Abstract

Aim: The International Consultation on Incontinence Questionnaire-Urinary Incontinence Short Form (ICIQ-UI SF) and the Wexner Scale have been included in the International Consortium for Health Outcomes Measurement core outcome set for pregnancy and childbirth, to measure urinary and anal incontinence. The reliability and validity of these instruments have not been fully evaluated in a maternity population. The aim of this study was to conduct a psychometric evaluation of the ICIQ-UI SF and Wexner Scale during pregnancy and postpartum.

Methods: Consecutive pregnant women (n = 309) who booked for maternity care at one Australian birthing facility between August 2017 and April 2018 completed online surveys. Women who screened positive for urinary and/or anal incontinence were administered the ICIQ-UI SF and/or Wexner Scale during pregnancy (<27- and 36-weeks) and postpartum (6- and 26-weeks). Scale internal consistency, construct validity, and responsiveness were evaluated.

Findings: In women with urinary-incontinence the ICIQ-UI SF demonstrated good internal consistency reliability at each time point (Cronbach’s α .87, .84, .89, and .82, respectively), recorded significant change across three time-points, and was sensitive to group differences in age and obesity during pregnancy. Wexner Scale data was unsuitable for psychometric analysis.

Conclusions: The ICIQ-UI SF is a valid and reliable instrument to measure urinary incontinence during pregnancy and postpartum. The findings support the inclusion of the ICIQ-UI SF in the International Consortium for Health Outcomes Measurement core outcome set for use during the perinatal period. Psychometric analysis of the Wexner Scale in maternity populations is recommended.
Title: Perinatal incontinence: Psychometric evaluation of the International Consultation on Incontinence Questionnaire-Urinary Incontinence Short Form and Wexner Scale

1. INTRODUCTION

Urinary and anal incontinence are common conditions affecting women, with prevalence reported to increase with age and associated with a significant impact on women’s long-term quality of life. Urinary incontinence is any involuntary loss of urine. The condition is generally sub-classified as either stress, urgency or mixed urinary incontinence. Stress urinary incontinence is common in both during pregnancy and postpartum period.

Research findings for maternity populations are limited by inconsistent measurement. Heterogeneous outcomes, measurement instruments and definitions inhibit the ability to pool data or draw accurate conclusions. To address these inconsistencies the International Consortium for Health Outcomes Measurement (ICHOM) developed a core outcome set of outcomes and measurement tools to measure value in maternity care. Using consensus-based methodology, the international multidisciplinary working party of experts and consumers, identified incontinence as an outcome important to women. Following a comprehensive review of all available patient-reported outcome measures of urinary and anal incontinence, and a quality assessment of each tool, the International Consultation on Incontinence Questionnaire-Urinary Incontinence Short Form (ICIQ-UI SF) and the Wexner Scale (also known as the Cleveland Clinic Florida Fecal Incontinence Score) were included in the set.

The reliability and validity of the ICIQ-UI SF and the Wexner Scale in maternity populations is under-investigated. To address calls from the ICHOM working party to validate the tools in the core outcome set, the current study sought to conduct a psychometric evaluation of the ICIQ-UI SF and the Wexner scale during pregnancy and postpartum.
2. MATERIAL AND METHODS

2.1 Study design

The current study is part of a larger program of work. The MoMeNT Study (MoMeNT Study (Models Meeting Needs Over Time), is a prospective, longitudinal cohort study which aims to compare maternal outcomes according to model of maternity care and assess the feasibility of the ICHOM core outcome set in an Australian maternity population.

2.2 Setting and participants

Participants were consecutive women presenting for maternity care at one publicly-funded tertiary, maternity facility in South East Queensland, Australia who completed or partially completed a baseline survey. Women were invited to participate if they met the following inclusion criteria: aged 18 years or more; with good command of English, 27-weeks gestation or less and with access to email and mobile phone. Women with an existing mental health disorder under the care of a psychiatrist were excluded. Sample size was based on evaluating two broad models of maternity care on maternal health and wellbeing. Recruitment occurred between August 2017 and April 2018. Follow-up occurred using online surveys at four time point during pregnancy (≤27-weeks and 36-weeks) and postpartum (6- and 26-weeks).

*Figure 1* presents the recruitment and flow of participants through the study.

Insert Figure 1 about here

2.3 Measures

The ICHOM core outcome set was administered at each time-point. Measures included those related to health-related quality of life, incontinence, breastfeeding, social support, role transition, mental health and satisfaction with care. The baseline survey obtained information on age, gestation, parity, body mass index. Routinely collected hospital data were collected at 6-weeks postpartum for several birth outcomes including mode of birth, perineal trauma, episiotomy and infant birth weight.
2.3.1 ICIQ-UI SF

The ICIQ-UI SF is a brief questionnaire comprising three scored items and one unscored self-diagnosis item to assess the prevalence, frequency and severity of urinary leakage and its impact on quality of life. The total sum of three items are calculated to provide a summary score, ranging from 0 to 21 with higher scores indicative of increased severity and greater impact on quality of life. The scale has demonstrated high internal consistency reliability among British urology clinic patients and community-based adults (Cronbach’s \( \alpha = 0.92 \)). The unscored diagnostic item was not evaluated in this study.

2.3.2 Wexner Scale

The Wexner Scale comprises five scored items that assess subjective frequency and severity of anal incontinence. A sum of all scores was calculated. A score of zero indicates perfect continence and a score of 20 indicates complete incontinence.

2.4 Procedure

Eligible women receiving pregnancy care were approached by either the first author or their primary midwife and informed of the study aims and requirements. Consenting women provided written consent to participate. Online survey links were sent to women during pregnancy at baseline (day of recruitment) and 36-weeks, and at 6- and 26-weeks postpartum. Women who failed to complete two-consecutive surveys were deemed lost to follow-up. Response rates were 89.3% at 36-weeks of pregnancy, and 85.1% and 77.7% postpartum (6- and 26-weeks) (see Figure 1). Each survey posed a multiple-response incontinence screen question, ‘In the last month, have you leaked urine, leaked stool or passed gas by accident?’ Responses included (1) Yes, I leaked urine, (2) Yes, I leaked stool or passed gas by accident, (3) No I do not leak urine, leak stool or pass gas by accident’. Women who responded positively to any urinary leakage were asked to complete the ICIQ-UI. Women who
responded positively to involuntary passage of stool or gas were asked to complete the Wexner Scale.

2.5 Ethics approval

Ethical approval was obtained to conduct this study from the participating Hospital and Health Service Human Research Ethics Committee (HREC/17/QGC/127) and University (GU Ref No: 2017/625).

2.6 Approach to analysis

Psychometric evaluation was conducted according to guidelines provided by COSMIN (COnsensus-based Standards for the selection of health Measurement INstruments)\textsuperscript{11,12} and followed the framework of Prinsen and colleagues\textsuperscript{13} for evaluating a patient-reported outcome measure. Internal consistency, construct validity, and responsiveness are reported. To establish representativeness, sample characteristics were compared to the hospital (study site) and State perinatal data statistics. For descriptive purposes, dependent variables relating to incontinence were transformed into binary dummy variables. Participants who reported no incontinence were recorded as ‘0’ (absence of incontinence). The full cohort of participants are included in the descriptive analysis. Only participants who responded positively to the incontinence-screen question were included in the analyses for internal consistency, construct validity and responsiveness. Missing data were managed using listwise deletion for computing total scores on ICIQ-UI SF and Cronbach’s (\(\alpha\)). In all other analyses pairwise deletion was used. All participants who commenced the ICIQ-UI and/or Wexner Scale completed all questions. Preliminary assessment of data indicated a violation of the assumption for normal distribution of scores, therefore non-parametric tests were conducted. Chi square tests were used to compare group differences for continent and incontinent women. Prevalence of incontinence was estimated as the ratio of the number of women who
answered positively to the screening question by the total number of women who completed surveys at each time point and are presented as frequencies and percentages with 95 percent confidence interval using Clopper Pearson Exact Test for binary probability. Significance level was set as $p<.05$.

### 2.6.1 Internal consistency

Internal consistency was assessed at each time-point using Cronbach’s alpha coefficient. Preliminary factor analysis was initially conducted to assess suitability of data. Assumptions of unidimensionality, uncorrelated errors, and equal loadings were assessed. Coefficients exceeding 0.70 are considered acceptable\textsuperscript{14}, 0.8 is considered robust and 0.9 is considered high\textsuperscript{15}.

### 2.6.2 Construct validity

The inter-relationship of the three items of the ICIQ-UI SF, in terms of direction and strength, were explored using Spearman Rank Order Correlation. The strength of the relationship ($\rho$) was interpreted according to Cohen\textsuperscript{16}: $0.10–.29 = $ small, $0.30–.49 = $ medium, and $0.50–1.0 = $ large. Based on findings of previous studies (summarized in Table 1) we expected the ICIQ-UI SF to discriminate between women’s age, parity, and obesity status at baseline, and mode of birth at 6-weeks postpartum. Differences were evaluated using Mann-Whitney U tests. The Mann-Whitney U value and Standardized Test Statistic are reported. Effect size ($r = \frac{z}{\sqrt{N}}$) is interpreted using Cohen’s criteria\textsuperscript{16}: $0.10 = $ small effect, $0.30 = $ medium effect, and $0.50 = $ large effect. Based on our clinical understanding of the condition in a community sample of pregnant women, we expected to see a small effect size.

Insert Table 1 about here
2.6.3 Responsiveness

Responsiveness of the ICIQ-UI SF was assessed using a series of Wilcoxon Signed Rank Tests. Change in scores between each two time-points were conducted. Effect size is interpreted using Cohen’s criteria\(^{16}\).

3. RESULTS

3.1 Participants

Baseline characteristics of the 309 participants are presented in Table 2. The sample was comparable to hospital and state birthing populations in all characteristics except mode of birth. The study sample were more likely to birth vaginally and less likely to birth by caesarean section.

Insert Table 2 about here

3.2 Urinary incontinence

Prevalence, frequency, severity, and impact of urinary incontinence during pregnancy (baseline and 36-weeks) and postpartum (6-weeks and 26-weeks) are presented in Table 3 and summarized in Figure 2 (a-c). Urinary leakage was reported by 34.6\% (95\% CI 29.3 – 40.2) women at baseline, 50.0\% (95\% CI 44.0 – 56.1) women at 36-weeks, 22.7\% (95\% CI 17.8 – 28.3) women at 6-weeks postpartum and 21.4\% (95\% CI 16.4 – 27.2) women at 26-weeks postpartum. (Figure 2a). For women who experienced leakage at each time-point, the severity in terms of amount of leakage, was generally small (Figure 2b). Moderate leakage or more was rare. In terms of impact on quality of life, for women with urinary incontinence, impact of a little or more was reported by 59.4\% (95\% CI 49.5 – 68.9) at baseline, increasing to 65.2\% (95\% CI 56.7 – 73.1) at 36 weeks, 66.1\% (95\% CI 52.6 – 77.9) at 6-weeks postpartum and 66.8\% (95\% CI 52.1 – 79.2) at 26-weeks postpartum. For those women who did report an impact this was generally ‘a little’ (Figure 2c). Impact reported as moderate or
greater was experienced by 11 (3.6%) women at baseline, which doubled (n = 21, 7.6%) by 36 weeks of pregnancy. Moderate impact or worse was less postpartum and stable, reported by 10 (3.9%) women at 6 weeks postpartum and seven women (3.0%) at 26 weeks.

3.2.1 Internal consistency and construct validity
The ICIQ-UI SF demonstrated good internal consistency reliability. At baseline, the Cronbach’s \( \alpha \) was .87; at each subsequent time point (Cronbach’s \( \alpha \) .84, .89, and .82, respectively). At baseline there was a strong positive correlation between urine leakage frequency and severity (\( \rho = .98 \)), and frequency and impact of urinary incontinence (\( \rho = .74 \)) explaining 96% and 55% of variance, respectively. Similarly, severity demonstrated a strong, positive correlation with impact (\( \rho = .70 \)), explaining 49% of the variance.

Table 4 presents group differences for all women with and without urinary incontinence during pregnancy and postpartum. During late pregnancy (36-weeks) multiparous women were more likely to report urinary incontinence compared to their primiparous counterparts. Following birth, women who birthed vaginally were more likely to report urinary incontinence, compared to those who birthed by caesarean with a small to medium effect.

For 106 women who reported some form of urinary leakage at baseline, Mann-Whitney U Tests revealed ICIQ-UI SF scores differed for women in terms of age and obesity with a small to medium effect size (see Table 5). Older women (Md = 5.5, \( n = 22 \)) reported higher scores compared to younger women (Md = 4, \( n = 84 \), \( p = .04 \), \( r = .2 \)), while obese women
(Md = 5, n = 28) recorded higher scores compared to non-obese women (Md = 4, n = 70, p = .02). No difference in scores were seen between primiparous and multiparous women at baseline (see Table 5).

Insert Table 5 about here

For 59 women who reported some form of urinary leakage at 6-weeks postpartum, Mann-Whitney U Tests revealed no statistically significant difference in ICIQ-UI SF scores in terms of mode of birth or infant birth weight (see Table 5).

3.2.2 Responsiveness

Changes in ICIQ-UI SF scores for 106 women who reported urinary leakage are shown in Figure 3. Scores increased during pregnancy from baseline to 36 weeks, before improving steadily to values seen at baseline by 26 weeks postpartum. A series of Wilcoxon Signed Rank Tests revealed a statistically significant increase in ICIQ-UI SF scores from baseline to 36 weeks with a medium to large effect size (Md = 4 vs 5, z = -3.26, p = .001, r = .37), and a significant decrease from late pregnancy (36-weeks) to late postpartum (26-weeks), with a small-medium effect size (Md = 5 vs 4, z = -2.81, p = .005, r = .20). No significant differences were seen at any other time.

Insert Figure 3 about here

3.3 Anal incontinence

Prevalence, frequency, severity, and lifestyle alterations of anal incontinence during pregnancy (baseline and 36-weeks) and postpartum (6-weeks and 26-weeks) are presented in Table 6. Anal incontinence was reported by 36 women (11.8%, 95% CI 8.4 – 15.9) at baseline, 43 women (15.6%, 95% CI 11.5 – 20.4) at 36-weeks, 27 women (10.4%, 95% CI 7.0 – 14.7) at 6-weeks postpartum and 16 women (6.7%, 95% CI 3.9 – 10.7) at 26-weeks postpartum. The prevalence of involuntary passage of solid or liquid stool was rare at all time
points. The highest prevalence for passing solid stool was at 6-weeks postpartum (n = 4, 1.5%), while the highest rate for liquid stool was during late pregnancy (n = 6, 2.2%). For women who reported fecal incontinence, passage of stool occurred no more than once a week. The involuntary passage of flatus was more common, reported by 34 women (11.1%, 95% CI 7.8 – 15.2) at baseline, increasing during pregnancy to 36-weeks (n = 41, 14.9%, 95% CI 10.9 – 19.6), before steadily decreasing in the postpartum weeks (6-weeks = 9.2%, 95% CI 6.0 – 13.4; 26-weeks = 5.9%, 95% CI 3.3 – 9.7). The frequency of involuntary flatus occurred no more than once a week for most women. The incidence of both pad-wearing and lifestyle alteration was highest in late pregnancy and lowest at 26-weeks postpartum.

Insert table 6 about here

4. DISCUSSION

The International Consortium for Health Outcomes Measurement included the ICIQ-UI SF and Wexner Scale in their core outcome set to evaluate urinary and anal incontinence in maternity populations. Neither scale had been fully evaluated in samples of pregnant and postpartum women. This study sought to validate these tools in one Australian maternity sample. The ICIQ-UI SF recorded good internal consistency at all four time points and was responsive to change over time from early to late pregnancy and from late pregnancy to late postpartum. Furthermore, the scale differentiated between women with urinary incontinence in terms of age and obesity during pregnancy. The small sample prohibited psychometric testing of the Wexner Scale.

Findings of this study demonstrated the ICIQ-UI SF to be a valid and reliable instrument for use during pregnancy and postpartum and reflect the robustness shown in other populations. Despite being widely used to measure urinary incontinence during pregnancy and postpartum, only limited evidence exists regarding the validity and reliability of the
instrument in maternity populations. For example, Chang and colleagues\textsuperscript{18} studied the effects of episiotomy on urinary incontinence using the ICIQ-UI SF at four-time points postpartum only and found the scale to be reliable, valid and responsive. A paucity of findings relating to the psychometric performance of the ICIQ-UI SF during pregnancy prevent a comparison of findings.

ICIQ-UI SF scores differentiated between groups for urinary incontinence. During pregnancy older women and obese women reported significantly higher UCIQ-UI SF scores compared to younger and non-obese women with small to medium effects. Contrary to the evidence\textsuperscript{1,2,19}, the findings demonstrated no significant difference in scores in terms of parity during pregnancy or mode of birth or infant birth weights postpartum. These findings might be explained by the small number of women with urinary incontinence. While studies generally report prevalence rates of urinary incontinence in community samples of pregnant and postpartum women, the current study reported ICIQ-UI SF change scores for incontinent women only. When comparing prevalence of incontinence for all women however we found a difference in parity during pregnancy, and parity and birth mode postpartum. Consistent with the literature birth mode has the greatest effect. Further, while severe perineal trauma in terms of third- and fourth-degree lacerations are associated with urinary incontinence, small numbers prevented meaningful comparison.

This study identified only a small number of women with anal-incontinence and psychometric analysis of the Wexner Scale was not feasible. Despite the Wexner Scale being the most widely used tool to measure anal incontinence, evidence regarding its psychometric performance is limited\textsuperscript{9}. One study with older women with fecal incontinence by Vaizey and colleagues\textsuperscript{19} compared a newly developed scale with three other grading scales including the Wexner Scale and clinical assessment. Although the scale correlated well with clinical
assessment \( r = 0.78, p < .001 \), demonstrated good test-retest reliability, and was sensitive to change, their findings were in direct contrast to those in a large Australian community sample. Findings from the 2004 South Australian Health Omnibus Survey\(^{20}\) showed the Wexner Scale to be unreliable in the older population. Furthermore, the scale has been criticized for the inclusion of a pad wearing question\(^{20, 21}\) and no question to assess fecal urgency\(^21\). Though acknowledging the Wexner Scale’s omissions and limitations, authors of a comprehensive review of continence outcome instruments recommended the use of the scale for clinical and research purposes\(^15\).

The recently released ICHOM Standard Set of Outcome Measures for Pregnancy and Childbirth\(^6\) was developed to assess outcomes that are important to women. The psychometric evaluation of these two measures is an essential process for refinement of the set and contribute to risk-adjusted international benchmarking that enables comparative evaluations and improvements of health care systems within and across countries. The findings are strengthened by the prospective, longitudinal design which recorded data over four time-points traversing the peripartum period from early pregnancy to 26-weeks postpartum. The findings address an important gap in knowledge as the reliability, validity and responsiveness of the ICIQ-UI SF in maternity populations was previously largely unknown. Our findings support some evidence-based recommendations for inclusion of the measures and use in the ICHOM core outcome set for pregnancy and childbirth\(^6, 7\).

This study was conducted with a community sample of 309 pregnant women from one maternity population in Australia. While the statistical techniques were appropriate for the sample size, the number of participants with incontinence was small. Urinary and anal incontinence data were collected from only those women reporting involuntary urinary leakage or the involuntary passage of fecal stool or flatus and no clinical diagnosis for
incontinence was conducted. The collection of data for all peripartum women in larger and more diverse samples would facilitate more accurate group comparisons and a better understanding of urinary and anal incontinence in maternity samples. Comprehensive psychometric evaluation of the Wexner Scale in community samples of maternity populations are recommended. Furthermore, to address the omission of the scale to measure fecal urgency, future researchers may consider a slight modification to the Wexner Scale (additional question) as described by Vaizey et al\textsuperscript{21}.

5. CONCLUSION

This study addressed an urgent need to validate two patient-reported outcome measures in maternity populations. The ICIQ-UI SF was found to be valid, reliable and responsive, supporting its inclusion in the ICHOM core outcome set for pregnancy and childbirth. Due to small numbers of women with fecal incontinence, the assessment of the psychometric performance of the Wexner Scale was not possible in this study. Future research is needed to evaluate the psychometric performance of the two scales in larger, and more diverse maternity populations. Until evidence becomes available to support the use of the total score of the Wexner Scale during pregnancy and postpartum, the reporting of individual items is recommended for clinical use.

ACKNOWLEDGEMENTS

We would like to thank the ICIQ study group for granting permission to use the ICIQ-UI SF questionnaire in this study. We thank the Gold Coast Hospital and Health Service Research Grants Committee for their financial support and declare, they played no role in the study design, data collection, data analysis or interpretation of findings. We would like to acknowledge Dr Julie Pallant for her ongoing support and statistical guidance during all
stages of analysis. We are particularly grateful to all the staff and midwives who supported the project and assisted in the recruitment of participants. Most importantly, we would like to acknowledge the overwhelming support of the women who so generously gave their time to participate in this study.

FUNDING

The MoMeNT study was supported by a grant awarded by the Gold Coast Hospital and Health Service Research Grants Committee (Ref: 015–01.02.17). The funding body played no role in the study design, data collection, data analysis or interpretation of findings.

REFERENCES