ABSTRACT

Objectives: To develop a conceptually and semantically valid Arabic version of a validated disease-specific instrument of quality of life (QoL) for women with abnormal uterine bleeding (AUB).

Methods: This is a prospective cohort study conducted at the Department of Obstetrics & Gynecology, King Saud University, Riyadh, Kingdom of Saudi Arabia between December 2010 and December 2011 following ethics approval. Forward translation of the Aberdeen Menorrhagia Severity Scale (AMSS) from English into Arabic was followed by backward translation of the consensus target (Arabic) version into the source (English) language. Subsequently, a final target (Arabic) language version was created. Sixty-one Arabic-speaking women of reproductive age participated in the study. The final Arabic questionnaire was administered to 41 women with self-perceived normal menses (Group 1) on two occasions 2 weeks apart. Agreement in the answers deems the questionnaire reliable. The final Arabic version was administered to 20 women with self-perceived AUB (Group 2), and their scores were compared with the first response of Group 1. A significant difference between the groups deems the questionnaire valid.

Results: For linguistic validation; intra-class correlation coefficient (ICC) of 0.87 and Kappa statistics of 0.56 to 0.87 indicated good to excellent agreement. For clinical validation, there was a significant difference between Group 1 and 2 (p=0.001).

Conclusion: The translated Arabic AMSS is a reliable and valid indicator of QoL in Saudi women with AUB.

From the College of Medicine (Abu-Rafea, Al Jasser, Al Anazy, Al Mandeel), King Khalid University Hospital, King Saud University, Riyadh, Kingdom of Saudi Arabia and the Department of Obstetrics and Gynecology (Vilos), Western University, Ontario, Canada.

Received 23rd May 2012. Accepted 8th July 2012.

Address correspondence and reprint request to: Dr. Basim Abu-Rafea, Obstetrics and Gynecology, King Khalid University Hospital, PO Box 7805, Riyadh 11472, Kingdom of Saudi Arabia. Tel. +966 (1) 4670818. Fax: +966 (1) 4679557. E-mail: baburafea@ksu.edu.sa
Abnormal uterine bleeding (AUB) is a common condition in women and can lead to physical, psychological, and social impairment in affected patients. It is the leading cause of iron deficiency anemia and a major indicator of hysterectomy worldwide.\(^1\) Traditionally, AUB has been defined as perceived changes or abnormality in one or more of frequency of menstruation, duration of menses, or volume of menstrual blood loss (MBL). The prevalence of AUB varies widely depending on the criteria used for its definition. When objective criteria for MBL >80 mL per cycle are used, the prevalence ranges between 9% and 14%, while when subjective criteria, such as women’s perception of menses are used, it ranges between 10% and 52% in developed countries and between 4% and 9% in developing countries.\(^2,5\) In the Arabian Gulf Region, the prevalence of anemia, from all causes, ranges from 15-48% in women of childbearing age while in Saudi Arabia, the overall country prevalence ranges from 30-56%. Abnormal uterine bleeding was found to be one of the most important factors for iron deficiency anemia and accounted for 38.8% of hysterectomies in Saudi women.\(^6,9\) The National Institute for Health and Clinical Excellence (NICE) in UK, defined heavy menstrual bleeding (HMB) based on the impact on quality of life (QoL) as “excessive menstrual blood loss which interferes with a woman’s physical, social, emotional and/or material quality of life”.\(^1\) Therefore, the primary goal for treating HMB should be to improve the patient’s QoL, and QoL should be the primary outcome of all therapies for HMB. Indeed, QoL has been a focus of interest by many agencies including the World Health Organization (WHO), which defined QoL as “an individualized perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards, and concerns.”\(^9\)

In order to measure QoL, different instruments have been developed by the WHO with 6 different broad domains including physical and psychological health, level of independence, social relationships, living environment, spirituality, religion, and personal beliefs.\(^9\) Some of these instrument, such as SF-36 and SF-12, are generic, while others, such as the Aberdeen menorrhagia severity scale (AMSS), menorrhagia utility scale, menorrhagia outcomes questionnaire, and menorrhagia impact questionnaire, are more disease-specific.\(^10-13\)

As a rule, QoL instruments of, so called, patient-based outcome measurements are used in clinical audits and research to provide subjective and objective assessment of clinical outcomes of therapeutic interventions.\(^14\) However, in every day clinical practice, health care providers are trained to employ the traditional approach of asking specific questions to elicit the patient’s perception of her menses and the incapacity she suffers as a result as measured by the above disease-specific instruments.

Since all of the above instruments were developed in the English language, and since we found no published QoL instrument for AUB translated into Arabic, the objective of our study was to develop a conceptually and semantically valid Arabic version of a known and validated English disease-specific instrument of QoL for women with AUB. We selected the AMSS, also referred to as the Ruta menorrhagia questionnaire, because of its simplicity, reliability and validity.\(^10\)

**Methods.** The study was conducted from December 2010 to December 2011 at the Department of Obstetrics & Gynecology, College of Medicine, King Saud University, Riyadh, Kingdom of Saudi Arabia. Ethics committee approval was obtained from the Institutional Review Board of King Khalid University Hospital. Permission to translate the Aberdeen menorrhagia severity scale (AMSS) and use it in our Arabic population was obtained from the principal author (Ruta DA), department of Public Health, University of Aberdeen Medical School. A committee was formed from the Department of Obstetrics & Gynecology to ensure accuracy of translation, reliability, and validity of the questionnaire.

**Study population.** After informed consent, 61 Arabic-speaking healthy women of reproductive age were invited to participate in the study. Of these, 41 reported self-perceived normal menses based on frequency, duration, and amount of bleeding (Group 1), and 20 attended clinics specifically for self-perceived AUB (Group 2). For Group 1, inclusion criteria included premenopausal women 20-50 years of age, with consistent (regular) menstrual cycles, with duration of 21-35 days, able and willing to give written informed consent, and reliable, compliant, and agreeable to cooperate with all study requirements. Women in Group 2 had similar characteristics to Group 1 with the addition of a self-reported history of HMB in at least 3 of the past 6 months. Women were excluded from both groups if: they were pregnant or lactating; had used any intrauterine device (IUD) including the levonorgestrel-containing intrauterine system (LNG-IUS) within 3 months, had used gonadotropin releasing agonist (GnRHa) or depomedroxyprogesterone acetate (DMPA) within 6 months; had history of clotting defects or bleeding disorders; had severe dysmenorrhea
secondary to confirmed endometriosis, suspected adenomyosis or chronic pelvic pain syndrome; had active genital or urinary tract infection.

Translation and linguistic validation: Forward translation. Translation from English into Arabic was conducted by 3 independent professional medical translators in the target country (Saudi Arabia).

Reconciliation: The committee created a consensus target (Arabic) version from the 3 forward translations.

Backward translation: Translation of the consensus target (Arabic) version back into the source (English) language was conducted by an independent professional medical translator. The source questionnaire was compared with the backward translated document to examine the conceptual content of the forward consensus version. Once the back-translated version was noted to be worded similarly to the source questionnaire, a final target (Arabic) language version was created.

Proof reading. The committee re-examined the final target (Arabic) version to ensure that no typing, spelling, or grammatical errors remained. To assess reliability, the final Arabic version questionnaire was administered to 41 healthy women of reproductive age (Group 1 only) on 2 different occasions roughly 2 weeks apart. If there would be a reasonable agreement in the answers to the same questions, the questionnaire would be deemed reliable.

Clinical validation. The final Arabic version was administered to 20 women complaining of AUB (Group 2) and their scores were compared with the scores of the first response of the 41 women who participated in the linguistic validation part of the study with self-perceived normal menses (Group 1). If a statistically significant difference was found between the 2 groups, the questionnaire would be deemed valid.

Statistical analysis. Data were analyzed on Statistical Analysis System (SAS 9.1) software (SAS Institute Inc., Cary, NC, USA) with Student’s t-test, Pearson correlation, Intra-class correlation coefficient (ICC) and Cohen’s Kappa coefficient. The ICC and Cohen’s kappa coefficients are commonly used to assess the consistency or reproducibility of quantitative measurements made by multiple observers measuring the same quantity. Since we were interested in agreement between each question in the interval of baseline and 2 weeks of administering the same questionnaire in Group 1, we used both the weighted Kappa statistic and the ICC. The total score was calculated and an unpaired t-test was used to compare the means at weeks 0 (baseline) and 2. Test-retest reliability of the total scores was evaluated using the ICC correlation. For the clinical evaluation comparing women complaining of AUB and women with self-perceived normal menses the AMSS scores were log-transformed to improve the homogeneity of the variances and compared using an unpaired t-test. A p value of <.05 was considered statistically significant.

Results. Linguistic validation (Test-retest reliability). The weighted Kappa statistics for all 13 questions of the AMSS given to 41 women, along with their associated 95% confidence intervals are presented in Table 1. All kappa values exceed 0.50. According to arbitrary guidelines for kappa, there was excellent agreement in questions 1, 2, 4, 7, and 13 (kappa >0.75), and fair to good agreement in all other questions (kappa >0.50). The mean±standard deviation of the total score at baseline was 11.23±4.42 and week 2 was 11.20±4.12, were not statistically significant (p=0.887). We also used the intra-class correlation coefficient (ICC) as the most appropriate measure of agreement for the total number of responses. The ICC was 0.87 (95% CI: 0.77-0.93). Intra-class correlation coefficient of 0.87 together with the kappa statistics ranging from 0.56 to 0.87 indicated good to excellent agreement. Therefore, the translated AMSS was deemed to be reliable.

Clinical validation. Descriptive statistics of Groups 1 and 2 are listed in Table 2. Since the data were not paired we did not look for kappa agreement or ICC coefficients. The mean total AMSS score in women complaining of AUB was 18.50 while the mean for women with self-perceived normal menses was 12.15. This difference of 6.35 (95% CI; 2.93-9.77)

Table 1 • Weighed Kappa statistics for the 13 questions of the Aberdeen Menorrhagia Severity Scale given to 41 women.

<table>
<thead>
<tr>
<th>Question number</th>
<th>Weighted kappa</th>
<th>95% confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.87</td>
<td>0.71-1.00</td>
</tr>
<tr>
<td>2</td>
<td>0.76</td>
<td>0.49-1.00</td>
</tr>
<tr>
<td>3</td>
<td>0.57</td>
<td>0.24-0.89</td>
</tr>
<tr>
<td>4</td>
<td>0.77</td>
<td>0.58-0.96</td>
</tr>
<tr>
<td>5</td>
<td>0.68</td>
<td>0.40-0.96</td>
</tr>
<tr>
<td>6</td>
<td>0.59</td>
<td>0.40-0.78</td>
</tr>
<tr>
<td>7</td>
<td>0.85</td>
<td>0.70-1.00</td>
</tr>
<tr>
<td>8</td>
<td>0.62</td>
<td>0.42-0.83</td>
</tr>
<tr>
<td>9</td>
<td>0.69</td>
<td>0.51-0.87</td>
</tr>
<tr>
<td>10</td>
<td>0.56</td>
<td>0.33-0.78</td>
</tr>
<tr>
<td>11</td>
<td>0.64</td>
<td>0.18-1.00</td>
</tr>
<tr>
<td>12</td>
<td>0.64</td>
<td>0.44-0.84</td>
</tr>
<tr>
<td>13</td>
<td>0.79</td>
<td>0.57-1.00</td>
</tr>
</tbody>
</table>
was statistically significant \( (p=0.001) \). Therefore, the translated AMSS was deemed to be valid.

**Discussion.** In a 2011 systematic review, 8 outcomes were found to be important in clinical decisions for patients with AUB: bleeding, QoL, pain, sexual health, bulk-related symptoms, patient satisfaction, need for additional treatment, and adverse events. The authors recommend assessing these outcomes before and after treatment to improve the quality, consistency, and utility of future AUB studies.\(^7\) Although there are multiple QoL instruments used by health care providers in clinical practice and research to provide subjective and objective assessment of clinical outcomes of therapeutic interventions, there is no one found to be superior and standardized for the evaluation of AUB.\(^14,18\) We chose to translate and validate the AMSS in Arabic, as the senior authors (BA and GAV) had extensive experience with the English version of AMSS from prior clinical audits and studies and found it to be user friendly, reliable, and valid.\(^9\) To our knowledge, this translated AMSS is the first one reported to assess the impact of AUB on QoL in Arabic speaking women. The committee found the 3 professionally translated versions in close agreement with each other and there were no major difficulties in reaching a consensus for the final target document. Furthermore, the questionnaire was easily understood (user friendly) by the target participants and no major questions or ambiguities were encountered.

The Ruta questionnaire as a measure of the effects of menorrhagia on health was developed by Ruta et al at the University of Aberdeen.\(^10\) Its validity was confirmed by its high correlation with all the varied aspects of perceived health status measured by the SF-36 which already had been deemed reliable and valid by various authors.\(^20,21\) The questions from the actual ruta menorrhagia severity scale used in our study are appended in Appendix 1.

Reliability of an instrument is the extent to which measurements on the same individual under different circumstances are similar, and test-retest is the most appropriate method for assessing reliability. In the original study by Ruta et al,\(^10\) the test-retest reliability was assessed by questionnaires mailed to 155 women of whom 104 (64%) returned completed questionnaires. Of the 73 patients reporting no change in health status since completing the first questionnaire, a significant correlation of 0.88 was achieved between the 2 sets of scores \( (p<0.001) \). Our intra-class correlation coefficient (ICC) of 0.87 derived from the 41 patients answering the questionnaire 2 weeks apart is certainly in agreement with the original Ruta questionnaire, indicating that our Arabic-translated version is at least as reliable as the original English version.

We also used Cohen's kappa coefficient as a statistical measure of inter-rater agreement, since it is generally thought to be a more robust measure than simple percent agreement calculation. It is also thought that kappa takes into account the agreement occurring by chance, and it may be an overly conservative measure of agreement.\(^22\) Arbitrary guidelines characterize kappas over 0.75 as excellent, 0.40 to 0.75 as fair to good, and below 0.40 as poor.\(^23\) Therefore, our kappa coefficients ranging from 0.56 to 0.87 again reinforce good to excellent agreement indicating that our Arabic-translated AMSS is reliable. Validation is the process to test the extent to which an instrument measures what is intended to do. The translated questionnaire was administered to a small group of patients with self-perceived AUB (N=20), and their scores were compared with the scores of the 41 women used in Group 1. Therefore, the questionnaire was shown to be valid.

**Study limitations.** The small number of patients is one of the major limitations of this study. Other limitations may be the application of the questionnaire to other regions and subgroups of the Arab world. Therefore, only time, larger number of patients, and wider administration will allow us to assess the true value and potential of the translated instrument by auditing its acceptance and utility by patients, health care providers, and researchers in Arabic speaking women with AUB. Administering the questionnaire to larger groups of women, in different regions and areas of the Arab world should be the next and main focus of future research.

### Table 2 - Comparison of the final Arabic version of the Aberdeen Menorrhagia Severity Scale between women with self-perceived normal menses (n=41) and abnormal uterine bleeding (n=20).

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>SEM</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>41</td>
<td>12.15</td>
<td>3.80</td>
<td>0.85</td>
<td>6.00</td>
<td>20.00</td>
</tr>
<tr>
<td>Group 2</td>
<td>20</td>
<td>18.50</td>
<td>6.81</td>
<td>1.52</td>
<td>8.00</td>
<td>33.00</td>
</tr>
<tr>
<td>Difference (1-2)</td>
<td>-6.35</td>
<td>5.51</td>
<td>1.74</td>
<td>2.93 - 9.77*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*confidence interval
In conclusions, based on a small number of patients, the translated Arabic AMSS was shown to be a reliable and valid indicator of QoL in Saudi women with AUB. Administering the questionnaire to larger groups of women, in different regions and areas of the Arab world should be the next and main focus of future research.

References

5. Harlow SD, Campbell OM. Epidemiology of menstrual disorders in developing countries: a systematic review. BJOG 2004; 54: 6-16.
### Appendix 1 • Aberdeen menorrhagia severity scale (AMSS).

<table>
<thead>
<tr>
<th>Question</th>
<th>Options</th>
</tr>
</thead>
</table>
| 1. On average, during the last three months, for how many days did your period last? | - No, always less than 1 day
  - Yes, between 1 and 5 days
  - Yes, between 6 and 10 days
  - Yes, between 11 and 15 days
  - Yes, between 16 and 20 days
  - Yes, between 21 and 30 days
  - Yes, more than 30 days |
| 2. On average, during the last three months, were your periods regular or irregular? | - No
  - Yes |
| 3. On average, during the last three months, how many days were there from the first day of a period to the first day of the next period? | - Less than 10 days
  - Between 11 and 20 days
  - Between 21 and 30 days
  - More than 30 days |
| 4. On average, during the last three months, have you leisure activities prevented you from carrying out your work, housework or other daily activities? | - No
  - Yes |
| 5. On average, during the last three months, have you been confined to bed with each period? | - No
  - Yes |
| 6. On average, during the last three months, have your periods been associated with any pain? | - No
  - Yes |
| 7. On average, during the last three months, did you require more than one form of protection at the same time (not including mini pads or mini pant-liners)? | - No
  - Yes |
| 8. On average, during the last three months, have your periods been associated with any pain? | - No
  - Yes |
| 9. On average, during the last three months, for how many days did you require more than one form of protection at the same time (not including mini pads or mini pant-liners)? | - No
  - Yes |
| 10. On average, during the last three months, have you had any problems with soiling/staining any of the following because of your period? | - No
  - Yes |
| 11. On average, during the last three months, has your sex life been affected by your heavy periods? (Including sports, hobbies, social life) | - No
  - Yes |
| 12. On average, during the last three months, how many pads might you use on the heaviest day of your period? | - Less than 5 pads
  - Between 5 and 10 pads
  - Between 11 and 15 pads
  - Between 16 and 20 pads
  - More than 20 pads |
| 13. At any time during the last three months, did you require more than one form of protection at the same time (not including mini pads or mini pant-liners)? | - No
  - Yes |
| 14. On average, during the last three months, were your periods regular or irregular? | - No
  - Yes |
| 15. On average, during the last three months, has your sex life affecting your social life? | - No
  - Yes |

Note: The Aberdeen Menorrhagia Severity Scale (AMSS) is a tool used to assess the severity of menorrhagia, a condition where menstruation is excessively heavy. The scale includes questions to determine the duration of periods, the amount of protection needed, the impact on daily activities, and the associated pain. The Arabic translation of the ruta menorrhagia severity scale is also provided, ensuring accessibility to non-English speaking populations.