Kincaid grade level signifies the average grade, ranging from 0 to 12, that an individual would have to have completed to be expected to comprehend the written text. The formula variables include total words, total sentence, and total syllables.

We also documented the Flesch-Kincaid reading ease test, which signifies how easy the written text is to read on a scale from 0 to 100. The higher the score, the easier it is to read. For example, a score of 90 to 100 suggests that an average 11-year-old student should understand the material; a score of 60 to 70 should be understood by a 13- to 15-year-old student, whereas scores of 0 to 30 are more difficult and would be understood by university graduates.

For each instrument, we measured readability by calculating the median Flesch-Kincaid grade level, the median Flesch-Kincaid reading ease test, the percentage of questionnaire items below an eighth-grade reading level, the mean number of characters per word, and the mean number of words per sentences.

RESULTS

We identified 26 instruments for UI, 5 instruments for POP, 10 instruments for FI, and 7 evaluating combined pelvic floor symptoms. Table 1 shows the median reading levels of the most common HRQOL instruments used to study PFDs. The r correlation value was found to be negative 0.8 between the Flesch-Kincaid reading ease and reading grade level.

The total number of questionnaire items per survey varied widely from 3 items in the Symptom Severity Index to 85 items in the Australian Pelvic Floor Questionnaire. Most of the instruments used a Likert-type response scale. Likert-type response scales ask an individual their level of agreement or disagreement on a series of statements. Usually, it is based on a 5- to 7-point scale, with responses ranging from strongly disagree, disagree, neutral, agree, to strongly agree.

For UI, the median Flesch-Kincaid grade reading level was 7.2. Of the 26 instruments, 9 (35%) were identified above an eighth-grade reading level. The median reading ease was 62.7. The outliers were the Incontinence Impact Questionnaire (IIQ) with a reading level of 11 and 0% of the content below the eighth-grade level, the IIQ-7 (Incontinence Impact Questionnaire Short Form) with a reading level of 11 and 0% of the content below the eighth-grade level, the Urological Distress Inventory (UDI), with a reading level of 10.5 and only 3% of the content below the eighth grade, the UDI-6 (Urological Distress Inventory Short Form), with a reading level of 8.4 and only 50% of the content below the eighth-grade level, the Urinary Incontinence Questionnaire (UIQ) with a reading level of 10 and only 13% of the content below the eighth-grade level, the UIQ-7 (Urinary Incontinence Questionnaire-7) with a reading level of 12 and 0% of the content below the eighth-grade level, the Urinary Incontinence Quality of Life Questionnaire (I-QOL) with a reading level of 8.9 and only 58% of the content below the eighth-grade level, the Stress-Related Leak, Emptying Ability, Anatomy, Protection, Inhibition, Quality of Life, Mobility and Mental Status (SEAPI-QMM) with a reading level of 9.4 and only 59% of the content below the eighth-grade level, the Urinary Incontinence Handicap Inventory (UIHI) with a reading level of 9.2 and only 60% of the content below the eighth-grade level, and finally the Urinary Incontinence Severity Score (UISS) with a reading level of 9.2 and only 60% of the content below the eighth-grade level.

For POP, the median Flesch-Kincaid reading level was 10.1, with 80% of questionnaires above the eighth-grade level. The median reading ease was 48.7. The outliers included the Pelvic Organ Prolapse Distress Inventory-6 (POPDI-6) with a reading level of 10, the Pelvic Organ Prolapse Distress Inventory (POPDI) with a reading level of 10.1 with 13% of the content written below an eighth-grade level, the Pelvic Organ Prolapse Impact Questionnaire (POPIQ) with a reading level of 12, and the Pelvic Organ Prolapse Impact Questionnaire-7 (POPIQ-7) with a reading level of 12. The POPQI, POPIQ-7, and the POPDI-6 all had zero items below the eighth-grade level.

For FI, the median Flesch-Kincaid reading level was 7.55, with 30% above the eighth-grade reading level. The median reading ease was 67.5. The outliers were the Fecal Incontinence Quality of Life-Scale (FIQOL-scale) with a reading level of 9.5 and only 22% below the eighth-grade level, colorectal-anal impact questionnaire (CRAIQ) with a reading level of

10.1 and 13% of the content below the eighth-grade level, and the colorectal-anal impact questionnaire-7 (CRAIQ-7) with a reading level of 12 and 0% of the content below the eighth-grade level.

For the combined pelvic floor symptom instruments, the median Flesch-Kincaid reading level was 9.7, with 86% written above the eighth-grade level. The outliers included the Pelvic Floor Distress Inventory (PFDI) with a reading level of 9.9 and only 24% of the content below the eighth-grade level, Pelvic Floor Impact Questionnaire (PFIQ) with a reading level of 12 and only 1% below the eighth-grade level, the Pelvic Floor Distress Inventory-20 (PFDI-20) with a reading level of 9.5 and only 35% of the content below the eighth-grade level, the Pelvic Floor Impact Questionnaire-7 (PFIQ-7) with a reading level of 12 and 0% of the content below the eighth-grade level, the Prolapse Incontinence Sexual Questionnaire-12 (PISQ-12) with a reading level of 9.2 and only 33% of the content below the eighth-grade level, and the Prolapse Incontinence Sexual Questionnaire-Long (PISQ-Long) with a reading level of 9.7 and only 26% of the content below the eighth-grade level.

Finally, it was found that many of the long forms had a lower reading level than their shorter versions. This was true for the UIQ (the urinary subscale of the PFIQ) and the UIQ-7, with a reading level of 10 versus 12, respectively. The same was true for the POPDI (9.7) versus the POPDI-6 (9.9), the POPIQ (11.5) versus the POPIQ-7 (12), and the CRAIQ (10) versus the CRAIQ-7 (12).

DISCUSSION

Our study had several findings. Of all 48 instruments studied, 54% were written at or below the recommended reading level. In fact, the majority (65%) of the incontinence questionnaires as well as the FI questionnaires (70%) were literacy appropriate. This is in contrast to the POP questionnaires with only 20% that were literacy appropriate. A possible explanation for the differences in literacy could be in the sentence length of these questionnaires. The average words per sentence for all incontinence questionnaires were only 10.5 versus 14.3 and 15.2 for FI and POP, respectively.

Another interesting finding is that some of the long forms had lower reading levels than their shorter form. The long form of the UIQ (the urinary subscale of the PFIQ) had a lower reading level than the shorter version, the UIQ-7. This can be explained by the longer version being easier to read (reading ease, 48.3 vs 28.6), in addition to its shorter sentence structure (13.1 vs 22.3). Although the shorter form used the same questions as the longer form, the specific questions used in the short form all had an individual reading level of 12 versus 6.7–12 in the long form. These slight variations between the long and short forms can account for the differences in reading levels as well as reading ease. The same was true for the POPDI, POPDI-6, POPIQ, POPIQ-7, CRAIQ, and CRAIQ-7 in that the long forms had lower reading levels than their corresponding short forms.

We also found that most of the POP and combined pelvic floor symptoms instruments were above the appropriate reading level, with a few of them requiring some college education for understanding. This finding is consistent with most of patient health information presented

to patients as well as medical information presented on the Internet requiring at least some high school to comprehend.^{11,12,17} As suggested by Bergman et al,¹³ HRQOL surveys are encouragingly more readable than most medical information given to patients, although only 54% of the questionnaires in our study were appropriate. Given this fact, we should as clinicians try to improve our patient teachings to better help them understand the information given to them. We also should consider choosing only a few HRQOL questionnaires as our standard to better standardize research data and facilitate better communication among colleges.

In 2008, SUFU carried out the Outcome Measures for Incontinence Treatment survey.¹⁸ According to this survey, most of SUFU members used the IIQ-7 and the UDI-6 as their preferred questionnaires when analyzing treatment outcomes.¹⁸ In our study, these 2 commonly used questionnaires had reading levels above the eighth grade (8.4 for UDI-6 and 11 for IIQ-7). Given that questionnaires may play an important role in individualizing treatment plans and outcomes measurements, patients are best served with literacy appropriate questionnaires. It is our impression that higher reading levels would be more difficult to read, especially for the underserved and less literate patient population. However, we did not administer questionnaires directly to patients to test this hypothesis and this could be an area of future study. We also realize the importance of a thorough history, physical examination, and follow-up to measure successful treatment outcomes and that the treatment of PFDs does not require administration of HRQOL instruments. However, many institutions and physicians involved in research use these as objective data tools to collect subjective data.

There were limitations to our study findings secondary to methodological considerations. First, the Flesch-Kincaid readability statistics were calculated through Microsoft Word software, but there are other methods of determining readability such as the Dale-Chall formula, the Gunning Fog formula, the Fry Readability Graph, the McLaughlin's SMOG formula, and the FORCAST formula. Using a different formula to calculate readability could potentially alter some of our results and findings. In addition, readability scores do not assess aspects of comprehensibility. Comprehension can only be inferred from the language and sentence structure of the items, which were not evaluated in this study. In addition, we were required to restructure our sentence fragments into complete sentences and add proper punctuation. Although this served the study purpose, this is not the traditional method of performing Flesch-Kincaid readability statistics and could have affected the accuracy of the study. Finally, we were unable to evaluate and quantify other factors that could have influenced comprehension such as graphics, layout, or learning stimulation.

Overall, most of the commonly used questionnaires addressing PFDs have reading levels that are appropriate for the average adult. However, some of the most frequently used questionnaires have high reading levels, and may not be ideal for all patients. As specialty societies focus on standardizing questionnaires for the purpose of research on PFDs, the reading level of such questionnaires should be considered. Questionnaires with an eighth-grade or lower reading level will be more generalizable to larger populations of women with PFDs.

Acknowledgments

This study was supported by the NIDDK (1 K23 DK080227-01, Dr Anger).

References

1. World Health Organization. [Accessed September 27, 2011] Definition of health. [WHO Web site] Available at: <https://apps.who.int/aboutwho/en/definition.html>
2. Haylen BT, de Ridder D, Freeman RM, et al. An International Urogynecological Association (IUGA)/International Continence Society (ICS) joint report on the terminology for female pelvic floor dysfunction. *Neurourol Urodyn*. 2010; 29:4–20. [PubMed: 19941278]
3. Weibe S, Guyatt G, Weaver B, et al. Comparative responsiveness of generic and specific quality-of-life instruments. *J Clin Epidemiol*. 2003; 56:52. [PubMed: 12589870]
4. Brubaker L, Cundiff G, Fine P, et al. A randomized trial of colpopexy and urinary reduction efforts (CARE): design and methods. *Control Clin Trials*. 2003; 24:629–642. [PubMed: 14500059]
5. Barber MD, Kuchibhatla MN, Pieper CF, et al. Psychometric evaluation of 2 comprehensive condition-specific quality of life instruments for women with pelvic floor disorders. *Am J Obstet Gynecol*. 2001; 185:1388–1395. [PubMed: 11744914]
6. Rogers RG, Kammerer-Doak D, Villarreal A, et al. A new instrument to measure sexual function in women with urinary incontinence or pelvic organ prolapse. *Am J Obstet Gynecol*. 2001; 184:552–558. [PubMed: 11262452]
7. Wyman JF, Harkins SW, Choi SC, et al. Psychosocial impact of urinary incontinence in women. *Obstet Gynecol*. 1987; 70:378–381. [PubMed: 3627585]
8. Shumaker SA, Wyman JF, Uebersax JS, et al. Health-related quality of life measures for women with urinary incontinence: the Incontinence Impact Questionnaire and the Urogenital Distress Inventory. Continence Program in Women (CPW) Research Group. *Qual Life Res*. 1994; 3:291–306. [PubMed: 7841963]
9. Rockwood TH, Church JM, Fleshman JW, et al. Fecal Incontinence Quality of Life Scale: quality of life instrument for patients with fecal incontinence. *Dis Colon Rectum*. 2000; 43:9–16. [PubMed: 10813117]
10. DeWalt DA, Berkman ND, Sheridan SL, et al. Literacy and health outcomes: a systematic review of the literature. *J Gen Intern Med*. 2004; 19:1228–1239. [PubMed: 15610334]
11. National Center for Education Statistics. [Assessed November 26, 2011] National Assessment of Adult Literacy. Available at http://nces.ed.gov/naal/kf_demographics.asp
12. Williamson JM, Martin AG. Analysis of patient information leaflets provided by a district general hospital by the Flesch and Flesch-Kincaid method. *In J Clin Pract*. 2010; 64:1824–1831.
13. Bergman J, Gore JL, Singer JS, et al. Readability of health related quality of life instruments in urology. *J Urol*. 2010; 183:1977–1981. [PubMed: 20303519]
14. Steinbrook R. Searching for the right search—reaching the medical literature. *N Engl J Med*. 2006; 354:4. [PubMed: 16394296]
15. Chall, JS. *Readability: An Appraisal of Research and Application*. Columbus, Ohio: Ohio State University Press; 1958.
16. Klare GR. A second look at the validity of the readability formulas. *J Read Behav*. 1976; 8:159–152.
17. Berland GK, Elliot MN, Morales LS, et al. Health information on the Internet: accessibility, quality, and readability in English and Spanish. *JAMA*. 2001; 285:2612. [PubMed: 11368735]
18. *Neurourol Urodyn; Abstracts from the 2008 SUFU (Society for Urodynamics and Female Urology) Annual Meeting; 2008; Miami, Florida, USA. 2008. p. 101*

TABLE 1

Readability of HRQOL Questionnaires for PFDs

Disease/Questionnaire	Median Reading Level (Range)	Items G8th-Grade Level, %	Median Reading Ease	Mean Words per Sentence	Mean Characters per Word
UI					
IIQ	11 (8.5–12.0)	0	47.8	19.6	4.8
IIQ-7	11 (9.5–12.0)	0	31.9	10.2	5.4
UIQ	10 (6.7–12.0)	13	50.5	13.1	5.4
UIQ-7	12 (12–12)	0	30	22.3	5.2
UDI	10.5 (7.7–12.0)	3	41	11.7	4.6
UDI-6	8.4 (4.8–12.0)	50	33.9	11.8	4.8
ICIQ-FLUTS	1.0 (0.5–9.3)	96	90.9	9.2	4.1
PRAFAB-Q	6.7 (2.2–11.9)	65	64.1	10.5	4.4
I-QOL	8.9 (4.3–12.0)	58	81.7	10.2	4.3
KHQ	4.5 (0.7–12.0)	84	77.9	7.2	4.6
BFLUTS	7.1 (2.2–12.0)	77	74.8	10.9	4.4
BFLUTS-SF	6.3 (3.2–12.0)	81	84.3	9.7	4.3
U-IIQ	4.1 (0.5–12.0)	86	57.5	10.1	4.2
U-UDI	4.9 (1.8–12.0)	83	57.2	9.2	4.9
SEAPI-QMM	9.4 (3.6–12.0)	59	18.4	13.8	5.2
SSI	3.6 (2.2–6.2)	100	61.6	9.8	4.4
SII	7.3 (4.9–10.8)	67	74.8	8.5	4.7
Contilife	5.6 (0.8–12.0)	73	63.8	11.3	4.6
EPIQ	7.5 (1.0–12.0)	78	58.3	10.5	4.2
MUDI	6.2 (0.5–12.0)	82	69.7	12.4	4.7
MUSIQ	7.7 (2.2–12.0)	80	52.1	7.9	4.3
UIHI	8.4 (3.1–12.0)	71	69.8	10.2	5.1
UISS	9.2 (2.8–12.0)	60	47.1	9.4	4.9
SIFUI	5.8 (0.8–10.7)	50	90.9	10.3	4.3
SUI-QOL	5.2 (2.4–12.0)	82	69.8	9.8	4.1
ICIQ	4.6 (0.7–8.7)	75	90.4	11.2	4.4
POP					

Disease/Questionnaire	Median Reading Level (Range)	Items G8th-Grade Level, %	Median Reading Ease	Mean Words per Sentence	Mean Characters per Word
POPDI-6	9.9 (8.7–12.0)	0	48.7	15.3	4.4
POPDI	9.7 (5.8–12.0)	13	53.6	11.5	4.3
POPDI-Q-7	12.0 (12.0–12.0)	0	25.7	22.3	5.2
POPDI-Q	11.45 (9.4–12.0)	0	44.6	19.1	4.9
P-QOL	6.7 (0.0–12.0)	74	76.6	14.5	4.4
FI					
FIQOL-Scale	9.2 (2.3–12.0)	22	58.4	17.2	4.2
FISI	7.9 (6.0–12.0)	64	60.7	10.1	4.5
Simple QOL	2.5 (0.8–6.7)	100	92.9	11.4	3.4
Q to Assess FI	7.3 (1.4–12.0)	62	76.7	17.4	4.3
CRAIQ-7	12.0 (12.0–12.0)	0	29.3	22.3	5.1
CRAIQ	10 (6.7–12.0)	13	50.4	11.5	4.3
CRADI-8	6.15 (2.4–9.3)	75	75.7	13	4.2
CRADI	8.1 (2.4–12.0)	47	61.6	11.8	4.2
Manchester	5 (0.0–9.1)	84	78.2	9.7	4.5
Modified Manchester	6.5 (3.8–9.5)	77	73.4	6.4	4.4
Combined pelvic floor symptoms					
Australian Pelvic Floor Q	4.8 (0.0–12.0)	80	78.2	9.1	4.3
PFDI	10.1 (4.1–12.0)	24	45.5	12.2	4.8
PFIQ	12.0 (5.8–12.0)	1	35.5	18.3	5
PFDI-20	9.6 (4.1–12.0)	35	58.2	12.5	4.8
PFIQ-7	12.0 (12.0–12.0)	0	28	22.2	5.1
PISQ-12	9.2 (5.6–12.0)	33	54.9	13.7	4.9
PISQ	9.5 (4.9–12.0)	26	46.4	13.3	4.9