Reliability of a German Questionnaire about General Practitioners' Handling of Female Urinary Incontinence

Elke Faust — Department of Obstetrics and Gynecology, Klinik Preetz, Am Krankenhaus 5, 24211 Preetz, Germany

Martina Juergensen — Institute for Social Medicine and Epidemiology, University of Luebeck, Ratzeburger Allee 160, 23562 Luebeck, Germany

Susanne Elsner — Institute for Social Medicine and Epidemiology, University of Luebeck, Ratzeburger Allee 160, 23562 Luebeck, Germany

Achim Niesel — Department of Obstetrics and Gynecology, Klinik Preetz, Am Krankenhaus 5, 24211 Preetz, Germany

Annika Waldmann[*] — Institute for Social Medicine and Epidemiology, University of Luebeck, Ratzeburger Allee 160, 23562 Luebeck, Germany

Abstract: Background: Recently our group performed a cross-sectional study in which 930 general practitioners (GP) in Germany and Denmark received a newly developed questionnaire concerning lower urinary tract symptoms. We developed the questionnaire on the basis of cognitive interviews with GPs and tested the reliability of the German version of the questionnaire in a test-retest process. Methods: 16 GPs took part in the test-retest process and completed the questionnaire twice with a time period of about four weeks between each attempt. The questionnaire consists of 28 questions. The given-answer categories and description fields sum up to a total of 60 items (answers). We assessed the reliability of answers by calculating and interpreting the absolute agreement and Cohen´s Kappa with a 95%-confidence interval for data that were nominal scaled and Pearson´s correlation coefficient for data that were interval scaled, respectively. Results: A total of 27 questions with 59 answer items were included in the analysis. Of them, 13 questions dealt with "management of UI", six questions addressed the "communication about UI", four questions asked for the "structure of the practice", and five questions assessed personal data of the GP. Each more than 50% of the items in the subject areas "management of UI" (53.1%), "communication about UI" (66.6%), "structure of the practice" (57%), and "personal data" (100%) were rated as having high reliability. In summary, 35 of the analyzed items were rated a having a high reliability and 22 items were rated as having a moderate reliability. Conclusion: Given the low number of study participants our results have to be interpreted with caution, but is seems that the developed questionnaire is – for the vast majority of items – a reliable tool for assessing the communication about and the management of female urinary incontinence during a general practitioners' consultation hour. Before application in future studies we recommend revising one item of the questionnaire in order to gain a higher reliability of this item.

Keywords: Questionnaire design; Reliability and validity; Methodology; Urinary incontinence.

1. Introduction

The prevalence of urinary incontinence [urinary incontinence (UI)] in women in Germany is said to range between 7% and 42% [1, 2] depending on the age and the type of the studied population as well as the assessment method. Although UI was found to increase with age in both sexes, women in general are affected two to four times more than men across most age-groups [3].

UI treatments include pelvic floor muscle training, bladder training, medication and/or surgery which can be prescribed or initiated by physicians. The German health care system offers patients a free choice as to whether they consult a general practitioner [general practitioner (GP)] or a specialist, such as gynecologists, uro-gynecologists or urologists [4]. UI seems to be a common topic in GP practices as a recent survey with 2,530 German GPs and gynecologists indicated. In that study nearly 60% of the GPs declared that they are often involved with patients

*Corresponding Author

presenting with UI during consultation hours. However, knowledge about the management of UI in GP practices including communication about the topic and factors which influence the treatment of UI in women is scarce [5].

The cross-sectional study „Lower Urinary Tract Symptoms (LUTS)" aims at adding to the scare knowledge about the communication between women with UI and their GPs and the management of UI in GP practices. It includes all 929 GPs practicing in the Fehmarnbelt region which covers one Danish district (Zealand) and three German districts (Luebeck, Ploen and Ostholstein). Due to the lack of validated instruments, the LUTS-study had to be conducted with a questionnaire which had to be developed by the study group [6] and had to be tested for test-retest reliability to assess the accordance of answers (stability) in repeated measurements.

The GP-questionnaire includes questions about frequency of the topic UI in GP's consultation hours, communication about UI symptoms, and factors influencing communication, diagnostic and therapeutic methods. The results of the repeatability testing are presented in this article.

2. Materials and Methods

The cross-national study „Lower Urinary Tract Symptoms (LUTS)" is supported by the European Interreg IVa program. It aims at adding to the scare knowledge about the communication between women with UI and their GPs and the management of UI in GP practices.

Ethics approval was given by the Medical Ethics Committee at the University of Luebeck.

2.1. Recruitment of Participants

In the first phase of the LUTS-study we aimed to test the test-retest reliability of a newly developed questionnaire for GPs (for more information on the development refer to Juergensen*, et al.* [6]. The participants were asked to answer the GP-questionnaire twice with an interval of four weeks between the two attempts.

Despite using several ways of motivating GPs to participate in our repeatability testing (personal contact, information via the local association of statutory health insurance physicians and advertisement in their newspaper, newsletter of the university, financial incentive for study participation), the recruitment of study participants was very difficult and time consuming. Within a given time limit, only 16 GPs took part in the repeatability testing and answered the questionnaire twice.

2.2. Questionnaire

The GP-questionnaire includes 28 questions (original questionnaire is available at http://www.sozmed.uni-luebeck.de/Forschung/FB+I+_+Sozialmedizin+und+Evidenzbasierte+Medizin/LUTS.html). Of them, 13 questions deal with the management of UI, six with communication about the topic, four with the structure of the GP practice and five with individual characteristics of the GP (e.g., age, sex, professional experience).

The answer options include five questions where GPs have to choose only one answer and three questions allowing for multiple answers, eight questions ask for estimations of proportions, five questions ask for years or numbers (e.g., age, number of patients), and nine visual analogue scales/Likert scales (range: 0 to 10). On the latter the lowest value (zero) indicated the lowest level of satisfaction, accordance or perceived difficulty, and the highest value (ten) indicated highest level of satisfaction, accordance or perceived difficulty depending on the question. The questionnaire does not allow calculating a summary score.

Taken altogether, the questionnaire includes 60 options/items to answer with one skipping sequence.

2.3. Statistical Analysis

For data analysis, the Statistical Package for Social Science SPSS, version 22, was used. Basic descriptive statistics were used to describe the study participants. For the analysis of the repeatability testing one item about the management of UI had to be excluded from the statistical analysis as it was answered only by two participants because of the skipping sequence instruction. Thus, the analysis included 27 questions with 59 items (answering categories) covering different thematic topics:

- Management of UI in the practice (32 items)
- Communication about UI (15 items)
- Structure of the practice (seven items)
- Personal data (five items)

Because data were scaled on nominal and interval scale, respectively, we used different approaches to assess the reliability. Pearson's correlation coefficient was calculated for data on an interval scale. The absolute accordance [7] and Cohen's Kappa with a 95% confidence interval [8] were calculated for data on a nominal scale.

The Pearson's correlation coefficient can reach values between +1 and −1. Where +1 means total positive correlation, 0 means no correlation, and −1 stands for a total negative correlation. Kappa values can range between +1 and -1. The value 1 implies the complete agreement between the two ratings, the value -1 means complete disagreement, and the value 0 implies that there is no relationship between the two ratings which cannot be explained by chance. Because Kappa cannot be calculated in asymmetric cross-tables and because the calculation of the absolute accordance is known to be more robust in small samples with low numbers of cases in the contingency tables, the final interpretation was based on the results of the absolute accordance and not on Cohen's Kappa.

The algorithm for rating the test-retest results as high, moderate or low reliable is shown in Table 1.

3. Results

3.1. Participants

16 GPs took part in the repeatability testing. Participants ranged in age between 37 and 60 years (Mean 49, SD 7): nine of them were females and seven were males. Three GPs worked in a single/solo practice, eight in a group practice with one additional physician, and five in a group practice with two or more additional physicians.

3.2. Repeatability Testing

A total of 38 out of the 59 items were nominally scaled. In these cases, absolute accordance and Cohen's Kappa was analyzed. In nine items it was not possible to evaluate the Kappa because of low numbers in some cells of the contingency tables (see Tables 2-5). The other 21 items were intervally scaled and Pearson's correlation coefficient was calculated.

A total of 13 questions (32 items) dealt with „management of UI". Of them, 17 items (53.1%) showed a high, 14 items (43.8%) a moderate, and one item (3.1%) a low reliability (Table 2). The item that was rated as low reliable was the question "Are you satisfied with the UI management that is offered by your practice?" which had to be answered on a scale from zero (not at all) to ten (very satisfied). A total of four GPs scored exactly the same score at both time points, while six GPs scored with a minor deviation of one point difference between the ratings. The remaining six scored the following combinations: time point one (t1)=2 and time point two (t2)=4, t1=3 and t2=5, t1=5 and t2=3, t1=6 and t2=3 (two GPS had this combination), and t1=7 and t2=3.

Six questions (15 items) covered the topic "communication about UI". Of them, ten items showed a high (66.6%), four items showed a moderate (26.6%), and one item (6.7%) showed a low reliability (Table 3). The item that was rated as low reliable was the answer option "lack of suitable situations or special occasions" to the superordinate question in which GPs were asked for potential barriers to addressing the topic of UI during the consultation hour.

In the subject area "structure of practice" (four questions, seven items) four items (57%) showed a high and three items (43%) a moderate reliability (Table 4).

All five questions/items about personal data of the participants GP (e.g., age, sex, professional experience) showed a high reliability (Table 5).

In summary, 35 items (59.3%) showed a high reliability, 22 items (37.3%) showed a moderate reliability, and two items (3.4%) showed a low reliability.

4. Discussion

The data presented in this article are part of the Danish-German cross-sectional "LUTS-study" which consists of a survey with 8,000 women living and a survey with 930 GPs practicing in the Fehmarnbelt region. The surveys use two newly developed questionnaires – one for women and one for GPs. The GP-questionnaire was developed on the basis of one focus groups interview and eleven one-to-one interviews, first tested in cognitive interviews in Germany and Denmark [6], and after that tested in Germany to verify the accordance of answers in repeated measurement (repeatability).

4.1. Repeatability Testing

Altogether, the results of the repeatability testing indicate a high reliability for the vast majority of items – with only two of the 59 tested items being rated as having a low reliability. These are the question "Are you satisfied with the UI management that is offered by your practice?" and the answer option "lack of suitable situations or special occasions" to the superordinate question on potential barriers to addressing the topic of UI during the consultation hour.

Maybe the low reliability for the latter item ("lack of suitable situations or special occasions") is due to the fact that this item is unclear or vague verbalized, and, therefore, prone to different interpretations. However, during the cognitive interviews with GPs before the repeatability testing, no major or minor problems were reported for this item. Another reason – and in our view the more important reason – for the low reliability of this item may be a study effect. Participants had to answer the questionnaire twice and it can be assumed that having answered the questionnaire the first time had an effect upon GP's awareness and perception afterwards. This might have changed the interpretation of "adequate situations or occasions" at the second time point. However, we do not recommend rewording of the item before application in future studies as study participants in upcoming studies will answer this items only once.

In contrast, we recommend revising the second item ("How satisfied are you with the UI management that is offered by your practice?"). First, during the cognitive interviews with GPs no major problems regarding the words used in this question or the general understandability of this question were reported. But GPs working in group practices mentioned that it was unclear whether their own or the usual management in their practice was meant. Thus, an explanatory sentence for GPs working in group practices may facilitate a better understanding of this item. Second, GPs were asked to rate their satisfaction on a scale with a range from zero (not at all) to ten (very satisfied). At both time points of the repeatability testing ten of the 16 GPs scored either the value two or three, and the remaining six GPs scored values in the range from four to seven. A reduction of the 11-point scale to a 5-point or 7-point scale might help reducing these deviations. Another explanation for the low reliability of this item – again – may be seen in a study effect as GPs had to answer the questionnaire twice. It is possible that the study participation

prompted the GPs to think about their mode and their opportunities to counsel and manage women with UI. This might have changed their rating of "their satisfaction with UI management" offered by their practice.

Aside from these two items, each more than 50% of the items in the subject areas "management of UI" (53.1%), "communication about UI" (66.6%), "structure of the practice" (57%) were rated as having high reliability and the others were rated as having a moderate reliability.

4.2. Recruitment of GPs

Like many other study teams we experienced problems to recruit physicians for this health care research project [9]. The plan to involve 50 GPs in the repeatability testing failed because of the low motivation of GPs to participate and limited time resources. Therefore, an important limitation of the study is the small sample size, which restricts statistical options and explanatory power.

Especially during the cognitive interviews with GPs, it became obvious that UI as compared to other diseases is a topic of only minor interest to many GPs, although they are quite aware of the high frequency of UI among their patients.

5. Conclusion

In the course of a bi-national, cross-sectional study we developed a questionnaire about general practitioners' handling of female urinary incontinence. Due to the small sample size of only 16 GPs our results should be interpreted with caution. Nevertheless, it seems that the GP-questionnaire is a reliable instrument which can be used to assess the communication about and the management of UI in practices of general physicians. Before application in future studies we commend to revise one item in order to increase its reliability.

Acknowledgements

Conflicts of Interest

The authors declare that they have no competing interests.

References

[1]   Beutel, M. E., Hessel, A., Schwarz, R., and Braehler, E., 2005. "Prevalence of urinary incontinence in the German population." *Urologe A,* vol. 44, pp. 232-238.

[2]   Hunskaar, S., Lose, G., Sykes, D., and Voss, S., 2004. "The prevalence of urinary incontinence in women in four European countries." *BJU Int,* vol. 93, pp. 324-330.

[3]   Niederstadt, C. and Gaber, E. l., 2007. "Gesundheitsberichterstattung des Bundes. In: Robert Koch Institut, ed. Statistisches Bundesamt." *Heft 39,* Available: http://www.rki.de/DE/Content/Gesundheitsmonitoring/Gesundheitsberichterstattung/GBEDownloadsT/harn inkontinenz.pdf?__blob=publicationFile

[4]   Monz, B., Hampel, C., Porkess, S., Wagg, A., Pons, M. E., Samsioe, G., Eliasson, T., Chartier-Kastler, E., Sykes, D*., et al.*, 2005. "A description of health care provision and access to treatment for women with urinary incontinence in Europe -- a five-country comparison." *Maturitas,* vol. 52, pp. S3-S12.

[5]   Wiedemann, A. and Fuesgen, I., 2009. "Therapy for urinary incontinence in general practice." *Aktuel Urol,* vol. 40, pp. 242-246.

[6]   Juergensen, M., Elsner, S., Schreiber, P. L., Faust, E., Kold, M. L., Lueckert, J., Niesel, A., Rudnicki, M. P., and Waldmann, A., 2016. "Development of a bilingual questionnaire about general practitioners' handling of female urinary incontinence." *Int J Healthcare Med Science,* vol. 2, pp. 20-27.

[7]   Armitage, P., Berry, G., and Matthews, J., 2001. *Statistical methods in medical research*. 4th ed. Oxford: Blackwell Publishing.

[8]   Altman, D. G., 1991. *Practical statistics for medical research*. London: Chapman and Hall.

[9]   Asch, S., Connor, S. E., Hamilton, E. G., and Fox, S. A., 2000. "Problems in recruiting community-based physicians for health services research." *J Gen Intern Med,* vol. 15, pp. 591-599.

Tables

Table-1. Algorithm of rating the repeatability/reliability

Absolute Agreement (%)	Pearson's Correlation Coefficient	Reliability
<60%	<0.3	low
60-85%	0.30-0.70	moderate
>85%	>0.70	high

Table-2. Reliability of answers in the subject area "Management of UI"

Items (bullet point style)	Absolute agreement (%)	Kappa [95% confidence interval]	Pearson's correlation coefficient	Interpretation
Who should be mainly responsible for the management of UI?	77%	0.61 [0.29; 0.93]		moderate reliability
Is the management of UI influenced by written guidelines from medical associations?	75%	-0.10 [-0.27; 0.06]		moderate reliability
Reasons interfering with an adequate management of women with UI in the practice *(Multiple ticks are allowed)*				
additional illnesses which are more important	75%	0.51 [0.1; 0.91]		moderate reliability
bad compliance	87.5%	0.59 [0.08; 1]		high reliability
diagnostics too time-consuming	100%	n.c.		high reliability
treatment options are insufficient	93.7%	0.77 [0.33; 1]		high reliability
I feel unsure of UI management	68.8%	0.36 [-0.11; 0.82]		moderate reliability
many patients to take care of	62.4%	0.20 [0.27; 0.67]		moderate reliability
high workload	87.5%	0.71 [0.36; 1]		high reliability
monetary reasons	100%	n.c.		high reliability
other reasons	100%	n.c.		high reliability
nothing interferes with an adequate management of UI	87.4%	0.43 [-0.23; 1]		high reliability
Opinion about the Urinary Incontinence guidelines? *(Multiple ticks are allowed)*				
Unavailable	85.7%	0.44 [-0.16; 1]		high reliability
good availability	100%	n.c.		high reliability
too long	84.6%	0.44 [-0.16; 1]		high reliability
not relevant	100%	n.c.		high reliability
support counseling of patients	75%	0.44 [-0.16; 1]		high reliability
helpful for clinical decision making	92.3%	0.76 [0.31; 1]		high reliability
they do not apply to all women	69.2%	0.13 [-0.45; 0.72]		moderate reliability
I do not know them in detail	69.2%	0.13 [-0.45; 0.72]		moderate reliability
other opinion	84.6%	-0.08 [-0.20; 0.03]		moderate reliability
How useful are the guidelines for you?	*n.c. [skipping sequence]*			
Suitability of micturation diaries			0.525	moderate reliability

Suitability of pelvic floor exercises			0.785	high reliability
Suitability of medication			0.522	moderate reliability
Frequency of prescription of micturation diaries				
In per cent			0.899	high reliability
Answer: not applicable	81.3%	0.46 [-0.05; 0.97]		moderate reliability
Frequency of prescription of physiotherapy				
In per cent			0.579	moderate reliability
Answer: not applicable	93.7%	0.77 [0.33; 1]		high reliability
Frequency of prescription of medication				
In per cent			0.625	moderate reliability
Answer: not applicable	87.5%	0.6 [0.13; 1]		high reliability
Satisfaction with the UI management offered by the practice?			0.243	low reliability
How satisfied do you think your patients are with the management of UI offered by your practice?			0.426	moderate reliability

n.c.= not calculated

Table-3. Reliability of answers in the subject area "Communication about UI"

Items (bullet point style)	Absolute agreement (%)	Kappa [95% confidence interval]	Pearson's correlation coefficient	Interpretation
Do you need a special occasion for asking female patients about UI?	87.5%	-0.03 [-1; 0.46]		high reliability
Who should be mainly responsible for raising the topic of UI during the consultation hour?	62.5%	0.36 [-0.04; 0.76]		moderate reliability
Which reasons might hinder you from asking women about UI during the consultation hour? (*Multiple ticks are allowed*)				
time-related reasons	87.5%	0.71 [0.36; 1]		high reliability
monetary reasons	100%	n.c.		high reliability
it is not in my routine to ask about UI	87.5%	0.71 [0.33; 1]		high reliability
lack of suitable situations or special occasions	56.2%	0.13 [-0.36; 0.61]		low reliability
I feel uncomfortable asking about UI	93.7%	0.64 [-0.01; 1]		high reliability
women may feel uncomfortable being asked about UI	87.5%	0.71 [0.36; 1]		high reliability
new patient	100%	n.c.		high reliability
women do not have to talk about everything	87.5%	0.59 [0.08; 1]		high reliability
other reason	100%	n.c.		high reliability
nothing hinders me	81.2%	0.29 [-0.30; 0.89]		high reliability

			0.594	moderate reliability
How uncomfortable do you feel asking women about UI?				
Do you think women feel uncomfortable raising the issue of UI during the GP's consultation hour in general?			0.426	moderate reliability
Is Urinary Incontinence among women a frequent issue that you have to deal with in your daily practice?			0.533	moderate reliability

n.c.= not calculated

Table-4. Reliability of answers in the subject area "Structure of the practice"

Items (bullet point style)	Absolute agreement (%)	Kappa [95% confidence interval]	Pearson's correlation coefficient	Interpretation
Frequency of women with UI symptoms in the practice			0.648	moderate reliability
Type of practice	93.8%	0.90 [0.70;1]		high reliability
Age distribution of patients (in per cent)				
younger than 30 yrs			0.715	high reliability
30-60 yrs			0.727	high reliability
older than 60 yrs			0.375	moderate reliability
Number of patients				
total number			0.894	high reliability
proportion of female patients			0.661	moderate reliability

Table-5. Reliability of answers in the subject area "Personal data"

Items (bullet point style)	Absolute agreement (%)	Kappa [95% confidence interval]	Pearson's correlation coefficient	Interpretation
Year: License to practice medicine			0.989	high reliability
Year: Start working as GP			0.999	high reliability
Sex	100%	n.c.		high reliability
Age in years			0.999	high reliability
Postal code (5 digits)	93.7%	0.86 [0.67; 1]		high reliability

n.c.= not calculated