Adaptation and validation of the Michigan Incontinence Severity Index in a Turkish population

Mehmet Akif Sargin
Murat Yassa
Bilge Dogan Taymur
Emrah Ergun
Gizem Akca
Niyazi Tug

Department of Obstetrics and Gynecology, Fatih Sultan Mehmet Training and Research Hospital, Istanbul, Turkey

Objective: To translate and validate the Michigan Incontinence Severity Index (M-ISI) for its use in Turkish-speaking women with urinary incontinence.

Methods: The translation and cross-cultural adaptation were based on international guidelines. Content validity by content validity ratio/content validity index, internal consistency by Cronbach’s alpha, test–retest reliability by Pearson’s correlation, and construct validity by using Spearman rank correlations to show the relationship between individual items and the relevant domains and subdomains were analyzed in 100 female participants with a chief complaint of urinary incontinence. Correlations between the relevant scores of M-ISI and The International Consultation on Incontinence Questionnaire – Short Form scores were analyzed to indicate convergent validity. The Varimax rotation method was used to conduct exploratory factor analysis in order to investigate the factor structures/distribution of M-ISI items.

Results: Content validity index and content validity ratio values increased to 0.97 and 1.00, respectively, showing sufficient content validity of the Turkish version of the M-ISI. The analysis formed three factors which was slightly different from original developers. In our proposed three-factor construct, all of the ten items demonstrated high correlations with their subdomains and lower correlations with the other domains, indicating good construct validity. Correlations between stress urinary incontinence and urge urinary incontinence (UUI) scores and The International Consultation on Incontinence Questionnaire – Short Form scores were found high, which indicated convergent validity ($r$: 0.953, $P<0.001$). Good internal consistency of the scores for each subdomain was observed (stress urinary incontinence, 0.787; UUI, 0.862; pad usage and bother, 0.832). Test–retest reliability was shown for each subdomain (stress urinary incontinence, 0.973; UUI, 0.973; pad usage and bother, 0.979).

Conclusion: The translated and cross-culturally adapted M-ISI showed good validity, reproducibility, and reliability that allow its use in Turkish-speaking populations with urinary incontinence. Its comprehensive structure means that it has become a practical instrument that is available for utilization in the primary health care setting, clinical research, and epidemiological trials in Turkey.

Keywords: bother, questionnaires, pad use, reliability and validity, translations, urinary incontinence

Introduction

Urinary incontinence (UI) is a very common complaint and results in a great deal of distress and embarrassment as well as significant costs to both individuals and the health system. Estimates of prevalence vary according to the definition of incontinence, different survey methods, response options, and the population studied. Estimates of prevalence of urgency urinary incontinence (UUI) also vary, ranging from 1.8% to 30.5% in European populations. Recent publications have highlighted the substantial economic burden of UI, and it is thought that this will increase over time, in parallel...
with the projected 25% increase in UI in the next decade, as a result of population aging.\textsuperscript{2} Since there is universal agreement regarding the importance of the problem of UI in terms of human suffering and high economic cost, it is essential to use a validated and appropriate questionnaire when standardized assessment is required in order to reduce the costs and time spent.

Patients usually do not reveal their UI problem unless it is questioned or it substantially bothers them. The use of self-report questionnaires can increase the rate of UI diagnosis; however, these questionnaires may not always be comprehensive and may not aid in differentiating the type of incontinence. Very few questionnaires, such as the Questionnaire for Urinary Incontinence Diagnosis and the 3-Incontinence Questions, have the potential to discriminate types of UI that occur in women.\textsuperscript{3,4} It is recommended that clinicians evaluate existing tools, used either alone or in combination, in the assessment and monitoring of treatment outcome.\textsuperscript{1,5} At present, there is no one questionnaire that fulfills all requirements for evaluating individuals with UI.\textsuperscript{1} However, combining existing measurements may cause difficulties for patients in the daily clinical setting.

The Michigan Incontinence Severity Index (M-ISI), which Suskind et al\textsuperscript{9} have recently developed, is prominent among other questionnaires as it aids clinicians and researchers in discerning type, severity, and bother related to UI, and also addresses the patients’ pad usage (PU). The developers also suggested that the M-ISI can be used with high sensitivity and specificity in the screening of clinically relevant UI in women.\textsuperscript{7} Therefore, we aimed to adapt and validate this questionnaire for use in a sample of Turkish women with UI.

**Methods**

**Cross-cultural adaptation procedure**

The recommendations of the translation and cultural adaptation group were followed for the development and cross-cultural adaptation of the Turkish version of the M-ISI.\textsuperscript{6-10} Permission to translate the M-ISI into Turkish and then validate this instrument was obtained from the developers of the original questionnaire on September 21, 2015. The sequence of steps used was as follows: forward-translation, synthesis, back-translation, expert committee review, pretesting and cognitive debriefing, proofreading, and final report and validation of psychometric properties.

Two experts in the original language of the M-ISI, English (Translator 1 and Translator 2), were aware of the concepts under examination in the questionnaire. The adaptation of Translator 1 was intended to provide reliable equivalence from a more clinical perspective, in terms of measurement. Translator 2 was naive to the concepts being quantified and had no medical or clinical background. It was aimed that Translator 2 would detect differences in meaning between the original text and the first translation under less influence of academic objectives and to provide a translation that reflects the language used by the target population, highlighting any ambiguous meanings in the original index. The two versions were compared, and a consensus on synthesis translation was reached by both translators to form a preliminary version. A panel of two forward translators and two bilingual authors critically reviewed the translation to form the reconciled forward-translation draft of the Turkish version of the M-ISI. The literal and conceptual back-translations were conducted by two outsourced bilingual translators who had no association with the research group, have no medical background, and had no knowledge of the content of the M-ISI and its first translations, in order to elicit unexpected meanings of the items in the translated index. This version was shared with the developers of the M-ISI, as recommended by Beaton et al.\textsuperscript{9}

Discrepancies among the original, forward-translation version, and the back-translated versions were discussed by a committee of experts composed of seven professionals (three gynecologists, one urologist, one postgraduate nurse working in the UI setting, and one language professional). Only minor revisions were required to eliminate the discrepancies and the previous versions were thus reconciled to produce a final Turkish version of the M-ISI.

A total of ten individuals were involved in the pretesting and cognitive debriefing stage in order to test alternative wording and check the understandability, interpretation, and cultural relevance of the translation. The respondents were native speakers who adequately represented the target population in terms of age, education, and diagnosis.\textsuperscript{9,10} Less than 5 minutes were required for self-administration of the questionnaire.

The cognitive debriefing was assessed, and the content validity was graded by the expert committee, with an analysis of the relevance of each item of the M-ISI. Questions that had a misunderstanding rate of 20% or higher were reviewed.\textsuperscript{9} One item (the fourth question) was reworded according to respondents’ comments, justifying such a change. Face validity measured the researchers’ and patients’ comprehension and acceptance of items of the pretest sample.\textsuperscript{8,9} Proofreading of the final translation was carried out, and no errors were found. The final report, which clearly explains the reasons for all translation decisions and wording choices for cultural adaptation, was written by the head of the expert committee.
in order to inform other future translations of the same instrument, such that they can be harmonized with the previously developed language versions.10

Patients who took part in the pretest and posttest stages were asked to return to the hospital after 14 days for readministration of the M-ISI. Reproducibility was assessed in 20 individuals.8–10

Study design and population

The study was conducted at the Department of Obstetrics and Gynecology of Fatih Sultan Mehmet Training and Research Hospital, Istanbul, Turkey, between October 2015 and January 2016. The M-ISI was administered to 100 female participants who were selected from women who attended the outpatient clinic of gynecology with a chief complaint of UI. Participants were prescreened to ensure that their UI was not due to a secondary condition, such as pregnancy or chronic urinary tract infection. Women with a history of any psychiatric condition, multiple sclerosis, spinal cord injury, congenital or acquired anatomic abnormalities of the urogenital tract, bladder malignancy, or age <18 were excluded. These criteria were similar to the exclusion criteria used for the development of the original M-ISI. The Istanbul Fatih Sultan Mehmet Training and Research Hospital, Ethics Committee of Clinical Researches approved the study (approval number: FSMEAHA-KAEK 2015/52), and written informed consent was obtained from all of the participants.

Instruments

M-ISI

The M-ISI used has proven reliability and validity. It has ten items, consisting of a total M-ISI domain (the sum of items 1–8) and a distinct Bother domain (the sum of items 9 and 10). The total M-ISI score consists of three subdomains (items 1–3 for stress urinary incontinence [SUI], items 4–6 for UUI, and items 7 and 8 for PU). The responses for each item range from 0 to 4 on a Likert-type scale, with higher values representing greater symptoms and greater bother. The total M-ISI domain ranges from 0 to 32, the Bother domain ranges from 0 to 8, the SUI and UUI subdomains range from 0 to 12, and the PU subdomain ranges from 0 to 8. Total domain and subdomain scores are obtained by simply adding the respective answers. The minimally important difference has been determined for the following domains/subdomains: total M-ISI (4 points), SUI (2 points), UUI (2 points), and PU (1 point).6 The original developers of the M-ISI questionnaire found that the sensitivity and specificity for stress, urgency, and total UI were 77% and 73%, 86% and 76%, and 84% and 75%, respectively.7

ICIQ-SF

The International Consultation on Incontinence Questionnaire—Short Form (ICIQ-SF) is a self-administered questionnaire designed for evaluating individuals with UI.11 It includes three scored items that evaluate urinary frequency and leakage, as well as perceived impact on patients’ daily life, with scores ranging from 0 (low bother) to 21 (maximum bother). Cetinel et al12 validated the Turkish version of the ICIQ-SF.

Statistical analysis

Number Cruncher Statistical System 2007 (Kaysville, UT, USA) software was used for statistical analyses. All clinical parameters were summarized by descriptive statistics. Data quality was assessed by mean, standard deviation, ranges, and percentages of patients scoring minimum (floor) and maximum (ceiling) possible M-ISI scores. The content/face validity, which indicates whether the questionnaire makes sense to the patients and experts and whether all of the important and relevant domains are included, was assessed through expert committee decision via analysis of the content validity ratio (CVR) and content validity index (CVI) before and after the changes. Internal consistency was assessed by Cronbach’s alpha, and a value of between 0.70 and 0.95 was considered to reflect good internal consistency. Pearson’s correlation test was used to verify the test–retest reliability. Convergent validity was confirmed by the range of correlation coefficients between the relevant items belonging to the severity domain of the M-ISI and the items of the ICIQ-SF, which is a widely used and validated instrument. The Varimax rotation method was used to conduct exploratory factor analysis in order to investigate the factor structures/distribution of M-ISI items. Construct validity was established using Spearman rank correlations to show the relationship between individual items and the relevant domains and subdomains. The significance level was set at P<0.05.

Results

The demographic characteristics of the participants are shown in Table 1. For good content validity, it was determined that the CVR must be over 0.99, and the CVI must be above 0.80, based on the ratings of item clarity and relevance provided by the seven experts on our committee. The CVI and CVR were 0.89 and 0.14, respectively, as assessed by these experts. A correction was made to the fourth question, so the CVI and CVR values increased to 0.97 and 1.00, respectively, showing sufficient content validity of the Turkish version of the M-ISI.
Floor or ceiling effects are considered to be present if >15% of respondents achieve the lowest or highest possible scores. None of the respondents in the present study had scores of 0 (floor) or 40 (ceiling), indicating that no floor and ceiling effect was present.

The Kaiser–Meyer–Olkin value was high at 0.825, and the Bartlett’s Test of sphericity was significant (P<0.001), confirming the suitability of using exploratory factor analysis. The analysis formed three factors: the first included the three originally proposed SUI items, the second factor included three UUI items, and the third factor consisted of the two originally proposed PU items and two Bother items, with the latter four items being mixed together to form a new domain. The factor loadings were 0.848, 0.803, and 0.782, respectively, for the three SUI items; 0.835, 0.855, and 0.873, respectively, for the three UUI items; and 0.734, 0.542, 0.869, and 0.769, respectively, for the PU and Bother items, as shown in Table 2.

Table 2 The exploratory factor analysis

<table>
<thead>
<tr>
<th>Questions</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1</td>
<td>0.848</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 2</td>
<td>0.803</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 3</td>
<td>0.782</td>
<td>0.835</td>
<td></td>
</tr>
<tr>
<td>Item 4</td>
<td>0.855</td>
<td>0.873</td>
<td></td>
</tr>
<tr>
<td>Item 5</td>
<td>0.873</td>
<td>0.734</td>
<td></td>
</tr>
<tr>
<td>Item 6</td>
<td>0.542</td>
<td>0.869</td>
<td>0.769</td>
</tr>
</tbody>
</table>

Notes: Factor 1: Originally named SUI, Factor 2: Originally named UUI, Factor 3: Combination of originally named pad usage and bother subdomains. Items 1–3 define the questions belonging to the SUI domain. Items 4–6 define the questions belonging to the UUI domain. Items 1 and 10 define the questions belonging to the bother domain.

Abbreviations: SUI, stress urinary incontinence; UUI, urgency urinary incontinence.

Use of the four-factorated structure proposed in the original article revealed high correlations between individual items of the SUI and UUI subdomains with their respective subdomains, and lower correlations with the other domains, indicating sufficient construct validity; however, PU subdomain items showed high correlations with the Bother domain and vice versa. In our proposed three-factor construct, all of the ten items demonstrated high correlations with their subdomains and lower correlations with the other domains, indicating good construct validity (Table 3).

Correlations between the SUI, UUI, and SUI + UUI scores and the ICIQ-SF scores were high, indicating convergent validity, as shown in Table 4.

Internal consistency was assessed using Cronbach’s alpha, and high values were obtained for each subdomain (SUI, 0.787; UUI, 0.862; PU and Bother, 0.832), indicating good consistency of the scores.

Test–retest reliability coefficients were calculated for each subdomain: SUI (0.973), UUI (0.973), and PU and Bother (0.979), indicating excellent reliability of the scores.

The successfully adapted and validated Turkish version of M-ISI is shown in Table S1.

Table 4 Correlations between the SUI, UUI, and SUI + UUI scores and ICIQ-SF scores for convergent validity

<table>
<thead>
<tr>
<th>Scores</th>
<th>ICIQ-SF</th>
<th></th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUI</td>
<td>0.664</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>UUI</td>
<td>0.751</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>SUI + UUI</td>
<td>0.953</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Pad use</td>
<td>0.766</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Severity</td>
<td>0.946</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Bother</td>
<td>0.636</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>M-ISI</td>
<td>0.929</td>
<td>&lt;0.001</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: ICIQ-SF, Incontinence Questionnaire—Short Form; M-ISI, Michigan Incontinence Severity Index; SUI, stress urinary incontinence; UUI, urgency urinary incontinence.
Discussion

The present study aimed to translate, cross-culturally adapt, and validate the M-ISI instrument for the Turkish population by analyzing content validity, reliability, and construct validity and exploring the factor structure of the M-ISI.

The costs of routine and nursing home care for UUI are counted as major contributors to overall health costs, and early diagnosis and correct management are important. The establishment of public health and clinical management programs is recommended to improve patient and clinician awareness of UUI. The most common method used to show the presence of UI is patient selection by patient-reported outcome (PRO) instruments. Patients benefit from simple interventions, such as modification of lifestyle factors, behavioral changes, and bladder and pelvic floor muscle training, which can all be easily managed in the primary care setting. It is widely known that implementation of these factors may improve UI. Moreover, primary care physicians often encounter common problems in Turkey, including limited time to spend with a patient, the high cost of advanced diagnostic tools, and limited access to such tools, and the need to determine whether the UI of a patient is clinically relevant. Shaw et al found that very few women with urinary leakage seek care for treatment of their UI due to lack of knowledge of their condition, lack of available treatments, feelings of embarrassment, and misconceptions around whether they are faced with a medical problem, and they primarily consider UI as part of the normal aging process or something that is a normal consequence of childbirth. Svihra et al found that the most frequent barriers to seeking care were patients’ fear of health care practitioners and long waiting times for appointments. They also found that one of the most important determinants of the probability of seeking health care was type and duration of UI; those patients with severe UI are less likely to seek care. Acedokun et al found that the most common reasons for not accessing health care services to help with UI were a belief that UI is not treatable, not life-threatening, and a normal condition. Therefore, a validated tool that can rapidly identify women with clinically relevant UI and can aid physicians in referring patients for available treatments should be effective in the primary health care setting.

The M-ISI is a beneficial tool for screening patients to determine the presence of clinically relevant UI. The original developers of the M-ISI demonstrated high sensitivity and specificity for stress, urgency, and total UI among women aged 35–64 years. This instrument was selected for validation in the present study as a result of it being an ideal, simple questionnaire for both clinical and research purposes.

It is believed that enrolling patients without UI in studies evaluating the validity and reliability of questionnaires regarding UI limits the scope of these PRO instruments when it comes to administration in patients with UI. Thus, we included only patients with complaints of urinary leakage in the present study.

Liberal usage of urodynamics appears to be a more objective and quantitative method compared to the use of questionnaires and diaries; however, it is not routinely used, due to its technical difficulties and relatively high cost. In addition, it has been frequently shown that clinical diagnosis and urodynamic findings do not correlate, and women without any complaints may have urodynamic abnormalities. The benefits of urodynamic evaluation in avoiding incontinence surgery have been shown in a Cochrane review; however, no evidence was found with regard to aiding clinical decision making in order to change the clinical outcome of treatment. The variability of the findings, even within the same session, limits the clinical usefulness of routine use of urodynamics.

PROs that are validated for the language in which they are being used are more simple, cheap, and effective methods to assess treatment success and early diagnosis. While urodynamic studies were included in a baseline evaluation by the original developers, we preferred not to include this method in our study, for the aforementioned reasons.

Quantitative pad testing can be used to measure the amount of urine loss over a set period of time, or during physical exercise, as well as to assess the severity and presence of UI and treatment responsiveness. It is known that a proper pad test can diagnose UI with high accuracy; however, Al Afraa et al addressed the conflicting data regarding the usefulness of pad tests in quantifying the severity of incontinence. We believe that it is not a practical method with regard to the actual daily practice of an outpatient gynecology clinic, where a longer amount of time and a standardized protocol for the pad test are required. We consider that using PRO instruments, such as the M-ISI, which addresses both PU and bother related to UI, is easier and more accessible than pad weight tests with a similar clinical contribution.

The Questionnaire for Urinary Incontinence Diagnosis is responsive to change and can discriminate the type of UI; however, it does not include items of bother and does not address PU. The ICIQ-SF is another short questionnaire that evaluates the prevalence and frequency of UI and its effect on everyday life; however, it also does not address PU and does not reliably distinguish the type of UI. The Incontinence Screening Questionnaire was developed to successfully identify the presence of UI among the general population in the primary care setting and includes 48-hour pad weights in
preference to clinical diagnosis or patient bother; however, it may potentially identify UI for which treatment is not necessary.\textsuperscript{22} We believe that combining two or more long questionnaires to achieve an improved understanding of a patient’s incontinence often results in confusing patients with similar questions and may result in careless or imprecise reporting with attempts to finish in a hurry.

The results of our factor analysis showed that the PU and Bother domains of the Turkish version of the M-ISI did not load on the original factor structure. Although the original developers had maintained four factors from the items of their respective version, our exploratory factor analysis formed three domains from the Turkish version of the M-ISI. As highlighted in the results, our proposed three-factor construct showed high correlations with subdomains and lower correlations with the other domains for all of the ten items, which indicates good construct validity. We think this is plausible because the questions around PU were closely associated with impacts on the daily activities of the Turkish population, such as the ninth item “Overall, how often have you needed to change your daily activities because of your urinary incontinence?” We hypothesize that the PU subdomain may be placed under the Bother domain, rather than the total M-ISI domain.

Botelho et al\textsuperscript{23} found that urologic symptom severity and bother did not directly correspond in their qualitative study. They observed that varying perceptions of symptoms (causal, relative, and uncertainty) might appease the bother of the patients. Bother was higher for patients who viewed their symptoms with uncertainty, in situations where the etiology was unknown. Therefore, we think that questionnaire that assess the bother, as well as the severity, of UI have a substantial advantage over other questionnaires. Moreover, we consider that using a questionnaire, such as the M-ISI in the primary care setting, can aid physicians in identifying patients with low urologic symptom severity, but with high bother, due to perception of uncertainty, and in referring them for secondary health care. The high correlation between the PU subdomain and the Bother domain that we observed may be due to the low admission rate to health care units among patients with urinary complaints in Turkey. A low hospital admission rate may result in an elevated perception of uncertainty, and therefore increased bother, among women who already have increased PU to manage their urinary leakage.

Teunissen et al\textsuperscript{24} observed that patients seek help regarding their UI in cases of increased severity of incontinence together with distress, primarily because of the presence of comorbidity. They stated that urine loss was not the main concern of older patients. In Turkey, older patients tend to wait until they develop worrying comorbidities before seeking medical help. We think these cultural and socioeconomic considerations that differ in Turkish population compared to others may contribute to explaining the high correlation between the PU and Bother domains in the present study. Evaluation of the burden of UI from the patient’s perspective and its impact on a patient’s quality of life, using versatile instruments, such as the M-ISI, is essential for management.

According to MacCallum et al,\textsuperscript{25} judgment of sample size adequacy depends more on commonalities and factor-to-variable ratio. If commonalities are high, recovery of population factors in sample data is normally very good, almost regardless of sample size. The commonalities in the present study were consistently high, and the mean level of commonality was 0.746 (standard deviation: 0.06), which is higher than the recommended threshold of 0.70.\textsuperscript{25} There is a consensus that a minimum of three variables per factor is critical. In the present study, we achieved three factors with ten items, representing a sufficient variable-to-factor ratio. Thus, we are confident that our sample size was adequate for factor analysis, based on the recommendations.

The strength of the present study is its design, which enabled us to address factor analysis. Although the Turkish version of the M-ISI showed a different factor structure, we think that this is not actually an inadequacy of the instrument, but a minimal cultural variation. Similarities of the original and Turkish version of the M-ISI add confidence to the construct validity of the instrument with regard to high internal consistency, test–retest reliability, and convergent validity. A limitation of the present study is a lack of assessment of M-ISI response to treatment and responsiveness will be tested in future investigations. Future studies are recommended to analyze responsiveness and design national surveys to determine the prevalence and characteristics of UI and analyze its associated risk factors by using the M-ISI.

**Conclusion**

The M-ISI instrument was successfully adapted and validated for a Turkish population. The M-ISI is a reliable, consistent, and valid instrument for the multidimensional assessment of UI. Its substantiality and comprehensive structure mean that it has become a practical instrument that is available for utilization in the primary health care setting, clinical research, and epidemiological trials in Turkey.
Acknowledgment
All data were analyzed by EMPIAR Statistical Consulting Service (http://www.empiaristatistik.com).

Author contributions
The authors declare that they are qualified for authorship and have participated sufficiently in the work to take public responsibility for appropriate portions of the content. All authors contributed toward data analysis, drafting and critically revising the paper and agree to be accountable for all aspects of the work.

Disclosure
The authors report no conflicts of interest in this work.

References