Validation and Clinical Translation of the Revised Continence and Patient Satisfaction Tools: Final Report

UNIVERSITY OF WOLLONGONG
Centre for Health Service Development
July, 2011
Jan Sansoni\(^1\)

Graeme Hawthorne\(^2\)

Nick Marosszeky\(^3\)

Kate Moore\(^4\)

Glenn Fleming\(^5\)

Elizabeth Owen\(^6\)

\(^1\) Director, Australian Health Outcomes Collaboration and Associate Professor, Centre for Health Service Development, University of Wollongong

\(^2\) Associate Professor, Department of Psychiatry, University of Melbourne

\(^3\) Psychometrician / Data Analyst, Educational Assessment Australia, UNSW Global, Rosebery NSW

\(^4\) Head, Department of Urogynaecology, The St George Hospital, Sydney and Associate Professor, University of NSW

\(^5\) Research Assistant, Australian Health Outcomes Collaboration, Centre for Health Service Development, University of Wollongong

\(^6\) Research Assistant, Australian Health Outcomes Collaboration, Centre for Health Service Development, University of Wollongong

Suggested citation:
# Table of Contents

1 Executive Summary ........................................................................................................... 9

2 Introduction ......................................................................................................................... 13
   2.1 General Context .............................................................................................................. 13
   2.2 General Background to the Project ............................................................................... 13
   2.3 Specific Background and Context .................................................................................. 15

3 Methods and Study Details: ............................................................................................... 18
   3.1 Outline of Study Design ................................................................................................. 18
   3.2 Research Aims and Hypotheses .................................................................................... 19
   3.3 Components of the Validation Study ............................................................................. 19
       3.3.1 Patient and Clinic Payments ............................................................................... 20
       3.3.2 Participating Clinics ............................................................................................ 20
       3.3.3 Project Establishment and Management Activities .............................................. 21
       3.3.4 Literature Searches ............................................................................................ 25

4 Results: Revised Incontinence Tools ............................................................................... 27
   4.1 Data Analysis: Introduction .......................................................................................... 27
       4.1.1 Methodology ....................................................................................................... 27
       4.1.2 Patients Recruited ............................................................................................... 27
       4.1.3 Issues Concerning Data Analysis ......................................................................... 27
       4.1.4 Data Quality ........................................................................................................ 28
   4.2 RUIS and Urinary Incontinence: Results ..................................................................... 29
       4.2.1 Characteristics of the Urinary Incontinence Sample ........................................... 29
       4.2.2 RUIS Scores ....................................................................................................... 30
       4.2.3 RUIS Scores at Baseline ..................................................................................... 33
       4.2.4 Pre-treatment to Post-treatment Change Scores ............................................... 36
       4.2.5 Reliability Data .................................................................................................... 38
       4.2.6 Ceiling and Floor Effects ...................................................................................... 39
       4.2.7 Responsiveness: Capacity to Detect Change ........................................................ 39
       4.2.8 Internal Structure: Principal Components Analysis .......................................... 42
       4.2.9 Validity: Correlations with Other Measures ......................................................... 43
       4.2.10 Cutpoints for Interpretation ............................................................................... 44
       4.2.11 Effect of Treatment ........................................................................................... 45
       4.2.12 Alternative Solutions .......................................................................................... 46
       4.2.13 Interpretation of Results ..................................................................................... 48
       4.2.14 Conclusions and Recommendations: Urinary Incontinence .............................. 49
   4.3 RFIS and Faecal Incontinence: Results ....................................................................... 49
       4.3.1 Characteristics of the Faecal Incontinence Sample ............................................. 49
       4.3.2 RFIS Scores ......................................................................................................... 50
       4.3.3 Item Characteristics and Endorsement Patterns .................................................. 50
6.4 Development of a Technical Manual for the Revised Tools ........................................ 100
6.5 Brochures .................................................................................................................. 101
6.6 Presentations at Continence Foundation of Australia (CFA) Conference .............. 101
6.7 Materials Placed on Relevant Web Sites .................................................................. 101

7 Conclusions .................................................................................................................. 102

8 References ...................................................................................................................... 110

Attachment A: Technical Manual: Incontinence and Patient Satisfaction Tools... 117
Attachment B: Brochures for the Incontinence and Patient Satisfaction Tools... 153
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Project logic model: evaluation plan</td>
</tr>
<tr>
<td>4.1</td>
<td>Demographic details of participants</td>
</tr>
<tr>
<td>4.2</td>
<td>Health and incontinence status at baseline in comparison with RUIS scores</td>
</tr>
<tr>
<td>4.3</td>
<td>RUIS in relation to other incontinence variables at follow-up</td>
</tr>
<tr>
<td>4.4</td>
<td>Principal component matrix for RUIS Items: urinary incontinence patients</td>
</tr>
<tr>
<td>4.5</td>
<td>Rotated factor matrices for urinary incontinence items</td>
</tr>
<tr>
<td>4.6</td>
<td>Demographic details of participants</td>
</tr>
<tr>
<td>4.7</td>
<td>Health and incontinence status at baseline in comparison with RFIS scores</td>
</tr>
<tr>
<td>4.8</td>
<td>RFIS and other incontinence variables at follow-up</td>
</tr>
<tr>
<td>4.9</td>
<td>Principal components analysis of RFIS items</td>
</tr>
<tr>
<td>4.10</td>
<td>Rotated factor matrix for faecal incontinence (2006)</td>
</tr>
<tr>
<td>4.11</td>
<td>Principal components analysis: RFIS with pad, flatus and urge Items; faecal patients (2011)</td>
</tr>
<tr>
<td>4.12</td>
<td>Prevalence of incontinence in Australia (without and with comorbidities)</td>
</tr>
<tr>
<td>4.13</td>
<td>Utility value (i.e. quality of life), using the AQoL, by incontinence and comorbid status</td>
</tr>
<tr>
<td>4.14</td>
<td>Utility value (i.e. quality of life), using the AQoL, by incontinence status</td>
</tr>
<tr>
<td>5.1</td>
<td>Demographic characteristics of patients providing patient satisfaction data</td>
</tr>
<tr>
<td>5.2</td>
<td>Health status at baseline of patients providing patient satisfaction data</td>
</tr>
<tr>
<td>5.3</td>
<td>SAPS item frequency response</td>
</tr>
<tr>
<td>5.4</td>
<td>Internal structure of the SAPS using factor and Rasch analysis</td>
</tr>
<tr>
<td>5.5</td>
<td>Impact of change in item wording for four SAPS items</td>
</tr>
<tr>
<td>5.6</td>
<td>Discrimination evaluation of the SAPS</td>
</tr>
<tr>
<td>A1</td>
<td>Health and incontinence status at baseline in comparison with RUIS scores</td>
</tr>
<tr>
<td>A2</td>
<td>RUIS in relation to other incontinence variables at follow-up</td>
</tr>
<tr>
<td>A3</td>
<td>Principal components analysis of RUIS items</td>
</tr>
<tr>
<td>A4</td>
<td>Rotated factor matrices for urinary incontinence items</td>
</tr>
<tr>
<td>A5</td>
<td>Health and incontinence status at baseline in comparison with RFIS scores</td>
</tr>
<tr>
<td>A6</td>
<td>RFIS and other incontinence variables at follow-up</td>
</tr>
<tr>
<td>A7</td>
<td>Principal components analysis of RFIS items</td>
</tr>
<tr>
<td>A8</td>
<td>Clinical and community sample statistics for the RUIS</td>
</tr>
<tr>
<td>A9</td>
<td>Clinical and community sample statistics for the RFIS</td>
</tr>
</tbody>
</table>
List of Figures

Figure 4.1: Items RUIS1 and RUIS2 ................................................................. 31
Figure 4.2: Items RUIS3 and RUIS4 ................................................................. 32
Figure 4.3: Items RUIS5 .................................................................................. 33
Figure 4.4: RUIS pre-treatment scores ............................................................. 33
Figure 4.5: RUIS pre-treatment and post-treatment scores ............................. 38
Figure 4.6: RUIS Change Scores .................................................................... 39
Figure 4.7: RUIS change scores by Patient Global Improvement ...................... 41
Figure 4.8: RUIS change and the clinical impression of improvement. ............. 42
Figure 4.9: Items RFIS1 and RFIS2 ................................................................. 51
Figure 4.10: Items RFIS3 and RFIS4 ................................................................. 51
Figure 4.11: Items RFIS5 .................................................................................. 52
Figure 4.12: RFIS pre-treatment scores ............................................................ 53
Figure 4.13: RFIS pre-treatment and post-treatment scores ............................. 57
Figure 4.14: RFIS change scores .................................................................... 59
Figure 4.15: RFIS change by Patient Global Rating of Improvement ................ 59
Figure 4.16: RFIS by changes in pad use, pre-post treatment ......................... 64
Figure 4.17: SF-36 profile scores for urinary incontinence at baseline ............ 69
Figure 4.18: SF-36 profile scores for faecal incontinence at baseline .................. 70
Figure 4.19: Incontinence expenses by severity group .................................... 81
Figure 4.20: Total personal costs vs. incontinence severity St Mark’s score ....... 82
Figure 5.1: SAPS item data distribution ............................................................ 88
Figure 5.2: SAPS score distribution ................................................................ 89
Figure 5.3: Scatterplot of SAPS by GUTTS scores .......................................... 93
Figure A1: RUIS pre-treatment scores ............................................................. 120
Figure A2: RUIS pre-treatment and post treatment scores ............................. 124
Figure A3: RUIS change scores .................................................................... 127
Figure A4: RFIS scores at baseline ................................................................. 130
Figure A5: RFIS pre-treatment and post treatment scores ............................. 133
Figure A6: RFIS change scores .................................................................... 137
Figure A7: RFIS by changes in pad use, pre-post treatment ............................ 139
Abbreviations

AAG Australian Association of Ageing
ACT Australian Capital Territory
ACTH Australian Capital Territory Health Department
AIHW Australian Institute of Health and Welfare
AQoL Assessment of Quality of Life
CFA Continence Foundation of Australia
CHSD Centre for Health Service Development
DBICI Dowell Bryant Incontinence Cost Index
df Degrees of freedom for statistical analyses
DoHA Department of Health and Ageing
ICIQ-N International Consultation on Incontinence Questionnaire - Nocturia
IIQ Incontinence Impact Questionnaire
ISI Incontinence Severity Index
ISOQOL International Society of Quality of Life
NCMS National Continence Management Strategy
NEAF National Ethics Application Form
PGI Patient Global Impression
PRO Patient Reported Outcome
PSG Project Steering Group
RFIS Revised Faecal Incontinence Scale
RPA Royal Prince Alfred Hospital
RUIS Revised Urinary Incontinence Scale
RWH Royal Women’s Hospital
SAHOS South Australian Health Omnibus Survey
SAPS Short Assessment of Patient Satisfaction
SF-36V2 Short Form 36 Version 2
SSA Site Specific Assessment (form)
UDI Urogenital Distress Inventory
UDI-6 Urogenital Distress Inventory - 6 Items
UOW University of Wollongong
WEI Wei Incontinence Symptom Index (Urinary)
Wexner Wexner Faecal Continence Grading Scale
1 Executive Summary

The National Continence Management Strategy (NCMS) has funded a number of major research studies relevant to the outcomes evaluation of continence conditions. Details of these studies are provided as background to this project.

This study builds on research undertaken for the project Refining Continence Measurement Tools (Sansoni et al., 2006). From the analysis of the urinary and faecal incontinence items and scales included in a community survey of 2004, the South Australian Health Omnibus Survey (SAHOS), the 2006 study developed some revised scales for the assessment of urinary and faecal incontinence (Revised Urinary Incontinence Scale (RUIS), Revised Faecal Incontinence Scale (RFIS)). These scales improved the assessment of incontinence when compared with the original measures (Sansoni et al., 2006). Both the revised scales were found to have excellent internal consistency reliability (RUIS 0.91, RFIS 0.89) in a large community sample.

A limitation of community survey data is its derivation from subjective reports of incontinence symptoms collected in face to face interviews, rather than from confirmed clinical diagnoses. This tends to restrict the range of responses to incontinence items, particularly those pertaining to more severe levels of symptoms. Thus it was necessary to trial the revised continence measures in a range of clinical settings in follow-up field trials.

Although the research activities of these projects to date have shown that the RUIS and the RFIS are highly desirable for evaluation and epidemiological research and have produced revised measures suitable for these uses, the generalisability of these revised instruments is circumscribed by the population samples in which they were developed. It was essential this work be replicated in clinical samples prior to these instruments being widely promoted, adopted and used. Sansoni et al. (2006) therefore recommended further testing of the revised incontinence instruments to obtain a wider range of responses and to examine the clinical outcomes from incontinence patients. The current study was designed to address this issue by field testing the revised scales in eleven clinical settings across Australia.

A related study (Hawthorne, Sansoni, Hayes and Marosszeky, 2006) Measuring Patient Satisfaction with Incontinence Treatment, also made recommendations concerning the further assessment of the newly developed Short Assessment of Patient Satisfaction (SAPS) scale and other patient satisfaction items for the assessment of incontinence treatment and services. The current study further examines the performance of this scale in relation to a number of other clinical indicators including the rating of patient improvement by the treating clinicians.

This project was designed to:
1. validate the revised continence and patient satisfaction tools for assessment and outcome evaluation in both clinical and community settings
2. increase the capacity of health care professionals to use validated tools for their clinical and research assessments
3. improve the level of information and evidence available.

People experiencing incontinence will benefit through better assessment tools becoming available to assess and monitor their condition. Through the project component that addressed clinical translation, activities have been implemented to disseminate the tools to health professionals. For example, presentations about the work have been made at relevant conferences such as the Continence Foundation of Australia (2009) conference.

Overall it was envisaged that the project would:
1. assist in raising awareness of continence issues
2. facilitate the measurement of treatment outcomes of incontinence interventions
3. assist in improving practice and the evidence base for continence management.

This project addresses the needs of major stakeholders which include the National Continence Management Program, the Continence Outcomes Section of the Department of Health and Ageing, peak groups such as the Continence Foundation of Australia, health professionals and researchers working in the field of continence and patients experiencing continence conditions. The target groups...
addressed by this project include the general public, people with incontinence, health service providers, researchers, clinicians and health care workers.

A Project and Evaluation Plan was submitted to the Continence Outcomes Section of the Department of Health and Ageing in August 2008. This outlined the study design, data analysis strategies and patient and clinician protocols. A Project Steering Group was convened, meeting initially in August 2008, with subsequent meetings in February and November 2009, August 2010 and May 2011. Its role was to ensure that the project addressed the needs of major stakeholders and target groups. This was facilitated by feedback received from practitioners and patients, and through project information dissemination activities.

The study commenced in mid June 2008. Since that time eleven continence clinics across 4 States and Territories of Australia have recruited patients to the study. Additional clinics were recruited in mid 2008 and during 2009 in an endeavour to increase patient recruitment. Ethics approval for all sites was obtained and staff at all participating clinics were trained in the study implementation procedures. Patient recruitment commenced in November 2008 at two clinics but most clinics commenced patient recruitment in March 2009.

The study recruited patients from a range of practice settings across Australia, particularly specialist and community continence clinics where patients seek and receive incontinence care. The study protocol contained the revised continence instruments (RUIS, RFIS), patient satisfaction measures (SAPS), health status and health related quality of life instruments (e.g. Short Form 36 Version 2 (SF-36V2)), Assessment of Quality of Life (AQoL)), some items from continence specific health related quality of life and/or impact questionnaires (e.g. Incontinence Impact Questionnaire) and patient and clinician ratings of severity and improvement. It also examined the relationship between these instruments and individual medical conditions, co-morbidity, gender and age.

The study examined clinical and patient definitions of treatment outcomes and success across four different treatment types:

1. continence advising (this includes nurse continence advisors and other staff undertaking advising roles)
2. physiotherapy
3. surgery (note: surgery is usually only undertaken on those whom physiotherapy has failed)
4. mixed/combined treatments and other treatments.

The relationships between the revised instruments, type of treatment, clinical feedback and patient satisfaction were examined. Additional reliability data were collected from post-test patients in order to examine the test-retest reliability of the instruments over a two week period. Although it had already been established that the internal consistency reliability of the revised tools was excellent (Sansoni et al., 2006), to facilitate clinical uptake of the tools it was necessary to show that the test-retest reliability for these instruments was also acceptable.

Although all clinics provided detailed estimates of patient recruitment for their site, many of these were found to be overestimates of the likely recruitment rate. There was a slower rate of patient recruitment than expected and advised by the clinics at a number of sites. This led to the recruitment of additional clinics in 2008 and 2009 in an endeavour to boosts patient recruitment numbers. The project duration was also extended from June 2010 to June 2011 to allow for a longer period of patient recruitment. As of May 2011 full pre-treatment data were available for 256 patients (195 urinary and 61 Faecal) and follow up data available for 138 patients (100 urinary and 38 faecal).

Data analyses indicated that the RUIS and the RFIS have excellent psychometric properties. The RUIS and the RFIS performed well in clinical settings demonstrating:

1. adequate to good internal consistency reliability and test-retest reliability
2. high and significant correlations with other measures of incontinence
3. good evidence that these instruments were sensitive to changes in continence status as a result of treatment, making them suitable for outcome evaluation.

The RUIS and the RFIS items and the scale total scores also discriminated well in relation to other clinical indications of severity and between people with differing levels of incontinence severity.
(discriminant validity). In terms of examining the severity of incontinence conditions across clinical and population groups the instruments appeared to be functioning as expected, with the clinical sample having higher scores (reflecting greater incontinence) than the community population sample.

The principal component analyses undertaken indicated the internal structure of the instruments was appropriate. The RUIS appears to be a uni-dimensional measure with all RUIS items loading highly (0.64 or above) on the primary urinary incontinence factor extracted (accounting for 49% variance) which is characterised as urinary leakage. The RFIS also appears to be a uni-dimensional measure with all items loading above 0.65 on the primary faecal incontinence factor which accounts for 54% of the variance.

With regard to the SAPS items, some changes were adopted to make the wording of the items and their response categories more consistent. These modified items were found to have equivalent psychometric properties. The findings concerning the Short Assessment of Patient Satisfaction (SAPS) also indicate that it is sensitive to changes in patient status as a result of treatment with those patients reporting the greatest levels of improvement also reporting higher levels of satisfaction with their treatment. It also has good internal consistency (alpha = 0.85) reliability. At post-treatment the RUIS correlation with untransformed patient satisfaction scores (SAPS) was r = -0.44 (p<0.00) which indicated there was an association between higher RUIS scores (indicating greater incontinence) and lower patient satisfaction scores (less satisfaction). At post treatment the RFIS correlation with SAPS was r = -0.28 (p = 0.09).

The impact of incontinence on health related quality of life for incontinence was also examined. Sansoni et al. (2006) reported that urinary and faecal incontinence measures had negative but significant correlations with measures of health related quality of life reflecting the impact of this condition (e.g. the higher the incontinence score the lower the physical functioning or health status score). At pre-treatment the RUIS correlation with the Physical Function Scale of the SF-36V2 was r = -0.15 (p<0.05) and the RFIS correlation with the Physical Function Scale of the SF-36V2 was r = -0.27 (p<0.05) which confirmed this finding. The Short Form 36 Version 2 (SF-36) profiles for urinary and faecal incontinence patients were examined. For both faecal incontinence and urinary incontinence these show that most profile means are in the vicinity of 42-43 and thus are about 7 points or 0.7 standard deviations below the Australian population norms (50). The Physical Summary component Score for faecal incontinence patients was 43.59 (urinary incontinence is 45.59) and the Mental Health Component Score was 43.41 (urinary incontinence is 43.43). This also confirmed that patients with incontinence conditions have poorer physical and mental health status scores compared to the Australian population.

A burden of disease analysis was undertaken using population prevalence estimates from the 2004 SAHOS dataset which contained a number of incontinence items and scales (e.g. UDI-6, ISI, Wexner). Based on disutility due to incontinence this study has shown that the excess burden of disease associated with incontinence in Australia is considerable, with the estimate, from a societal perspective, at $25 billion per annum. These findings, showing that society carries considerable incontinence excess burden of disease, contrasts with other estimates. Earlier estimates (Mathers et al., 1999; Goss, 2008) did not include those with slight or moderate urinary incontinence in their analyses yet the current study found that these cases were responsible for three-quarters of all excess costs. These earlier studies also did not consider faecal incontinence.

The Faecal Cost of Illness Sub-study provides estimates of the costs of faecal incontinence in Australia. It was found that the Total Costs (which includes personal items and investigations) did not relate directly to severity largely because the costs of investigation are largely fixed in Australia. However, the total Personal Costs of hygiene items (pads, creams) increased with the severity of incontinence.

**Research Recommendations**

A number of further research activities could be considered by the Continence Outcomes Section.

The Australian Longitudinal Study of Women’s Health has included the RUIS in the latest surveys for the cohort of older women (N = 5000) across two time points. An analysis of these data would be extremely useful in helping to confirm the reliability and validity of the RUIS in a large population
sample of elderly women as well as providing detailed information on prevalence, severity and the correlates of urinary incontinence (function, health status, social isolation) in this population group.

It should be noted that the focus of this project was to develop validated continence assessment tools for Australian adults. It was not proposed to customise these continence instruments for use by other target groups such as people from Culturally and Linguistically Diverse Backgrounds, Aboriginal and Torres Strait Islander Groups, Children, Proxies or Carers - within the confines of this project. It is recognised that the tools may need to be modified for such groups and this could form a potential follow up project.

Although the RFIS was found to perform well in clinical settings and to have superior measurement properties when compared with other faecal incontinence instruments in this study these findings are based on a sample of 61 patients (48 females and 13 males at pre-treatment and 31 females and 7 males at post-treatment). It would be desirable if a follow up study could continue to collect more data on faecal incontinence (particularly in males) using the same study protocols to analyse gender and type of treatment aspects for faecal incontinence in more detail. Given the lower prevalence of faecal incontinence, particularly for males, the cooperation of a large number of clinics would be required. However, there are few studies in the research literature addressing male faecal incontinence in particular and this would help to further address this gap. It would also be desirable if further data collection for males with urinary incontinence could be undertaken. An online collaboration (see below) between clinics may be a way to identify and address these research gaps.

General Recommendations

With only 5 items each the RFIS and the RUIS are short and simple to use and score and continence clinics treating incontinence patients should be encouraged to use them both as assessment measures and as an outcome evaluation measures in routine practice. The use of such measures can provide effective feedback to clinicians concerning the effectiveness of their treatments, can facilitate the systematic review and monitoring of patients, and assist in identifying ways to improve practice. Similarly the SAPS is a short and effective measure of patient satisfaction with treatment which can readily identify patient concerns for the clinic or practice.

It would be desirable if these data were to be collected routinely both prior to and following treatment. These are simple tools with which clinics can report on the effectiveness of their treatments. In the longer term the Continence Outcomes Section could consider the development of a continence outcomes data collaborative as has occurred with rehabilitation and palliative care (Australian Rehabilitation Outcomes Centre, the Palliative Care Outcomes Collaboration) although it is suggested that an online, real time framework is used rather than static warehousing given recent developments in this field. For example, the Diabetes Educators Association of Australia and a number of State Asthma groups use an online community to manage patients, share data, practices, and recruit patients to major conjoint research initiatives. Through such online collaborations clinics can compare their treatment outcomes and patterns of practice with other related clinics. Such organizations are useful change agents in promoting best practice within the field.

The data analyses confirm that the revised incontinence and patient satisfaction tools have demonstrated good psychometric properties in clinical settings and this confirms findings from earlier community sample data (Sansoni et al., 2006). Papers about this study have been presented at the Australian Health Outcomes Conference in 2008, the Continence Foundation of Australia Conference in 2009, the International Society for Quality of Life Research (2007, 2008, and 2010) and the International Continence Society (2007) conferences. An additional article on the SAPS has been prepared for submission to a relevant journal. Although outside the confines of this project further presentations will be given at the 2011 International Continence Society and Continence Foundation of Australia Conferences to update practitioners on the final results of this study.
2 Introduction

2.1 General Context

Incontinence is a significant health issue across the lifespan with physical, social and economic implications for the individual, their carers and the community. Almost four million Australians are estimated to have some degree of incontinence with the prevalence of incontinence increasing with age in both men and women (AIHW, 2006).

The National Continence Management Strategy (NCMS) was established in 1998 by the Australian Government Department of Health and Ageing to provide funding for research and service development initiatives aimed at prevention and treatment of incontinence. The aims, objectives, and guiding principles of Phase 3 have been outlined in the National Continence Management Strategy Phase 3 Action Plan: 2006-2010 (Australian Government, 2006a).

The aim of Phase 3 of the NCMS was to improve continence awareness, management and treatment so that more Australians can live and participate in their community with confidence and dignity. The desired outcomes of the NCMS were to:

1. improve the level of information and evidence available
2. raise awareness of continence issues and promote continence health
3. increase the capacity of primary healthcare professionals
4. improve the evidence base for continence management.

A guiding principle of the NCMS and the more recent National Continence Management Program (NCMP) is also that support to people with continence issues should be accessible from appropriately trained and informed health professionals.

This project, which sought to validate improved tools for continence assessment and outcome evaluation in both clinical and community settings, increases the capacity of health care professionals to use validated tools for their clinical and research assessments and will improve the level of information and evidence available. People experiencing incontinence will benefit through better assessment tools becoming available to assess and monitor their condition.

Through the project component that addresses clinical translation, activities have been implemented to disseminate the tools to health professionals. For example, presentations about this work have been made at relevant conferences such as the Continence Foundation of Australia (2007, 2009) conference and the Australian Health Outcomes Conference (2008). Overall it is hoped that the project will assist in raising awareness of continence issues, facilitate the measurement of treatment outcomes of incontinence interventions and assist in improving practice and the evidence base for continence management.

The project addresses the needs of major stakeholders which include the NCMS, the Continence Outcomes Section of the Department of Health and Ageing, peak groups such as the Continence Foundation of Australia, health professionals and researchers working in the field of continence and patients experiencing continence conditions. The role of the Project Steering Group has been to ensure the needs of such major stakeholders and target groups are addressed throughout the project processes. These processes include feedback from practitioners and patients and through project information dissemination activities. The target groups addressed by this project include the general public, people with incontinence, health service agencies, researchers, clinicians and health care workers.

2.2 General Background to the Project

The continence outcomes research funded by the Commonwealth’s National Continence Management Strategy, has achieved remarkable success. Key reports to date, which provide the background to the current study include:


In addition to these major reports describing incontinence in Australia and informing policy, several journal articles have been prepared:


The findings from these studies have been presented at both national and international conferences by the project team, as outlined below:


Ageing Australia Conference (2007) – Adelaide

There has also been significant collaboration with other projects funded by the NCMS such as the Women's Health Australia studies and projects assessing incontinence in residential aged care (O'Connell, 2007). At the CFA conference in Adelaide in 2009, Dariah Mohd Yusoff presented a paper ‘Urinary incontinence and cultural practices amongst postnatal women in Kelantan, Malaysia’ which reported on the use of the Revised Urinary Incontinence Scale in Malaysia and a postgraduate student of Professor Jan Patterson is using the faecal incontinence protocol in Indonesia. There has also been interest expressed in considering the inclusion of this tool as part of a toolkit for the Victorian implementation of the Council of Australian Governments Long Stay Older Patients Initiative (COAG LSOP). The revised continence tools are also recommended as follow up continence assessment tools for the proposed standardised ACAT assessment (Sansoni et al., 2010) for use by Aged Care Assessment Teams (ACATs).

Graeme Hawthorne has reported international interest in the Short Assessment of Patient Satisfaction (SAPS) including requests for translation. Given feedback received at international conferences as part of this project some minor changes to item wording for this scale were tested in this study.

The Australian Longitudinal Study of Women’s Health has also included the RUIS in the latest surveys for the cohort of older women (N = 5000) across two time points. An analysis of this data would be extremely useful in helping to confirm the reliability and validity of the RUIS in a large population sample of elderly women (N = 5000+) as well as providing information on prevalence, severity and correlates of urinary incontinence in this population group.

This body of work, however, is unfinished. First, there is an urgent need for the results of this work to be placed in the hands of health workers who provide services for those at risk of or with incontinence, and to be widely disseminated to policy makers. Second, it has raised a number of additional research questions which need to be investigated.

This project addresses the following two issues:

1. the validation and field testing of the revised incontinence instruments
2. the clinical translation of the revised incontinence instruments.

2.3 Specific Background and Context

A Continence Outcome Measurement Suite Project (COMS: Thomas et al., 2006) was commissioned by the Australian Government Department of Health and Ageing, National Continence Management Strategy Research Program with the goal of recommending a suite of continence outcome measures to be used by clinicians and researchers in Australia. This project was finalised in early 2006 (Thomas et al., 2006).

Recommendations from this report led to a related project Measuring Incontinence in Australia (Hawthorne, 2006). Measuring Incontinence in Australia (2006) assessed a number of the recommended measures (Urogenital Distress Inventory 6, Incontinence Severity Index, and the Wexner Faecal Continence Grading System) by including them in the autumn 2004 South Australian Health Omnibus Survey (SAHOS: Harrison Health Research, 2004), which was a community population survey. This study provided Australian prevalence estimates for both faecal and urinary incontinence based on this community survey.

For urinary incontinence, the results suggested that the preferred urinary incontinence measure was the Incontinence Severity Index (ISI; Sandvik et al., 1993; Sandvik et al., 1995; Sandvik et al., 2000). It was found to possess superior measurement properties in comparison with the Urogenital Distress Inventory (UDI-6; Uebersax et al., 1995). As the UDI-6 contains some items that may be endorsed by those without urinary incontinence (e.g. pain in the lower abdominal region), the UDI-6 may overstate incontinence prevalence and the impact of this on peoples’ lives. Given its psychometric properties, there was a case for major revision of the UDI-6. Although the ISI was the preferred measure, because it violated the assumptions of classic psychometric theory relating to scale stability as it contained only two items, further research into its properties was also recommended.

Regarding the measurement of faecal incontinence, as Hawthorne (2006) indicated the current definition of faecal incontinence by the International Continence Society (Abrams et al., 2008) excludes flatus, yet an item on flatus is included in the Wexner Faecal Continence Grading System (Jorge and
Wexner, 1993). In addition to this definitional inconsistency, the evidence from Hawthorne (2006) and AIHW (2006) suggested that the inclusion of the flatus item led to overestimates of faecal incontinence prevalence. It was recommended that further work on the Wexner be undertaken to remove the flatus item and to improve the measurement properties of this scale.

A follow up project *Refining Continence Measurement Tools* (Sansoni et al., 2006) was undertaken to further analyse the SAHOS dataset to develop better instruments for the assessment of incontinence in Australia. This project made use of both Classic Test Theory and Modern Test Theory approaches for the analysis of data. From the analysis of the urinary and faecal incontinence items and scales included in the 2004 SAHOS this study developed some revised scales for urinary and faecal incontinence (Revised Urinary Incontinence Scale (RUIS); Revised Faecal Incontinence Scale (RFIS)) which improved the screening assessment of incontinence when compared with the original measures (Sansoni et al., 2006). Both scales were found to have excellent internal consistency reliability (RUIS 0.91; RFIS 0.89) in a community sample. A study using the RUIS in a clinical sample (Hawthorne, et. al., 2006) indicated the RUIS could describe more severe cases of incontinence than would be found in a population survey sample and that it was sensitive to change / improvement arising from treatment.

One of the limitations of using community survey data is that, as the data is collected in face to face interviews the data are at the level of subjective reports of incontinence symptoms rather than confirmed diagnoses. These considerations mean that in a community survey there will be a limited range of responses to incontinence items particularly those pertaining to more severe levels of symptoms. Thus it was necessary to trial the revised continence measures in a range of clinical settings in follow-up field trials. This study was designed to address this issue by field testing the revised scales in eleven clinical settings across Australia.

There were some psychometric issues raised by the *Refining Continence Measurement Tools* project (Sansoni et al., 2006) concerning some items from the UDI (e.g. stress incontinence; emptying bladder). The analysis of the urinary incontinence items using modern test theory (MTT) approaches (Item Response Theory) indicated that some UDI items did not fit the best model due to differential item functioning as it was found that males and females respond to some items quite differently. This aspect is further explored in the current study. The response categories for the frequency of urinary leakage item from the ISI were also found to be problematic as the response thresholds were disordered, meaning that there was not a graded relationship between incontinence status and endorsed response levels in the community sample. Given this it was thought that some additional items concerning stress incontinence and the frequency of urinary leakage could be included in the clinical dataset. It was also thought that to gain acceptability by clinicians, the psychometric properties of these items needed to be checked against comparable items.

With regard to faecal incontinence Sansoni et al. (2006) found the flatus item included in the Wexner had a low corrected item-total correlation and that the internal consistency reliability (Cronbach’s alpha = 0.57) would be improved if this item were removed from the scale. Both the Hawthorne (2006) and AIHW (2006) reports also recommended this item should be removed from the Wexner as it confounded prevalence estimates and is outside the current ICS definition of faecal incontinence (Norton et al., 2002; 2005; Abrams et al., 2008). As a result the item on flatus was not included in the Revised Faecal Incontinence Scale. However, clinical feedback indicated that it may be worth testing some further items concerning flatus as they felt this was an important aspect to cover, so some additional items on flatus were included in the faecal incontinence protocol.

Additional questions were also been included in the clinical field trials. These included patient-rated global assessments of treatment benefit, satisfaction and willingness to continue treatment (Pleil et al., 2005) and the Patient Global Impression of improvement and severity for incontinence (Yalcin and Bump, 2003). These items could provide useful global indices for routine use in clinical practice.

The Hawthorne, Sansoni, Hayes and Marosszeky (2006) study *Measuring Patient Satisfaction with Incontinence Treatment* also made recommendations concerning the further assessment of the newly developed Short Assessment of Patient Satisfaction (SAPS) scale and other patient satisfaction items.

Although the research activities of these projects to date have shown that the revisions to the incontinence instruments were highly desirable for evaluation and epidemiological research, and produced revised measures suitable for these uses, the generalisability of these revised instruments is
circumscribed by the population samples in which they were developed. It was essential this work was replicated in clinical samples, prior to these instruments being widely promoted, adopted and used. It is for this reason that Sansoni et al. (2006) recommended that further testing should be undertaken on the refined incontinence instruments in order to obtain a wider range of responses and to examine the clinical outcomes from incontinence patients.

This examined the revised measures in a range of clinical settings across Australia and was conducted to further establish the validity, reliability and suitability of these instruments for Australians.
3 Methods and Study Details:

3.1 Outline of Study Design

This study recruited patients from a range of practice settings across 4 Australian States: particularly specialist and community continence clinics where patients seek and receive incontinence care. The study protocol contained the revised incontinence instruments (RUIS, RFIS), related measures of incontinence, patient satisfaction measures (SAPS) and included health status and health related quality of life instruments (e.g. SF-36V2, AQoL and items from continence specific health related quality of life and/or impact questionnaires). It also examined the relationship between these instruments and individual medical conditions, co-morbidity, gender and age.

This study examined clinical and patient definitions of treatment outcomes and success, across four different treatment types:

1. continence advising
2. physiotherapy
3. surgery (note: surgery is usually only given to those whom physiotherapy has failed)
4. mixed / combined treatments or other treatments.

The relationship between clinical indicators (e.g. clinical ratings of severity), the revised instruments and patient satisfaction was examined. Additional reliability data were collected in order to examine the test-retest reliability of the instruments at two weeks post-treatment. Although it had already been established in a community survey that the internal consistency reliability of the revised tools was excellent (Sansoni et al., 2006), to facilitate clinical uptake of the tools it was also necessary to show that the test-retest reliability for these instruments was acceptable in clinical settings.

This study has built on the work completed by Hawthorne et al. (2006) where patient satisfaction items (SAPS) and the refined continence assessment tools were tested in a retrospective study utilising a clinical sample of women with urinary incontinence. This study extended this work by examining patients' responses pre and post treatment for the revised incontinence and patient satisfaction tools, as well as attempting to obtain a larger sample of males with incontinence and patients with faecal incontinence. The SAPS is administered post-treatment only. Sensitivity estimates for the SAPS by incontinence type and treatment group were examined.

The responsiveness of the instruments was explored through an examination of patients' responses pre and post treatment. It is important that the final instruments are shown to be sensitive to patient experiences and to clinical improvement.

Both the Continence Outcomes Measurement Suite Project and the Framework for Economic and Cost Evaluation for Continence Conditions Project, from the National Continence Management Strategy, concluded that information about the economic consequences of faecal incontinence and the burden of disease for this condition is severely lacking for the Australian population. A recent search of the international literature revealed only one cost of illness study regarding faecal incontinence that was conducted in the Netherlands (Deutekom et al., 2005). This validation study provided an opportunity to collect further cost of illness data.

A component of the validation study examined the costs of illness of faecal incontinence. This included interview time for 50 faecal incontinence patients and costs were assessed by the Dowell Bryant Incontinence Cost Index (DBICI). The DBICI is a detailed questionnaire that measures all direct personal costs of managing incontinence over the preceding week, including pads, linen, laundry costs of washing soiled clothes, dry cleaning costs etc., as well as all medical costs for treatment over the previous 12 months (Dowell et al., 1999). It is designed for administration by a Continence Nurse Advisor – self-administration is not successful (Moore et al., 2006). The test-retest reliability of the test has been established (Moore et al., 2006) and the direct costs correlate strongly with the severity of incontinence, on objective measures. The DBICI is also responsive to change, i.e. reduction in incontinence after treatment is mirrored by reduced direct costs of incontinence.
A number of the colorectal clinics (e.g. St George, Royal Prince Alfred Hospital) collected the additional data required from their faecal incontinence patients. For this component of the validation study, the desired outcomes were:

- the cost of faecal incontinence was determined based upon a sample of Australian men and women
- additional estimates of the burden of disease for this condition were determined (AIHW 2006 estimates excluded persons with mild and moderate urinary incontinence and those with faecal incontinence).

### 3.2 Research Aims and Hypotheses

The following research questions were investigated by this study:

- have responsive instruments to assess incontinence been developed that are sensitive to patient experience and clinical improvement?
- have the psychometric properties of the additional items tested in the study protocols led to any additional items being included in the revised incontinence tools?
- has the test-retest and internal consistency reliability of the instruments been assessed and do the revised measures have adequate reliability?
- has the validity of the revised instruments been assessed and did these instruments show evidence of adequate validity?
- did the DBICI provide the necessary cost data required to inform on the costs of faecal incontinence in Australia?
- was sufficient information gained from the quality of life measures to estimate the impact and burden of disease for incontinence?

In considering these aims the research examined key psychometric parameters for the revised continence assessment instruments which included:

- test–retest and internal consistency reliability estimates for the RUIS and RFIS
- validity estimates for the RUIS and RFIS in comparison with other standard tests and measures (e.g. ISI, UDI–6, ICIQ-SF and Wei Incontinence Symptom Index for urinary incontinence; Wexner and St Mark’s Incontinence Score for faecal incontinence) from clinical samples
- estimates of sensitivity to incontinence type, and estimates of responsiveness to change in the patient’s condition over time
- sensitivity estimates for the SAPS by incontinence type and treatment group
- examination of the internal (factor) structure of the revised incontinence and patient satisfaction instruments
- examination of the relationship between demographics, health status, incontinence status, treatment type, co-morbidity and treatment outcomes as assessed by the RUIS/RFIS and SAPS from the patient perspective
- examination of the relationship between clinical and patient definitions of treatment outcome and success, across different types of incontinence and treatment types
- comparison of this clinical data with the large scale population data already collected in previous work (Sansoni et al., 2006)

### 3.3 Components of the Validation Study

With advice from the former National Continence Management Strategy Implementation Subcommittee, the Project Steering Group, and the Continence Outcomes Section eleven continence centres were recruited to the study. The aim was to recruit 250 patients by November 2010 when patient recruitment ended. Two hundred and fifty six patients were recruited to the study. There was a sample of 195 urinary incontinence patients with pre-treatment data available and there were 100
urinary incontinence patients with post-treatment data available. For faecal incontinence data was available for 61 patients at pre-treatment and 39 patients at post treatment but for one patient their pre-treatment data was unavailable.

Analyses considered the following major factors:
1. type of incontinence (faecal, urinary); patients with double incontinence were also identified
2. type of treatment (continence advising, physiotherapy, surgery, mixed/other treatments)
3. gender (male and female).

As part of the study, patients completed the relevant incontinence scales (depending on type of incontinence) and health status and/or related quality of life scales both before and following treatment. Patients completed the patient satisfaction items and scales at the end of their treatment.

Additional reliability data were collected on a sub-sample of 78 post-test patients in order to examine the test-retest reliability of the instruments over a two week period. A sub-sample of 54 patients with faecal incontinence was interviewed using the Dowell Bryant Incontinence Cost Index to gain data on the cost of illness of faecal incontinence.

The participating clinics ascertained patient willingness to consent to the mailing out of the relevant pre-treatment protocol, and consent forms were sent to patients when the initial appointment for treatment was made. Patients then returned the protocols with their consent forms to the participating clinics. The clinics forward the protocols to the research team.

The post treatment protocols (including a pre-paid return envelope) were posted to participating patients, after the completion of the patient’s treatment.

Relevant ethics approvals were prepared for the University of Wollongong and participating clinics in NSW, the ACT and Victoria (refer Section 3.3.2).

3.3.1 Patient and Clinic Payments
Adequate funds were allocated to pay and recruit staff to collect and survey patients at the recruitment sites. Each clinic was paid a $40 recruitment fee for each patient who completed the study. The senior clinician coordinating the study at each site was also involved in ethics approvals, patient coordination and reviewing the final report. To cover these time commitments a coordination fee ($90) for every patient completing the study was also paid to the clinics.

For those clinics participating in the collection of cost of illness data additional funds were allocated for patient recruitment and for the payment of Continence Nurse Advisors to undertake interviews concerning the cost of illness (this was based on $100 per hour per patient interview).

Feedback received from the clinics was that these fees were adequate and the clinics were pleased that they were being reimbursed for the use of their staff time. However, the impression gained was that operational aspects, such as adequate extant staffing during the project period, were the main factors which determined whether a clinic would participate in the study or not and this also influenced the degree to which clinics could expedite patient recruitment.

A small reward was provided to patients who participated in and completed the study. For example, a shopping voucher to the value of $25 was provided to participating patients on completion of the main study; an additional $15 was provided for patients who participated in the test-retest reliability component.

Following discussion with the clinics recruitment targets were initially set with regard to the aim of recruiting over 400 patients. However, it was only possible to recruit 256 patients given the slower than expected rates experienced for patient recruitment at the clinics (refer to Sections 4.1.2 and 4.1.3).

3.3.2 Participating Clinics
The project team contacted a range of multidisciplinary continence services across Australia to participate in this study. The following continence clinics collaborated in this study:
1. Pelvic Floor Unit at the St George Hospital (SGPFU), Sydney. This included the participation of surgeons and the physiotherapy and Nurse Continence
Advising (NCA) staff attached to this unit. St George Hospital offers Nurse Continence Advising, Physiotherapy and Surgical interventions for patients with urinary and faecal incontinence.

2. Surgical patients (urinary incontinence associated with the SGPFU).

3. Anorectal Clinic, St George Hospital, Sydney.
   This included the participation of surgeons and the NCA staff attached to this unit. St George Anorectal Clinic offers Nurse Continence Advising and Surgical interventions for patients with faecal incontinence.

4. NSW Biofeedback and Continence Centre, Royal Prince Alfred Hospital Sydney.
   This included participation from two specialist Nurse Continence Advisors in the area of faecal incontinence.

5. Royal Women’s Hospital (RWH) Victoria - Surgery.

6. Royal Women’s Hospital (RWH) Victoria - Physiotherapy.
   This included physiotherapy patients from the RWH associated Continence Clinic.

7. ACT Continence Promotion Centre.
   This Centre provides Nurse Continence Advising and Physiotherapy services to patients.

8. Colorectal Surgery South Australia.
   This is a group of four colorectal surgeons within private practice in South Australia.

9. Women’s and Men’s Health Physiotherapy, Melbourne.
   This private clinic provides a range of physiotherapy services to patients with faecal and urinary incontinence.

10. Lemongrove Community Health Centre.
    This Centre mainly provides Nurse Continence Advising services to older patients.

11. Patients from a Colorectal Surgeon (Private Clinic) Fitzroy, Victoria.

Two other clinics in Victoria and South Australia were initially recruited but had to withdraw due to unexpected staff shortages arising at these clinics before patient recruitment commenced.

### 3.3.3 Project Establishment and Management Activities

A number of activities were completed as part of the project’s establishment. These include the development and revision of the patient pre-treatment and post-treatment surveys; ethics applications and approval processes; the development of a project and evaluation plan; the formation of the Project Steering Group; the training of clinic staff in study implementation procedures; the provision of relevant study materials to clinics; the establishment of a hot-line for patients; clinical consultation and the processes for coordinating and monitoring patient recruitment. These activities are outlined below.

1. **Develop the patient protocols/surveys.**

   The pre-treatment patient protocols included the revised incontinence and patient satisfaction tools, other measures of incontinence, a number of socio-demographic and health condition items, measures of the impact of incontinence and health related quality of life and patient and clinician ratings of incontinence severity. The post-treatment protocols also included the Short Assessment of Patient Satisfaction scale, some additional patient satisfaction items and clinician and patient ratings of severity and improvement.

   To ensure best practice measurement of incontinence a number of additional incontinence items were included in the research questionnaires. This allowed for detailed comparisons to be made between selected items in terms of their endorsement levels and clinical validity.

   Additional items included those from the Wei et al. (2003) Incontinence Symptom Index which were included in the Urinary Incontinence Protocol to provide further assessment of stress and urge incontinence. Additional items from the International Consultation on Incontinence Questionnaire - Nocturia (ICIQ-N; Donovan et al., 1996) were included to cover this domain. The Incontinence Impact Questionnaire (Uebersax et al., 1995) was also included in the Urinary Incontinence Protocol to assist with the validation of the revised instruments.

   An alternative item on flatus from the Colorectal Anal Distress Inventory (Barber et al., 2005) was included in the Faecal Incontinence Protocol as well as additional items on incontinence impact, type
and severity from the Defecation Distress Inventory (Roovers et al., 2001), the St Mark’s Incontinence Score (Vaizey et al., 1999), the Faecal Incontinence Severity Index (Rockwood et al., 1999), the Faecal Incontinence Quality of Life Scale (Rockwood et al., 2000) and the Faecal TyPE Specification (Wexner et al., 2002). The selection of additional items was guided by liaison with clinical practitioners participating in the study and feedback received from the Project Steering Group.

2 Develop an ethics application for the University of Wollongong. Prepare and coordination of the ethics approvals for collaborating clinics/centres.

The ethics application for the University of Wollongong (UOW) was submitted on 29 July 2008. Approval to conduct the study was granted in mid September. The Site Specific Assessment (SSA) process was completed at St George Hospital, Sydney on 4 November 2008 and patient recruitment commenced on 24 November 2008. At Royal Prince Alfred Hospital, the SSA approval was received on 6 March 2009. The private clinics in South Australia (Colorectal South Australia) and Victoria (Women’s and Men’s Health Physiotherapy) associated with the UOW ethics application began patient recruitment in early December 2008.

The ethics application for ACT Health (ACTH) was developed and notification of approval was received from the Ethics Committee on 26 November 2008. The ethics application for Royal Women’s Hospital (RWH) Melbourne was submitted on 27 October 2008. Following detailed review and correspondence with the RWH Research Committee and the RWH Human Research Ethics Committee a few minor changes were made to the standard Victorian Participant Information and Consent Form pending final approval of the project which was received in February 2009.

Additional private clinics like (Colorectal South Australia) were covered by the UOW ethics application (see above). Site approval for the Lemongrove Community Health Centre was obtained on 12 June 2009.

3 Develop an evaluation plan in line with the NCMS Evaluation Framework and Guidelines.

The Project and Evaluation Plan was submitted to the Continence Outcomes Section of DoHA and was considered by the Project Steering Group at its first meeting on the 29 August 2008. Project Briefings and Progress Reports provided feedback in relation to these evaluation questions and Key Performance Indicators throughout the duration of the project. The Project Logic Model: Evaluation Plan is depicted below.
### Table 3.1: Project logic model: evaluation plan

**Aim:** to validate and field test the revised Instruments and translate them for use by clinicians and health care workers

<table>
<thead>
<tr>
<th>Inputs</th>
<th>Activities</th>
<th>Process Outcomes</th>
<th>Project Outcomes</th>
<th>Long term Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Revised Continence &amp; Patient Satisfaction Tools</td>
<td><strong>Validation and field testing</strong></td>
<td>Ethics approvals</td>
<td>Internal consistency and test-retest reliability of measures</td>
<td>Establish the validity, reliability and suitability of instruments for Australians</td>
</tr>
<tr>
<td><strong>Project Team</strong></td>
<td></td>
<td>Recruit and train up to 11 clinics</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Recruit approx. 200-250 patients</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Clinics</strong></td>
<td><strong>Pre and post treatment survey and interviews</strong></td>
<td>Establish test-retest reliability of the incontinence instruments on 60 patients</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Interview 50 faecal incontinent patients</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Test and consider the inclusion of additional items in the revised tools</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Examine key psychometric parameters (validity, responsiveness) for revised continence instruments</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gain clinical feedback</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Health Professionals</strong></td>
<td><strong>Clinical translation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Academic publications</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Technical manual for revised tools</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Brochures and website tools</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Conference presentation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Project Steering Group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>NCMS COMs projects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>NCMS funding</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Updated version of the Revised Instruments</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>NCMS Economic Framework Project</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Continence Tools Dissemination Strategy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
As can be seen from the Project Logic Model above, information on activities and process outcomes for the project has been provided in the previous Progress Reports. Both process and project outcomes are also reflected in the major deliverables such as this Final Report, the Technical Manual for the Validated Instruments (refer Appendix A), brochures (web and hard copy; refer Attachment B), peer reviewed articles and conference presentations.

As part of summative evaluation it is also desirable to reflect on longer term outcomes and impact assessment as the longer term goal is that the validated instruments are widely promoted, adopted and used. Although all materials will be placed on relevant websites as soon as they are completed, it should be noted that it will be difficult to ascertain the uptake of the validated instruments until some months post project completion.

4 Establish and convene the Project Steering Group.

The recruitment of members for the Project Steering Group began in June 2008. The Project Steering Group contained representatives from relevant health professional groups, academic experts and includes a representative of the Continence Outcomes Section of the Department of Health and Ageing. A representative from the Independent Evaluation team was invited to attend meetings as required. The first meeting of this group was convened on 29 August 2008 and further meetings were held in February 2009, November 2009, and August 2010.

The following persons were members of the project steering group:

- Dr Michael Murray, Geriatrician (Chairperson)
- Mrs Irmina Nahon, Physiotherapist, ACT Continence Promotion Centre
- Dr Susan Hunt, Senior Nurse Advisor, Office of Aged Care Quality and Compliance, Department of Health and Ageing
- Ms Anne Sargeant, CNC Continence, Lemongrove Community Health Centre
- Professor Julie Byles, Director of the Research Centre for Gender, Health and Ageing and co-Director of the Newcastle Institute of Public Health
- Dr Nick Rieger, Colorectal Surgery South Australia
- Mr Mark Gaukroger, Continence Outcomes Section, Department of Health and Ageing
- Ms Lesley Millar, Representative of Independent Evaluation Team for NCMS

Associate Professor Janet Sansoni represented the project team and the participating clinics. Mr Nick Marosszeky, Research Fellow, acted as Secretary to this group.

The role of the Project Steering Group was to ensure that the expertise of the project team was informed by a range of multidisciplinary expertise across the field of incontinence research and practice. The Project Steering Group also monitored project processes and assisted the project team in project level evaluation. Although the project team used an evidence based approach in reporting on the findings from this research, the Project Steering Group served as an independent body to further ensure that impartial and evidence based reporting occurred throughout the project.

5 Convene project briefing meetings to train all continence centres in the use of patient protocols and study implementation procedures. Tailor implementation to local sites.

Project team members visited all sites (ACT, VIC, NSW and SA) to provide training for staff at the participating clinics. A number of minor changes to the process, which did not affect the standardisation of data collection procedures, were agreed as part of tailoring the project implementation to local sites. Changes related to how patients of the clinic were made aware of and recruited into the study; and how patients were given the questionnaire to fill out (e.g. sent via the post to be completed before their first clinic appointment or handed to them at their first clinic appointment to be filled out at home prior to treatment).

At sites where there were a lot of staff changes the project team revisited these clinics to provide additional training. Such an activity was undertaken with the ACT Continence Promotion Centre in March 2010.
6 Undertake clinical consultations with centre staff to obtain clinical feedback on the revised incontinence and patient satisfaction instruments.

Feedback was obtained from clinical staff and the Project Steering Group concerning the patient protocols and the inclusion of items in these protocols.

7 Ensure all participating clinics have easy access to the patient instrument protocols and information concerning the incontinence assessment tools.

Key staff members at the clinics were provided with a copy of the latest patient protocols. All clinics were posted the prepared patient protocol packages prior to the commencement of patient recruitment. Training concerning both the contents and the processes of administration occurred.

8 Establish a hotline for patients and clinics to obtain any further information about the study and to respond to queries.

A hotline was established and the phone number and an email address were provided in the Participant Information Sheet for patients and in the Instructions Guide for clinical staff. Use of the hotline was monitored. Only a few calls were received. These calls mainly related to completion of the follow-up form – usually to advise that it had been sent.

9 Coordinate and monitor patient recruitment.

Considerable liaison concerning patient recruitment processes occurred with the clinics prior to the commencement of patient recruitment. Liaison continued throughout the project. Regular audits were undertaken to streamline information gathering and to examine data collection and research processes. The Project Steering Group was updated concerning these matters at each meeting.

10 Data management.

Personal identifiers were removed at the start of the data entry process. A single paper copy containing the key to the codes was kept in a locked filing cabinet by the principal researchers. This was necessary in order to allow the principal researchers to withdraw any data if a participant requested that their data be removed.

Computer files were password protected and restricted to the principal researchers. All the paper questionnaires, bladder or bowel diaries and any other relevant clinical information received was converted to electronic format and placed in a password protected folder. The data was then collated and coded to form a single electronic file. Once this file was checked the paper questionnaires were then securely shredded. All electronic files had personal identifying information removed and had password protection.

The paper survey form was converted to an electronic version via the SurveyMonkey online survey system. This approach provided a check for transcription accuracy in order to minimise errors being entered onto the research database. The research database was created in SPSS V 17 (2009). Final data checking and a cleaning check of the database were undertaken. Particular emphasis was placed on patient disease, treatment and clinical severity codes, patterns in missing data and any written responses the patients may have made on their survey forms.

3.3.4 Literature Searches

At project commencement in 2008 a search of MEDLINE, CINAHL and PsycINFO was undertaken using the terms faecal incontinence and urinary incontinence crossed with assessment, outcomes, evaluation, instrument, measure, questionnaire and scale. Separate searches were undertaken on specific instrument names for those instruments or items proposed for use in the patient protocols (e.g. Wexner, St Mark’s Incontinence Score, UDI–6, etc). Given previous searches had been undertaken for earlier related projects (Sansoni et al 2006; Hawthorne et al., 2006) results were limited to those published from 2001 onwards.

Further searches were undertaken in relation to faecal and urinary incontinence and patient satisfaction, health related quality of life, burden of disease, cost of illness, prevalence and incidence. These searches were replicated in 2011 to seek papers published since 2007. Although
a systematic literature review was not the primary purpose of this project the relevant literature identified was important for the analysis and discussion sections of this validation study.
4 Results: Revised Incontinence Tools

4.1 Data Analysis: Introduction

Data were analysed according to type of incontinence (faecal, urinary), treatment type (continence advising, physiotherapy, surgery) with data points at pre-treatment and post treatment. The analyses also examined relevant background variables such as gender, age, educational level and BMI. The relationship between the RUIS and RFIS with other measures of incontinence, incontinence impact and health status and health related quality of life was also examined. All data were analysed using SPSS Version 17 (2009).

4.1.1 Methodology

The statistical methods used include both Classic Test Theory analyses (e.g. item descriptive statistics, item endorsement and discrimination, item-total correlations, internal consistency, test-retest reliability, and exploratory factor analysis) and Modern Test Theory approaches (e.g. Item Response Theory, Rasch Analyses).

4.1.2 Patients Recruited.

Patients (N = 255) were a convenience sample recruited consecutively from 11 continence clinics (specialist and community) across 4 regions in Australia. Study eligibility criteria were attending a clinic to receive treatment for urinary incontinence, age between 18-85, and having sufficient English to complete a self-report questionnaire.

There was a sample of 195 urinary incontinence patients with pre-treatment data available and there were 100 urinary incontinence patients with post-treatment data available. There were 167 females in this sample and 28 males at pre-treatment. The most common treatment was physiotherapy (57%), followed by continence advising (20%) and surgery (19%).

For faecal incontinence data was available for 61 patients at pre-treatment and full data was available for 38 patients at post treatment. There were 51 females and 10 males in the pre-treatment sample. The most common treatment was continence advising (36%) followed by physiotherapy (27%) and surgery (27%).

4.1.3 Issues Concerning Data Analysis

Actual patient recruitment indicated that it was difficult to recruit the number of faecal patients that would be required to achieve more substantial cell counts and thus sub analyses of the faecal dataset, as for example by gender, had less than 80% power to avoid the possibility of a Type 2 error – that is a potential failure to detect a difference when in fact there may be one. The prevalence of faecal incontinence is less than that for urinary incontinence and for this reason targets were set with the faecal incontinence clinics at the start of this study to help to address this issue. However, a number of faecal continence clinics reported a lower than expected number of patients attending these clinics for treatment during 2009-2010. The clinics could suggest no reason for this other than that it may have been associated with the recession with patients less likely to consider treatment at a time of financial constraint. Similarly, overall it was harder to recruit male patients as compared with female patients and it is suspected this again reflects the lesser prevalence of both faecal and urinary incontinence amongst males (Hawthorne 2006, AIHW 2006) with fewer males attending clinics for treatment.

It was anticipated there would be some variation in cell counts reflecting treatment pathways that are occurring amongst continence clinics in Australia. For example, preliminary liaison with clinics had indicated that relatively few male faecal incontinence patients are treated by physiotherapists. It also appears that there are relatively few males that require surgery for urinary incontinence as their primary condition. Many males will receive surgery for the primary condition of prostate cancer and thus would only enter the study if they received physiotherapy or continence advice for urinary incontinence that may occur as a result of this surgery.
Although the various generalist clinics contacted indicated they would have a number of patients receiving nurse continence advising as a treatment, the final recruitment numbers for this treatment form are somewhat lower than initially advised by the clinics for patients with urinary incontinence (20% of urinary incontinence patients received the continence advising treatment). Some of this reflects staffing issues that have resulted in fewer clinics with NCAs participating in the study than was initially anticipated. An additional clinic, the Lemongrove Community Health Centre, was recruited to boost recruitment of patients receiving NCA treatment.

A number of clinics reported that there is very little difference in the treatments provided by nursing and physiotherapy staff – as both professionals provide physical exercises, bladder training and advice. It was suggested at the November 2009 Project Steering Group meeting that these treatments could be combined and classed as ‘conservative treatment’ for the purposes of data analysis. Both combined and non combined analyses are presented where sample size permits.

Although sufficient patients were recruited to allow an analysis by treatment type, type of incontinence and to a lesser extent gender, low numbers in some cells limit detailed analysis. For example, it was not possible to break down faecal incontinence simultaneously by both treatment type and by gender. However there were sufficient patient numbers across most cells to undertake key analyses.

### 4.1.4 Data Quality

Data quality issues were monitored with reported rates considered to be acceptable for missing data, attrition and refusal. The patient initial refusal rate for the study to date was low (under 20 percent) across the sites. Within either the pre-treatment or post-treatment datasets about 5% of the dataset contains a missing survey form component (e.g. a pre-treatment or post treatment clinician form). The exception to this was a slightly lower rate of return of post treatment clinician forms for the faecal patients (@ 18% missing). As for most variables the rate of missing data was low no imputation of missing data was performed unless this has been identified in the text for the relevant analysis. One example is that as 3 post-treatment item scores were missing for the St Mark’s Incontinence Score an imputation using the last value carried forward (LVCF) was used (Eliot and Hawthorne, 2005) so that the numbers for this comparative analysis of instruments were more equivalent. No attempt was made to impute categorical missing data (e.g. gender).

Initially the project team had hoped to use bowel and bladder diaries provided by the clinics as clinical indicators. However, there was a higher rate of missing data for these items (40%) and the data quality was very poor – clearly some patients experienced difficulties in completing these forms. As a result of this little of this data was able to be used in the data analysis. It would be useful both for future research, as well as routine practice, if simpler bowel and bladder diaries could be designed for patient completion.

The attrition rate for those patients who completed pre-treatment assessment but did not complete the follow up assessment was higher than expected for the urinary sample. For this sample only 51.3% of patients completed the post treatment forms whereas for the faecal sample the return rate was 63%. A 60% return rate is often reported (Hawthorne et al., 2006). This issue was discussed with the clinics and they indicated that in a number of cases urinary patients (approximately 15%) opted out of their incontinence treatment prior to completion at the clinic and that these cases would be lost to follow up. A number of post forms for patients still completing treatment are still outstanding as they were recruited late in the study. Approximately 3% of patients also indicated they did not wish to continue participation in the study.

Given the lower than expected rate of return of the post-treatment patient forms for the urinary incontinence sample a number of statistical analyses were undertaken to see if there are any systematic differences between those patients that returned post-treatment forms as against those that completed pre-treatment forms only.

For the urinary incontinence sample a number of analyses indicated that there were no relevant differences between patients with post forms as against those with pre-treatment forms only. There were no significant differences between these groups at baseline on RUIS scores, patient assessed incontinence severity, gender, age group, educational background, the number of co
morbidities, health status and body mass index. The post-treatment sample was older \( (p<0.05) \) in years but it is thought this may be due to a higher than expected rate of post-treatment returns for retired persons and a lower than expected rate of return for those in full time employment \( (X^2 = 8.28, df 2; p<0.02) \).

For the faecal incontinence sample a number of analyses indicated that there were no relevant differences between patients with post forms as against those with pre-treatment forms only. There were no significant differences between these groups at baseline on RFIS scores, pre-treatment patient assessed incontinence severity, clinician assessed pre-treatment severity, gender, educational background, the number of co-morbidities, health status and body mass index.

Given the above analyses it is clear that there is no systematic difference in the characteristics of the pre-treatment and post-treatment samples.

### 4.2 RUIS and Urinary Incontinence: Results

During the study 195 patients seeking treatment for urinary incontinence (excluding those with substantial missing data) were recruited to the study and 100 patients submitted post-treatment protocols.

The following sections outline the basic analyses of the data. Once patients were identified, the analyses examined the following four issues: (1) the mean performance (descriptive statistics) of the clinic samples comparing them to community population data; (2) the level of missing data in the sample; (3) differences or change scores between patients pre and post treatment; and (4) an examination of the internal structure of the instruments (Cronbach’s alpha for internal consistency reliability; principal component analyses) in order to see how the individual items hang together.

#### 4.2.1 Characteristics of the Urinary Incontinence Sample

Table 4.1 depicts the basic demographics of the sample. Most participants were Australian born, were middle aged, had achieved completion of high school and/or a trade qualification, were working (PT/FT) and were female. There were no significant differences in RUIS scores by country of birth, age group, gender, education of labour force status.

Physiotherapy was the most common treatment (57%), followed by continence advising (20%) and surgery (19%). With regard to type of urinary incontinence 80% of patients were identified by the clinicians as having stress incontinence symptoms and 63% had urge incontinence symptoms and many of these patients (49%) had mixed incontinence or a combination of these types. Only 5% of patients were identified by clinicians as having overflow incontinence and 1.5% of patients were identified as having functional incontinence (associated with a physical or mental impairment impeding the capacity to get to the toilet).

Many of the patients had experienced incontinence symptoms for more than 2 years (65%) and only 16% had developed their incontinence symptoms within the last year. Only 10% of the sample had no other illnesses or health conditions, 17% of patients had one other health condition and 73% of the sample had 2 or more co-morbidities. The most common co-morbidities were problems with the neck, back or spine (47%), arthritis (44%), high blood pressure (34%) and migraine headaches (28%)
Table 4.1: Demographic details of participants

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country of birth</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>131</td>
<td>68%</td>
</tr>
<tr>
<td>Other</td>
<td>63</td>
<td>32%</td>
</tr>
<tr>
<td>Age group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than 40 years</td>
<td>18</td>
<td>9%</td>
</tr>
<tr>
<td>40-49 years</td>
<td>36</td>
<td>19%</td>
</tr>
<tr>
<td>50-59 years</td>
<td>47</td>
<td>24%</td>
</tr>
<tr>
<td>60-69 years</td>
<td>62</td>
<td>32%</td>
</tr>
<tr>
<td>70+ years</td>
<td>32</td>
<td>16%</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>28</td>
<td>14%</td>
</tr>
<tr>
<td>Female</td>
<td>167</td>
<td>86%</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary and some High School</td>
<td>59</td>
<td>30%</td>
</tr>
<tr>
<td>Completed Higher School Certificate</td>
<td>47</td>
<td>24%</td>
</tr>
<tr>
<td>Trade Certificate/TAFE Diploma (a)</td>
<td>43</td>
<td>22%</td>
</tr>
<tr>
<td>College/University</td>
<td>45</td>
<td>23%</td>
</tr>
<tr>
<td>Labour force status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Labour force (i.e. working, student, unemployed)</td>
<td>80</td>
<td>44%</td>
</tr>
<tr>
<td>Homemaker</td>
<td>32</td>
<td>18%</td>
</tr>
<tr>
<td>Retired/Sickness benefit</td>
<td>70</td>
<td>38%</td>
</tr>
</tbody>
</table>

Notes:
All percentages are rounded up to the nearest integer, so percentages may not round up to 100%.
Excludes all missing cases. The base number was 195 cases.
a = Technical and Further Education.

4.2.2 RUIS Scores

A copy of the RUIS can be found in Attachment A. The RUIS contains the following items:

1. Urine leakage related to the feeling of urgency? (UDI-6)
2. Urine leakage related to physical activity, coughing or sneezing? (UDI-6)
3. Small amounts of urine leakage (drops?) (UDI-6)
4. How often do you experience urine leakage? (ISI)
5. How much urine do you lose each time? (ISI)

Scoring is by simple summation and the range is 0-16 with 0 indicating no incontinence. This analysis of the research database identified 167 females and 28 males suitable for analysis of their RUIS data at pre-treatment.

RUIS scores could be calculated for the entire sample as the rate for missing data was 0%. At baseline for all 195 cases the mean RUIS score was 10.92 (SD = 3.33; N = 195), the mode was 12 and the median was 11. The mean RUIS score for the sample of females at pre-treatment was 10.90 (SD = 3.16, N = 167) and the mean RUIS score for the sample of males at pre-treatment was 11.07 (SD = 4.18, N = 28) (NB: Similar median results were found for both genders). There was no significant difference between these total scores when analysed by gender (t = 0.25, df 193; p>0.05). There were no significant gender differences found between the means for any of the individual items (p>0.05).

These results are in the expected direction as a clinical sample should give higher scores when compared with a community sample. The community sample results reported a mean RUIS score of 2.47 for women (N = 1712) and a mean RUIS score of 0.70 for men (N = 1203) (SAHOS: Harrison Health Research, 2004; Sansoni et al., 2006). In the community sample there was a significant difference in RUIS scores between men and women with women having a higher total and item scores reflecting the higher prevalence of urinary incontinence symptoms amongst women in the general community.
**Item characteristics and Endorsement Patterns**

For most RUIS items 62-68% of the sample experienced these symptoms moderately or greatly at pre-treatment. The response levels for each item were examined by RUIS total scores, and by total scores on the Wei Symptom Index (which shares no common items with the RUIS), as indices of severity.

With regard to endorsement patterns for the urge incontinence item (RUIS1) 22 patients (11.3%) indicated they did not experience urine leakage associated with urgency, 45 patients (23.1%) indicated they experienced this symptom slightly and 128 patients (65.7%) experienced urge incontinence moderately (30.3%) or greatly (35.4%). When examined by RUIS and WEI symptom total scores, as measures of severity, the response levels are ordered with those endorsing the more severe levels of this symptom also having higher mean RUIS ($F = 53.00$, df 3, 191; $p<0.00$) and WEI total scores ($F = 23.96$, df 3, 191; $p<0.00$). For those that endorsed ‘not at all’ on this item the mean RUIS total score was 7.73 as contrasted with those that endorsed ‘greatly’ where the mean was 13.55.

**Figure 4.1: Items RUIS1 and RUIS2**

For the stress incontinence item (RUIS2) 20 patients (10.3%) did not experience this symptom, 45 patients (23.1%) experienced it slightly and 130 patients (66.6%) experienced stress incontinence moderately (21%) or greatly (45.6%). When examined by RUIS total scores ($F = 61.86$, df 3, 191; $p<0.00$) and WEI symptom total scores ($F = 15.43$, df 3, 189; $p<0.00$) the response levels are ordered with those endorsing the more severe levels of this symptom also having higher mean RUIS and WEI total scores and this is true for both males and females. For those that endorsed ‘not at all’ on this item the mean RUIS total score was 7.50 as contrasted with those that endorsed ‘greatly’ where the mean was 13.22. On both of these analysis there was little difference between the RUIS means for the ‘not at all’ and slightly’ categories.

Although there is no significant difference between the item means by gender ($t = -0.50$, df 193, $p>0.05$) there is a significant difference (Pearson Chi Square = 8.52, df 3; $p<0.05$) in the pattern of response categories endorsed by men and women. Men are more likely than women to endorse ‘not at all’ (21.4% vs. 8.4%) and are less likely to endorse ‘slightly’ (7.1% vs. 25.7%). Men and women have more similar levels of endorsement for ‘moderately’ and ‘greatly’ (42.9% vs. 46.1%). This might suggest that some differential item functioning is occurring for males (e.g. they might have a ‘response’ tendency to endorse some response categories more/less than females rather than this actually reflecting a true difference in the symptom itself). However, RUIS1 and RUIS3 make use of these identical response categories and no significant differences in the pattern of responses were found for these items which would be expected if differential item functioning was occurring if, for example, men were generally less likely to choose ‘slightly’ vs. ‘never’ as a response option. The data suggests that where males have this symptom they experience it moderately or greatly rather than slightly whereas for females they experience all levels of severity of the symptom. However, it is noted that the sample size of males for this analysis is small (N =
28) so it would be desirable to confirm this finding with a larger group of males with urinary incontinence.

For the item concerning whether one experiences and is bothered by small amounts of urine leakage (RUIS3) 15 patients (7.7%) indicated ‘not at all’, 59 patients (30.3%) responded with ‘slightly’, 121 patients (62.1%) responded with ‘moderately (31.3%)’ or ‘greatly’ (30.8%). When examined by RUIS total scores (F = 110.99, df 3, 191; p<0.00) and WEI symptom total scores (F = 22.68, df 3, 189; p<0.00) the response levels are ordered with those endorsing the more severe levels of this symptom also having higher mean RUIS and WEI total scores. For those that endorsed ‘not at all’ on this item the mean RUIS total score was 7.07 as contrasted with those that endorsed ‘greatly’ where the mean was 14.15.

**Figure 4.2: Items RUIS3 and RUIS4**

**RUIS3: Small amounts of urine leakage (drops)**

**RUIS4: How often do you experience urine leakage?**

For the item concerning the frequency of leakage (RUIS4) 3 patients (1.5%) indicated they never experienced this, 14 patients (7.2%) indicated they experienced leakage less than once per month, 19 patients (9.7%) experienced leakage a few times a month, 49 patients (25.1%) experienced leakage a few times per week and 110 patients (56.4%) experienced leakage of small amounts of urine every day/night. When examined by RUIS and WEI symptom total scores the response levels are ordered with those endorsing the more severe levels of this symptom also having higher mean RUIS (F = 51.95, df 3, 190; p<0.00) and WEI total scores (F = 30.92, df 3, 189; p<0.00). However, on the Wei Symptom Scale analysis there was little difference between the means for the ‘never’ and ‘less than once a month’ and ‘a few times a month’ categories whereas there was a mean difference of 2.86 between ‘never’ and ‘less than once a month’ categories for the RUIS analysis. For those that endorsed ‘never’ on this item the mean RUIS total score was 3.00 as contrasted with those that endorsed ‘every day/night’ where the mean was 12.61.

With regard to the amount or volume of leakage (RUIS5) 2 patients (1%) indicated they had none; 63 (32.3%) patients leaked drops, 87 patients (44.6%) leaked small splashes and 43 patients (22.1%) experienced more leakage than small splashes. When examined by RUIS and Wei total scores the response levels are ordered with those endorsing the more severe levels of this symptom also having higher mean RUIS (F = 42.74, df 3, 191; p<0.00) and WEI total scores (F = 15.14, df 3, 189; p<0.05). For those that endorsed ‘none’ on this item the mean RUIS total score was 1.5 as contrasted with those that endorsed ‘more’ where the mean was 13.74.
The above analyses provide evidence that the RUIS items all discriminate well by different levels of urinary incontinence severity. Although the sample size of males is small there were no significant differences in item means by gender and there were no significant differences in the response patterns by gender for four of these items.

4.2.3 RUIS Scores at Baseline

Figure 4.4 shows the distribution of RUIS pre-treatment scores.

Table 4.2 provides a summary of patient health and incontinence status indicators at baseline in comparison to RUIS scores and the statistics relating to these comparisons.

Clinical ratings of pre-treatment severity were grouped into 3 groups (mild, moderate and severe) and these were examined by RUIS scores. The mean RUIS score for the mild group was 9.22, for the moderate it was 11.79 and for the severe group it was 12.13. The overall analysis was significant (F = 16.99, df 2,191; p = 0.000) and there were significant differences in RUIS group means between those with mild incontinence and those with moderate and severe incontinence (p <0.05).

Similarly, when examined by the pre-treatment patient rating of the severity of urinary incontinence the RUIS mean for those in a mild state was 8.36, for those with moderate incontinence it was 11.60 and for those with severe incontinence it was 14.03; this analysis was significant (F Welch = 80.46, df 2,109.07; p = 0.000) and all group comparisons were significant (p <0.05).
The mean RUIS scores were significantly higher ($t = -2.10$, $df = 70.49$; $p = 0.007$) for those that received surgical vs. conservative treatments – this might be expected as generally surgery is used with patients with more severe incontinence or for whom conservative treatment has failed. For this analysis those that were receiving mixed treatments and other treatments such as medicines were excluded ($N = 13$). The pre-treatment RUIS mean for the surgical group was 12.03 and for the conservative treatment group (physiotherapy and continence advising) it was 10.71.

At pre-treatment there was a significant differences ($F = 9.83$, $df = 2,176$; $p = 0.000$) in RUIS mean scores by type of urinary incontinence. The mean for those with urge incontinence only was 9.00, for those with stress incontinence only it was 10.94 and for those with mixed incontinence it was 11.79. There was also a significant difference ($t = -2.71$, $df = 193$; $p = 0.007$) in RUIS mean scores between those that had both faecal and urinary incontinence (double incontinence) as against urinary incontinence only. The mean for those with urinary incontinence only was 10.60 and for those with double incontinence it was 12.18.

The RUIS does not include any items concerning pad use but pad use should reflect the severity of urinary leakage. Patient scores on the Wei items concerning the number and the size of pads were compared to RUIS scores. Patient scores on the number of pads used were grouped as less than 1 per day and 1 per day or more. The mean RUIS score for the low pad use group was 8.06 and for the high pad use group it was 11.55. This difference was statistically significant ($t = -6.74$, $df = 192$; $p = 0.000$).

For the pad size item three groups were classified as no pad/thin pad or tissue; medium or regular pad; and large/maxi pad or absorbent disposable undergarment. This analysis was significant ($F = 38.56$, $df = 2,191$; $p = 0.000$). The mean RUIS score for those with no pad/thin pad/tissue was 9.21; for those with medium pads it was 12.00 and for those using large pads it was 13.57. These group comparisons were statistically significant ($p < 0.05$).

The ICIQ-SF (Avery et al., 2004) contains items about whether particular symptoms associated with urinary leakage are present or absent (e.g. leaks before you can get to toilet, when you cough or sneeze, when asleep, when physically active or exercising, when finished urinating and are dressed, leaks for no obvious reason and leaking all the time). The number of these symptoms present (= 1) at baseline, plus an additional item concerning leaking without knowing about it, were added to form an index of leakage severity. From an examination of the data distribution the groups were classified as 2 or less symptoms; 3 and 4 symptoms, and 5 or more symptoms. There was a significant difference in RUIS pre-treatment scores for these groups. For those that had 2 or less symptoms the mean RUIS score was 8.60 for those with three and four leakage symptoms it was 11.46 and for those with 5 or more symptoms it was 13.21 ($F$ Welch = 43.84, $df = 2, 125.42$; $p = 0.000$). All group comparisons were significant. All the instruments discriminated significantly ($p < 0.05$) by the alternate index of severity but the RUIS was the most sensitive measure and had the highest F value whereas the ISI was the least sensitive measure for this analysis.

There was a significant difference in RUIS mean scores by type of treatment ($t = -2.76$, $df = 70.49$; $p = 0.0070$). Patients that were receiving conservative treatment (physiotherapy and continence advising) had a lower mean score ($M = 10.71$) compared with those that were receiving surgical treatment ($M = 12.03$).

These analyses reflect that the RUIS discriminates well between different levels of incontinence severity as measured by other clinical and symptom indicators.

When the RUIS was examined by baseline health status there were no significant differences between those in Excellent/Very Good, Good, or Fair/Poor Health ($F = 0.40$; $df = 2, 192$; $p > 0.05$); similarly there were no statistically significant associations by gender, age group, education level, work status, the number of co-morbidites or the length of time patients had experience incontinence symptoms. There was a statistically significant association between BMI and RUIS ($r_s = 0.20$, $p < 0.02$), but the common explained proportion of variance was very small (~4%). This would suggest the RUIS is assessing the underlying condition of urinary incontinence and this assessment appears to be independent of some common confounders.
### Table 4.2: Health and incontinence status at baseline in comparison with RUIS scores

<table>
<thead>
<tr>
<th>Variable</th>
<th>Classifications</th>
<th>N (%)</th>
<th>RUIS Mean (SD)</th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Male</td>
<td>N = 28; 16%</td>
<td>11.07 (4.18)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>N = 167; 86%</td>
<td>10.90 (3.81)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>N = 195</td>
<td>10.92 (3.33)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$t = 0.25, df 193; p = 0.80$</td>
</tr>
<tr>
<td>Age Groups</td>
<td>Less than 40 yrs</td>
<td>N = 18; 9%</td>
<td>11.72 (3.39)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>40-49</td>
<td>N = 36; 19%</td>
<td>10.92 (3.18)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>50-59</td>
<td>N = 47; 24%</td>
<td>10.94 (3.20)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>60-69</td>
<td>N = 62; 32%</td>
<td>10.69 (3.47)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>70 yrs or more</td>
<td>N = 32; 16%</td>
<td>10.91 (3.52)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$F = 0.33, df 4, 190; p = 0.66$</td>
</tr>
<tr>
<td>Number of Co-morbidities</td>
<td>0/1</td>
<td>N = 54; 28%</td>
<td>11.30 (3.13)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2/3/4</td>
<td>N = 97; 50%</td>
<td>10.70 (3.60)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5+</td>
<td>N = 44; 23%</td>
<td>10.95 (2.97)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$F = 0.55, df 2, 192; p = 0.58$</td>
</tr>
<tr>
<td>General Health Status</td>
<td>Excellent/Very good</td>
<td>N = 79; 41%</td>
<td>10.71 (3.13)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>N = 70; 36%</td>
<td>10.94 (3.51)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fair/Poor</td>
<td>N = 46; 24%</td>
<td>11.26 (3.44)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$F = 0.40, df 2, 192;p = 0.67$</td>
</tr>
<tr>
<td>Symptom Duration</td>
<td>Less than 2 years</td>
<td>N = 61; 35%</td>
<td>11.05 (3.24)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 years +</td>
<td>N = 109; 65%</td>
<td>11.20 (3.10)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$t = -0.30, df 168; p = 0.76$</td>
</tr>
<tr>
<td>Clinician Rated Incontinence</td>
<td>Mild</td>
<td>N = 65; 34%</td>
<td>9.22 (2.97)</td>
<td></td>
</tr>
<tr>
<td>Severity</td>
<td>Moderate</td>
<td>N = 99; 51%</td>
<td>11.79 (3.12)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Severe</td>
<td>N = 30; 15%</td>
<td>12.13 (2.68)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$F = 16.99, df 2, 191; p = 0.000$</td>
</tr>
<tr>
<td>Patient Rated Incontinence</td>
<td>Mild</td>
<td>N = 66; 34%</td>
<td>8.36 (2.78)</td>
<td></td>
</tr>
<tr>
<td>Severity</td>
<td>Moderate</td>
<td>N = 94; 48%</td>
<td>11.60 (2.85)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Severe</td>
<td>N = 34; 18%</td>
<td>14.03 (1.66)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$F_{Welch} = 80.46, df 2, 109.07; p = 0.000$</td>
</tr>
<tr>
<td>Pad Use</td>
<td>&lt;1 Pad per day</td>
<td>N = 92</td>
<td>8.06 (3.46)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>≥ 1 Pads per day</td>
<td>N = 102</td>
<td>11.55 (2.94)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$t = -6.22, df 192; p = 0.000$</td>
</tr>
<tr>
<td>Pad Size</td>
<td>No Pad/ Thin Pad</td>
<td>N = 97</td>
<td>9.21 (3.14)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Medium Pad</td>
<td>N = 60</td>
<td>12.00 (2.54)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Large Pad</td>
<td>N = 37</td>
<td>13.57 (2.45)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$F = 38.56, df 2, 191; p = 0.000$</td>
</tr>
<tr>
<td>UI Type</td>
<td>Urge Only</td>
<td>N = 62</td>
<td>9.00 (2.69)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stress Only</td>
<td>N = 33</td>
<td>10.94 (2.93)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mixed UI and SI</td>
<td>N = 84</td>
<td>11.79 (3.28)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$F = 9.83, df 2,176; p = 0.000$</td>
</tr>
<tr>
<td>Double Incontinence</td>
<td>UI only</td>
<td>N = 155</td>
<td>10.60 (3.37)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>UI and Fi</td>
<td>N = 40</td>
<td>12.18 (2.89)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$t = -2.71, df 193; p = 0.007$</td>
</tr>
<tr>
<td>Type of Treatment</td>
<td>Conservative</td>
<td>N = 151</td>
<td>10.71 (3.46)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Surgical</td>
<td>N = 33</td>
<td>12.03 (2.23)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$t = -2.76, df 70.49; p = 0.007$</td>
</tr>
<tr>
<td>No. of UI leak symptoms</td>
<td>2 or less</td>
<td>N = 70</td>
<td>8.60 (3.25)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3-4</td>
<td>N = 72</td>
<td>11.46 (2.61)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5+</td>
<td>N = 52</td>
<td>13.21 (2.16)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$F_{Welch} = 43.84, df 2, 125.42; p = 0.000$</td>
</tr>
</tbody>
</table>
4.2.4 Pre-treatment to Post-treatment Change Scores

At follow-up the mean RUIS score was 6.95, (SD = 4.77, N = 100). For females the mean was 6.91 and for males the mean was 7.21 (refer Figures 4.5a and 4.5b).

Table 4.3 shows RUIS scores in relation to other incontinence indicators at follow up and the statistics relating to these comparisons. When RUIS post-treatment scores were examined by the clinician post-treatment rating of incontinence severity there was a significant difference (F = 28.68, df 2, 94; p = 0.000) in mean scores by level of severity with those that the clinicians considered now to be 'normal' having a mean RUIS score of 2.72 as against a mean score of 11.29 for those still rated as 'severe'.

Similarly, when examined by post-treatment patient rated severity the RUIS mean for those in a normal state was 2.68; for those with mild incontinence it was 6.48, and for those with moderate and severe incontinence it was 12.34. This finding was statistically significant (F = 122.85; df 2, 97; p = 0.000) and all group comparisons were significant.

The patient’s ratings concerning the outcome of their treatment was that 17% indicated their incontinence had not been helped by treatment, 23% reported some improvement, 39% considered themselves to be partly cured and 20% indicated that their continence was cured. There was a significant difference in RUIS post treatment scores between these groups (F = 41.41, df 3, 94; p = 0.000). For those that were 'not helped' the mean was 12.29, for those that had 'some improvement' the mean was 9.35, the mean for those that were 'partly cured' was 5.63 and for those who considered themselves 'cured' it was 1.70. All group comparisons were significant.

In the case of the follow-up item concerning whether the patient still had problems with incontinence there was a significant difference (F = 124.65, df 3, 96; p = 0.000), in RUIS post-treatment mean scores between those that considered that they had major problems (M = 12.85), some problems (M = 8.83) slight problems (M = 5.12) and those that had no problems at all (M = 1.27) and all group comparisons were significant.
### Table 4.3: RUIS in relation to other incontinence variables at follow-up

<table>
<thead>
<tr>
<th>Variables</th>
<th>Classifications</th>
<th>(N)</th>
<th>Mean (SD)</th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>RUIS</td>
<td>Pre-treatment</td>
<td>N = 100</td>
<td>11.02 (3.08)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Post-treatment</td>
<td>N = 100</td>
<td>6.95 (4.77)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean Change Score</td>
<td></td>
<td>4.07 (4.76)</td>
<td>t paired = 8.56, df 99; p = 0.000</td>
</tr>
<tr>
<td>ISI</td>
<td>Pre-treatment</td>
<td>N = 100</td>
<td>6.44 (3.26)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Post-treatment</td>
<td>N = 100</td>
<td>3.74 (3.65)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean Change Score</td>
<td></td>
<td>2.70 (4.06)</td>
<td>t paired = 6.65, df 99; p = 0.000</td>
</tr>
<tr>
<td>UDI-6</td>
<td>Pre-treatment</td>
<td>N = 100</td>
<td>8.82 (3.3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Post-treatment</td>
<td>N = 100</td>
<td>5.46 (4.06)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean Change Score</td>
<td></td>
<td>3.36 (3.73)</td>
<td>t paired = 9.00, df 99; p = 0.000</td>
</tr>
<tr>
<td>ICIQ-SF</td>
<td>Pre-treatment</td>
<td>N = 99</td>
<td>11.81 (4.25)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Post-treatment</td>
<td>N = 99</td>
<td>8.03 (4.98)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean Change Score</td>
<td></td>
<td>3.78 (5.18)</td>
<td>t paired = 7.25, df 96; p = 0.000</td>
</tr>
<tr>
<td>RUIS post scores by Patient Rated Severity (post)</td>
<td>Moderate/Severe</td>
<td>N = 32</td>
<td>12.34 (2.62)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mild</td>
<td>N = 31</td>
<td>6.48 (2.50)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Normal</td>
<td>N = 37</td>
<td>2.68 (2.56)</td>
<td>F = 122.85, df 297; p = 0.000</td>
</tr>
<tr>
<td>RUIS post scores by Clinician Rated Severity (post)</td>
<td>Severe</td>
<td>N = 14</td>
<td>11.29 (4.01)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Moderate and Mild</td>
<td>N = 54</td>
<td>7.98 (4.25)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Normal</td>
<td>N = 29</td>
<td>2.72 (2.74)</td>
<td>F = 28.68, df 294; p = 0.000</td>
</tr>
<tr>
<td>RUIS post scores by Patient Rated Outcome</td>
<td>Not Helped</td>
<td>N = 17</td>
<td>12.29 (2.91)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Some Improvement</td>
<td>N = 23</td>
<td>9.35 (3.33)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Partly Cured</td>
<td>N = 38</td>
<td>5.63 (3.19)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cured</td>
<td>N = 20</td>
<td>1.70 (3.06)</td>
<td>F = 41.41, df 394; p = 0.000</td>
</tr>
<tr>
<td>RUIS post scores by Current Incontinence Problems (post)</td>
<td>Major Problems</td>
<td>N = 26</td>
<td>12.85 (2.34)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Some Problems</td>
<td>N = 18</td>
<td>8.83 (2.53)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Slight Problems</td>
<td>N = 34</td>
<td>5.12 (1.92)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No Problems</td>
<td>N = 22</td>
<td>1.27 (2.10)</td>
<td>F = 124.65, df 396; p = 0.000</td>
</tr>
<tr>
<td>RUIS Change by Patient Rated Outcome</td>
<td>Not Helped</td>
<td>N = 17</td>
<td>-0.88 (2.85)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Some Improvement</td>
<td>N = 23</td>
<td>1.48 (1.62)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Partly Cured</td>
<td>N = 38</td>
<td>4.97 (3.68)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cured</td>
<td>N = 20</td>
<td>9.85 (3.69)</td>
<td>F = 41.88, df 343.37; p = 0.000</td>
</tr>
<tr>
<td>RUIS Change by Patient Rated Improvement</td>
<td>Worse/ No Change</td>
<td>N = 22</td>
<td>-0.41 (2.84)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A Little Better /Much Better</td>
<td>N = 44</td>
<td>2.80 (3.16)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Very Much Better</td>
<td>N = 33</td>
<td>8.82 (3.44)</td>
<td>F = 43.53, df 395; p = 0.000</td>
</tr>
<tr>
<td>RUIS Change by Clinician Rated Improvement</td>
<td>Worse/ No Change</td>
<td>N = 12</td>
<td>0.42 (3.26)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Little and Much Better</td>
<td>N = 53</td>
<td>2.83 (3.90)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Very Much Better</td>
<td>N = 31</td>
<td>7.97 (4.34)</td>
<td>F = 22.48, df 293; p = 0.000</td>
</tr>
<tr>
<td>Patient Rated Improvement by Improved by &gt; 2 point RUIS Change</td>
<td>Improved by &gt; 2</td>
<td>N = 67</td>
<td>5.15 (1.05)</td>
<td>t = 6.44, df 45.92; p = 0.000</td>
</tr>
<tr>
<td></td>
<td>Improved by &lt; 2</td>
<td>N = 32</td>
<td>3.25 (1.50)</td>
<td></td>
</tr>
<tr>
<td>Clinician Rated Improvement by Improved by &gt; 2 point RUIS Change</td>
<td>Improved by &gt; 2</td>
<td>N = 66</td>
<td>5.15 (0.92)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Improved by &lt; 2</td>
<td>N = 30</td>
<td>4.07 (1.17)</td>
<td>t = 4.29, df 95; p = 0.000</td>
</tr>
<tr>
<td>RUIS Change by Baseline health status</td>
<td>Excellent/Very Good</td>
<td>N = 41</td>
<td>5.73 (4.51)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>N = 35</td>
<td>3.91 (4.70)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fair/Poor</td>
<td>N = 24</td>
<td>1.46 (4.17)</td>
<td>F = 6.87, df 2.97; p = 0.002</td>
</tr>
<tr>
<td>RUIS Change by Patient Rated Severity at Baseline</td>
<td>Mild</td>
<td>N = 31</td>
<td>3.32 (3.4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Moderate</td>
<td>N = 51</td>
<td>4.22 (5.21)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Severe</td>
<td>N = 18</td>
<td>4.94 (5.45)</td>
<td>F Welch = 0.85, df 2.43; p = 0.50</td>
</tr>
</tbody>
</table>
Change scores on the RUIS (pre-post) ranged from an improvement of 15 points to a deterioration of 8 points. Examination of pre-post scores revealed a statistically significant improvement of 4.07 points (paired t-test, t = 8.56, CI 3.13-5.01, df 99; p = 0.000).

To illustrate this change, the frequency distributions of RUIS scores pre and post-treatment for the sample are provided in Figures 4.4 and 4.5.

**Figure 4.5: RUIS pre-treatment and post-treatment scores**

![Figure 4.5a: RUIS pre-treatment scores for post treatment sample](image)

![Figure 4.5b: RUIS post-treatment scores](image)

<table>
<thead>
<tr>
<th>Mean</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-treatment</td>
<td>11.92</td>
<td>3.08</td>
</tr>
<tr>
<td>Post-treatment</td>
<td>6.95</td>
<td>4.77</td>
</tr>
</tbody>
</table>

When examined by pre-treatment incontinence severity status there were no statistically significant differences in change scores between the groups. This indicates there was no significant difference in the degree of improvement from treatment based on the initial patient severity rating as all groups improved to a similar level. The mean change scores were 3.32 points for the mild incontinence group, 4.22 points for the moderate group and 4.94 points for the severe group (F Welch = 0.85, df 2, 43.25; p>0.05).

RUIS change scores did not significantly vary by gender, education or age. They did, however, vary by health status with those reporting at baseline that they experienced excellent/very good health obtaining the greater benefit from treatment than those in fair or poor health. The mean change scores were 5.73 points for those in excellent/very good health, 3.91 for those in good health and 1.46 for those in fair/poor health (F = 6.87, df 2, 97; p<0.01).

### 4.2.5 Reliability Data

The RUIS pre-treatment internal consistency reliability was Cronbach’s alpha = 0.73 (N = 195 urinary incontinence patients) which is considered adequate (Streiner and Norman, 2003). The pre-treatment alphas for UDI-6, a scale with 6 items, alpha = 0.64, for the ICIQ-SF it was 0.65 and for the ISI, which has only 2 items, alpha = 0.54. Streiner and Norman (2003) considered the latter levels to be inadequate as they are less than 0.70.

It has been suggested that Cronbach’s alpha can be influenced by outliers (Liu et al., 2010) the RUIS was checked for outliers in the urinary pre-treatment sample and it was noted there was 1 outlier identified that scored 0. The Cronbach’s alpha was calculated on 194 cases with the outlier removed and alpha = 0.72 which remains adequate (Streiner and Norman, 2003).

For the sample of all urinary and faecal incontinence patients combined there were no outliers and the RUIS pre-treatment alpha was 0.84 (N = 254) which is considered to be very good. The RUIS post-treatment alpha = 0.90, N = 100 which is similar to the post-treatment alpha (0.85) reported by Hawthorne et al. (2006).
Test-retest reliability was assessed at two weeks after the submission of post-treatment forms. Test-retest reliability was assessed using the Intra Class Correlation Coefficient. The RUIS test-retest ICC = 0.80 (N = 78; all incontinence retest sample) and ICC = 0.77 (N = 60; urinary incontinence retest sample). For the UDI-6 ICC = 0.74; for the ICIQ-SF ICC = 0.67 and for the ISI it was 0.76 (urinary retest sample). A test-retest reliability estimate over 0.70 is considered to be acceptable (Streiner and Norman, 2003) which indicates the RUIS test-retest reliability is adequate to good but the test-retest reliability of the ICIQ-SF could be considered inadequate.

The standard error of measurement (SEM) is related to reliability and estimates the extent, to which a test provides accurate scores (or the observed score reflects the true score), given there will always be some degree of variation in scores in repeated measures using same test. A lower SEM indicates greater precision in measurement. Using Cronbach’s alphas for the urinary sample the SEM for the RUIS = 1.73, for the UDI-6 it was 2.09, for the ISI it was 2.32 and for the ICIQ-SF it was 2.79.

### 4.2.6 Ceiling and Floor Effects

At pre-treatment scores on the RUIS ranged from 0 – 16 (1 outlier had a score of 0) indicating the full range of scores were used. There was 0.5% of the sample on the floor of the instrument (lowest possible score) and 5.6% of the sample at the ceiling of the scale (maximum possible score). Generally a figure of less than 15% of the sample at the floor or the ceiling is preferred (Streiner and Norman, 2003). It is clear that this instrument does not suffer from significant floor or ceiling effects.

### 4.2.7 Responsiveness: Capacity to Detect Change

RUIS change scores were calculated by subtracting the post-treatment score from the pre-treatment score and thus a reduction in scores on this scale is indicative of an improvement. Table 4.3 shows there was a significant improvement of 4.07 RUIS scores (SD = 4.76, paired t-test, t = 8.56, df 99; p = 0.000) following treatment. RUIS change scores are depicted in Figure 4.6 below which indicated the majority of patients improved following treatment.

![Figure 4.6: RUIS Change Scores](image)

**Figure 4.6: RUIS Change Scores**

![Image of bar chart showing RUIS change scores]

Dark Blue = Improved; White = Little or No Change; Light Blue = Worse

The UDI-6 scale, which is an 18 point scale vs. a 16 point scale, with 6 items had a mean improvement of 3.36 UDI-6 scores (SD = 3.73, paired t-test, t = 9.00, df 99; p = 0.000). These results indicate that the RUIS is similar to the UDI-6 in detecting change with a very similar t value
despite the lesser range of scores. The ICIQ-SF is a 21 point scale based on 3 items and the mean change is 3.74 scores (paired t test = 7.25, df 98; p = 0.000). The ISI, which is a 12 point scale based on 2 items, also showed a significant improvement of 2.7 ISI scores (SD = 4.06, paired t-test, t = 6.65, df 99; p = 0.000) but it is the least sensitive instrument for detecting change as a result of treatment - although this instrument is mainly used in population rather than clinical studies. The relative efficiency of the RUIS in comparison to the ISI was RE = 1.66 indicating it was 66% more likely to detect change than the ISI and the relative efficiency of the UDI-6 was 1.83. The UDI was slightly more sensitive than the RUIS but it is also a slightly longer scale. The RE for the ICIQ-SF = 1.19 indicating it is more sensitive to change than the ISI but less sensitive to change than the UDI-6 and the RUIS. The Wei Symptom Scale was not considered for these comparisons as it is a longer 8 item (32 point) scale containing a large degree of item redundancy by containing 3 items each for urge and stress incontinence which may inflate the estimates derived from this measure.

The Kazis effect size (ES) is the mean change score divided by the SD of the baseline score and it was (-1.32). This is classed as a large effect size as it is over 0.8 (Kazis et al., 1989; Cohen, 1988; Crosby et al., 2003). The Kazis ES for the UDI-6 was -1.02, for the ICIQ-SF it was -0.89 and for the ISI it was -0.82 (note for all these measures a reduction in scores indicates an improvement). This indicated that all measures were responsive over time but the RUIS had the largest effect size of all these measures.

Another estimate of effect size is the standardised response mean (Deyo et al., 1991) which is the change in mean scores divided by the SD of the change = 0.86 which is also considered good (Crosby et al., 2003). Pallant (2011) suggests the eta squared statistic can be used to estimate effect size and this (0.49) also indicated a large effect size (Cohen 1988) as it is greater than 0.14. These provide further evidence that the RUIS is responsive to change.

The responsiveness of the instrument is also concerned with the capacity of the instrument to detect change regardless of whether it is improvement or deterioration. Ignoring signs indicating the direction of change the average change score for the sample was 4.83 RUIS points SD = 3.97 N = 100 (as compared with UDI-6 = 4.12 points). Change scores on the RUIS (pre-post) ranged from an improvement of 15 points to a deterioration of 8 points in the context of a 16 point scale. These findings would also suggest that the instrument has the capacity to detect both an improvement and deterioration in patient incontinence status.

RUIS change scores were examined by the patient global rating of improvement (Yalcin and Bump, 2003) which is a 7 point scale. These categories were combined to form 3 groups (no change through much worse/ a little and much better/ very much better). There was a significant difference in RUIS mean change scores in relation to the patient rating of global improvement (F = 61.72, df 2, 96; p = 0.000). The no change to worse group had a RUIS mean change score of -0.41, the little to much better group had a mean change score of 2.80 and the very much better group had a mean change score of 8.82 and all these comparisons were significant at p< 0.05 level. Similarly there was a significant difference in clinician ratings of global improvement in relation to RUIS change scores (F = 22.48, df 2, 93; p = 0.000). The Spearman correlation between RUIS change scores and the patient global rating of improvement $r_s = 0.77$ (p = 0.000) and with the clinician ratings of global improvement it was $r_s = 0.61$ (p = 0.000).

RUIS change scores were also examined by the patient rating of the treatment outcome from incontinence not helped through some improvement, partly cured and cured and there was a significant difference (F Welch = 41.88, df 3, 43.37; p = 0.000) between the means. The mean RUIS change score for those ‘not helped’ was -0.88, for ‘some improvement’ it was 1.48, for ‘partly cured’ the mean RUIS change score was 4.97 and for ‘cured’ it was 9.85. The comparisons between the not helped, partly cured and cured groups were all significant at the p<0.05 level. The comparison between the not helped and the some improvement group showed a trend (p = 0.10) but it was not significant at the p< 0.05 level.
Figure 4.7: RUIS change scores by Patient Global Improvement

Figure 4.7 might suggest that an important improvement may lie between the RUIS change means for the ‘a little better’ (M = 1.65) and the ‘much better’ groups (M = 3.32). RUIS change scores were reclassified into 2 groups where there was an improvement greater than 2 RUIS scores and where there was an improvement of less than 2 RUIS scores (including deterioration). These RUIS change groups were analysed by the patient global rating of improvement (a seven point scale, refer Table 4.3) and there was a significant difference (t = 6.44, df 45.92). The mean on the PGI for the improved by 2 RFIS scores group was 5.15 and the mean for the less improved group was 3.25. This suggested that a change of >2 RFIS scores was associated with a change in the patients’ average rating of their improvement. Similarly Figure 4.7 indicates that those rating themselves much and very much worse have an average deterioration of 2 RUIS change scores although the data was sparse.

A similar analysis was undertaken using the clinician’s rating of patient global improvement. Figure 4.8 depicts RIUS change scores in relation to the clinician’s rating of improvement.

Figure 4.8 might also suggest that an important improvement may lie between the RUIS change means for the ‘little better’ (M = 1.10) and the ‘much better’ groups (M = 4.10). RUIS change scores were reclassified into 2 groups where there was an improvement greater than 2 RUIS scores and where there was an improvement of less than 2 RUIS scores (including deterioration). These RUIS change groups were analysed by the clinicians’ global rating of improvement (continuous scores) and there was a significant difference (t = 4.29, df 95; p = 0.000). The mean on the PGI for the improved by 2 RUIS scores group was 5.15 (SD 0.92, CI 4.93-5.38, N = 66) and the mean for the less improved group was 4.07 (SD 1.17, CI 3.63-4.50, N = 30) on the clinician rating of improvement.

These analyses might suggest that a change in the vicinity of <2 RUIS scores may be the minimal detectable difference concerning change in the patients’ and clinicians’ global perception of improvement, at the group level.
The SEM for the RUIS of 1.73 (95% CI SEM = 3.39) would suggest that a score change of 3-4 points is likely to be a more clinically and statistically reliable estimate for the purposes of patient monitoring although it is noted that the SEM reported is for the instrument as a whole and standard errors of measurement vary by score levels (they more be more or less for people scoring at the extremes of the scale) (Spratt, 2009).

Further analyses, using a larger post-treatment sample combined with the use of other clinical indicators as anchor points, will be required to further address this issue.

4.2.8 Internal Structure: Principal Components Analysis

The sample sizes were sufficient for analysis with a sample of 195 urinary patients at pre-treatment (Gaudagnoli and Velicer, 1988; Pallant, 2011).

For all analyses presented below the Kaiser-Meyer-Olkin value was >0.6 (Kaiser 1970, 1974) and Bartlett’s Test of Sphericity (Bartlett 1954) reached statistical significance, supporting the factorability of the correlation matrices.

The 5 items of the RUIS were analysed (N = 195, urinary incontinence sample) using a Principal Component Analysis which is presented in Table 4.4 below. There was only 1 factor/component extracted with an eigenvalue over 1 (eigenvalue = 2.43) and this could be described as a general urinary incontinence/leakage factor explaining 48.62% of the variance. All RUIS items had loadings of 0.64 or above on this factor indicating a moderate to high level (0.69) of component saturation (Gaudagnoli and Velicer, 1988). The analysis of the community survey data produces a similar structure with a similar pattern of item factor loadings. The one component extracted explains more of the variance (74%) but this may be explained by the greater homogeneity of this sample.
Table 4.4: Principal component matrix for RUIS Items: urinary incontinence patients

<table>
<thead>
<tr>
<th>RUIS Items</th>
<th>Urinary Incontinence Patients (N = 195)</th>
<th>Community Survey (N = 2915)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Factor 1 (49%)*</td>
<td>Factor 1 (74%)*</td>
</tr>
<tr>
<td>RUIS1 - Urge</td>
<td>0.64</td>
<td>0.80</td>
</tr>
<tr>
<td>RUIS2 - Stress</td>
<td>0.67</td>
<td>0.83</td>
</tr>
<tr>
<td>RUIS3 – leak small amounts</td>
<td>0.80</td>
<td>0.88</td>
</tr>
<tr>
<td>RUIS4 – leak frequency</td>
<td>0.72</td>
<td>0.91</td>
</tr>
<tr>
<td>RUIS5 – leak volume</td>
<td>0.64</td>
<td>0.90</td>
</tr>
</tbody>
</table>

* = proportion of variance explained

In 2006 (Sansoni et al., 2006) the authors analysed a number of urinary incontinence items which included the 5 RUIS items (3 from UDI-6 and 2 from ISI), the 3 other UDI items concerning the frequency of urination, pain, and difficulty emptying the bladder. This analysis was repeated using the clinical sample of urinary patients (N = 195).

The 2006 analysis (N = 2,915) produced a 2 factor/component solution accounting for 67% of the variance with all RUIS items loading highly (loadings of 0.6 or above) on factor/component 1 (eigenvalue = 4.29). This was described as a ‘urinary leakage factor’ and it explained 54% of the variance. The other items concerning lower abdominal pain and emptying bladder from the UDI loaded on factor/component 2 (explaining 13 % of the variance) and this was described as ‘other bladder symptoms’. The frequency of urination item had low to moderate loadings on both factors (0.48 on the urinary leakage factor and 0.49 on the other bladder symptoms factor) which suggested it did not warrant inclusion in the scale. In 2006 (Sansoni et al., 2006) the RUIS items were selected with regard to their loadings on the primary urinary incontinence factor/component. This pattern of factor/component loadings was replicated in this recent analysis of urinary incontinence patients (N = 195) which confirms the current descriptive system.

Table 4.5: Rotated factor matrices for urinary incontinence items

<table>
<thead>
<tr>
<th></th>
<th>Community Survey 2006 N = 2915</th>
<th>Clinical Survey 2011 N = 195</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale</td>
<td>Factor/Component 1 2</td>
<td>Factor/Component 1 2</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Freq Urination</td>
<td>0.48</td>
<td>0.49</td>
</tr>
<tr>
<td>Urgency Leakage</td>
<td>0.74</td>
<td>0.60</td>
</tr>
<tr>
<td>Stress Leakage</td>
<td>RUIS1</td>
<td>RUIS2</td>
</tr>
<tr>
<td>Leak Small Amount</td>
<td>RUIS3</td>
<td>0.77</td>
</tr>
<tr>
<td>Emptying Bladder</td>
<td></td>
<td>0.76</td>
</tr>
<tr>
<td>Pain Lower Abdominal</td>
<td></td>
<td>0.75</td>
</tr>
<tr>
<td>Leakage Frequency</td>
<td>RUIS4</td>
<td>0.89</td>
</tr>
<tr>
<td>Leakage Amount</td>
<td>RUIS5</td>
<td>0.89</td>
</tr>
</tbody>
</table>

Note: Principal Components Analyses with Varimax Rotation were employed in both studies.

4.2.9 Validity: Correlations with Other Measures:

The RUIS was correlated with a number of other measures of urinary incontinence. Pre-Treatment correlations between RUIS and UDI-6 r = 0.76 (p<0.01); RUIS and ISI r = 0.76 (p<0.01); RUIS and ICIQ-SF r = 0.74 (<0.01); RUIS and WEI Symptom Scale (urinary) r = 0.72 (p<0.01); WEI Bother Scale r = 0.66 p<0.01; and RUIS and Incontinence Impact Questionnaire r = 0.53 (p<0.01). These correlations were high, significant and in the expected directions providing good evidence of construct validity for the RUIS. As may be expected the RUIS correlations with other urinary incontinence symptom measures were higher than with the Incontinence Impact Questionnaire.
and the WEI Bother Scale which are scales that are more concerned with the impact of urinary incontinence on daily activities rather than the extent of the symptom itself.

At pre-treatment the correlation of the RUIS with patient rated severity was \( r_s = 0.62 \) (\( p<0.01 \)). However, the correlation between the RUIS and the pre-treatment clinician rating of severity, although significant \( p<0.01 \), was somewhat lower \( r_s = 0.37 \) although it had the highest correlation of the urinary incontinence measures. The correlations between the clinician rating and the WEI Symptom Scale \( r_s = 0.36 \); UDI-6 \( r_s = 0.29 \); ICIQ-SF \( r_s = 0.26 \) and the ISI \( r_s = 0.33 \). The correlation between patient rated severity and clinician rated severity was also low but significant at \( r_s = 0.36 \) (<.01). This pattern is replicated with the other urinary incontinence symptom measures – they all correlate highly with each other (0.72 or above) and with patient rated severity (0.59 – 0.68) but have lower correlations (0.29 -0.37) with clinician rated severity at pre-treatment. An analysis examined which urinary incontinence items correlated most highly with the clinician’s pre-treatment severity rating and it was found to be frequency of urinary leakage \( r_s = 0.36 \) and an item concerning how often one leaked urine without knowing about it \( r_s = 0.35 \). The latter item is not contained in the 4 scales that were examined above. The clinician pre-treatment ratings also have slightly higher and more significant associations with variables such as BMI, weight, the Physical Function Index of the SF-36 and a global assessment of health status than do the other incontinence scales suggesting that possibly the clinician’s ratings may be influenced by their overall perception of the health status of the client rather than only the presence of urinary incontinence symptoms. Clinicians also suggested that their severity ratings might be more influenced by their assessment of the biological cause of the incontinence rather than its extent.

For the urinary sample at pre-treatment the RUIS correlation with the Physical Function Scale (PFI) of the SF-36V2 was \( r = -0.15 \) (\( p<0.05 \)) (e.g. the higher the RUIS score the lower the PFI score). The correlation with the Physical Component Summary (PCS) scale was \( r = 0.11 \) (\( p = 0.18 \)) and for the Mental Component Summary (MCS) it was \( r = -0.17 \) (\( p = 0.03 \)). These are similar findings to the Sansoni et al. (2006) study which reported a low negative correlation between measures of health related quality of life and urinary incontinence.

At post-treatment the RUIS correlation with untransformed patient satisfaction scores (SAPS) was \( r = -0.44 \) (\( p<0.00 \)) which indicated there was an association between higher RUIS scores (which indicates greater incontinence) and lower patient satisfaction scores (less satisfaction).

The change scores on RUIS were classified into 2 groups to analyse SAPS data further. When change scores were classified into worse / no change (a deterioration of 8 to 0, \( N = 22 \)) and improved (an improvement of 1 point or more, \( N = 78 \)) there was no significant difference between SAPS scores and this classification of RUIS change (\( t = -1.71, df 98; p = 0.20 \)). However, if this classification of no change or worse was altered to include those that only had an improvement of 1, which could be considered little change (and thus ‘improved’ was classed as an improvement of 2 points or more) there was a significant difference between SAPS means for these groups (\( t = -2.26, df 98; p = 0.03 \)) with the RUIS improvers group having a higher SAPS mean (22.46, \( N = 68 \)) than the worse/little change group (20.06, \( N = 32 \)).

This provides further evidence that the RUIS can discriminate well between levels in related variables, in an ordered way and in the expected directions.

### 4.2.10 Cutpoints for Interpretation

It would be advisable to refer people who score 4 or more on this scale for further continence assessment. This is two standard deviations below the mean for the total clinical sample but one standard deviation above the average score (\( M = 1.74; \ SD = 2.93 \)) found in the community survey sample. To obtain this score one would need to endorse either ‘slightly’ or ‘rarely’ on most incontinence items.

An analysis of the distribution of the scores of the clinical sample indicated that a score of 4 to 8 is considered to reflect mild incontinence. A score of 8 was the 23rd percentile for the clinical sample and scores of 8 or below might be considered ‘mild’, a score of 10 was the 50th percentile which might be considered ‘moderate’ and a score of 14 represented the 75th percentile which might be considered ‘severe’. 
The clinician pre-treatment ratings indicated that a RUIS score of 9 (M = 9.22) or below was considered ‘mild’, a score of 10 -12 was considered ‘moderate’ (M = 11.79) and a score above 12 (M = 12.13) was classified as severe. The patient pre-treatment ratings indicated that a score of 8-9 or below was considered ‘mild’ (M = 8.36), a score of 10 through 12 was considered ‘moderate’ (M = 11.60) and a score of 13 or above was considered ‘severe’ (M = 14.03). At post–treatment a score of 3 or less was classified as ‘normal’ by the clinicians’ and by the patient’s rating of severity.

The ICIQ-SF scale contains 3 items (scale range is from 0-21) concerning the frequency and the volume of urinary leakage and the interference this causes in everyday life. The ICIQ-SF total scores, were tertiled to form three groups based on scores +/- 1 standard deviation from the mean (low scores 0-6, N = 37 / moderate scores 7-16, N =125 / high scores 17-21, N = 29). The RUIS means for these three groups were then examined. The low ICIQ-SF score group had a RUIS mean of 7.22 (SD = 3.01, CI 6.21-8.22), the moderate score ICIQ-SF group had a RUIS mean of 11.26 (SD = 2.72, CI 10.74-11.74) and the high score ICIQ-SF group had a RUIS mean of 14.21 (SD = 1.47, CI 13.65-14.77). This would suggest that RUIS scores of 5 through 8 could be considered to reflect ‘mild incontinence’; that scores from 9 through 12 might reflect ‘moderate incontinence’, and that scores of 13 or above might reflect ‘severe’ incontinence.

At pre-treatment the RUIS mean for those wore less than 1 pad per day was 8.06 and the mean for those that wore 1 pad per day or more was 11.55 ( t = -6.22, df 192, p = 0.000). At post-treatment the mean for those that did not wear pads was 3.74 and the mean for those that were still wearing pads at post-treatment was 9.69 ( t = -7.92, df 98; p = 0.000) This suggests that a useful cutpoint based on behavioural anchor points would be a RUIS score of <9 to distinguish between those patients whose incontinence was not sufficiently severe to wear pads (mild incontinence) and those who needed to wear pads (moderate incontinence).

Given the above the following RUIS score ranges are suggested to aide interpretation:

0-3: no urinary incontinence or extremely mild (‘slight’) or occasional (‘rare’) incontinence
4-8: mild urinary incontinence
9-12: moderate urinary incontinence
13-16: severe urinary incontinence (scores of 15 – 16 could be considered very severe).

The community survey dataset (SAHOS 2004) was reanalysed with respect to the proposed RUIS cutpoint of 4 to distinguish between those that have no incontinence and those with ‘mild’ symptoms of incontinence. Using this cutpoint the prevalence of urinary incontinence would be estimated at 18.4% overall; for females it would be 29.9% and for males it would be 6.3%. These are somewhat lower prevalence estimates than those provided from an analysis of ISI scores (refer Section 4.6 on Incontinence and the Burden of Disease) and in other reports (Hawthorne, 2006). However, it is interesting to note that if a RUIS cutpoint of 3 is used the prevalence estimates derived form ISI and RUIS are very similar (ISI: 24% overall, 38% for females and 10% for males) and this cutpoint may be preferred for prevalence studies. The RUIS cutpoint of 4 in routine practice, however, would exclude more cases with very few and slight symptoms of incontinence and it is noted above that both clinicians and patients rated RUIS scores of 3 and below at post-treatment as normal/no incontinence.

4.2.11 Effect of Treatment

As shown by Table 4.3 the examination of pre-post RUIS scores revealed a statistically significant improvement of 4.07 points (SD = 4.76, N = 100) (paired t-test, t = 8.56, df 99; p = 0.000). There was a statistically significant difference in the degree of improvement by treatment type (F = 9.35, df 2, 94; p = 0.000) and it is noted the 3 patients receiving medicine as treatment were excluded from this analysis. For those receiving continence advising the mean RUIS change score was 2.00 (SD = 2.68, CI 0.38-3.62, N = 13), for physiotherapy it was 3.09 (SD = 3.94, CI 2.03-4.16, N = 55) and for surgery it was 7.07 (SD = 5.76, CI 4.88-9.26 N = 29). There was a greater degree of improvement in post-treatment RUIS scores for those receiving surgical rather than conservative treatments although all treatment groups showed improvement. This is despite the fact one may expect some rate of complications (2% risk of voiding difficulty and 5% risk of urge incontinence).
for surgical treatments. However, it should be noted that the change in mean scores for all treatment types from pre-treatment to post-treatment was significant (e.g. NCA mean difference = 2.00; t = 2.69, df 12; p < 0.02).

4.2.12 Alternative Solutions

The RUIS and UDI items concerning stress incontinence, urge incontinence, and the leakage of small amounts of urine all ask the patient if they experience and how much they are bothered by the symptom and this could be considered to be a double barrelled question albeit that it is asking for a global assessment concerning the symptom. Some alternative wording of these questions was tested in the clinical protocol. For example RUIS 1 asks:

**Do you experience and if so how much are you bothered by:**

**Urine leakage related to the feeling of urgency?**

- Not at all
- Slightly
- Moderately
- Greatly

As an example the alternative wording for this item became:

1a. **Do you experience urine leakage related to the feeling of urgency?**

- Never
- Rarely, i.e. less than once in the past four weeks
- Sometimes, i.e. less than once a week, but once or more in the past four weeks
- Often or usually, i.e. less than once a day but once a week or more
- Always, i.e. once or more per day

1b. **If so, how much does this bother you?**

- Not at all
- Slightly
- Moderately
- Greatly

Overall, there are a number of issues that were considered to be problematic with the reworded alternative items. One issue is that these questions contain 2 components and thus to add them into the scale will effectively increase the weight of these items. Alternatively, one can combine the scoring of the items by adding the experience and bother components to form a seven point scale but this will also place a greater weight on these items in forming the total scale score than for the other items that were coded 0-3 or 0-4.

Another issue was that if a patient endorsed ‘never’ for the experience of the symptom then some patients did not complete the ‘bother’ aspect of the question thus missing data was a problem with these restructured items (17 elements of missing data applied to this issue). In order to facilitate data analysis of these items this missing data was recoded as 1 = not at all for such cases as the assumption was made that if the patient did not have the symptom they were not likely to be botherred by it. However, for use in routine practice the degree of missing data could well be of concern.

The correlation between the experience and bother components of these items were high (around 0.7) and significant (p<0.01) and the correlations between the experience component with the original RUIS items ranged from 0.65 – 0.77 and for the bother component they ranged from 0.75 - 0.80. When the scoring of experience and bother components was combined for the reworded items, by adding the scores for these two elements to form a 7 point response scale, the correlations with the corresponding original RUIS items ranged from 0.74 to 0.80 indicating there is a lot of overlap between these items, as might be expected. The internal consistency reliability was examined by including these three reworded and rescored items with RUIS item 4 and RUIS item 5 and this was named ALT RUIS which had a score range from 0-28. The Cronbach’s alpha that
resulted was the same as for the original RUIS. However the t test values for pre and post comparisons for ALT RUIS were much lower as given the alternative items have a 7 point scale the Standard Deviations were of course much higher and this is reflected in a much lower t value for the pre and post comparisons. The mean change score for ALT RUIS was 4.11, the SD = 9.67, t = 4.23 whereas for the original RUIS the mean change score was 4.07, the SD = 4.76 and the t value was 8.56. In smaller samples the increased variability in these item scores would lessen the likelihood of obtaining significant analysis values and this in effect reduces the sensitivity to change of the ALT RUIS as compared with the original RUIS.

The internal consistency reliability was also examined by including just the experience component of the reworded items (B1a, 2a, 3a) with RUIS 4 and 5. This resulted in a Cronbach’s alpha of 0.726 – much the same than for the original RUIS.

Thus there seems to be no advantage gained by substituting these items for the original RUIS items and there are disadvantages such as higher rates of missing data, a longer scale, more complex scoring and less sensitivity to change as discussed above. The original RUIS has performed well so there does not appear to be a good case to change it.

It was suggested by the Steering Group that some additional items concerning stress and urge incontinence and the frequency of urinary leakage could be included in the clinical dataset to check how well these items performed within the RUIS and to see if better items covering these aspects could be substituted.

The item on stress incontinence (RUIS 2) was significantly correlated (r = 0.36 - 0.62; p<0.01) with three stress items from the Wei Incontinence Symptom index (Wei et al 2006) and stress incontinence at pre-treatment as rated by the clinician. The RUIS 2 stress item had the highest correlation (r_s = 0.48, p< 0.01) of these ‘stress’ items with the clinician’s stress incontinence diagnosis. An analysis of variance indicated that there was a significant difference (t = -5.07, df 198; p = 0.000) in RUIS 2 scores by the clinician rating of stress-type (e.g. between does or does not have stress incontinence). A factor analysis of all the stress items contained in the protocol also indicated that the RUIS 2 item had a high loading on factor 1 which could be described as the stress incontinence factor.

Earlier an analysis had indicated that there was a significant gender difference in the pattern of responses to the RUIS stress item with men more likely to endorse ‘never’ and less likely to endorse ‘slightly’ compared with women, although there were no significant gender difference between the means for this item. The WEI stress incontinence items did not exhibit this response pattern feature. Using internal consistency reliability analyses the three WEI stress items were substituted in turn for the RUIS stress item and it was found the alpha coefficients for the scale dropped from 0.73 to 0.70 to 0.71. Thus there was no real benefit in substituting any of these items for RUIS 2.

The item on urge incontinence (RUIS 1) was correlated with an alternatively worded RUIS 1 item, the urge items from the Wei Incontinence Symptom Index and the clinician identification of urge incontinence. All items had significant correlations (p<0.01) with the clinician diagnosis of urge incontinence but the original RUIS urge item had a lower correlation (r_s = 0.24) than the best Wei item B28 (r_s = 0.33) and the reworded RUIS 1 item (r_s = 0.31). However, it should be noted the latter items have a greater score range (0-4) than the RUIS 1 urge item (0-3). Given this both of these items in turn replaced RUIS1 in an internal consistency reliability analysis. When the alternatively worded RUIS item (B1a and b) was included in the RUIS the Cronbach’s alpha dropped from 0.73 to 0.68. Similarly when the best Wei urge item (B28) was included the Cronbach’s alpha dropped to 0.69. As Streiner and Norman (2003) indicate an alpha less than 0.7 is considered inadequate so there appears that there is little to be gained by changing this item in the RUIS. A factor analysis of all the urge items in the patient protocol also indicated that the RUIS urge item had a high loading of 0.8 on factor 1 which can be described as the urge incontinence factor. The original RUIS urge item also had a higher correlation with patient and clinician rated severity at pre-treatment than the alternative RUIS urge item.

Given the relatively low correlation between the clinician’s patient severity rating at baseline and the three standardised measures of urinary incontinence an analysis above had examined which
urinary incontinence item correlated most highly with the clinician’s pre-treatment severity rating and it was found to be an item concerning how often one leaked urine without knowing about it. Given the correlation of the item on ‘leaking urine without knowing it’ (LNK) and the clinician’s pre-treatment severity rating ($r_s = 0.35$) some additional analyses were undertaken to see if the inclusion of this item would improve the RUIS. The first issue concerning this variable is that 44.3% of the sample do not experience the symptom at pre-treatment which contrasts with other RUIS items where 0 -11% of the sample do not experience the particular symptom.

An internal consistency analysis was undertaken with the other 5 RUIS items and this variable which would make the RUIS a 6 item scale with a range from 0-20. This version was called RUIS-LNK. The mean and the standard deviation for the new scale would be 11.88 (SD = 3.64) versus 10.82 (SD = 3.33) for the original RUIS. The alpha for this scale would be 0.75 which is slightly higher than for the original RUIS (0.73). However, it should be noted that alpha is not only dependent on the magnitude of the correlations among items, but also on the number of items in the scale. A scale can be made to look more homogenous simply by the addition of extra items, even though the average correlation remains the same (Streiner and Norman, 2003). In this case an additional item was added and so the slight difference in the alphas might be influenced by just the addition of an extra item. To see the effect on alpha in a 5 item scale we included this item with only 4 of the original RUIS items – replacing RUIS 3 (leak small amounts), on logical grounds as it is also a leakage item, with the LNK item. In this analysis the alpha dropped to an unacceptable level of 0.67.

The RUIS-LNK was also examined with regard to the capacity to detect change. The mean change score for the 6 item scale is 4.31 (SD = 4.8) and the $t = 8.84$ (df 95; $p<0.001$) as compared with the RUIS a 5 item scale where the mean change score is 4.07 (SD = 4.76) and the $t = 8.56$ (df 99; $p<0.001$) and thus there is a negligible difference between these alternative scales concerning their capacity to detect change.

Given consideration of all the analyses, and given that only 40% of the clinical sample experience the symptom it was considered there was little to be gained by adding this item to the RUIS. RUIS-LNK would also have adequate psychometric properties but it does not provide sufficiently better psychometric properties to warrant the change.

4.2.13 Interpretation of Results

The RUIS demonstrated good internal consistency reliability and test-retest reliability. The RUIS was sensitive to change as a result of treatment for both males and females indicating it can be used to assess patient outcomes. The RUIS appears to be a uni-dimensional measure and it has an appropriate internal structure with all items loading highly on the primary urinary leakage factor. These findings have been confirmed in both population and clinical settings.

There were no gender differences detected in total RUIS scores at pre-treatment and post-treatment suggesting it may be appropriate for use with both genders although it is noted that the sample size of the males is small and this needs to be confirmed with further samples.

The RUIS possessed evaluative discrimination in relation to both clinician and patient assessed incontinence severity and also in relation to other clinical indications of incontinence severity such as pad use. It did not discriminate by some unrelated health or socio-demographic variables (e.g. education, work status, co morbidity number etc.) whereas it did discriminate by general health status and BMI. Similarly, the RUIS appeared to be responsive over time to changes in incontinence status. These two findings suggested it has both content and construct validity; i.e. it assesses the underlying condition of urinary incontinence and this assessment appears to be independent of some possible common confounders.

The correlations between the RUIS and other urinary incontinence measures were significant and in the expected directions providing validation for this measure. With regard to health related quality of life the RUIS has a low negative correlation with Physical Function Index of the SF-36V2 which is consistent with findings in the literature showing a negative association between urinary incontinence and generic measures of HRQOL.
4.2.14 Conclusions and Recommendations: Urinary Incontinence

Initial indications are that the RUIS performed well in clinical settings and demonstrated: adequate to good internal consistency reliability; correlations with other measures were in the expected directions indicating construct validity, and it discriminated well between groups varying in incontinence severity. There is good evidence that it was sensitive to changes in continence status. It is a short, valid and reliable instrument suitable for use in routine practice in clinical settings and should also be considered by researchers and epidemiologists in population health settings.

4.3 RFIS and Faecal Incontinence: Results

A copy of the RFIS is at Attachment A. At pre-treatment, a total of 61 patients (10 males and 51 females) have RFIS scores entered on the research database and there are 39 patients for whom post-treatment data was available. The RFIS consists of the following self-report items:

In the past 4 weeks:
- Do you leak, have accidents or lose control with solid stool? (Wexner)
- Do you leak, have accidents or lose control with liquid stool? (Wexner)
- Do you leak stool if you don’t get to the toilet in time?
- Does stool leak so that you have to change your underwear?
- Does bowel or stool leakage cause you to alter your lifestyle? (Wexner)

Scoring is by simple summation, and the range is 0-20 with 0 indicating no incontinence.

4.3.1 Characteristics of the Faecal Incontinence Sample

Table 4.6 shows the demographic details of the participants. Most participants were Australian born, spoke English as their primary language, were middle aged, had achieved completion of high school or a trade qualification, were retired and were female. The average age of patients was 60.97 years – a somewhat older group than for urinary incontinence.

With regard to type of faecal incontinence, 58% experienced passive incontinence symptoms, 50% experienced urge incontinence, 50% experienced faecal seepage and many of these patients (55%) had mixed incontinence or a combination of these types. Clinicians identified 47% of patients as having flatus incontinence symptoms usually in combination with other forms of faecal incontinence but 4 patients had flatus incontinence only.

Continence Advising was the most common treatment (39%), followed by physiotherapy (28%) and surgery (26%). Many of the patients had experienced incontinence symptoms for more than 1 year (79%), a large number of patients had experienced symptoms for 3 years or more (51%) and only 21% had first received treatment for their faecal incontinence symptoms within the last year. Only 8% of the sample had no other illnesses or health conditions, 16% had one co-morbidity and 76% of the sample had 2 or more co-morbidities which compares to 63% of patients in the urinary incontinence sample. The most common co morbidities were problems with the neck, back or spine (53%), arthritis (57%), high blood pressure (31%) and migraine headaches (25%).
### Table 4.6: Demographic details of participants

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Country of birth</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>45</td>
<td>74%</td>
</tr>
<tr>
<td>Other</td>
<td>16</td>
<td>26%</td>
</tr>
<tr>
<td><strong>Age group</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than 40 years</td>
<td>7</td>
<td>12%</td>
</tr>
<tr>
<td>40-49 years</td>
<td>5</td>
<td>8%</td>
</tr>
<tr>
<td>50-59 years</td>
<td>10</td>
<td>16%</td>
</tr>
<tr>
<td>60-69 years</td>
<td>23</td>
<td>38%</td>
</tr>
<tr>
<td>70+ years</td>
<td>16</td>
<td>26%</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>10</td>
<td>16%</td>
</tr>
<tr>
<td>Female</td>
<td>51</td>
<td>84%</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary and some High School</td>
<td>17</td>
<td>28%</td>
</tr>
<tr>
<td>Completed Higher School Certificate</td>
<td>14</td>
<td>23%</td>
</tr>
<tr>
<td>Trade Certificate/TAFE Diploma (a)</td>
<td>13</td>
<td>21%</td>
</tr>
<tr>
<td>College/University</td>
<td>17</td>
<td>28%</td>
</tr>
<tr>
<td><strong>Labour force status</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Labour force (i.e. working, unemployed)</td>
<td>17</td>
<td>29%</td>
</tr>
<tr>
<td>Homemaker</td>
<td>8</td>
<td>14%</td>
</tr>
<tr>
<td>Retired/Sickness benefit</td>
<td>34</td>
<td>58%</td>
</tr>
</tbody>
</table>

**Notes:**
All percentages are rounded up to the nearest integer, so percentages may not round up to 100%.
Excludes all missing cases. The base number was 61 cases.
a = Technical and Further Education.

#### 4.3.2 RFIS Scores

At baseline for all 61 cases the mean RFIS score was 9.66 (SD = 4.66) and scores ranged from 0-20. The mean RFIS score for the samples at pre-treatment was 9.10 for males and 9.76 for females and there was no significant difference by gender (p>0.05) albeit there are few males in the sample (N = 10). This is in contrast to the very low mean scores for the RFIS in the community sample of 0.43 (females) and 0.26 (males) respectively (SAHOS: Harrison Health Research, 2004; Sansoni et al., 2006).

#### 4.3.3 Item Characteristics and Endorsement Patterns

For most RFIS items 50-68% of the sample experienced these symptoms ‘sometimes’ through to ‘always’.

With regard to endorsement patterns for the leakage of solid stool (RFIS1) 14 patients (23%) indicated ‘never’ for this symptom, 16 patients (26.2%) indicated they experienced this symptom rarely, 12 patients (19.7%) indicated they experienced this sometimes and 19 patients (31.1%) experienced this often or usually (21.3%) or always (9.8%). When examined by RFIS and Wexner total scores, as measures of severity, the response levels are ordered with those endorsing the more severe levels of this symptom also having higher mean RFIS ($F = 11.59$, df 4, 56; $p<0.00$) and Wexner total scores ($F = 6.68$, df 4, 53; $p<0.00$). For those that endorsed ‘never’ on this item the mean RFIS total score was 5.25 as contrasted with those that endorsed ‘greatly’ where the mean was 15.33.

For the item concerning leakage of liquid stool (RFIS2) 12 patients (19.7%) endorsed ‘never’ for this symptom, 10 patients (16.4%) endorsed ‘rarely’, 14 patients (23%) endorsed ‘sometimes’ and 25 patients (41%) endorsed ‘often/usually’ (23%) and ‘always’ (18%). When examined by RFIS and Wexner total scores the response levels were ordered with those endorsing the more severe levels of the symptom also having higher mean RFIS scores ($F = 18.89$, df 4, 56; $p<0.00$) and Wexner total scores ($F = 22.14$, df 4, 53; $p<0.00$). For those that endorsed ‘never’ on this item the mean RFIS score was 5.18 and for those that endorsed always it was 14.8 although it should be noted there was little difference in RFIS total scores between those endorsing ‘never’ and ‘rarely’.
With regard to the item concerning leaking stool if you don’t get to the toilet in time (RFIS3) 6 patients (9.8%) endorsed ‘never’, 10 patients (16.4%) endorsed ‘rarely’, 26 patients (42.6%) endorsed ‘sometimes’, 19 patients (31.2%) endorsed ‘often/usually’ (16.4%) and ‘always’ (14.8%). When examined by RFIS and Wexner total scores the response levels were ordered with those endorsing the more severe levels of the symptom also having higher mean RFIS scores (F = 26.78, df 4, 56; p<0.00) and Wexner total scores (F = 10.16, df 4, 53; p<0.00). For those that endorsed ‘never’ on this item the mean RFIS score was 3.67 and for those that endorsed always it was 16.67. For the Wexner total score comparison there was no difference between the ‘never’ and ‘rarely’ categories although there was a minimum difference of 2.2 RFIS mean scores between these categories for the RFIS comparison.

RFIS4 is an item concerning whether stool leaks so that you have to change your underwear. Only 5 patients (8.2%) endorsed ‘never’, 14 patients (23%) endorsed ‘rarely’, 17 patients (27.9%) endorsed ‘sometimes’ and 25 patients (41%) endorsed ‘often/usually’ (23%) and ‘always’ (18%).
When examined by RUIS and Wexner total scores the response levels were ordered with those endorsing the more severe levels of the symptom also having higher mean RFIS scores (F = 19.72, df 4, 56; p<0.00) and Wexner total scores (F = 8.62, df 4, 53; p<0.00). For those that endorsed ‘never’ on this item the mean RFIS score was 3.80 and for those that endorsed always it was 15.00.

**Figure 4.11: Items RFIS5**

**RFIS5: Does bowel or stool leakage cause you to alter your lifestyle?**

![Graph showing frequency distribution of RFIS5 responses]

With regard to the item concerning whether bowel or stool leakage causes you to alter your lifestyle (RFIS5) 17 patients (27.9%) endorsed ‘never’ for this item. Eight patients (13.1%) endorsed ‘rarely’, 22 patients (36.1%) endorsed ‘sometimes’ and 14 patients (22.9%) endorsed ‘often/usually’ (13.1%) and ‘always’ (9.8%). When examined by RFIS and Wexner total scores the response levels were ordered with those endorsing the more severe levels of the symptom also having higher mean RFIS scores (F = 13.45, df 4, 56; p<0.00) and Wexner total scores (F = 13.91, df 4, 53; p<0.00). For those that endorsed ‘never’ on this item the mean RFIS score was 5.35 and for those that endorsed always it was 15.33. For the Wexner total score comparison there was no difference between the ‘often/usually’ and ‘always’ categories although there was a minimum difference of 2.2 mean scores between these categories for the RUIS total mean scores.

The above analyses provide evidence that the RFIS items all discriminate well by different levels of faecal incontinence severity. Although the sample size of males is small there were no significant differences in item means by gender and there were no significant differences in the response patterns by gender for these items.

### 4.3.4 RFIS Scores at Baseline

Figure 4.12 shows the distribution of RFIS scores at baseline. Table 4.7 shows health and incontinence status at baseline in comparison with RFIS scores and the statistical data relating to these analyses.

When examined by the baseline clinician rating of the severity of faecal incontinence there was a significant difference between RFIS scores for the more and less severe faecal incontinence groups (F = 3.17, df 2, 56; p = 0.05). The mean for those with mild incontinence was 8.05, for those with moderate incontinence it was 9.76 and for those with severe incontinence it was 12.08. Similarly, when examined by the baseline patient rating of the severity of faecal incontinence the mean for those in a mild state was 7.12, for those with moderate incontinence it was 10.64 and for those with severe incontinence it was 15.14 (F = 13.11, df 2, 58; p = 0.000).
Faecal incontinence patients who had experienced faecal incontinence symptoms for two years or more had a significantly higher ($t = -2.23$, df 55; $p = 0.03$) mean RFIS score (M = 10.57) than those patients who had experienced incontinence for less than 2 years (M = 7.75). Excluding flatus (as anal incontinence) patients diagnosed with mixed faecal incontinence types had a higher RFIS mean score than those diagnosed as having only one type of faecal incontinence ($t = -3.26$, df 59, $p = 0.002$). Bowel diary data indicated that those who used pads 10 or more times per week had higher RFIS mean scores than those who used pads 9 or less times per week ($t = -2.31$, df 38, $p = 0.03$).

There was a trend for those having surgical treatment to have higher RFIS scores than those receiving conservative treatment ($t = 1.87$, df 56; $p = 0.067$) but this did not attain significance at the $p<0.05$ level.

The RFIS does not include any items concerning pad use but pad use should reflect the severity of faecal leakage. Patient scores on the Wexner and St Mark’s items concerning pad use were compared to RFIS scores. Patient scores on the number of pads used (Wexner item) were grouped as ‘never or rarely’; or ‘sometimes, often and always’. The mean RFIS score for the low pad use group was 6.48 and for the high pad use group it was 11.89. This difference was statistically significant ($t = -5.10$, df 56; $p = 0.000$). For the St Mark’s Incontinence Score pad the mean RUIS score for the ‘no pad’ group was 7.23 and for the ‘wears pad’ group it was 11.28 and this difference was significant ($t = -3.45$, df 56; $p = 0.001$).

The study protocol contains a number of other (non RFIS) items about whether particular faecal leakage symptoms were present (e.g. leak mucus, incomplete bowel emptying, faecal urgency, faecal incontinence at night, seepage following a bowel movement, leaking without knowing it and pad use). These symptom scores were added to form an alternate index of severity and then formed into 3 groups (+/- 0.5 SD units from the mean). There was a significant difference in RFIS pre-treatment scores for these groups ($F = 14.19$, df 2, 53; $p = 0.000$) For the low score group the RFIS mean was 5.50, for the moderate scores group it was 9.96 and for the higher scores group the RFIS mean was 13.31. For comparisons with the other measures the pad use item was deleted from the alternate index (as both Wexner and St Mark’s contain pad items) and the groups dichotomized. For all three instruments there was a significant difference in pre-treatment scores in relation to the alternate index of severity (RFIS $t = -3.05$, df = 54, $p = 0.002$; Wexner $t = -2.75$, df = 54, $p = 0.008$; SMIS $t = -3.09$, df = 53, $p = 0.003$). The RFIS and the SMIS were slightly more sensitive measures for this comparison.

These analyses reflect that the RFIS discriminates well between different levels of incontinence severity as measured by other clinical indicators. By contrast when examined at baseline by a general item on health status there were no significant differences between those in Excellent/Very Good, Good, or Fair/Poor Health ($F = 0.80$; $p>0.05$); similarly there were no statistically significant associations by gender, age group, co morbidity number, type of treatment or BMI.
Table 4.7: Health and incontinence status at baseline in comparison with RFIS scores

<table>
<thead>
<tr>
<th>Variable</th>
<th>Classifications (N)</th>
<th>RFIS (Mean (SD))</th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Male N = 10; 16%</td>
<td>9.10 (3.60)</td>
<td>t = -0.41, df 59; p = 0.68</td>
</tr>
<tr>
<td></td>
<td>Female N = 51; 84%</td>
<td>9.76 (4.87)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>All N = 61</td>
<td>9.66 (4.66)</td>
<td></td>
</tr>
<tr>
<td>Age Groups</td>
<td>Less than 50 yrs N = 12; 20%</td>
<td>7.42 (5.35)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>50-59 N = 10; 16%</td>
<td>10.00 (3.83)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>60-69 N = 23; 38%</td>
<td>11.17 (4.69)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>70 yrs or more N = 16; 26%</td>
<td>8.94 (4.09)</td>
<td>F = 1.97, df 3,57; p = 0.13</td>
</tr>
<tr>
<td>Number of Co-morbidities</td>
<td>0/1 N = 15; 25%</td>
<td>9.93 (5.50)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2/3/4 N = 32; 46%</td>
<td>9.28 (4.73)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5+ N = 14; 23%</td>
<td>10.21 (3.68)</td>
<td>F = 0.22, df 2,58; p= 0.80</td>
</tr>
<tr>
<td>Symptom Duration</td>
<td>Less than 2 years N = 20; 35%</td>
<td>7.75 (4.95)</td>
<td>t = -2.23, df 55; p = 0.03</td>
</tr>
<tr>
<td></td>
<td>2 years + N = 37; 65%</td>
<td>10.57 (4.34)</td>
<td></td>
</tr>
<tr>
<td>Body Mass Index</td>
<td>Less than 30 N = 41; 79%</td>
<td>9.78 (4.830)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>30 or more N = 11; 21%</td>
<td>8.00 (4.12)</td>
<td>t = 1.12, df 50; p = 0.27</td>
</tr>
<tr>
<td>General Health Status</td>
<td>Excellent/Very good N = 18; 30%</td>
<td>8.50 (4.64)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Good N = 24; 39%</td>
<td>10.00 (4.44)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fair/Poor N = 19; 31%</td>
<td>10.32 (4.99)</td>
<td>F = 0.80, df 2,58; p = 0.45</td>
</tr>
<tr>
<td>Clinician Rated Incontinence Severity</td>
<td>Mild N = 22; 37%</td>
<td>8.05 (4.48)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Moderate N = 25; 42%</td>
<td>9.76 (4.31)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Severe N = 12; 20%</td>
<td>12.08 (4.89)</td>
<td>F = 3.17, df 2,56; p = 0.05</td>
</tr>
<tr>
<td>Patient Rated Incontinence Severity</td>
<td>Mild N = 26; 43%</td>
<td>7.12 (3.80)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Moderate N = 28; 46%</td>
<td>10.64 (3.97)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Severe N = 7; 11%</td>
<td>15.14 (4.34)</td>
<td>F = 13.11,df 2,58; p = 0.000</td>
</tr>
<tr>
<td>Diagnosis Type: Faecal Incontinence</td>
<td>1 Type only N = 27</td>
<td>7.63 (4.21)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mixed Types Fl N = 34</td>
<td>11.26 (4.42)</td>
<td>t = -3.26, df 59; p = 0.002</td>
</tr>
<tr>
<td>Bowel Diary Soiling</td>
<td>&lt; 9 per week N = 29</td>
<td>8.41 (4.73)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt; 10 per week N = 11</td>
<td>12.09 (3.78)</td>
<td>t = -2.31, df 38; p = 0.03</td>
</tr>
<tr>
<td>Type of Treatment</td>
<td>Conservative N = 42</td>
<td>8.88 (4.89)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Surgery N = 17</td>
<td>11.35 (3.79)</td>
<td>t = -1.87, df 57; p = 0.067</td>
</tr>
<tr>
<td>Wexner Pad Use Item</td>
<td>Never/Rarely N = 23</td>
<td>6.48 (3.47)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sometimes-Always N = 35</td>
<td>11.89 (4.23)</td>
<td>t = -5.10, df 56; p = 0.000</td>
</tr>
<tr>
<td>Alternate FI Symptoms Index by RFIS</td>
<td>Low N = 12</td>
<td>5.50 (2.75)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Moderate N = 28</td>
<td>9.96 (4.23)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High N = 16</td>
<td>13.31 (3.65)</td>
<td>F = 14.19, df 2,53, p= 0.000</td>
</tr>
</tbody>
</table>

4.3.5 Pre-treatment to Post-treatment Change Scores

Table 4.8 shows the changes in RFIS, Wexner and St Mark’s mean scores from pre-treatment to post-treatment. It should be noted that for the SMIS analyses 3 items of missing data at post treatment were imputed using the last value carried forward (LVCF) as suggested by Eliot and Hawthorne (2005).
Table 4.8 also shows RFIS scores in relation to other incontinence variables at follow up and the statistics relating to these comparisons. At follow up the mean RFIS score was 6.64 (SD = 4.76, N = 39). Of these cases, 38 had pre and post treatment scores. When examined by patient post treatment severity ratings there was a significant difference (t = 4.33, df 37; p = 0.000) in RFIS mean scores between the less severe (M = 4.60) and the more severe incontinence groups (M = 10.29). Table 4.8 shows the RFIS and the Wexner were marginally more sensitive instruments for this comparison (p = 0.000) with the St Marks being slightly less sensitive (p = 0.006).

The patient’s ratings concerning their treatment outcome was that 7.7% indicated their incontinence had not been helped by treatment, 35% reported some improvement, 41% considered themselves to be partly cured and 15% indicated that their continence was cured. There was a significant difference in RFIS post treatment scores between these groups (t = 4.01, df 37; p = 0.000). For those that were not helped/ had little improvement, the RFIS mean score was 9.59 and for those that were partly cured /cured the mean was 4.36. The RFIS and the Wexner were the most sensitive instruments for this comparison (p = 0.000) with the St Marks being slightly less sensitive (p = 0.002).

There were also significant differences (t = 3.80, df 37; p = 0.001), in RFIS post-treatment mean scores between those that considered that they still had major problems/ some problems following treatment and those that had slight problems/no problems.

RFIS scores discriminated between levels of pad use at post–treatment (p = 0.000). The mean for those that used pads never or rarely at post-treatment was 4.29 whereas for those that used pads sometimes/always the mean was 10.40. RFIS scores also discriminated between levels of the Alternate Index of Severity (comprised of non RFIS other faecal items) at post-treatment (F = 16.14, df 2, 36; p = 0.000).
**Table 4.8: RFIS and other incontinence variables at follow-up**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Classifications</th>
<th>N</th>
<th>Mean (SD)</th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>RFIS Scores</td>
<td>Pre-treatment</td>
<td>N = 38</td>
<td>9.79 (4.68)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Post-treatment</td>
<td>N = 38</td>
<td>6.68 (4.82)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean Change Score</td>
<td>3.11 (4.92)</td>
<td>t paired = 3.89, df 37; p = 0.000</td>
<td></td>
</tr>
<tr>
<td>Wexner Scores</td>
<td>Pre-treatment</td>
<td>N = 38</td>
<td>9.58 (4.74)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Post-treatment</td>
<td>N = 38</td>
<td>6.89 (4.83)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean Change Score</td>
<td>2.68 (5.18)</td>
<td>t paired = 3.19, df 37; p = 0.003</td>
<td></td>
</tr>
<tr>
<td>St Mark’s Scores</td>
<td>Pre-treatment</td>
<td>N = 37</td>
<td>12.46 (5.40)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Post-treatment</td>
<td>N = 37</td>
<td>8.97 (5.56)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean Change Score</td>
<td>3.49 (5.08)</td>
<td>t paired = 4.17, df 36; p = 0.000</td>
<td></td>
</tr>
<tr>
<td>RFIS post scores by Patient Rated Severity (post)</td>
<td>Moderate/Severe</td>
<td>N = 14</td>
<td>10.29 (4.55)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Normal/Mild</td>
<td>N = 25</td>
<td>4.60 (3.56)</td>
<td>t = 4.33, df 37; p = 0.000</td>
</tr>
<tr>
<td>Wexner post scores by Patient Rated Severity (post)</td>
<td>Moderate/Severe</td>
<td>N = 14</td>
<td>10.43 (4.99)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Normal/Mild</td>
<td>N = 25</td>
<td>4.72 (3.36)</td>
<td>t = 4.27, df 37; p = 0.000</td>
</tr>
<tr>
<td>SMIS post scores by Patient Rated Severity (post)</td>
<td>Moderate/Severe</td>
<td>N = 14</td>
<td>12.07 (5.27)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Normal/Mild</td>
<td>N = 24</td>
<td>7.13 (4.82)</td>
<td>t = 2.95; df 36; p = 0.006</td>
</tr>
<tr>
<td>RFIS post scores by Patient Rated Outcome (post)</td>
<td>Not Helped/ Little Improvement</td>
<td>N = 17</td>
<td>9.59 (4.37)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Partly Cured/Cured</td>
<td>N = 22</td>
<td>4.36 (3.75)</td>
<td>t = 4.01, df 37; p = 0.000</td>
</tr>
<tr>
<td>RFIS post scores by Current Incontinence Problems (post)</td>
<td>Major/ Some Problems</td>
<td>N = 18</td>
<td>9.33 (4.60)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No Problems/ Slight Problems</td>
<td>N = 21</td>
<td>4.33 (3.61)</td>
<td>t = 3.80, df 37; p = 0.001</td>
</tr>
<tr>
<td>RFIS post scores by Pad Use</td>
<td>Sometimes to Always</td>
<td>N = 15</td>
<td>10.40 (4.76)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Never or Rarely</td>
<td>N = 24</td>
<td>4.29 (2.96)</td>
<td>t = 4.96, df 37; p = 0.000</td>
</tr>
<tr>
<td>RFIS post scores by Alt. Index Severity</td>
<td>Low</td>
<td>N = 15</td>
<td>3.53 (2.67)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Moderate</td>
<td>N = 16</td>
<td>6.69 (4.22)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>N = 8</td>
<td>12.38 (3.54)</td>
<td>F = 16.14, df 2,36; p = 0.000</td>
</tr>
<tr>
<td>RFIS Change by Patient Rated Outcome</td>
<td>Not Helped/Little Improvement</td>
<td>N = 17</td>
<td>0.59 (4.39)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Partly Cured/Cured</td>
<td>N = 22</td>
<td>5.14 (4.43)</td>
<td>t = 3.16, df 37; p = 0.003</td>
</tr>
<tr>
<td>RFIS Change by Patient Rated Improvement</td>
<td>Worse</td>
<td>N = 5</td>
<td>-0.80 (4.32)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No Change/A Little Better</td>
<td>N = 11</td>
<td>0.64 (4.08)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Much/Very Much Better</td>
<td>N = 21</td>
<td>5.29 (4.53)</td>
<td>F = 6.34, df 2,34; p = 0.005</td>
</tr>
<tr>
<td>Patient Rated Improvement by 2 point RFIS Change</td>
<td>Improved by &gt; 2</td>
<td>N = 22</td>
<td>5.27 (1.08)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Improved by &lt; 2</td>
<td>N = 15</td>
<td>4.13 (0.99)</td>
<td>t = 3.26, df 35; p = 0.002</td>
</tr>
<tr>
<td>RFIS Change by Patient Rated Severity at Baseline</td>
<td>Normal/Mild</td>
<td>N = 19</td>
<td>2.84 (6.08)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Moderate/Severe</td>
<td>N = 19</td>
<td>3.37 (3.56)</td>
<td>t = -0.11, df 36; p = 0.75</td>
</tr>
</tbody>
</table>

* There were 39 patients that submitted post-treatment forms but for 1 subject no pre-treatment data was available.
Change scores on the RFIS (pre-post) ranged from an improvement of 18 points to a deterioration of 8 points. Examination of pre-post mean scores revealed a statistically significant improvement of 3.11 RFIS scores (paired t-test, t = 3.89, df 37; p = 0.000). To illustrate this change, the frequency distributions (histograms) of RFIS data pre and post-treatment for the sample are provided in Figure 4.13.

**Figure 4.13: RFIS pre-treatment and post-treatment scores**

![Fig 4.13a: RFIS pre-treatment scores for post treatment sample](image1)

![Fig 4.13b: RFIS post-treatment scores](image2)

RFIS change scores were examined by Patient Rated Improvement. The RFIS mean change score for those that considered themselves worse was -0.80, for those that experienced no change or were a little better it was 0.64 and for those who were much/very much better it was 5.29 (F = 6.34, df 2,34; p = 0.005).

RFIS change scores were examined by Patient Rated Outcome. The RFIS mean change score for those that considered themselves not helped/little improved was 0.59, and for those that considered themselves cured/partly cured it was 5.14 (t = 3.16, df 37; p = 0.003).

When examined by the pre-treatment patient rating of the severity of faecal incontinence there was no statistically significant difference in RFIS change scores (t = -0.11, df 36; p > 0.05) for those with normal/mild incontinence (M = 2.84) and for those with moderate/severe incontinence (M = 3.37). This indicates there was no significant difference in the degree of improvement during treatment based on the initial patient severity rating – both groups improved. When these scores were dichotomized (improvers vs. no change/deterioration) and examined there were no significant differences by patient severity at baseline, health status, gender, education, age group or BMI.

### 4.3.6 Reliability Data

The internal consistency reliability of the RFIS pre-treatment is Cronbach’s alpha = 0.78, (N = 61 faecal incontinence patients). Lui et al (2010) suggests that Cronbach’s alpha can be inflated by outliers but the box plot analysis in SPSS indicated there were no outliers.

The pre-treatment alpha for the Wexner was 0.65 (N = 58) and for the St Mark’s it was 0.65 (N = 57) which are both considered marginal as they are less than 0.70 (Streiner and Norman, 2003). The RFIS pre-treatment alpha for a combined sample of urinary and faecal incontinence patients was 0.91 (N = 254) which is the similar to the reliability estimate for the community population sample (Sansoni et al., 2006). Post-treatment RFIS alpha = 0.86 (faecal incontinence sample, N = 39) compared with the Wexner = 0.77 and the St Mark’s = 0.66. Post-treatment RFIS alpha for all incontinence patients = 0.92 (N = 138).

Test-retest reliability was assessed at 2 weeks after completion of treatment and the submission of post-treatment forms. For the faecal retest sample (N = 19) the ICC = 0.79 and for the Wexner
test-retest ICC = 0.74 and for the St Mark’s ICC = 0.68. For the total sample of incontinence patients (N = 78) the ICC = 0.80 (CI 0.70-0.87). These reliability estimates are considered to be adequate to good (Streiner and Norman, 2003).

The standard error of measurement (SEM) is related to reliability and estimates the extent, to which a test provides accurate scores (or the observed score reflects the true score), given there will always be some degree of variation in individual scores in repeated measures using same test. A lower SEM indicates greater precision in measurement. Using Cronbach’s alphas for the faecal sample (N = 61) the SEM for the RFIS = 2.19, for the Wexner it was 2.70 and for the St Mark’s it was 3.06.

Ceiling and Floor Effects
At pre-treatment scores on the RFIS ranged from 0 – 20 (two people with faecal incontinence only received a score of 0 on the RFIS as it does not include an item on flatus) indicating the full range of scores were used. There was 3.3% of the sample on the floor of the instrument (lowest possible score) and 1.6% of the sample at the ceiling of the scale (maximum possible score). Generally a figure of less than 15% of the sample at the floor or the ceiling is preferred (Streiner and Norman, 2003). It is clear that this instrument does not suffer from significant floor or ceiling effects.

4.3.7 Responsiveness: Capacity to Detect Change
Change scores were calculated by subtracting the post-treatment score from the pre-treatment score. For this scale a reduction in scores post-treatment indicates that the patient has improved.

As indicated in Table 4.8 there was a significant improvement of 3.11 RFIS scores (SD = 4.92, CI 1.49-4.72, paired t-test, t = 3.89, df 37; p = 0.000) following treatment. Table 4.8 shows that for the Wexner, also a 20 point scale, the t paired value was 3.19. The St Mark’s Incontinence Score is a 24 point scale and the t paired value = 4.17. The relative efficiency (RE) of the measures was Wexner RE = 1.00, RFIS RE = 1.49 and the SMIS RE = 1.71 (Liang et al., 1985). These results indicate that the RFIS is 49% more sensitive to detecting change than the Wexner. The SMIS showed slightly more sensitivity to change than the RFIS but far more sensitivity than the Wexner. However, It should be noted that for the SMIS analyses 3 items of missing data at post treatment were imputed using the last value carried forward (LVCF) as suggested by Eliot and Hawthorne (2005). If missing data is not imputed the SMIS t = 3.82 and the RE for the SMIS is 1.43 —similar to the RFIS.

The Kazis effect size (ES) for the RFIS was -0.66 (95%CI: -2.15 - +0.82), and this is classed as a moderate effect size as it is over 0.5 (Kazis et al, 1989; Cohen, 1988). All three measures were responsive over time and expressed as the Kazis’ ES the score changes for the Wexner and the SMIS were ES = -0.56 (95%CI: -2.08 - +0.94) and -0.65 (95%CI: -2.39 - +1.09), respectively.

Another estimate of the effect size is the standardised response mean (Deyo et al., 1991) which is the change in mean scores divided by the SD of the change = 0.63 which is also considered adequate (Crosby et al., 2003). Pallant (2011) suggests the eta squared statistic can be used as an estimate of effect size and this (0.29) also indicated a large effect size as it is over 0.14 (Cohen 1988). These provide further evidence that the RUIS is responsive to change.

The responsiveness of the instrument is also concerned with the capacity of the instrument to detect change regardless of whether it is improvement or deterioration. Ignoring signs indicating the direction of change the average change score for the sample was 4.32 RFIS points SD = 3.98 N = 36. Change scores ranged from an improvement of 18 points to a deterioration of 8 points in the context that the scale range is from 0-20 points. These findings would suggest that the instrument has the capacity to detect both an improvement and deterioration in patient incontinence status.
Figure 4.14: RFIS change scores

Figure 4.15: RFIS change by Patient Global Rating of Improvement

Figure 4.15 show RFIS change scores by the Patient Global Rating of Improvement (post-treatment). The patient’s global impression of improvement was reclassified into 3 groups of worse to no change (M = -0.80, N = 5), a little better (M = 0.64, N = 11), and much and very much better (M = 5.3, N = 21). There was a significant difference in RFIS change scores (F = 6.34, df 2, 34; p = 0.005) between these groups. There was a significant difference between the no change/worse group and the much and very much better group (p = 0.023), and between the little better and the much better group (p = 0.020) but there was no difference between the little better and worse/no change group. It should be noted there is sparse data for this analysis as only 5 subjects considered themselves worse/unchanged. The Spearman rho correlation between RFIS change and the patient’s global impression of improvement was 0.62 (p = 0.000).
Figure 4.15 might suggest that an important improvement for the group of patients may lie between the RFIS change score means for the much better (M = 1.67) and very much better groups (M = 6.73). RFIS change score were reclassified into 2 groups where there was an improvement of equal to or greater than 2 RFIS scores and where there was an improvement of less than 2 RFIS scores (including deterioration). These RFIS change groups were analysed by the patient global rating of improvement (seven point scale) and there was a significant difference (t = 3.26, df 35; p = 0.002). The mean for the improved by 2 RFIS scores group was 5.27 and the mean for the less improved group was 4.13. These analyses might suggest that a change of ≥2 RFIS score may be the minimal detectable difference required concerning change in the patients’ perception of improvement, at the group level.

The SEM for the RFIS of 2.19 (95% CI SEM = 4.29) would suggest that a score change of 4 points or more is likely to be a more clinically and statistically reliable estimate for the purposes of patient monitoring although it is noted that the SEM reported is for the instrument as a whole and standard errors of measurement vary by score levels (they may be more or less for people scoring at the extremes of the scale) (Spratt, 2009). However, the following anchor based analysis also suggests that a change of 4 points represents a clinically important change.

Change scores pre- and post-treatment were also anchored against changes in patient behaviour. Pad use was examined across pre-treatment to post-treatment. Four groups were identified (no pads used at pre or post treatment, reduced pad use from pre to post-treatment, stopped using pads after treatment, and no change/increased use of pads). RFIS post scores were examined by these pad groups with RFIS baseline scores treated as a covariate to control for any pre-existing differences between the groups (F = 5.59, df = 1,3,4,33; p<0.01). This analysis suggested that a change in score of >4.00 points on the RFIS was associated with a reduction in the use of incontinence pads (the mean RFIS reduction for those no longer using pads was 8.20 RFIS points and for those who had reduced their use of pads it was 4.78). A change of ≥4 points appears to represent a clinically important change in that this score change was associated with a change in patient behaviour.

4.3.8 Internal Structure: Principal Components Analysis

The faecal incontinence sample (N = 61) is small for a principal components analysis (PCA) – a much larger sample of 100 – 300 cases is generally recommended (Guadagnoli and Velicer, 1988; Pallant, 2011) in determining the comparability between sample and population component patterns. However, Pallant (2011) reports that Nunnally (1978) recommends 10 cases for each item to be analysed and Tabachnick and Fiddell (2007) suggest that 5 cases per item may be sufficient. Guadagnoli and Velicer (1988) in a simulation study found that sample size as a function of the number of variables was not an important factor in determining analysis stability; and that a stable analysis could be achieved with smaller samples where the component saturation (e.g. the magnitude of the correlation between the observed variables and the components) was high (average loading =+/+ 0.70) but even so it is estimated that a sample size of 77 would be required in this instance. In order to address the sample size limitations this analysis is compared with the same analysis conducted on the community survey sample (N = 2,915).

For all analyses presented below the Kaiser-Meyer-Olkin value was >0.6 (Kaiser 1970, 1974) and Bartlett’s Test of Sphericity (Bartlett 1954) reached statistical significance, supporting the factorability of the correlation matrices.

The 5 items of the RFIS were analysed (N = 61, faecal incontinence patients) using a Principal Component analysis. Table 4.9 below shows there was only one component with an eigenvalue over 1 (eigenvalue = 2.71) and this could be described as a general faecal incontinence factor explaining 54.22% of the variance. All RFIS items had loadings on the first component of 0.65 or above indicating a moderate to high level of component saturation (average factor loading = 0.73). The analysis of the community survey data produces a similar structure with a similar pattern of item factor loadings. The one component extracted explains more of the variance (63%) but this may be explained by the greater homogeneity in this sample.
Table 4.9: Principal components analysis of RFIS items

<table>
<thead>
<tr>
<th>RFIS Items</th>
<th>Faecal Incontinence Patients (N = 61)</th>
<th>Community Survey (N = 2917)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Factor 1 (54%)*</td>
<td>Factor 1 (63%)*</td>
</tr>
<tr>
<td>Leak solid</td>
<td>0.65</td>
<td>0.76</td>
</tr>
<tr>
<td>Leak liquid</td>
<td>0.74</td>
<td>0.85</td>
</tr>
<tr>
<td>Leak stool/urgency</td>
<td>0.82</td>
<td>0.84</td>
</tr>
<tr>
<td>Leak/change underwear</td>
<td>0.76</td>
<td>0.81</td>
</tr>
<tr>
<td>Alter Lifestyle</td>
<td>0.69</td>
<td>0.70</td>
</tr>
</tbody>
</table>

* = proportion of variance explained

In 2006 the authors used PCA to analyse a number of faecal incontinence items which included the 5 RFIS items, the 2 other Wexner items concerning wearing a pad and flatus, items on bowel patterns/movements and a faecal urge item. The results of this PCA analysis indicated that the RFIS items and pad item loaded highly on Factor/Component 1 (explaining 40% variance) and urge and flatus items loading on Factor/Component 2 (explaining 11% variance) with only the bowel movement item loading on Factor/Component 3 (explaining 10% variance).

Table 4.10: Rotated factor matrix for faecal incontinence (2006)

<table>
<thead>
<tr>
<th>Items</th>
<th>Factor 1 (40%*)</th>
<th>Factor 2 (11%*)</th>
<th>Factor 3 (10%*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bowel Pattern</td>
<td>0.27</td>
<td>0.59</td>
<td>-0.12</td>
</tr>
<tr>
<td>Bowel Movements</td>
<td>0.08</td>
<td>-0.03</td>
<td>0.95</td>
</tr>
<tr>
<td>Urgency</td>
<td>0.20</td>
<td>0.70</td>
<td>0.33</td>
</tr>
<tr>
<td>Leak Solid (RFIS 1)</td>
<td>0.71</td>
<td>0.22</td>
<td>0.07</td>
</tr>
<tr>
<td>Leak Liquid (RFIS 2)</td>
<td>0.75</td>
<td>0.31</td>
<td>0.10</td>
</tr>
<tr>
<td>Leak Gas</td>
<td>0.08</td>
<td>0.74</td>
<td>-0.08</td>
</tr>
<tr>
<td>Leak Stool / Urgency (RFIS3)</td>
<td>0.77</td>
<td>0.25</td>
<td>0.06</td>
</tr>
<tr>
<td>Wear Pad</td>
<td>0.71</td>
<td>-0.03</td>
<td>-0.06</td>
</tr>
<tr>
<td>Leak / Change Underwear (RFIS 4)</td>
<td>0.78</td>
<td>0.18</td>
<td>0.06</td>
</tr>
<tr>
<td>Alter Lifestyle (RFIS 5)</td>
<td>0.70</td>
<td>0.15</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Note: Matrix derived from items W1 – 10 2004 SAHOS Survey (N = 2,917)
* % variance explained

In 2006 (Sansoni et al., 2006) the RFIS items were selected with regard to their loadings on the primary faecal incontinence factor (defined as soiling and leakage). The flatus item had a low loading on this factor. It was not included as the international definition of faecal incontinence did not include flatus and an analysis of internal consistency reliability of the Wexner indicated the item should be deleted. Item 4 of the RFIS (Does stool leak so that you have to change your underwear?) had higher loadings on the primary faecal incontinence than the pad item from the Wexner, and given the overlap between these items only the RFIS 4 item was selected for inclusion.

Although not quite of sufficient size for a stable analysis (Pallant, 2011) this analysis was repeated with the sample of faecal incontinence patients (N = 61) to re-examine the pattern of item loadings - although it is noted that no bowel pattern/movement items are included in the 2011 clinical dataset. This produced a 2 component solution which accounted for 58% of the variance. The primary faecal incontinence factor accounted for 40% of the variance (eigenvalue = 3.17) and the other bowel symptoms factor/component accounted for 19% of the variance (eigenvalue = 1.48) which is very similar pattern to the 2006 analysis. All RFIS items had loadings of 0.50 or above on
the primary faecal incontinence factor which confirmed their selection in the RFIS. This analysis confirms the current descriptive system. One minor difference concerned the lifestyle item which had moderate loadings on both factors for the faecal vs. the community sample. However, it is noted that this item had quite high loadings on the primary faecal incontinence factor in the analyses of the RFIS items.

Table 4.11: Principal components analysis: RFIS with pad, flatus and urge items; faecal patients (2011)

<table>
<thead>
<tr>
<th>Items</th>
<th>Factor/Component</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 (40%*)</td>
</tr>
<tr>
<td>Leak solid (RFIS 1)</td>
<td>0.73</td>
</tr>
<tr>
<td>Leak liquid (RFIS 2)</td>
<td>0.69</td>
</tr>
<tr>
<td>Leak stool/ urgency (RFIS 3)</td>
<td>0.75</td>
</tr>
<tr>
<td>Leak/change underwear (RFIS 4)</td>
<td>0.76</td>
</tr>
<tr>
<td>Alter lifestyle (RFIS 5)</td>
<td>0.53</td>
</tr>
<tr>
<td>Leak gas</td>
<td></td>
</tr>
<tr>
<td>Wear pad</td>
<td>0.71</td>
</tr>
<tr>
<td>Urgency</td>
<td></td>
</tr>
</tbody>
</table>

* % variance explained

4.3.9 Validity: Correlations with Other Measures:

The correlation of the pre-treatment RFIS with the Wexner was $r = 0.88$ (p<0.01); and with the St Mark’s $r = 0.85$ (p<0.01). Correlations between RFIS with Wexner Type Specification (impact) items were all significant at the p<0.01 level save for sexual relations (p>0.05) and church attendance (p>0.05) but there were higher rates of missing data for the latter 2 items. The correlation of RFIS with four Faecal Incontinence Quality of Life (FIQL) coping items was $r = 0.57$ (p = 0.000), and with the three embarrassment items $r = 0.30$ (p = 0.03).

Sansoni et al. (2006) reported that urinary and faecal incontinence measures had negative but significant correlations with measures of health related quality of life reflecting the burden of disease for this condition (e.g. the higher the incontinence score the lower the physical functioning or health status score). The RFIS correlation with the Physical Function Scale of the SF-36V2 was $r = -0.27$ (p = 0.037) at pre-treatment for the faecal sample. It is interesting that this is a higher correlation than was found for urinary incontinence suggesting that the impact of faecal incontinence may be greater although it should be noted that the faecal incontinence sample was somewhat older and with a greater level of co-morbidities than the urinary incontinence sample.

For the faecal sample (N = 50) the Physical Component Summary (PCS) score was 43.59 and for the Mental Component Summary (MCS) score it was 43.41 which was lower than the Australian population norm (tpcs = -4.54, tMCS = -3.75, df 49; p < 0.01) (Hawthorne et al., 2007). These results were consistent with findings in the literature showing a small negative association between faecal incontinence and generic measures of HRQOL (Sansoni et al., 2006).

The RFIS has a correlation of $r_s = 0.50$ (p<0.01) with patient rated severity at pre-treatment and a correlation $r_s = 0.29$ (p<0.02) with the clinicians’ pre-treatment severity rating. The correlation between the clinician’s and patient’s severity ratings at pre-treatment is very low ($r_s = 0.13$) and not significant. This pattern is replicated with the other faecal incontinence symptom measures – they all correlate highly with each other (0.86 or above) and with patient rated severity (0.50 – 0.57) but have lower correlations (0.27-0.51) with clinician rated severity at pre-treatment as was the case with urinary incontinence. The scale that has the highest correlation with the pre-treatment clinical rating is the St Mark’s Incontinence score and it is interesting that this is the only scale to heavily weight the inability to defer defecation by 15 minutes (e.g. score 4 if cannot defer, score 0 if can defer). There is a significant correlation between the clinician’s rating and this recoded item at pre-treatment 0.34 (p<0.01).Thus it seems that the clinician’s rating of faecal incontinence severity may be influenced by the consideration of this particular symptom and other
broader aspects of health status such as BMI and weight, physical function and general health status as was the case for urinary incontinence. The clinician’s rating may also be influenced by the severity of the biological cause of the condition. The value of considering the inclusion of this item in the RFIS is also considered further below.

At post-treatment the RFIS correlation with untransformed patient satisfaction scores (SAPS) was $r = -0.28$ ($p = 0.09$) which indicates there was a trend ($p< 0.10$) for an association between higher RFIS scores (indicating greater incontinence) and lower patient satisfaction scores (less satisfaction). The change scores on RFIS were dichotomized to analyse the SAPS data. When change scores were classified into worse / no change (a deterioration of 8 to 0, N = 13) and improved (an improvement of 1 point or more, N = 25) there was no significant difference between SAPS scores and this classification of RFIS change ($t = 0.68$, df 36; $p>0.05$). This would suggest that SAPS scores appear to reflect more the current incontinence state of the patient at post-treatment as measured by RFIS scores.

4.3.10 Cutpoints for Interpretation

It would be advisable to refer people who score 4 or more on this scale for further continence assessment. This is more than one standard deviation below the mean for the total clinical sample and represents the 10th percentile of the clinical sample. It is 2 standard deviations above the mean for the community sample. To obtain this score one would need to endorse ‘slightly’ or ‘rarely’ on most incontinence items.

An analysis of the distribution of the clinical sample scores indicated that a score of 6 was the 25th percentile for the clinical sample which might be considered ‘mild’, a score of 10 was the 50th percentile which might be considered ‘moderate’ and a score of 13 represented the 75th percentile which might be considered ‘severe’.

Other faecal leakage items that are not contained in the RFIS (leak mucus, incomplete bowel empty, faecal urge, leak sleep/night, seepage following bowel movement, leak not knowing and pad use) were summed to form an alternate index of the severity of faecal leakage. This index was used to assist in determining cutpoints for the RFIS. The scores were classified into 3 groups of low, moderate and high scores ($+/- = 0.5$ SD units from the mean). Table 4.7 shows there was a significant difference in RFIS pre-treatment scores for these groups. For the mild/low score group the RFIS mean was 5.50, for the moderate group it was 9.96 and for the more severe/ high scores group the RFIS mean was 13.31. All group comparisons were significant.

Clinician ratings of ‘mild’ equated to a RFIS mean score of 8.05 (N = 22), for ‘moderate’ this equated to an RFIS mean score of 9.76 (N = 25) and for severe this equated to a mean score of 12.08 (N =12). Patient pre-treatment severity ratings indicated that the mean for a rating of ‘mild’ was 7.12 (N = 26) – which was consistent with the likelihood of pad use (Figure 4.12), the mean for a rating of moderate was 10.64 (N = 28), and the mean for a rating of severe was 15.14 (N =7). At post treatment the clinical ratings suggested a score of 3 or less would be considered ‘normal’ as did the patient severity ratings.

As described earlier pad use was examined across pre-treatment to post-treatment. Four groups were identified (no pads used at pre or post treatment, reduced pad use from pre to post-treatment, stopped using pads after treatment, and no change/increased use of pads). Figure 4 16 suggested that a critical cutpoint score on the RFIS was at 7-points.
Most participants at both baseline and follow up with scores <7-points had faecal incontinence insufficiently severe to warrant the use of incontinence pads. In contrast, most of those with scores ≥7-points reported using pads. Based on a cutpoint of 7 RFIS points, at baseline the relative risk of a participant using an incontinence pad sometimes/often/always was almost three times that of a person with a RFIS score of ≤6 (RR = 2.79; 95%CI: 1.26 – 5.81). At follow up they were 8-times more likely to be using incontinence pads based on the same cutpoint (RR = 8.41; 95%CI: 2.19 – 32.76).

Given these considerations the following RFIS score ranges are suggested to aide interpretation:

- **0-3:** no faecal incontinence or extremely mild incontinence symptoms
- **4-6:** mild faecal incontinence
- **7-12:** moderate faecal incontinence
- **13-20:** severe faecal incontinence (RFIS scores above 17 could be considered very severe)

The community survey dataset (SAHOS 2004) was reanalysed with respect to the proposed RFIS cutpoint of 4 to distinguish between those that have no faecal incontinence and those with ‘mild’ symptoms of incontinence. Using this cutpoint the prevalence of faecal incontinence would be estimated at 3% overall; for females it would be 4.4% and for males it would be 1.5%. These are somewhat lower prevalence estimates than those provided from an analysis of Wexner scores - excluding the flatus item - which has been observed to confound prevalence estimates (refer Section 4.6 on Incontinence and the Burden of Disease and AIHW, 2006; Hawthorne, 2006; Sansoni et al., 2006). The Wexner data indicated a prevalence estimate at 7.9% overall; for females it was 10% and for males it was 6%. It should be noted the Wexner estimates are based on 4 items (leakage of solid and liquid stool, pad use and effect on lifestyle) and include any endorsement of a faecal incontinence item (e.g. scores of 1 or more). The RFIS contains 5 items
validation and clinical translation of the revised continence and patient satisfaction tools

(Leakage of liquid and solid stool, faecal urgency, soiling of underwear and effect on lifestyle) and the prevalence estimates for RFIS when a total score of 1 or more is included are 10.2% overall (11.7% for females and 8.6% for males). However, the suggested RUIS cutpoint of 4 in routine practice would exclude more of the cases with very few and very slight symptoms of incontinence and it is noted above that both clinicians and patients rated RFIS scores of 3 and below as normal/no incontinence at post-treatment.

It is noted that the RFIS does not contain an item on flatus. If flatus is an issue of concern the flatus item from the Wexner Faecal Continence Grading Scale or the CRADI – 8 could be used as a separate exercise (refer to the discussion in section 4.2.1.9 below).

4.3.11 The Effect of Treatment

As indicated above there was a significant improvement of 3.11 RFIS scores following treatment. An analysis of RFIS change scores by type of treatment was undertaken. For this analysis the four cases receiving medicinal treatments were excluded. There was a significant difference by type of treatment F = 4.79; df 2, 33; p = 0.015. The mean change score for continence advising was 4.60 (SD = 5.55, CI 1.53-7.67, N = 15), for physiotherapy it was −0.89 (SD = 4.26, CI -4.16-2.38, N = 9) and for surgery it was 4.50 (SD = 3.20, CI 2.46-6.54) N = 12). Further analyses using t tests indicate this difference is brought about by significant differences in change scores between physiotherapy; and surgery and continence advising (p<0.01). There was no significant difference in the change scores between continence advising and surgery (p>0.05).

The sample sizes for these comparisons are small and it is noted the physiotherapy sample (N = 9) contained two patients who experienced substantial deterioration during treatment and in the context of a small sample such outliers can have a major effect on the results. It was also found that the physiotherapy group had significantly lower RFIS scores at baseline (F = 4.26, df 2, 54; N = 56) than the other 2 treatment groups (physiotherapy mean = 6.88 as compared with continence advising mean = 10.25 and surgery mean = 11.06). The net effect of this would be to reduce the amount of change possible within the physiotherapy group which may also assist in explaining the differences detected above.

As the sample size for physiotherapy was small (N = 9) and given the treatment provided was reported to be very similar to that provided by Nurse Continence Advisors these groups were combined for further analyses and classed as ‘conservative therapy’. When this reclassification occurs there is no significant difference in RFIS scores at baseline between the treatment groups (t = -1.87, df 57; p>0.05). The mean change score for conservative treatment was 2.52 RFIS scores (t = 2.26, df 24; p = 0.033) and for surgery it was 4.50 RFIS scores (t = 4.86, df 11; p = 0.000). For this sample size the difference in improvement by treatment type was not significant – both treatment groups improved (t = -1.14; df 35; p = 0.26).

4.3.12 Alternative Solutions

The RFIS has been designed for use in both population health and clinical settings and for this reason an item concerned with flatus has not been included in the scale. It has been found that when the flatus item is included in population health settings, such as community surveys, this item will confound prevalence estimates (AIHW 2006, Hawthorne 2006, Sansoni et al., 2006;). The Project Steering Group was keen to further assess some flatus items in the clinical context and 3 flatus items (from the Wexner and the Colorectal Anal Distress Inventory - 8) were included in the pre and post treatment protocols for patients with faecal incontinence. These items were:

B.11 Do you leak, have accidents or lose control with gas (flatus or wind)? (Wexner)

- Never
- Rarely, i.e. less than once in the past four weeks
- Sometimes, i.e. less than once a week, but once or more in the past four weeks
- Often or usually, i.e. less than once a day but once a week or more
- Always, i.e. once or more per day or whenever you have a bowel movement
B.26a Do you lose gas from your rectum beyond your control? (CRADI-8)

Never
- Rarely, i.e. less than once in the past four weeks
- Sometimes, i.e. less than once a week, but once or more in the past four weeks
- Often, i.e. less than once a day but once a week or more
- Usually, i.e. once per day
- Always, i.e. several times per day

B.26b If so, how much does this bother you?

- Not at all
- Somewhat
- Moderately
- Quite a bit.

These items had low and non-significant correlations (p>0.05) with RFIS pre-treatment scores (used as an indicator of overall faecal incontinence severity with flatus excluded) and pretreatment clinician and patient ratings of severity. The internal consistency reliability of the Wexner at baseline was improved from 0.65 to 0.71 if the flatus item was removed. If the flatus item is added to the RFIS the internal consistency reliability drops from 0.78 to 0.75. A factor analysis of the seven items contained in the Wexner and RFIS scales indicates the gas item did not load highly on the first factor extracted (which could be classed as faecal soiling or leakage – explaining 45% of the variance). For these reasons there was no good case to include a flatus item within the RFIS particularly as it will confound prevalence estimates when used in population health settings.

For clinicians wishing to check on flatus either the Wexner item or the CRADI-8 items could be used as a separate question to consider. There is little to choose between these items. For simplicity the single item from the Wexner might be the preferred one for use although it is noted that its correlation with the presence of flatus type of incontinence (clinician rating) is $r_s = 0.26$ which only just reaches significance ($p = 0.05$). The other alternative could be to sum the scores of the two CRADI-8 items (gas and bother) and this derived score has a correlation of 0.3 with the flatus type of incontinence ($p = 0.023$).

Other items were examined to see if they would be useful to include in a revision of the RFIS. The item distributions for a number of faecal symptoms were examined to identify items that were widely endorsed amongst the sample. An item on urge incontinence was tested but its inclusion would reduce the internal reliability of the scale from 0.78 to 0.76. The resulting scale would, however, be slightly more sensitive to detecting change between pre-treatment and post treatment. When included in a factor analysis with the other RFIS items – it loaded on its own factor and did not load on the primary faecal incontinence factor and thus its inclusion did not appear warranted.

A question on soiling following a bowel motion was also tested. In this case the internal consistency was similar (0.77) but the analysis identified that the internal consistency would be improved if the item were deleted. Similar results were found concerning an item on the leakage of mucus but it was noted that there was a higher rate of missing data for this item and that 54% of the sample had never experienced this symptom.

Given the moderate correlation ($r_s = 0.34$) between the clinician’s pre-treatment severity rating and the item concerning the capacity/incapacity to delay defecation (for 15 minutes), analyses were undertaken to see if the inclusion of this item would improve the RFIS. This item is a dichotomous item – if defecation cannot be delayed it is scored 4 and if it can be delayed it is scored 0. This is a highly endorsed item, at pre-treatment 76.8% of the sample can’t defer defecation. There are 6 cases of missing data for this item = 10% of the pre-treatment sample compared with 0% for RFIS items. This item was added to the RFIS items for the purposes of an analysis of internal consistency (RFIS-DEF). The alpha for was RFIS-DEF = was 0.76 compared with 0.78 for the RFIS. Change scores on RFIS-DEF, a 6 item scale, were examined and the mean change score was 3.59 (SD = 4.88, df 31) and the t value was 4.17, p<0.001. This compares with a change
score on the original RFIS, a 5 item scale, of 3.11 RFIS scores (t paired = 3.89, df 37; p< 0.001). The marginal increase in t scores noted here is probably due to the effect of adding an additional item which can increase the magnitude of change scores as was discussed earlier. When included in a factor analysis with the other RFIS items (faecal sample N = 61) – its loading on the one faecal incontinence factor extracted (accounting for 48% of the variance) is low (0.44) compared with other RFIS items (0.63-0.82). Given this and given the reduction in internal reliability consistency its inclusion does not appear warranted.

Earlier an issue had been raised as to whether the item concerning the effect of incontinence on lifestyle should be included in the RFIS and the Wexner and the St Mark’s faecal grading scales. In the 2006 community survey study Item 5 of the Wexner and the RFIS (Item 5: Does bowel or stool leakage cause you to alter your lifestyle?), loaded highly on the ‘general faecal incontinence factor’ and thus was included in the RFIS (Sansoni et al., 2006). Thomas et al (2006) queried whether this was including a consequence rather than a symptom of incontinence. The project teams views this as an important aspect of the severity of faecal incontinence and for this reason feel the item should be retained. If the lifestyle item is dropped from the RFIS the internal consistency reliability of the scale drops from 0.78 to 0.76. For these reasons this item has been retained in the RFIS.

4.3.12.1 Interpretation of Results

The RFIS demonstrated good internal consistency reliability at pre-treatment as compared with the internal consistency reliability of the Wexner and the St Mark’s which are classed as marginal. It also showed excellent test-retest reliability which was superior to the other instruments. Good internal consistency reliability at pre-treatment is essential to identify patients who require treatment and to assess the severity of their condition. The RFIS was also sensitive to change as a result of treatment indicating it can be used to assess patient outcomes. Scores on the RFIS also discriminate well between different levels of incontinence severity as rated by both patients and clinicians. The RFIS appears to be a uni-dimensional measure of faecal incontinence (as defined by leakage) and all items load highly on this factor/component which explains 54% of the variance. It has an appropriate internal structure which has been confirmed in both population and clinical settings.

The RFIS possessed evaluative discrimination by patient and clinician assessed incontinence severity but did not discriminate by unrelated health or socio-demographic variables (e.g. height, number of co morbidities, work status, education level etc.). Similarly, the RFIS appeared to be responsive over time to changes in incontinence status. These findings suggest the RFIS has both content and construct validity; i.e. it assesses the underlying condition of faecal incontinence and this assessment appears to be independent of possible confounders.

These findings regarding discrimination and responsiveness were observed despite the fact that this was an analysis with a small sample size (N = 61) which placed some limitations on the sub analyses that could be conducted (e.g. gender and type of treatment analyses). It would be desirable if a follow up study could continue to collect more data on faecal incontinence (particularly in males) using the same study protocols to analyse these aspects in more detail. In this study consecutive patients requesting treatment were recruited by the clinics, and although a number of faecal incontinence clinics participated the number of males recruited did not achieve the rates initially identified by the participating clinics. Also given the low prevalence of faecal incontinence amongst males in Australia in order to achieve a more substantive study sample of males (e.g. 84 at pre-treatment) it is thought a much larger number of clinics would need to be recruited.

No flatus item is included in the RFIS as it confounds prevalence estimates for epidemiological research. Internal consistency reliability of the Wexner at pre-treatment is also improved (from 0.65 to 0.69) if this item is removed. However, clinicians could use the Wexner or CRADI 8 flatus item for flatus assessment as a separate exercise.
4.3.12.2 Conclusions and Recommendations: Faecal Incontinence

Indications are that the RFIS performed well in clinical settings (as well as population settings) demonstrating good internal consistency reliability; correlations with other measures were in the expected directions providing evidence of validity; and there is evidence that it was sensitive to changes in continence status as a result of treatment making it suitable for outcome evaluation. The RFIS has also been shown to have superior psychometric properties to other commonly used faecal incontinence scales such as the Wexner and the St Mark’s Incontinence Score.

The above findings are based on a sample of 61 patients (48 females and 13 males at pre-treatment and 31 females and 8 males at post-treatment). It would be desirable if a follow up study could continue to collect more data on faecal incontinence (particularly in males) using the same study protocols to analyse gender and type of treatment aspects in more detail.

The RFIS is short and simple to use and score and continence clinics treating faecal incontinence patients should be encouraged to use the RFIS both as an assessment measures and as an outcome evaluation measure in routine practice.

4.4 Recommendations: Incontinence Assessment and Monitoring

With only 5 items each the RFIS and the RUIS are short and simple to use and score and continence clinics treating incontinence patients should be encouraged to use them both as assessment measures and as an outcome evaluation measures in routine practice.

It would be desirable if this data were to be collected routinely both prior to and following treatment. These are simple tools with which clinics can report on the effectiveness of their treatments. In the longer term the Continence Outcomes Section could consider the development of a continence outcomes data collaborative as has occurred with rehabilitation and palliative care (Australian Rehabilitation Outcomes Centre, the Palliative Care Outcomes Collaboration) although it is suggested that an online, real time framework is used rather than static warehousing given recent developments in this field. For example, the Diabetes Educators Association of Australia and a number of State Asthma groups use an online community to manage patients, share data, practices, and recruit patients to major conjoint research initiatives. Through such online collaborations clinics can compare their treatment outcomes and patterns of practice with other related clinics. Such organizations are useful change agents in promoting best practice within the field.

4.5 Incontinence Impact on Health Related Quality of Life

Sansoni et al. (2006) reported that urinary and faecal incontinence measures had negative but significant correlations with measures of health related quality of life reflecting the burden of disease for this condition (e.g. the higher the incontinence score the lower the physical functioning or health status score). At pre-treatment the RUIS correlation with the Physical Function Scale of the SF-36V2 was $r = -0.15$ (N = 190; $p = 0.05$) for the urinary incontinence sample and $r = -0.17$ (N = 248; $p = 0.007$) for the total sample of incontinence patients. The RFIS correlation with the Physical Function Index of the SF-36V2 was $r = -0.27$ (N = 59; $p = 0.05$) for the faecal sample and $r = 0.23$ (N = 248; $p = 0.001$) for the total sample at pre-treatment which, is also consistent with findings in the literature showing a negative association between faecal incontinence and generic measures of HRQOL. It is interesting that this was a higher correlation than was found for urinary incontinence suggesting that the impact of faecal incontinence may be greater although it should be noted that the faecal incontinence sample was slightly older and with a greater level of co-morbidities than the urinary incontinence sample.

The SF-36 profiles for urinary and faecal incontinence patients can be found below. The SF-36 contains eight health profiles (Physical Function–PF; Role Physical–RP, Bodily Pain–BP, General Health–GH, Vitality–Vit, Social Function–SF, Role Emotional–RE, Mental Health–MH) and two summary scores (Physical Component Summary Score–PCS and Mental Health Component Summary Score–MCS). The scores for the SF-36 are transformed scores to provide an average of 50 and a Standard Deviation of 10 based on the norms for the Australian population.
It can be seen that the transformed profile scores for urinary incontinence range from 36 to 46 for males and from 40-46 for females reflecting a slightly poorer health status and health related quality of life for these patients compared with Australian normative data. There was a significant gender difference between males and females for the Role Physical Scale (t = -3.58, df 189; p = 0.000) and Social Functioning Scale (t = -2.39, df 190; p = 0.019) with the male scores reflecting poorer health status on these variables. This would suggest that urinary incontinence for males is associated with greater limitations associated with work and other activities and the degree of interference that urinary incontinence may place on their social activities. The total score for both males and females combined on the Physical Component Summary (PCS) scale is 45.59 and for the Mental Component Summary (MCS) it is 43.43 reflecting a small but significant negative impact of urinary incontinence on these domains of health status.

The correlation between the PCS and the RFIS for the total sample of incontinence patients was $r = -0.23$ (p = 0.001; N = 248). The correlation between the MCS and the RFIS for the total sample was $r = -0.23$ (p = 0.001; N = 217).

For faecal incontinence the table below shows that most profile means were in the vicinity of 42-43 points and thus are about 0.7 of a standard deviation below the Australian population norms (t pcs = -4.54, t mcs = -3.75; df = 49, p < 0.01). The Physical Summary Component Score for all patients is 43.59 and the Mental Health Component Score is 43.41. This also confirms that patients with faecal incontinence conditions have poorer physical and mental health status scores compared to the Australian population and the burden of faecal incontinence is marginally greater than that for urinary incontinence. There are also low negative associations between incontinence and domain scores for the AQoL as was the case for the 2006 study (Sansoni et al., 2006).
Figure 4.18: SF-36 profile scores for faecal incontinence at baseline

4.6 Incontinence and the Burden of Disease

Incontinence is sometimes referred to as the ‘last taboo’ (Chiarelli, 2004) – a condition that is frequently overlooked and under researched. When estimating the burden of disease, incontinence is often subsumed under other health conditions or – particularly in the case of faecal incontinence – simply ignored (Begg et al., 2007, Goss, 2008, Mathers et al., 1999). Accordingly, this chapter provides an estimate of the excess burden of disease attributable to faecal and urinary incontinence.

4.6.1 Method

Although there are several methods of estimating the burden of disease, where detailed data are not available societal estimates can be provided from utility scores where population prevalences are known. This is the approach taken in this study.

4.6.1.1 Study perspective

Although there are many viewpoints from which evaluations can be undertaken, in economic evaluation the societal perspective is preferred because it represents the public interest (Drummond et al., 1998, Manning, 1999, Weinstein et al., 1996), particularly where there may be high patient, family or other sector costs. Utility (loss of quality of life) is presented by a global estimate from the Assessment of Quality of Life (AQoL) utility measure (Hawthorne et al., 1999b, Hawthorne et al., 2001). This estimate provides the broadest perspective since it is assumed it reflects all losses associated with ill health.

4.6.1.2 Incontinence prevalence

The prevalence of both faecal and urinary incontinence in Australian society is subject to widely varying estimates. Different researchers have used different definitions, different measures and different sampling methods to obtain their estimates (AIHW, 2006, Avery et al., 2004, Begg et al., 2007, Chiarelli et al., 1999, Goss, 2008, Hawthorne, 2006, Hughes et al., 2000, Kalantar et al., 2002, Mathers et al., 1999, Muscatello et al., 2001).

Given this situation in the interests of compatibility with the current study when calculating the excess burden of disease attributable to incontinence we used the participant reported outcome (PROs) measures the Incontinence Severity Index (ISI (Sandvik et al., 1993)) estimates for urinary
incontinence and Wexner estimates (Jorge and Wexner, 1993) for faecal incontinence, but excluding flatus (Sansoni et al., 2006). The source of the population prevalence data was the 2004 South Australian Health Omnibus Survey (SAHOS) (Hawthorne, 2006) extrapolated to the 2010 Australian Bureau of Statistics (ABS) population data for Australians over the age of 14 years (ABS, 2010).

The SAHOS is an annual user-pays population-based survey for health organisations comprising both metropolitan and rural samples. The metropolitan sample was based on the collectors’ districts used by the ABS 2001 census. Sampling was based on the probability of selection proportional to size, based on every fourth household. The rural sample was selected from towns with a population of 1000 or more, using the same procedure as for the metropolitan area. One interview was conducted per household and, when more than one person over the age of 15 resided in that household, the respondent was the person who last had a birthday. Interviews were conducted by trained and experienced interviewers. For reliability purposes re-interviews for selected questions were conducted on a random 10% of each interviewer’s interviews. A full description of the SAHOS methodology can be found elsewhere (Wilson et al., 1992).

For an extensive review of the ISI see Sansoni et al. (2006). Briefly, it comprises two items assessing frequency of urine leakage and the amount leaked. After Sansoni et al, the items and response scales are:

**How often do you experience urine leakage?**

- Never: 0
- Less than once a month: 1
- A few times a month: 2
- A few times a week: 3
- Every day and/or night: 4

**How much urine do you lose each time?**

- None: 0
- Drops: 1
- Small splashes: 2
- More: 3

Scoring is multiplicative with scores falling on a scale of 0-12. There is an issue, however, with this scoring because it will fail to detect cases who report incontinence on one of the items but not the other (e.g. consider a person who reports urine leakage a few times a week but who reports that he lost no urine; this gives 3*0 = 0 = continent; or a person who reports never leaking and small splashes; which gives 0*2 = 0 = continent). To overcome this difficulty, for this study the scoring was slightly modified such that a person reporting urine leakage (<monthly/ few times a month/ few times a week/ daily) and quantity (drops/ small splashes/ more) was not penalized. This was achieved through treating all 0 scores as missing data and then imputing the missing values based on the original responses. Given the uncertainties of PROs and issues with sparse data, ISI scores were collapsed into four categories: those with no symptoms (score range 0), those with slight symptoms (1-2), moderate symptoms (3-6) and those with severe/very severe symptoms (7-12) (Hawthorne, 2006, Sandvik et al., 1993).

For an extensive review of the Wexner see Sansoni et al. (2006). Briefly, the Wexner comprises five items assessing solid stool, liquid stool, flatus, soiling, and lifestyle alterations. Response categories are Never/ Rarely/ Sometimes/ Usually/ Always. Scoring is through simple summation. Consistent with the current International Continence Society definition of faecal incontinence (Norton et al., 2005), Sansoni et al. (2006) recommended removal of the flatus item and that has been done in this study. The available score range was 0 to 16. Given the uncertainties of PROs and issues with sparse data, Wexner scores were collapsed into four categories based on item response severity: those with no symptoms (reporting they never had any symptoms), those with
rare symptoms (reporting one episode in the past month), sometimes symptoms (2-4 faecal episodes in the past month), and those with weekly/daily symptoms (those reporting faecal incontinence weekly or daily) (Hawthorne, 2006, Jorge and Wexner, 1993).

4.6.1.3 Utility loss due to incontinence

The Assessment of Quality of Life (AQoL) instrument is a multi-attribute utility (MAU) instrument (Hawthorne et al., 1999a, Hawthorne et al., 2001). Although it consists of 15 items, 12 items are used in the utility scoring algorithm. These form 4 dimensions with 3 items each: Independent Living, Social Relationships, Physical Senses and Psychological Well-being. For scoring, individual item responses are replaced with community preference values which were obtained from a representative sample of the population using time-trade off (TTO). A multiplicative model combines these into the 4-dimension scores, again weighted by community preferences obtained through TTO. The resulting 4 dimension scores are then combined into a single score which is re-weighted (again from community sample based on TTOs) and presented as a utility score on a life-death scale where the endpoints are –0.04 (worse than death HRQoL equivalent state), 0.00 (death equivalent HRQoL state) to 1.00 (best HRQoL).

4.6.1.4 The impact of comorbidities

Previous Australian burden of disease studies reporting incontinence have either subsumed incontinence within other disease categories or have discounted incontinence on the basis of its relationship with other health conditions or comorbidities (Begg et al., 2007, Goss, 2008, Mathers et al., 1999). Where there are relationships between incontinence and comorbidities, utility estimates will be confounded, unless adjusted for during data analyses.

The relationship between incontinence and comorbidities was investigated through examination of 14 disability areas of life: restrictions in mobility, vision, hearing, breathing, sleeping, ingestion, speaking, cognition, pain, depression, distress and sexual activity. Each was dichotomized into no/slight disability versus moderate/severe disability and the number of disabilities summed (the lower limit was 0 and the upper limit was 10 disabilities). The resulting disability index was dichotomized into those with no/one disability and those with two or more disabilities.

4.6.1.5 The value of life

There is no agreement on the value of life in the literature. Generally, after Hawthorne et al (Hawthorne et al., 2003), life values are (a) estimates extracted from revealed behaviours (eg. observing wage differentials as a function of occupational risk) or (b) estimates based on discounted loss of lifetime earnings. This second approach is reflected in insurance payouts, court awards etc. Actual life-value estimates used in Australia in the past 15 years have ranged between $3-4 million or $151,000pa, which is based on allowing for 40 years of life lost and a discount rate of 3% (Abelson, 2007). This estimate is consistent with earlier estimates from the environment industry where the value of life ranges between AUD$1–5 million (Knieser, 1991). The transport sector uses $151,000 per life lost averted (quoted in Abelson, 2007).

In contrast, medical costs are lower. The UK medicine cost-effectiveness threshold has been set at £20,000 - £30,000; the range in terms of cost/QALY (quality adjusted life year) has been reported to be between £12,000 and £45,000 (Appleby et al., 2007). The Australian Pharmaceutical Benefits Advisory Committee (PBAC) will, in general, fund pharmaceuticals where the cost-per-life-saved is up to AUD$76,000 (George et al., 2001, Makarounas-Kirchmann et al., 2007, Sweeny, 2007). Abelson (2007) estimated that for many chronic morbidities the range for relief from condition would be between $60,000 to $80,000pa.

We accepted the value of $65,000pa which is consistent with Abelson’s (Abelson, 2007) estimate for chronic health conditions, the PBAC’s upper limit for funding pharmaceuticals (George et al., 2001, Makarounas-Kirchmann et al., 2007, Sweeny, 2007) and Appleby’s upper estimate for the UK (Appleby et al., 2007).
4.6.2 Results

Prevalence estimates for incontinence based on the 2004 SAHOS and extrapolated to the 2010 Australian population are shown in Table 4.13. Although not shown in the table for reasons of sparse data, the prevalence of urinary (faecal) incontinence was 38% (10%) for females and 10% (6%) for males. The table shows the proportion of participants with incontinence only (72% for urinary and 61% for faecal) and incontinence plus comorbidities (28% and 39%, respectively). For both urinary and faecal incontinence there were higher proportions with comorbidities, although in terms of absolute numbers these were a minority of those with incontinence. The figures are dominated by those who reported slight or moderate symptoms (23% for urinary and 7% for faecal); the overall prevalence of severe incontinence was 2% for urinary and 1% for faecal. The extrapolated data suggest that just over 4 million Australians suffer from any urinary incontinence symptoms as do just over 1 million Australians for faecal incontinence and that ~300,000 suffer from severe urinary and ~200,000 from severe faecal incontinence.
Table 4.12: Prevalence of incontinence in Australia (without and with comorbidities)

<table>
<thead>
<tr>
<th>Type</th>
<th>Comorbidities (c)</th>
<th>All</th>
<th>Australian population estimate (15+years) (b)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No</td>
<td>Yes</td>
<td>N = 3015</td>
</tr>
<tr>
<td>Urinary (d)</td>
<td>None</td>
<td>79.0%</td>
<td>57.4%</td>
</tr>
<tr>
<td>Slight</td>
<td>16.7%</td>
<td>26.3%</td>
<td>18.2%</td>
</tr>
<tr>
<td>Moderate</td>
<td>3.6%</td>
<td>10.4%</td>
<td>4.7%</td>
</tr>
<tr>
<td>Severe/Very severe</td>
<td>0.7%</td>
<td>6.0%</td>
<td>1.5%</td>
</tr>
<tr>
<td>Faecal (e)</td>
<td>None</td>
<td>94.3%</td>
<td>80.4%</td>
</tr>
<tr>
<td>Rare</td>
<td>4.6%</td>
<td>10.6%</td>
<td>5.5%</td>
</tr>
<tr>
<td>Sometimes</td>
<td>0.8%</td>
<td>4.7%</td>
<td>1.4%</td>
</tr>
<tr>
<td>Weekly/Daily</td>
<td>0.3%</td>
<td>4.3%</td>
<td>1.0%</td>
</tr>
</tbody>
</table>

Notes:
- b = ABS Catalogue 3101, Table 9, September 2010. Numbers are the calculated numbers from the proportion reporting incontinence status on the SAHOS.
- c = No = No/One comorbidity reported; Yes = 2 – 10 comorbidities reported.
- d = Estimate based on ISI.
- e = Estimate based on Wexner (excluding flatus).

Utility value for health-related quality of life (HRQoL) as measured by the AQoL is shown in Table 4.14. As shown there is a monotonic loss of HRQoL with increasing incontinence severity, regardless of incontinence type. The very large 95% CIs for those with the most severe levels of incontinence are due to the very small numbers (see Table 4.13). A notable feature of the table is the difference in utilities between those with no comorbidities and those with comorbidities. This difference, for example, is 0.34 utilities for both those with no urinary and faecal incontinence symptoms. This finding suggests that comorbidities play an extremely important role in incontinence.
Table 4.13: 
Utility value (i.e. quality of life), using the AQoL, by incontinence and comorbid status

<table>
<thead>
<tr>
<th>Type</th>
<th>No comorbidities</th>
<th>With comorbidities</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>95%CI</td>
<td>Mean</td>
</tr>
<tr>
<td>Urinary (a)</td>
<td>None</td>
<td>0.88</td>
<td>0.87-0.89</td>
</tr>
<tr>
<td></td>
<td>Slight</td>
<td>0.84</td>
<td>0.83-0.85</td>
</tr>
<tr>
<td></td>
<td>Moderate</td>
<td>0.81</td>
<td>0.77-0.85</td>
</tr>
<tr>
<td></td>
<td>Severe/Very severe</td>
<td>0.74</td>
<td>0.64-0.84</td>
</tr>
<tr>
<td>Faecal (b)</td>
<td>None</td>
<td>0.87</td>
<td>0.86-0.88</td>
</tr>
<tr>
<td></td>
<td>Rare</td>
<td>0.80</td>
<td>0.77-0.83</td>
</tr>
<tr>
<td></td>
<td>Sometimes</td>
<td>0.80</td>
<td>0.74-0.86</td>
</tr>
<tr>
<td></td>
<td>Weekly/Daily</td>
<td>0.75</td>
<td>0.70-0.80</td>
</tr>
</tbody>
</table>

Notes:
a = Estimate based on ISI.
b = Estimate based on Wexner (excluding flatus).

The estimated excess burden of disease associated with incontinence is shown in Table 4.15, by incontinence type, level of severity and comorbid status. The loss in utility, derived from Table 4.14, suggests that the relative impact of urinary and faecal incontinence is similar by incontinence severity. The data also suggest that the loss in utility is similar by the presence or absence of comorbidities, other than for those with severe incontinence where those with comorbidities suffer utility losses that are just under double those suffered by those without comorbidities. This pattern is also reflected in the costs borne by individuals. For those with slight incontinence the excess costs are higher among cases with no comorbidities; a situation that is reversed for those with severe incontinence. At the aggregate level, the highest excess cost to society is borne by those with slight/rare incontinence because of the number of people suffering this level of incontinence.

The very broad 95%CIs reflect the uncertainties of the estimates, particularly where there are small numbers of cases (see Table 4.13) and where there were large variations in AQoL utility scores (Table 4.14). As shown in the table, the excess burden of disease associated with urinary incontinence was estimated at approximately ~$17 billion per annum and for faecal it was approximately $8 billion per annum. The total burden of disease was estimated at approximately $25 billion per annum. As the 95%CIs show, this estimate could be as low as $9 billion or as high as $46 billion per annum.
<table>
<thead>
<tr>
<th>Incontinence type and level</th>
<th>Incontinence prevalence by comorbid status (b)</th>
<th>Value (c)</th>
<th>Utility loss by comorbid status (95%CI) (d)</th>
<th>Annual per case value lost by comorbid status (95%CI) (e)</th>
<th>Estimated total value lost/year by comorbid status (millions; 95%CI) (f)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Urinary Slight</td>
<td>2836439</td>
<td>838219</td>
<td>$65,000</td>
<td>0.04</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>(0.02-0.06)</td>
<td>(0.00-0.10)</td>
<td>($1,300-$3,900)</td>
<td>(0-6,500)</td>
<td>($3,687-$11,062)</td>
</tr>
<tr>
<td>Moderate</td>
<td>611448</td>
<td>328276</td>
<td>$65,000</td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>(0.00-0.12)</td>
<td>(0.00-0.17)</td>
<td>($1,300-$7,800)</td>
<td>($0-$11,050)</td>
<td>($795-$4,769)</td>
</tr>
<tr>
<td>Severe/Very severe</td>
<td>118893</td>
<td>191229</td>
<td>$65,000</td>
<td>0.14</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>(0.03-0.25)</td>
<td>(0.12-0.34)</td>
<td>($1,950-$16,250)</td>
<td>($7,800-$22,100)</td>
<td>($232-$1932)</td>
</tr>
<tr>
<td>S/Total</td>
<td>781295</td>
<td>337837</td>
<td>$65,000</td>
<td>0.07</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>(0.03-0.11)</td>
<td>(0.00-0.12)</td>
<td>($1,950-$7,150)</td>
<td>($0-$7,800)</td>
<td>($1,524-$5,596)</td>
</tr>
<tr>
<td>Faecal Rare</td>
<td>135877</td>
<td>149796</td>
<td>$65,000</td>
<td>0.07</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>(0.00-0.14)</td>
<td>(0.00-0.22)</td>
<td>($0-$9,100)</td>
<td>($0-$14,300)</td>
<td>($0-$1,236)</td>
</tr>
<tr>
<td>Sometimes</td>
<td>50954</td>
<td>137047</td>
<td>$65,000</td>
<td>0.12</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>(0.06-0.18)</td>
<td>(0.07-0.33)</td>
<td>($3,900-$11,700)</td>
<td>($4,550-$21,450)</td>
<td>($199-$596)</td>
</tr>
<tr>
<td>S/Total</td>
<td>4,571</td>
<td>3,439</td>
<td>$8,010</td>
<td>$4,714-$17,763</td>
<td>($1,492-$13,302)</td>
</tr>
</tbody>
</table>

Notes:

a = Assumes independence in costs (i.e. that persons with dual incontinence experience separate costs).
b = Calculated from Table 4.13. Because of the small numbers of males with incontinence these have been collapsed.
c = Value of life. See text for derivation of this life value.
d = Calculated from Table 4.14.
e = Column b x Column c.
f = Column d x Column a.
4.6.3 Discussion

Based on disutility due to incontinence this study has shown that the excess burden of disease associated with incontinence in Australia is considerable, with the estimate, from a societal perspective, at $25 billion per annum.

This estimate is subject to several caveats. The SAHOS sample was based on a single Australian state – South Australia. Although the data were weighted by the 2004 ABS Census to ensure representativeness, the extent to which South Australia is representative of all Australia is unknown. The classification of cases to incontinence type and severity is based on self-reported PROs. The extent to which this reflects the clinical situation is unknown.

The value of life estimate is also debatable. Although it was based on a review of the literature we have no evidence confirming that $65,000 is an appropriate estimate and it may well be that this should be a much higher or lower figure. It was assumed that urinary and faecal cases would have independent costs; for those cases with dual diagnosis (both urinary and faecal incontinence) should the value of life and costs be related then the excess costs will be slightly overestimated. We also made the assumption that utility scores on the AQoL were solely due to incontinence; although we adjusted the model for comorbidities and presented the calculated adjusted utilities these estimates are open to discussion. We have also assumed that AQoL utilities represented all aspects of life (e.g. time off from work). With respect to this point, several other estimates have reported on treatment costs only (i.e. the studies have adopted the health service sector perspective), hence the costs in this report are much higher due to the societal perspective of the study. For example, Subak et al. (2008) reported that women with stress incontinence paid US$750pa for treatment and were willing to pay US$1,400pa for treatment costs which relieved their incontinence; other costs were not considered.

In addition to these considerations, the estimate from this study is likely to be an underestimate because the SAHOS recruitment method excludes those living in residential or nursing home care. The literature suggests that among older adults around 44% of those living in residential or nursing homes suffer urinary or faecal incontinence (Peet et al., 1995). Regarding the prevalence of faecal incontinence, it is reassuring to note that for those with moderate or severe faecal incontinence (those reporting faecal incontinence sometimes or weekly/daily; Table 4.13) these were reasonably similar to UK estimates of 1.7% and 1.4% (Perry et al., 2002).

Subject to these caveats, the study findings showing that society carries considerable incontinence excess burden of disease stands in stark contrast with other estimates. In 1999 Mathers et al. (Mathers et al., 1999) reported that the years of life lost due to disability (YLD) for urinary incontinence was just 1.1% of all YLDs in Australia and was ranked the 40th leading condition contributing to disability adjusted life years (DALYs) for females and was not ranked at all for males. The updated Begg et al. (2007) burden of disease study did not report separately on incontinence, but subsumed it within the genitourinary classification which was responsible for 2.5% of DALYs in 2003. The more recent Goss report (Goss, 2008) projecting health care costs for Australia included urinary incontinence costs with other genitourinary disorders and reported that expenditure for 2012-13 was predicted to be $4.9 billion rising to an expected $10.9 billion for 2032-33 – placing these increases 11th behind diseases such as diabetes, dementia, dermatological conditions and sense disorders. The value of these estimates of the incontinence burden of disease rests on the assumptions behind the calculations.

In the current study, based on extrapolation of the 2004 SAHOS population survey, the number of female urinary incontinent cases was estimated at 4,925,000 cases of which 72% did not have comorbidities (Table 4.13). This figure is far higher than in previous reports. Mathers et al. (1999) only reported on severe female urinary incontinence cases; the prevalence was estimated to be 307,210 with an annual incidence rate of 12,985 cases – a figure very similar to that of the current study for males and females with severe urinary incontinence (310,000). Although the similarity of the two estimates is reassuring, the main point is that the Mathers et al. estimates ignored those with slight or moderate incontinence yet the current study found that these cases were responsible for three-quarters of all excess costs. The Begg et al. (2007) study reported 195,000 female
incontinent cases – a figure which may have reflected the study assumptions regarding co-morbidities described below. The Goss (2008) report into Australia’s future health care costs did not define urinary incontinence, but it is likely the Mathers et al. definition was used since the Mathers et al. (1999), Begg et al. (2007) and Goss (2008) estimates were drawn from the Australian Bureau of Statistics Survey of Disability, Ageing and Carers (1998) and the Australian Longitudinal Study of Women’s Health Survey (1996-2002) (the Begg et al. (2007) study also used estimates from the Lea 1993 study). Regarding male incontinence, this was not considered separately in the Goss (2008) study. Mathers et al. and Begg et al. both considered male incontinence, but only in the context of outcomes from prostate cancer surgery.

In part the different prevalence levels may also be a function of definition. The current study defined urinary incontinence according the internationally developed and widely used Incontinence Severity Index; adjustment was made for the presence of 14 different co-morbidities. The Mathers et al. study definition of urinary incontinence was that urine had to leak ‘often’ and that this was not due to neurological disorders, stroke, prostate problems or other diseases or injury. Begg et al. (2007) did not provide a definition, but discounted urinary incontinence by 50% on the grounds that half of all incontinence was due to multifactorial causes and was captured within other disabilities (e.g. prostate cancer for males, or severe stroke). In the current study we observed that comorbidities were present in only 28% of all urinary cases. Regarding faecal incontinence, in the current study this was classified by the modified Wexner scale (excluding flatus), whereas Mathers et al., Begg et al. and Goss did not consider faecal incontinence at all.

A third difference relates to the disutility or disability weights used in the different studies. This study used the AQoL which is weighted with Australian values based on the time-trade off. The Mathers et al. disability weights were derived from the EQ5D+ instrument using British weights for moderate incontinence and a regression weight for severe incontinence. No weight was assigned for those with mild urinary incontinence.

The findings from this study are based on the population prevalence of incontinence in 2004 extrapolated to the 2010 Australian population using utility losses and a best estimate of the value of life. The results show that the excess burden of disease is higher than previously estimated at $25 billion per annum, although there are very broad confidence intervals around this figure given the uncertainties of the study. With the aging of the Australian population (Productivity Commission, 2005) and that incontinence is associated with ageing, with increases up to middle age, a plateau until the age of about 70 years and then further increases in the proportion of people with incontinence (Hunskaar et al., 2005, Perry et al., 2002), it is likely that, ceteris paribus, the excess burden of disease associated with incontinence will increase substantially over the next 15-20 years (Fonda et al., 2005).

4.6.4 References


4.7 Faecal Cost of Illness Sub-Study

This section is concerned with the Faecal Cost of Illness sub-study that was conducted by the St George Hospital Pelvic Floor Unit. This short paper was submitted and accepted by the International Continence Society (2011).

4.7.1 Aims

Although there have been a few of studies regarding the cost of faecal incontinence in the past decade, these studies mainly employed previously collected databases (Dunivan et al., 2010, Sung et al., 2010) or focused upon inpatients subjects (Morris, 2005) with constipation as the main complaint (Dunivan et al., 2010), women with faecal incontinence after obstetric injury (Mellgren, 1999). Few conducted face to face direct enquiry of the personal costs of leakage of faeces (Morris, 2005). Most did not employ a validated measure of severity (Dunivan et al., 2010, Sung et al., 2010, Morris, 2005, Dowell et al., 1999). Thus this study aimed to conduct personal interviews with a broad sample of ambulatory home dwelling patients who presented with faecal incontinence to a tertiary Unit, and to collect cost data prior to the onset of treatment in relation to baseline severity.

4.7.2 Study design materials and methods

A consecutive series of patients attending a tertiary outpatient clinic with a main complaint of faecal incontinence were interviewed, using a 3 page questionnaire, modelled on the DBIC/CI questionnaire for urinary incontinence (Dowell et al., 1999). Interviews were conducted face to face, taking approximately 15 minutes each.

The information collected included basic personal hygiene costs (pads, laundry, wipes, and cleansers), medication costs (loperamide, creams & stool bulking agents etc) and diagnostic costs (including medical attendance, anorectal physiology, and colonoscopy). Also at this visit the patients completed a St Mark’s score (Vaizey et al., 1999) to gauge severity of faecal incontinence (max score 24).

Following each interview, the personal hygiene items used by patients were costed from known tables compiled by visiting local pharmacies and suppliers. These costs were further broken down...
into personal “out of pocket” expenses, Medicare subsidised costs, and health fund rebated expenses, over the last 12 months. Costs were recorded in AU dollars.

As shown in Table 4.16, the category of Total Patient Expenses comprised pads, medication, reusable items, laundry, miscellaneous additional expenses and the patients’ “out of pocket” costs for diagnostic tests. The cost attributed to the Medicare bulk billing system (“government expense”) and any rebates provided by patients health fund (“Health Fund Expenses”) were then added in to yield the Total Incontinence Expenses.

### 4.7.3 Results

A sample of 54 consecutive patients [5 male, 49 female; age 35 – 91, median 69.5, IQR 61.5-74.0] performed the Faecal Cost Questionnaire (Dowell et al., 1999) and the St Mark’s score. Figure 4.19 shows breakdown of all personal and investigation costs (blue) and the values for Medicare subsidies (green), private health rebates (cream).

*Figure 4.19: Incontinence expenses by severity group*

![Figure 4.19](image)

In terms of patient “out of pocket”, costs the major expense overall were for pads and personal hygiene items (median 70.89 per annum, IQR 0.63-310.68). The bulk of Medicare costs included medical consultation and rebates for physiology testing and endoscopy (median $576.92 IQR). Figure 4.19 above showed no relationship between these overall costs and incontinence severity.

Therefore we analysed the subset of Personal Costs alone; we hypothesized that these costs should directly relate to severity.
The Spearman rank correlation showed no significant relation between total personal costs and severity ($r_s = 0.21$), as this cost category includes diagnostic tests that are relatively uniform, regardless of severity. When we drilled down into the costs just for pads and creams alone, a relationship became apparent ($r = 0.34$, $p = 0.05$).

### Table 4.15: Summary of all costs by category, and summation

<table>
<thead>
<tr>
<th>Category</th>
<th>N</th>
<th>Range</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Incontinence Expenses</strong></td>
<td>54</td>
<td>1,248</td>
<td>0</td>
<td>1,248</td>
<td>209</td>
<td>291</td>
</tr>
<tr>
<td>Pads</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Medication Expenses</strong></td>
<td>54</td>
<td>1,745</td>
<td>0</td>
<td>1,745</td>
<td>110</td>
<td>262</td>
</tr>
<tr>
<td>Prescription meds</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Prescription meds</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Creams</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Reusable Expenses</strong></td>
<td>54</td>
<td>42</td>
<td>0</td>
<td>42</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Chair Pads</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Washable pants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bedcovers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Laundry Expenses</strong></td>
<td>54</td>
<td>109</td>
<td>0</td>
<td>109</td>
<td>18</td>
<td>23</td>
</tr>
<tr>
<td>Washing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Additional Expenses</strong></td>
<td>54</td>
<td>1,500</td>
<td>0</td>
<td>1,500</td>
<td>62</td>
<td>208</td>
</tr>
<tr>
<td>Extra toilet paper</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wet wipes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extra personal cleaning Equip</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Patient diagnostic expense</strong></td>
<td>54</td>
<td>1,858</td>
<td>0</td>
<td>1,858</td>
<td>210</td>
<td>385</td>
</tr>
<tr>
<td>Colonoscopy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sigmoidoscopy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total Patient Expense</strong></td>
<td>54</td>
<td>2,806</td>
<td>1</td>
<td>2,807</td>
<td>609</td>
<td>597</td>
</tr>
<tr>
<td>Government expense</td>
<td>54</td>
<td>1,522</td>
<td>135</td>
<td>1,657</td>
<td>670</td>
<td>300</td>
</tr>
<tr>
<td>Health fund expense</td>
<td>54</td>
<td>1,628</td>
<td>0</td>
<td>1,628</td>
<td>506</td>
<td>710</td>
</tr>
<tr>
<td><strong>Total Incontinence Expense</strong></td>
<td>54</td>
<td>3,482</td>
<td>243</td>
<td>3,725</td>
<td>1,386</td>
<td>989</td>
</tr>
</tbody>
</table>

### 4.7.4 Discussion

In this face to face study of personal and medical costs of patients presenting for the first visit with faecal incontinence, the Total Costs of all personal items and investigations (regardless of the
Payer) did not directly relate to severity. This arises because the costs of investigation are largely fixed.

However, the total Personal Costs of hygiene items increased with severity of incontinence, as would be expected. Patients often reported during their interviews that they wore pads “just in case”.

4.7.5 Concluding message

To our knowledge, this is the first report of detailed costs for faecal incontinence in ambulatory home-dwelling men and women, obtained by personal interview. The relationship between severity and expense is indeed multifactorial, and reflects the sufferer’s fear of potential leakage rather than actual necessity.

4.7.6 References


5 Patient Satisfaction

5.1 Patient Satisfaction Instruments Used

This report is a validation of the SAPS (Short Assessment of Patient Satisfaction) (Hawthorne et al., 2006) for use with patients receiving treatment for their incontinence. In addition to the SAPS, participants were also administered the GUTSSV2 (Version 2; Hawthorne et al., 2006), a specialist incontinence patient satisfaction measure, as a reference patient satisfaction measure.

Because neither the GUTSS nor SAPS have been published in the academic literature, although both are derived from earlier versions, they are described as the GUTSS Version 2 (GUTSSV2) and SAPS. The earlier versions are described as the GUTSS Version 1 (GUTSSV1) and draft SAPS.

5.1.1 The SAPS

The SAPS consists of seven items assessing the core domains of patient satisfaction (Hawthorne, 2006b, Hawthorne et al., 2006). The items assess treatment satisfaction, explanation of treatment results, clinician care, participation in medical decision-making, respect by the clinician, time with the clinician, and satisfaction with hospital/clinic care. Responses scales are 5-point scales. To limit the potential for acquiescent response set bias (ARSB) – which is where a person gives the same endorsement regardless of item content (Couch and Keniston, 1960, Trott and Jackson, 1967) – the SAPS was designed to be a balanced scale (Crowne and Marlowe, 1960, Miller and Cleary, 1993), that is a scale in which there are equal numbers of both positive and negative items. The SAPS contains three positively worded items and four negatively worded items. Prior to scoring, responses on the negatively worded items are reversed. Scores are obtained by simply summing the items. A copy of the SAPS can be found in Appendix D.

The internal consistency of the draft SAPS has been previously reported at Cronbach’s alpha = 0.85. In the same study the SAPS correlated with other measures of patient satisfaction 0.73 with the Consultation Satisfaction Questionnaire, 0.78 with the Client Satisfaction Questionnaire-18 and 0.83 with the Patient Satisfaction Index (Hawthorne et al., 2006).

The SAPS used in this study was slightly modified from the draft SAPS through changes to the wording of three items. The original items assessed happiness with treatment effects (#1), time with the clinician was not long enough (#6) and happy with received care (#7) (Hawthorne et al., 2006). Following consultation with international researchers after presentation of the draft SAPS at the International Quality of Life Conferences in Budapest (Hungary, 2007) and Montevideo (Uruguay, 2008), items #1 and #7 were revised in the interests of greater cross-cultural applicability given that ‘happiness’ with medical care may have been colloquial and not reflect patients’ constructs relating to health cross-culturally. Item #6 was revised to make the language simpler. The revised items were satisfied with treatment effect (#1) that time with the clinician was too short (#6) and satisfied with received care (#7). In addition to these modified items, the original draft items were also asked. Finally, the item probing satisfaction with the clinician’s explanations of the treatment results was asked in two forms: as a negative item (SAPS #2) and as a positive item from the GUTSSV2.

5.1.2 The GUTSSV2

The GUTSSV2 (Hawthorne et al., 2006) consists of eight items assessing happiness with treatment, continuing problems with incontinence, satisfaction with treatment outcome, disappointment with outcomes, pre-treatment information, attitude/behaviour of the treating clinician, post-treatment explanations and satisfaction with hospital/clinic care. Responses scales are 5-point scales. The reliability of the GUTSSV2 has been previously reported at Cronbach’s alpha = 0.89. In the same study the GUTSSV2 correlated with other measures of patient satisfaction 0.48 with the Consult SQ, 0.70 with the CSQ-18 and 0.64 with the PSI (Hawthorne et al., 2006).
The GUTSSV2 used in this study was that reported by Hawthorne et al (Hawthorne, 2006b) and differs from the original GUTSS Version 1 (Hawthorne and Harmer, 2000) through removal of two filter items and modification of remaining items. The purpose of revision was to shorten the GUTSSV1, make it more user-friendly and easier to score (Hawthorne et al., 2006, Hawthorne, 2006b). The Spearman correlation between the GUTSSV1 and GUTSSV2 was reported in a urinary incontinence treatment sample as \( r_s = 0.97 \) (Hawthorne et al., 2006) – suggesting they were equivalent scales.

### 5.2 Method

Missing data were observed on two SAPS items; there was one datum missing from item #2 and two from #5. Horizontal mean imputation was used to impute these values.

Items were examined by descriptive statistics. Examination of SAPS items and scores showed that they were statistically skewed (significance of skew, \( p<0.05 \)). Data were therefore transformed prior to analysis. This is indicated in the text by the subscript ‘trans’. Untransformed data (e.g. means and standard deviations) are presented in the interests of readability.

The criteria for assessing the adequacy of item-level data were that all response levels were endorsed, that >5% of endorsements fell on any 2 adjacent categories, that <60% of endorsements were on any one category, and that there was evidence of monotonicity as defined by the absence of disordered thresholds under Rasch item response theory analysis (Bond and Fox, 2001, Pedhazur and Schmelkin, 1991, Streiner and Norman, 2003). Acquiescent response set bias (ARSB) was defined as satisficing (Couch and Keniston, 1960, Krosnick, 2000) and was deemed to be present where it was observed in >10% of cases (Ware et al., 1983). The criterion for detecting ARSB was identical extreme responses across three consecutive items including both positive and negatively worded items. Related to this, is whether respondents interpreted items by the item direction (positive/negative). Item rest of test correlations (IRTC) were examined where the criteria was \( r>0.40 \) after correction for overlap (McHorney et al., 1994).

Scale ceiling and floor effects were observed and regarded as important where these were observed in >19% of cases, which was the average percentage of cases obtaining ceiling scores on four patient satisfaction measures (Hawthorne et al., 2006).

To assess its validity, the SAPS was subjected to three examinations: (a) its data distribution; (b) analysis of its internal structure using two statistical approaches: classical test theory (factor analysis; reliability) and item response theory (Rasch analysis); and (c) tests of sensitivity (discriminatory power) to known groups and correlation with the GUTSSV2. Both factor and Rasch analyses require data distributions preventing undue leverage from sparse data. Where there are small sample sizes (<150) (Guadagnoli and Velicer, 1988) the convention is that where <10% of responses fall on any given response level, the level should be collapsed with an adjacent level. As shown in Table 5.3, four of the SAPS items had data distributions with <10% on the extreme two response levels. Accordingly SAPS item responses were recoded into three levels through collapsing the three most extreme response options prior to the factor, reliability and Rasch analyses.

Principal component analysis (PCA) has been recommended as a precursor to Rasch analysis for establishing unidimensionality and to inform the subsequent Rasch analysis (Bjorner et al., 2003, Luquet et al., 1996, Tennant and Pallant, 2006). Unidimensionality with PCA is established by meeting three criteria: that Kaiser’s lower boundary for the number of factors is met, viz., \( m/3 \) where \( m \) is the number of items (Kaiser, 1960) – for the SAPS scale this implies that no more than 2 factors with eigenvalues >1.00 would be identified; that the principal component would explain >40% of the variance (Carmines and Zeller, 1979) and that the eigenvalue ratio between the factors would be >2.00 (Lumden, 1961, Luquet et al., 1996). Regarding sub-scales, we assumed that if these were identified they would be interlacing and accordingly used exploratory factor analysis with an oblimin rotation in addition to the PCA.

The Rasch model rests on two axioms. The first requires that items measure a single underlying construct, thereby forming a unidimensional scale, which is tested through principal component analysis (described above). The second is that items have local independence; viz., that
responses to one item are not dependent on responses to another after controlling for a person's ability (Hambleton et al., 1991). These two axioms are related: where unidimensionality exists local independence also exists, but evidence of local independence does not imply unidimensionality. From this it follows that where data are consistent with the Rasch model the person responses are expected to be consistent with the order of item difficulty which is indicated by a total-item \( \chi^2 \) probability greater than 0.05. Thus local independence can be tested through examination of item residuals. A finding of no association between residuals for individual items has been argued as evidence of local item independence. High positive correlation between residuals provides evidence of local item dependence, and high negative correlation violation of local independence (Smith, 2002).

In addition, item thresholds were studied to investigate the existence of disordered thresholds (i.e. response patterns for item categories that are not in the expected order). Item misfit was considered if the \( \chi^2 \) or \( F = \) statistic probability value was less than 0.05 or the fit residuals were greater than \( |2.50| \), which is equivalent to setting the test value to \( p<0.01 \). The reason for setting this stringent criteria was to avoid chance capitalization due to item parameter estimation errors which may occur where \( p<0.05 \) or \( |2.00| \) (Hambleton et al., 1993).

Differential item functioning (DIF) was investigated to identify items that operated differently for people of the same level of ability (i.e. have the same amount of the underlying trait of interest) who share another feature (variable). DIF = was examined by gender (male/ female), age group (<40 years/40-49 years/ 50-59/ 60-69/ 70+ years), and study cohort (urinary/faecal incontinence status). DIF = was considered significant if \( p<0.05 \). The person separation index (PSI) was reported to provide an indication of the ability of the SAPS to discriminate among known groups; a PSI 0.70-0.79 implied that the SAPS would be able to discriminate between two groups and >0.80 between three groups (Fisher, 1992).

A key axiom of psychometrics is that the selection of items in an instrument should constitute a sample of all possible items assessing the construct of interest. Under this postulate there are always alternative items to those selected for the descriptive system. An important test of construct validity, therefore, is whether alternate items provide improved isomorphism between the manifest descriptive system and the construct of interest. In the current study, three of the draft SAPS items were replaced with slightly differently worded items and a variation of a fourth was administered, as described above. The effect of replacement was investigated by examining the agreement between the new revised and draft item versions where the criterion was kappa >0.60 indicating at least substantial agreement (Landis and Koch, 1977) and by examining changes to the internal structure of the SAPS PCA and Rasch analyses following replacement of the new standard items with the original draft items.

Various tests of convergent/divergent discrimination by known groups were carried out. Tests on which it was hypothesized there would be no significant differences in mean SAPS scores were tests by gender (male/female), age group (<60 years/≥60 years), education attainment (primary/high/trade, TAFE (Technical and Further Education) and college diploma/university degree) and incontinence type (faecal/urinary). Tests for which it was hypothesized there would be statistically significant differences in mean SAPS scores were by general health at follow-up, incontinence treatment and by correlation between SAPS and GUTSSV2 scores. Regarding this correlation the criterion for establishing that the SAPS also assesses patient satisfaction was set at >0.75 but <0.85 (Kline, 2000). Finally, SAPS scores were examined by changes in incontinence status by clinician report, self-reported change in incontinence severity, and change on the RUIS and RFIS.

Data analysis was performed in SPSS Version 18 (SPSS 2008) and RUMM2020 (Andrich et al., 2005).

### 5.3 Characteristics of the patient satisfaction sample

The patient satisfaction sample comprised 136 patients who provided post-treatment data; these patients comprised 54% of the total study population, raising issues relating to sample bias; viz., that those who remained in the study were those with the more severe incontinence condition.
therefore biasing results through the overstatement of continuing health problems and the understatement of treatment effect (Moser and Kalton, 1971). This was investigated through examination of pre-treatment demographics and incontinence severity by status where study status was dichotomized into those providing pre-post data and those providing pre-treatment data only. The pre-treatment clinician rating of patient incontinence severity was reclassified into 2 groups (Group 1 = normal/mild/moderate incontinence, N = 211; Group 2 = severe incontinence, N = 41) and cross-tabulated with pre-post and pre-treatment status. The results showed that those providing pre-post data were statistically significantly more likely to have baseline severe incontinence ($\chi^2 = 4.27, \text{df} = 1; p = 0.04$) although this was not significant when the classification combined the mild and normal group and compared these with a combined moderate and severe group. The same classification and analysis with patient rated severity (Group 1 = normal/mild/moderate, Group 2 = severe) also did not confirm this ($\chi^2 = 0.06, \text{df} = 1; p = 0.81$). There were no significant differences by incontinence type (urinary/faecal), by gender, education attainment, age group, country of birth, the number of comorbidities or general health status.

Demographic patient details are provided in Table 5.1. This shows that 74% were urinary incontinence cases and 26% faecal incontinence cases. There were no statistically significant differences by incontinence type by any of the demographic variables.

### Table 5.1: Demographic characteristics of patients providing patient satisfaction data

<table>
<thead>
<tr>
<th>Incontinence type</th>
<th>Urinary (N = 100)</th>
<th>Faecal (N = 36)</th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>86%</td>
<td>83%</td>
<td>$\chi^2 = 0.15, \text{df} = 1; p = 0.70$</td>
</tr>
<tr>
<td>Male</td>
<td>14%</td>
<td>17%</td>
<td></td>
</tr>
<tr>
<td>Age group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;60 years</td>
<td>47%</td>
<td>34%</td>
<td>$\chi^2 = 1.70, \text{df} = 1; p = 0.19$</td>
</tr>
<tr>
<td>≥60 years</td>
<td>53%</td>
<td>66%</td>
<td></td>
</tr>
<tr>
<td>Country of birth</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>72%</td>
<td>80%</td>
<td>$\chi^2 = 0.92, \text{df} = 1; p = 0.34$</td>
</tr>
<tr>
<td>Other</td>
<td>28%</td>
<td>20%</td>
<td></td>
</tr>
<tr>
<td>Education attainment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>27%</td>
<td>37%</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>26%</td>
<td>23%</td>
<td></td>
</tr>
<tr>
<td>Trade/TAFE/Diploma</td>
<td>21%</td>
<td>17%</td>
<td></td>
</tr>
<tr>
<td>University</td>
<td>26%</td>
<td>23%</td>
<td>$\chi^2 = 1.24, \text{df} = 3; p = 0.75$</td>
</tr>
<tr>
<td>Labour force</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>participation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>36%</td>
<td>20%</td>
<td></td>
</tr>
<tr>
<td>Retired</td>
<td>40%</td>
<td>54%</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>24%</td>
<td>26%</td>
<td>$\chi^2 = 3.24, \text{df} = 2; p = 0.20$</td>
</tr>
</tbody>
</table>

The most commonly reported health conditions were neck, spine or back pain (49%), arthritis (49%), high blood pressure (36%), being obese (BMI $\geq$ 30 (WHO Expert Committee on Physical Status, 1995), 30%), and migraine (26%). Overall, the number of comorbid conditions ranged from 0 to 9. Due to sparse data these were collapsed into the three categories reported in Table 5.2. As shown, most participants reported two, three or four comorbid conditions.

In terms of self-reported general health, 39% reported being excellent or very good health, 33% in good health and 27% in fair or poor health. The mean T-scores on the SF-36 Version 2 suggested that participants obtained scores about half a standard deviation below the Australian population mean (50 points ±10 points) (Hawthorne et al., 2007), that their incontinence symptoms had been present for more than three years, and that for most participants their symptoms were moderate. On none of these health status indicators was there any statistically significant difference between those with urinary and faecal incontinence.
### Table 5.2: Health status at baseline of patients providing patient satisfaction data

<table>
<thead>
<tr>
<th>Incontinence type</th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Urinary (N = 100)</td>
</tr>
<tr>
<td>Number of comorbidities</td>
<td></td>
</tr>
<tr>
<td>0/1</td>
<td>24%</td>
</tr>
<tr>
<td>2/3/4</td>
<td>52%</td>
</tr>
<tr>
<td>5+</td>
<td>24%</td>
</tr>
<tr>
<td>General health status</td>
<td></td>
</tr>
<tr>
<td>Excellent/Very good</td>
<td>42%</td>
</tr>
<tr>
<td>Good</td>
<td>33%</td>
</tr>
<tr>
<td>Fair/Poor</td>
<td>25%</td>
</tr>
<tr>
<td>Physical health (SF-36 PCS)</td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>45.86 (10.60)</td>
</tr>
<tr>
<td>Mental health (SF-36 MCS)</td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>43.72 (12.89)</td>
</tr>
<tr>
<td>Time experienced incontinence symptoms</td>
<td></td>
</tr>
<tr>
<td>≤1 year</td>
<td>13%</td>
</tr>
<tr>
<td>1-&lt;2 years</td>
<td>20%</td>
</tr>
<tr>
<td>2-&lt;3 years</td>
<td>20%</td>
</tr>
<tr>
<td>3+ years</td>
<td>47%</td>
</tr>
<tr>
<td>Severity of symptoms</td>
<td></td>
</tr>
<tr>
<td>Normal/Mild</td>
<td>32%</td>
</tr>
<tr>
<td>Moderate</td>
<td>47%</td>
</tr>
<tr>
<td>Severe</td>
<td>21%</td>
</tr>
</tbody>
</table>

### 5.4 The SAPS – Descriptive Examination

The data distributions for the SAPS items were examined. As expected (since most patients are satisfied with their health care and treatment) this revealed that all items were statistically significantly skewed. As shown in Figure 5.1 all response categories were endorsed by participants, there were only two items where any two adjacent levels were endorsed by <5% of participants (items #5 and #7; levels 4 and 5), only one item had >60% endorse one response level (#5, level 1). In terms of data distribution, the item with the worst distribution was #5 (How much of the time were you respected by the clinician). The item with the best data distribution was #2 (How satisfied were you with the explanations about results of your treatment).

*Figure 5.1: SAPS item data distribution*
Regarding ARSB, the raw response data, prior to reversal, are shown in Table 5.3. To assess ARSB the extreme categories (values 1 and 5) were examined for three of the four items administered consecutively (items #2, #3, #4). This analysis showed that there were 3 cases with the pattern 1, 1, 1 and no cases with the pattern 5, 5, 5 across these three items. Thus just 2% of participants exhibited signs of ARSB on these three items. Related, but different to ARSB is item direction effects. The SAPS contains three positively worded items and four negatively worded items. The mean (SD) response for the three positive items was 1.10 (0.15) and for the negatively worded items 1.68 (0.24). The difference between the two item sets were statistically significant (t = 3.64, df = 5; p = 0.02) suggesting that respondents interpreted the negative and positive items in slightly different ways. Finally, after item reversal, IRTC after correction for overlap was examined and the range was from 0.44 (item #5) to 0.61 (#7).

Table 5.3: SAPS item frequency response

<table>
<thead>
<tr>
<th>Item</th>
<th>Item direction</th>
<th>Percentage of cases endorsing each response level (N = 136)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>SAPS1</td>
<td>-</td>
<td>40%</td>
</tr>
<tr>
<td>SAPS2</td>
<td>+</td>
<td>10%</td>
</tr>
<tr>
<td>SAPS3</td>
<td>-</td>
<td>49%</td>
</tr>
<tr>
<td>SAPS4</td>
<td>+</td>
<td>4%</td>
</tr>
<tr>
<td>SAPS5</td>
<td>-</td>
<td>75%</td>
</tr>
<tr>
<td>SAPS6</td>
<td>+</td>
<td>5%</td>
</tr>
<tr>
<td>SAPS7</td>
<td>-</td>
<td>53%</td>
</tr>
</tbody>
</table>

Figure 5.2 shows the data distribution of SPS scores after recoding. Consistent with the item response distributions (Figure 5.1) the data were significantly skewed towards participants being satisfied. The distribution suggests that there were three groups of participants – those who were particularly well satisfied (scoring in the range 26-28), those who were reasonably satisfied (20-25) and those who were mostly dissatisfied (6-19). Ceiling scores (value of 28 points) were observed for 13% of cases; no cases obtained floor scores (0 points). The overall mean score was 21.96 (SD = 4.91).

Figure 5.2: SAPS score distribution
5.5 **Internal Structure and Reliability of the SAPS**

Table 5.4 presents the results of the PCA and IRT analysis. For the PCA, all items loaded ≥0.60, just the one factor with an eigenvalue >1.00 was identified (the ratio of the first two eigenvalues was 4.33), and the proportion of explained variance was 53%. The Rasch analysis showed that no items were reported with disordered thresholds, that all item locations were within |2.00| of the centroid, and that no items exceeded the fit residual threshold of ±2.50. The item-trait interaction was not statistically significant, the person separation index exceeded the criterion, and the internal consistency of the SAPS was $\alpha = 0.85$.

In summary, on both the factor and Rasch analyses the SAPS items were shown to be unidimensional with none of the items displaying statistically significant misfit. The only issue which was identified was that on two of the items (#3 and #4) there was evidence of statistically significant DIF = by participants’ incontinence type (urinary/faecal).

### Table 5.4: Internal structure of the SAPS using factor and Rasch analysis

<table>
<thead>
<tr>
<th>Item</th>
<th>PCA factor loading (a)</th>
<th>Rasch analysis</th>
<th>Gender</th>
<th>Age group</th>
<th>Cohort (d)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Disordered threshold</td>
<td>Item location (b)</td>
<td>Fit residual</td>
<td>$\chi^2$</td>
<td>p</td>
</tr>
<tr>
<td>SAPS1</td>
<td>0.73</td>
<td>N</td>
<td>0.32</td>
<td>0.35</td>
<td>0.83</td>
</tr>
<tr>
<td>SAPS2</td>
<td>0.78</td>
<td>N</td>
<td>0.68</td>
<td>-0.85</td>
<td>4.12</td>
</tr>
<tr>
<td>SAPS3</td>
<td>0.73</td>
<td>N</td>
<td>-0.19</td>
<td>0.75</td>
<td>3.14</td>
</tr>
<tr>
<td>SAPS4</td>
<td>0.77</td>
<td>N</td>
<td>0.42</td>
<td>-0.64</td>
<td>3.16</td>
</tr>
<tr>
<td>SAPS5</td>
<td>0.60</td>
<td>N</td>
<td>-1.34</td>
<td>0.58</td>
<td>0.50</td>
</tr>
<tr>
<td>SAPS6</td>
<td>0.70</td>
<td>N</td>
<td>0.72</td>
<td>0.81</td>
<td>4.71</td>
</tr>
<tr>
<td>SAPS7</td>
<td>0.77</td>
<td>N</td>
<td>-0.61</td>
<td>-0.73</td>
<td>3.49</td>
</tr>
</tbody>
</table>

| Eigenvalue | 3.69 |
| % variance  | 52.67 |
| Item-trait interaction | $\chi^2 = 19.49$, df = 14; p = 0.15 |
| Person Separation Index | 0.84 |
| Internal consistency (Cronbach $\alpha$) | 0.85 |

Notes:
- a = PCA analysis with oblimin rotation.
- b = Logit scale
- c = Differential item functioning: N = non-significant; Y = statistically significant (p<0.05)
- d = Faecal/Urinary

Table 5.5 lists the standard SAPS items, the original draft item versions for three items and an item in which the response scale was reversed. In three cases the kappa agreement between the different versions was >0.60 indicating good agreement. The PCA loadings and IRTC correlations were almost identical, there was very little difference in item location or fit residual other than for the original draft version of item 6 where the fit residual exceeded the criterion suggesting that the original version of this item did not fit the standard SAPS. In short, there was no evidence that changing the three draft SAPS items had any deleterious effect on the structure of the SAPS.

For SAPS #2 (D2) and the alternate positively worded item (D15), however, there was agreement in endorsement by just 65% of respondents; the kappa, 0.52, was below the study criterion threshold and suggested just moderate agreement between these two items (Landis and Koch, 1977). Examination of the two items showed the data distribution was worse for the alternative positive form of the question than the original negative form (the proportions of cases, after recoding, were 38%, 33%, 12%, 7%, 10% for the original negative form and 52%, 27%, 13%, 6%, 2% for the alternative positive form). Neither form exhibited significant ARSB. When the factor loadings and IRTC were examined they suggested that the alternative positive form of the question was marginally superior to the original negative form, but that the opposite was the case.
for the Rasch IRT analysis – indeed regarding misfit on the Rasch analysis when the alternative positive form of the question was used it obtained the highest misfit value (although at -2.09 this was still within the statistical limit of |2.50|). The details are shown in Table 5.5.

Table 5.5: Impact of change in item wording for four SAPS items

<table>
<thead>
<tr>
<th>SAPS item</th>
<th>Item stem</th>
<th>Traditional statistics</th>
<th>Rasch IRT analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Kappa (a) PCA loading (b)</td>
<td>Item location</td>
</tr>
<tr>
<td>1.</td>
<td>How satisfied are you with the effect of your treatment/care?</td>
<td>0.73 0.60</td>
<td>0.32</td>
</tr>
<tr>
<td>Original wording</td>
<td>How happy are you with the effect of your treatment/care?</td>
<td>0.65 0.71</td>
<td>0.47</td>
</tr>
<tr>
<td>2.</td>
<td>How satisfied with explanations of results - negative</td>
<td>0.78 0.67</td>
<td>0.68</td>
</tr>
<tr>
<td>Alternative wording</td>
<td>How satisfied with explanations of results – positive (d)</td>
<td>0.52 0.84</td>
<td>0.04</td>
</tr>
<tr>
<td>6.</td>
<td>The time you had with the doctor/other health professional was too short</td>
<td>0.70 0.58</td>
<td>0.72</td>
</tr>
<tr>
<td>Original wording</td>
<td>The time you had with the doctor/other health professional was not long enough</td>
<td>0.75 0.67</td>
<td>1.50</td>
</tr>
<tr>
<td>7.</td>
<td>Are you satisfied with the care you received in the hospital/clinic?</td>
<td>0.77 0.66</td>
<td>-0.61</td>
</tr>
<tr>
<td>Original wording</td>
<td>Are you happy with the care you received in the hospital/clinic?</td>
<td>0.74 0.76</td>
<td>-0.45</td>
</tr>
</tbody>
</table>

Notes:
All analyses were after recoding into three response levels. See the methods section for an explanation.
a = Agreement between items.
b = Principal component analysis loading.
c = Item rest of test correlation
d = Simple reversal of item responses. In the original version the responses were Very dissatisfied/ Dissatisfied/ Neither/ Satisfied/ Very satisfied. In the alternative version the responses were Very satisfied/ Satisfied/ Neither/ Dissatisfied/ Very dissatisfied.

5.6 Discrimination Evaluation of the SAPS

There were no statistically significant differences in mean SAPS scores by gender, age group, education attainment or incontinence type. In contrast there were statistically significant differences in mean scores by follow-up general health status, type of incontinence treatment, current clinician and patient rated incontinence severity and patient perceived change in incontinence severity. There were no statistically significant differences by clinician rated change in incontinence status or by change scores on the RUIS and RFIS. The details are given in Table 5.6.
Table 5.6: Discrimination evaluation of the SAPS

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>t_trans</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>20</td>
<td>23.25</td>
<td>3.93</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>116</td>
<td>21.74</td>
<td>5.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Age group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;60 years</td>
<td>59</td>
<td>21.75</td>
<td>5.05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥60 years</td>
<td>76</td>
<td>22.07</td>
<td>4.82</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Education attainment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>40</td>
<td>21.63</td>
<td>4.32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>34</td>
<td>22.97</td>
<td>4.22</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trade/TAFE/Diploma</td>
<td>27</td>
<td>21.04</td>
<td>6.38</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>University</td>
<td>33</td>
<td>21.88</td>
<td>4.95</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Incontinence type</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urinary</td>
<td>100</td>
<td>21.69</td>
<td>5.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Faecal</td>
<td>36</td>
<td>22.72</td>
<td>4.51</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>General health status at follow-up</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excellent/Very good</td>
<td>48</td>
<td>24.25</td>
<td>3.64</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td>48</td>
<td>21.27</td>
<td>4.91</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fair/Poor</td>
<td>39</td>
<td>20.31</td>
<td>5.05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Incontinence treatment (a)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Counselling</td>
<td>26</td>
<td>20.42</td>
<td>4.61</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physiotherapy</td>
<td>64</td>
<td>21.77</td>
<td>5.46</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surgery</td>
<td>36</td>
<td>23.83</td>
<td>3.32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Current incontinence severity (clinician rated)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>38</td>
<td>24.18</td>
<td>3.25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mild</td>
<td>48</td>
<td>21.60</td>
<td>4.45</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate</td>
<td>22</td>
<td>22.50</td>
<td>5.53</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Severe</td>
<td>16</td>
<td>19.19</td>
<td>5.89</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Current incontinence severity (patient rated) (b)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>47</td>
<td>23.96</td>
<td>3.61</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mild</td>
<td>41</td>
<td>22.59</td>
<td>4.67</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate/Severe</td>
<td>47</td>
<td>19.38</td>
<td>5.22</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Change in incontinence severity (clinician rated)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improved</td>
<td>70</td>
<td>23.11</td>
<td>3.87</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No change</td>
<td>38</td>
<td>20.76</td>
<td>5.70</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Worse</td>
<td>15</td>
<td>21.80</td>
<td>5.49</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Change in incontinence severity (patient rated)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cured</td>
<td>24</td>
<td>24.92</td>
<td>3.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partly cured</td>
<td>50</td>
<td>23.20</td>
<td>4.16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Some improvement</td>
<td>40</td>
<td>21.23</td>
<td>4.35</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not helped</td>
<td>20</td>
<td>17.15</td>
<td>5.58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Change in RUIS (c)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improved</td>
<td>78</td>
<td>22.04</td>
<td>4.86</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No change/Worse</td>
<td>22</td>
<td>20.45</td>
<td>5.54</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Change in RFIS (c)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improved</td>
<td>23</td>
<td>22.91</td>
<td>3.91</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No change/Worse</td>
<td>13</td>
<td>22.09</td>
<td>5.53</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes:
- a = Excludes those with multiple treatments
- b = There were only 5 cases who rated their current incontinence ‘Severe’.
- c = Excludes faecal cases. Four cases reported no change, so these were included in the No change/Worse category. Improved describes those whose incontinence improved between baseline and follow up.
- d = Excludes urinary cases. Five cases reported no change, so these were included in the No change/Worse category. Improved describes those whose incontinence improved between baseline and follow up.

SAPS scores were correlated with those of the GUTSSV2. The correlation showed that the two measures of patient satisfaction shared 61% of common variance, which suggested the SAPS was assessing a similar construct to the GUTSS but was a different measure than the GUTSS. Figure 5.3 shows a scatterplot of this relationship.
5.7 Discussion

This validation of the standard SAPS showed that it was a reliable, valid and sensitive measure of patient satisfaction. This conclusion is based on three extensive tests: an examination of the data at the item and instrument levels, an assessment of the internal structure of the SAPS using both traditional and Rasch analyses, and tests of its discriminatory function by known groups selected to demonstrate convergent and divergent scores.

When item response data were examined against the study criteria, the items generally met these criteria. Of the seven SAPS items, five met the criterion for data distribution (>5% of endorsements fell on any 2 adjacent categories); the two items which did not were items #5 and #7 – both of which were at 5%. Six items met the criterion for the percentage of cases selecting any one response category; the item which failed this criterion was #5 where 72% of cases endorsed the best possible outcome. It is possible that the difficulties with #5 were a function of its content; viz., that the patient felt respected by the clinician. Certainly, the percentage of endorsements at 72% was higher than in the draft SAPS construction study where for the same item it was 44% (Hawthorne et al., 2006). It is possible this may reflect that the original version of this item in the Patient Satisfaction Index (PSI) (Guyatt et al., 1995) which was used in the earlier Hawthorne et al study had seven possible response categories compared with the five in the SAPS item. It is also possible that responses to this item were influenced by the item order within the questionnaire. The previous two items in the SAPS ask about the care taken during the physical examination (SAPS #3) and the patient’s involvement in medical decision-making (SAPS #4). These two preceding items may ‘prime’ respondents through providing contextual clues which may be used by the respondent to frame the response in an affirmative way for an immediately following item probing clinician respect for the patient (Foddy, 1993). In contrast, in the PSI the preceding item probes family involvement in medical decision-making.

The skew in the SAPS data was both obvious and normal: most people are satisfied with their health care (Cartwright, 1967, Gamagami et al., 1999). The mean SAPS score (22 points) was at the 79% mark of the SAPS total score – a finding almost identical with the Hawthorne et al. earlier work on patient satisfaction among those with incontinence (Hawthorne et al., 2006). In this context it should be noted that the reliability estimate was based on collapsing the 5-point response scales into 3-point response scales as a way of ensuring sufficient cases for stable reliability, factor and Rasch analysis estimates. In the case of the SAPS’s reliability, Cronbach α is a function of the variance of each item and the variance between items (the total composite score), i.e. the underlying correlations assume normality in item distributions. The effect of skewed
responses on these correlations is well known: it degrades the true relationship and the greater the skew the more degraded the correlations (Enders and Bandalos, 1999). Under these circumstances α will provide a poorer measure of reliability than where the item responses are normally distributed. This can be seen in the present study. The Cronbach’s α for the original 5-point response scales was 0.80 compared with 0.85 for the recoded response scales.

Comparison with the original construction sample for the SAPS showed almost identical reliability where the Cronbach’s α = 0.86 (Hawthorne et al., 2006). There are two implications. The obtained α for the SAPS using the 5-point response scales is likely to represent the lower boundary of reliability due to data skew (Bandalos and Enders, 1996, Greer et al., 2006); the 3-point response scale Cronbach’s α is likely to be a more accurate estimate of the SAPS’s true reliability. It also follows that the use of 5-point response scales allows greater discrimination in the skewed tail (Enders and Bandalos, 1999, Nunnally and Bernstein, 1994).

That 13% of cases obtained ceiling scores represents an excellent outcome for the SAPS as this was well under that reported for three other patient satisfaction instruments where the percentage of cases obtaining ceiling scores was over 20% (Hawthorne et al., 2006).

It is possible the low percentage of cases obtaining ceiling scores reflects that the SAPS was designed to be a balanced scale to overcome issues associated with acquiescence (Crowne and Marlowe, 1960, Miller and Cleary, 1993). Under these test circumstances two phenomenon have been observed. Factor analysis has commonly loaded positive and negative items on separate factors (Hawthorne and Hogan, 2002, Hawthorne, 2006a, Reiser et al., 1986). This, however, was not observed for the SAPS in this sample (Table 5.4), nor was it observed in the draft SAPS construction sample (Hawthorne et al., 2006). The implication is that respondents discriminated well by item direction (positive/negative). This is, however, subject to the finding that positive/negative items in the SAPS were scored significantly differently – the mean score for positive items was 1.10 (after reversal) and for negative items 1.68. This is consistent with previous reports that in balanced scales respondents are generally less inclined to reject a positive statement than to accept a negative one, regardless of item content (Reiser et al., 1986).

This issue is important when considering the findings regarding alternative items for the standard SAPS items. As shown in Table 5.4, alternatives were considered for four SAPS items. The finding that there was no real difference between the standard and alternative item forms for three of the items led to the conclusion that retaining the standard SAPS items was to be preferred to making changes (the possible exception to this would be to replace #6 with the original draft SAPS item version on the grounds that the standard current #6 obtained an item location on the IRT logit scale that was similar to the other items (Table 5.4) whereas on the original draft SAPS this item was located very differently (item location = 2.20) (Hawthorne et al., 2006); but this is, on balance, not sufficient evidence for such a change). Indeed, it is axiomatic of psychometrics that adding, removing or changing items based entirely on statistical results is highly undesirable as it can subtly corrupt the construct validity of the scale (Davidson, 2008, Nunnally and Bernstein, 1994), even where such changes marginally improve the psychometric properties (Kenna et al., 2005).

For example, the SAPS was designed to be a general measure of patient satisfaction based on a theoretical model of general patient satisfaction (Hawthorne et al., 2006). In contrast, the GUTSS was designed to assess the impact of surgery for urinary incontinence (Hawthorne and Harmer, 2000, Hawthorne et al., 2006). Hypothetically, if several GUTSS items were incorporated into the SAPS (or replaced SAPS items because they performed better – and since this was a study of incontinence it could be expected that they would perform better) – then the SAPS would become corrupted in the sense that its underlying general patient satisfaction model would become biased towards patient satisfaction in incontinence.

With regard to the alternative item for the SAPS #2, there was a slightly different situation. The only difference between the alternative and standard SAPS #2 item was in the reversal of the response scale (see Table 5.5). Although the evidence regarding the statistical properties was equivocal (suggesting retention of the standard item based on the axiom above), just 65% of participants provided responses that agreed across the two items. In addition, the data distribution for the alternative positive form of the item was worse than the standard form with a higher
proportion of respondents endorsing the best possible response (52% for the alternative version and 38% for the standard version).

It is likely that this discrepancy between the two forms of the same item was a function of the contextual issues raised above. The preceding items for the standard item SAPS #2 probing satisfaction with the clinician’s explanations of treatment effects were items covering the participant’s general health and an item of happiness with treatment effects. In contrast, the two questionnaire items preceding the alternative form of #2 were items probing information from the clinician and the attitude of the clinician towards the patient. This arrangement of items leading to a discrepancy of +15% endorsing the best health option is consistent with priming effects reported elsewhere and it has been argued that the unprimed responses are likely to be closer to participants’ real feelings (Bowling and Windsor, 2008, McFarland, 1981). A second possible explanation is that respondents provide very different opinions for reversed items which are concerned with human immediacy (health, loss of employment, money or loss of life) – a phenomenon that has been documented for over 60 years in the psychometric literature (Foddy, 1993, Payne, 1951, Tversky and Kahneman, 1981). Whichever of these possibilities applies – or a mixture of both – the implication is there is no necessary reason to expect respondents to provide consistent answers to items which differ only by positive/negative response scales and where for one form of the item there are clear possible priming effects. Given the very equivocal and marginal improvement in the SAPS with replacement of the standard #2 with the alternative item, there is probably no overall net benefit by inclusion of an item with response patterns that may reflect contextual issues related to its questionnaire position. Replacement of the standard #2 item with the positive version of the item is therefore unlikely to improve the scale in any way that is important and this was rejected.

That mean SAPS scores did not significantly vary by demographic or incontinence status (faecal/urinary) is important, as is the finding that 19/21 item tests revealed no statistically significant DIF = by age, gender or incontinence status (urinary/faecal). The demographic findings were consistent with the construction findings reported by Hawthorne et al (Hawthorne and Harmer, 2000). DIF = by study cohort (faecal/urinary) was observed for items #3 and #4. Although it is the convention that items with DIF = should be considered for removal from a descriptive system, there are two reasons this may not be appropriate. If these SAPS items were removed, then the SAPS would not longer maintain fidelity to the underlying model of patient satisfaction. These items are concerned with clinician care and patient choices in decision-making. It is possible this DIF = may reflect true differences in the role of the patient vis-a-vis the role of the clinician. In urinary incontinence the patient may have more real choices and a more interactive relationship with the clinician, whereas in faecal incontinence the patient may have fewer real choices regarding treatment options and be much more reliant upon the surgeon’s decision-making.

Subject to the caveat, the findings suggest that it is likely the SAPS can be used with confidence that its patient satisfaction scores are not unduly influenced by background variables, that it has equal applicability to faecal and urinary incontinence, and that scores from one study can be compared with those from another study; i.e. that SAPS scores are generalisable.

In contrast, the results of the discriminatory function of the SAPS by health and incontinence status were mixed (Table 5.6). Statistically significant differences were reported by treatment type (also reported in the construction study (Hawthorne 2006), current health status, clinician and patient estimates of current incontinence severity, and patient reported severity change over time (also reported in the construction study (Hawthorne 2006)). Non-significant findings were reported for changes in clinician estimate and changes in both urinary and faecal symptoms over time.

The interpretation is that patient satisfaction as measured by the SAPS is a function of current health status (generally and specifically) and patient belief about improvement in their condition as well as more objective measures of treatment effect (such as change in incontinence). Of the major patient satisfaction theories reviewed by Hawthorne (Hawthorne, 2006b), if this interpretation of the data is substantiated, then it would appear that the underlying model of patient satisfaction in incontinence is close to Linder-Pelz’s expectancy-value theory in which she postulated that the two conditions under which high satisfaction would be reported were (a) where positive expectations and positive experiences coincided and (b) where experiences were
perceived to be as good as or better than those of others (Linder-Pelz, 1982). Whilst there is no evidence in this study on the latter point, the findings are consistent with the first point.

The findings of this validation of the SAPS are subject to several caveats. Recruitment into the study involved sampling from clinics willing to participate. Although efforts were made to include clinics from across broad geographic areas and populations, only ten clinics participated. The recruitment rate within clinics widely varied; the proportion of participants ranging from 2% to 32%. There was also considerable variation between clinics in the proportion of recruited participants who also completed follow up questionnaires; the range was from 33% to 60%. All clinics recruited convenience samples of patients. It is possible that these issues may have implications for the study findings. The number of follow up study participants (N = 136) was 54% of the total study population. There are important sequelae. The finding that patients who provided follow up data were patients who were more likely to have severe incontinence at baseline as rated by clinicians (severe vs. normal/mild/moderate) may suggest that the study findings are circumscribed in their generalisability to patients with mild or moderate incontinence. However, it is noted that there was no difference in clinician rated severity between patients submitting post forms when this was grouped as normal/mild vs. moderate/severe. That this finding was also not confirmed by patient assessment of incontinence severity would imply both that we cannot be certain of the non-generalisability of the findings – hence our use of the word ‘circumscribed’ – and that there are different perceptions of incontinence severity between clinicians and patients. The sample size also has implications for the robustness of the data analyses. Although there is some evidence that stable factor and Rasch analyses can be carried out with small samples, the available sample size was under the generally accepted sample size for these procedures. The limited sample also restricted the interrogation of the data, particularly for sub-group analyses.

5.8 Conclusion

This study has validated the SAPS in a clinical sample of patients with faecal and urinary incontinence. The validation was based on three extensive tests: an examination of the data at the item and instrument levels, an assessment of the internal structure of the SAPS using both traditional and Rasch analyses, and tests of its discriminatory function by known groups selected to demonstrate convergent and divergent scores.

In general, all three tests confirmed that the SAPS is a valid, reliable and sensitive measure of patient satisfaction in incontinence. They also confirmed that the changes to SAPS items from the construction draft version to the standard version appeared to have no deleterious effects on either the structure or measurement properties of the SAPS.

In conclusion, the SAPS is a short, valid, reliable and sensitive measure of patient satisfaction suitable for use in incontinence studies. Further work is needed to investigate how the SAPS model relates to theories of patient satisfaction.

5.8.1 References


Bjorner, J. B., Kosinski, M. and Ware, J. E., Jr. (2003), 'Calibration of an item pool for assessing the burden of headaches: an application of item response theory to the headache impact test (HIT)', Quality of Life Research, 12, 8, 913-33.


McHorney, C. A., Kosinski, M. and Ware, J. E. (1994), *Comparisons of the costs and quality of norms for the SF-36 health survey collected by mail versus telephone interview: results from a national survey*, Medical Care, 32, 6, 551-567.


6 Clinical Translation of the Revised Instruments

6.1 Background to the Clinical Translation Component

To further maximise the value of recent work, a clear and ongoing dissemination strategy was included which covered dissemination to health service providers, local and international academics and researchers in the field, and policy makers. Achievement of this includes maximising clinicians’ involvement and potential practice change, across a number of media platforms. The literature on health care knowledge transfer suggests that successful strategies involve three inter-related activities: (a) informing and gaining the support of respected leaders in the field (e.g. key researchers and leading clinicians); (b) exposing clinicians and other health care workers to a critical mass of peer-reviewed literature and to face-to-face presentations or training; and (c) providing simple and straightforward messages, reminders and cues. Collectively, these activities increase health literacy on a topic and enhance the uptake of new practices and technologies in health care.

A number of dissemination activities were undertaken in 2007 as part of the Continence Outcomes Measures Dissemination project. These included:

- Presentations at the International Continence Society Conference (Rotterdam 2007; 2 presentations and 1 poster)
- Presentations at the CFA conference 2007 (Gold Coast 2007, presentation and workshop)
- Presentations at the AAG (Adelaide, 1 presentation and 2 posters)
- Presentation at the ISOQOL Clinical Conference (Budapest)

Although reports and conference presentations, such as those above, are extremely important for informing policy, the knowledge currency of clinicians and academics is dependent on articles published in peer-reviewed journals which provide instant access online around the world. To facilitate the translation of the current continence research into clinical practice the following activities outlined below were undertaken during the course of the project.

6.2 Peer Reviewed Articles and Presentations

To ensure access in this competitive marketplace, in addition to the papers above, it is highly desirable that a critical mass of papers are published reflecting the depth of the continence body of work. To achieve this critical mass, three papers, based on the latest project research, were prepared for submission to high impact journals and conferences covering:

- Urinary incontinence measurement
- Faecal incontinence assessment
- The measurement of patient satisfaction

Two peer reviewed publications have already been published (refer Section 6.1.2.1 below). Two abstracts (peer reviewed short papers) were submitted to the 2011 International Continence Society Conference. These were:

- A Prospective “Bottom up” Study of the Direct Personal and Investigation Costs of Faecal Incontinence in Ambulatory Men and Women, in Relation to Severity.
- The Clinical Validation of a New Short Faecal Incontinence Measure for Epidemiological and Outcomes Research

The paper on the Faecal Cost of Illness study was an oral poster presentation which means it will be published in the Journal of Neurourology and Urodynamics. The paper on the RFIS has been accepted as a non oral presentation poster but will be published in the ICS Conference Proceedings.

A draft short paper was prepared on the RUIS for presentation at conferences and for journal submission (draft) and a short paper/abstract was prepared (and accepted) for the next CFA conference.
An article on the SAPS was prepared for submission to a relevant journal and it is anticipated further journal articles will be submitted to journals but these are seen as additional to the deliverables for this project.

### 6.3 Activities Undertaken

From 2008 and 2011, in association with this project, the following dissemination activities were undertaken.

**Peer reviewed presentations and papers:**


**Other presentations and/or papers**


Presentations from the 2008 Health Outcomes Conference are available at: [http://chsd.uow.edu.au/ahoc/conferences/2008/program08.html](http://chsd.uow.edu.au/ahoc/conferences/2008/program08.html)


An additional (peer reviewed) paper on the Short Assessment of Patient Satisfaction was prepared for submission to a relevant journal.

### 6.4 Development of a Technical Manual for the Revised Tools

Development of a technical manual and professional training materials to assist clinicians and health care workers to use, score and understand the new incontinence instruments will serve to enhance the portability of the revised tools to clinics throughout Australia. A brief technical manual is provided at Attachment A.
6.5 **Brochures**

Clinicians and other health care workers respond to reinforcement of academic learning through the provision of simple, informative reminders and/or instant access to web-based materials. The model behind these will be sourced from the “train the trainer” model, and the aim was to develop simple tools and handouts which could be easily adapted and incorporated into everyday clinical practice. To further aid knowledge transfer we developed PDF documents and summary handouts/pamphlets to be attached to the Bowel and Bladder and to the Continence Foundation of Australia’s web-sites. These materials will be freely accessible and were prepared so they would encourage health workers to contribute to and participate in web-based discussion forums. The brochures can be found at Attachment B.

6.6 **Presentations at Continence Foundation of Australia (CFA) Conference**

To keep continence practitioners and researchers up to date with the evolving findings from this research Jan Sansoni and Nick Marosszeky also undertook two presentations at the CFA conference in November 2009 in Adelaide. These presentations were forwarded to the CFA for placement on the CFA website. These presentations were:


A further presentation about the final outcomes of this project will be made at the CFA conference in Melbourne in November 2011. A peer reviewed short paper on the project was submitted to the CFA for the 2011 conference and was accepted. This will be published in a forthcoming edition of the Australia and New Zealand Continence Journal.

6.7 **Materials Placed on Relevant Web Sites**

It is noted that upon completion, and with approval from the Department of Health and Ageing, all materials will be placed on relevant and recommended web sites such as [www.bladderbowel.gov.au](http://www.bladderbowel.gov.au) and at [www.continence.org.au](http://www.continence.org.au) to encourage access to these materials by health professionals and to enhance the sustainability of the research work undertaken. The placement of information on relevant web sites will, in most cases, take place following completion of the project. Attempts will be made to monitor access to information provided by the project, although it is noted that such data is more likely to be available post project completion.
7 Conclusions

This study commenced in mid June 2008. A Project and Evaluation Plan report was submitted on schedule to the Department of Health and Ageing in August 2008. Since project commencement eleven continence clinics across 4 States and Territories of Australia have recruited patients to this study. Additional clinics were recruited in mid 2008 and during 2009 in an endeavour to increase patient recruitment and details of the participating clinics can be found in Section 3.1.5.

Ethics approval for all sites was obtained and staff at all participating clinics were trained in the study implementation procedures. Patient recruitment commenced in November 2008 at two clinics but most clinics commenced patient recruitment in March 2009.

This study recruited patients from a range of practice settings across Australia: particularly specialist and community continence clinics where patients seek and receive incontinence care. The study protocol contains the revised continence instruments (RUIS, RFIS), patient satisfaction measures (SAPS) and health status and health related quality of life instruments (e.g. SF-36V2, AQoL) and some items from continence specific health related quality of life and/or impact questionnaires (e.g. Incontinence Impact Questionnaire). The study also examined the relationship between these instruments and individual medical conditions, co-morbidity, gender and age.

The study examined clinical and patient definitions of treatment outcomes and success, across four different treatment types:

- Continence Advising (this includes Nurse Continence Advisors and other staff undertaking advising roles);
- Physiotherapy;
- Surgery (note: surgery is usually only given to those whom physiotherapy has failed); and
- Mixed or other treatments.

The relationship between clinical indicators (e.g. clinician pre and post treatment forms), the revised instruments and patient satisfaction was examined. Additional reliability data was collected from post-test patients in order to examine the test-retest reliability of the instruments over a two week period. Although it had already been established that the internal consistency reliability of the revised tools was excellent in population health settings (Sansoni et al., 2006), to facilitate clinical uptake of the tools it was necessary to show that it is appropriate in clinical settings and that test-retest reliability for these instruments was also acceptable.

The details of all the following statistical analyses can be found in Section 4 of the report.

**RUIS**

At baseline for all 195 cases the mean RUIS score was 10.92 (SD = 3.33; N = 195). The mean RUIS score for the sample of females at pre-treatment was 10.90 (SD = 3.16, N = 167) and the mean RUIS score for the sample of males at pre-treatment was 11.07 (SD = 4.18, N = 28) (NB: Similar median results were found for both genders). There was no significant difference between these total scores when analysed by gender (t = 0.25, p>0.05). There were no significant gender differences were found for any of the individual items (p>0.05). These data suggested the scale may be appropriate for use with both genders. For most RUIS items at baseline 62-68% of the clinical sample experienced these symptoms moderately or greatly.

When examined by clinical ratings of pre-treatment severity there was a significant difference (F = 16.99, df 2, 191; p = 0.000) in RUIS means between those with mild incontinence (M = 9.22) and those with moderate (M = 11.79) and severe incontinence (M = 12.13). Similarly, when examined by the pre-treatment patient rating of the severity of urinary incontinence there was a statistically significant difference (F Welch = 80.46. df 2, 109.07; p = 0.000). The mean for those in a mild state was 8.36, for those with moderate incontinence it was 11.60 and for those with severe incontinence it was 14.03 and these group comparisons were all significant (p<0.05).

The mean RUIS scores were significantly higher for those that were receiving surgical vs. conservative treatments (t = -2.76, df 70.49; p = 0.007) as generally surgery is used with patients with more severe incontinence or for whom conservative treatment has failed. For this analysis
those that were receiving mixed treatments and other treatments such as medicines were excluded (N = 13). The pre-treatment RUIS mean for the surgical group was 12.03 and for the conservative treatment group (physiotherapy and continence advising) it was 10.71.

At pre-treatment there were significant differences in RUIS mean scores between those that had both faecal and urinary incontinence (double incontinence) as against urinary incontinence only (t = -2.71, 193; p = 0.007). The mean for those with urinary incontinence only was 10.60 and for those with double incontinence it was 12.18. Patients diagnosed with mixed types of incontinence (e.g. stress and urge) had higher RUIS scores than those with only urge incontinence (F = 9.83, df 2, 176; p = 0.000). The mean for those with urge incontinence only was 9.00, for those with stress incontinence only it was 10.94 and for those with mixed incontinence it was 11.79.

Patient scores on the number of pads used were grouped as less than 1 per day and 1 per day or more. The mean RUIS score for the low pad use group was 9.02 and for the high pad use group it was 11.55. This difference was statistically significant (t = -6.22, df 192; p = 0.000).

The RUIS does not include any items concerning pad use but pad use should reflect the severity of urinary leakage. Patient scores on the Wei items concerning the number and the size of pads were compared to RUIS scores. For the pad size item (F = 38.56; df 2, 191; p = 0.000) the mean RUIS score for those with no pad/thin pad was 9.21, for those with medium pads it was 12.00 and for those using large pads it was 13.57. All these comparisons were statistically significant (p<0.05).

The International Consultation Incontinence Questionnaire (ICIQ-SF) contains a number of items about whether particular symptoms associated with urinary leakage are present or absent (e.g. leaks before you can get to toilet, when you cough or sneeze, when asleep, when physically active or exercising, when finished urinating and are dressed, leaks for no obvious reason and leaking all the time). The number of these symptoms present (= 1) at baseline, plus an additional item concerning leaking without knowing about it, were added to form an index of leakage severity. The groups were classified as 2 or less symptoms; 3 and 4 symptoms, and 5 or more symptoms. There was a significant difference in RUIS pre-treatment scores for these groups. For those that had 2 or less symptoms the mean RUIS score was 8.60 for those with three and four leakage symptoms it was 11.46 and for those with 4 or more symptoms it was 13.21 (F Welch = 43.84 (df 2, 125.42; p = 0.000)) and all group comparisons were significant (p<0.05).

There was a significant difference in RUIS mean scores by type of treatment (t = -2.76, df 70.49; p = 0.007). Patients that were receiving conservative treatment (physiotherapy and continence advising) had a lower mean score (M = 10.71) compared with those that were receiving surgical treatment (M = 12.03).

These analyses reflect that the RUIS demonstrates discriminant validity between different levels of incontinence severity as measured by other clinical and symptom indicators.

By contrast when the RUIS was examined by baseline health status there were no significant differences between those in Excellent/Very Good, Good, or Fair/Poor Health (p>0.05); similarly, there were no statistically significant associations by gender, age group, educational level, work status, the number of co-morbidities or the length of time patients had experienced incontinence symptoms. There was a statistically significant association between BMI and RUIS (rs = 0.20; p<0.02), but the proportion of variance explained was very small (~4%). This would suggest the RUIS is assessing the underlying condition of urinary incontinence and this assessment appears to be independent of some common confounders.

At follow-up post-treatment the mean RUIS score was 6.95, (SD = 4.76, N = 100). For females the mean was 6.92 and for males the mean was 7.21. This shows a significant improvement for the patient group.

When RUIS post-treatment scores were examined by the clinical post-treatment rating of incontinence severity there was a significant difference (F = 28.68, df 2, 94; p = 0.000) in mean scores by level of severity with those that the clinicians considered now to be ‘normal’ having a mean RUIS score of 2.72 as against a mean score of 11.29 for those still rated as ‘severe’. Similarly, when examined by patient rated post-treatment severity the mean for those in a normal
state was 2.68; for those with mild incontinence it was 6.48, and for those with moderate and severe incontinence it was 12.34. These differences were statistically significant ($F = 122.85; df 2, 97; p = 0.000$) and all group comparisons were significant.

The patient’s ratings concerning the outcome of their treatment was that 17% indicated their incontinence had not been helped by treatment, 23% reported some improvement, 39% considered themselves to be partly cured and 20% indicated that their continence was cured. There was a significant difference in RUIS post treatment scores between these groups ($F = 41.41; df 3, 94; p = 0.000$). For those that were ‘not helped’ the mean was 12.29, for those that had ‘some improvement’ the mean was 9.35, the mean for those that were ‘partly cured’ was 5.63 and for those who considered themselves ‘cured’ it was 1.70. All group comparisons were significant.

In the case of the follow-up item concerning whether the patient still had problems with incontinence there was a significant difference ($F = 124.65; df 3, 96; p = 0.000$), in RUIS post-treatment mean scores between those that considered that they had major problems ($M = 12.85$), some problems ($M = 8.83$) slight problems ($M = 5.12$) and those that had no problems at all ($M = 1.27$) and all group comparisons were significant.

**RFIS**

At baseline for all 61 cases the mean RFIS score was 9.66 (SD = 4.66) and scores ranged from 0 -20. For most RFIS items at baseline 50-68% of the sample experienced these symptoms ‘sometimes’ through to ‘always’. The mean RFIS score for the samples at pre-treatment was 9.10 for males and 9.76 for females. There was no significant difference between these total scores when analysed by gender ($t = -0.41; df 59; p>0.05$) and there were no significant gender differences found for any individual items or in the pattern of responses by gender for any items. This may suggest the scale is suitable for use with both genders although this will need to be confirmed with a larger sample of males.

When examined by the clinician rating of the severity of faecal incontinence at baseline there was a significant difference between RFIS scores for the more and less severe faecal incontinence groups ($F = 3.17; df 2, 56; p = 0.05$). The mean for those with mild incontinence was 8.05, for those with moderate incontinence it was 9.76 and for those with severe incontinence it was 12.08. Similarly, when examined by the baseline patient rating of the severity of faecal incontinence the mean for those in a mild state was 7.12, for those with moderate incontinence it was 10.64 and for those with severe incontinence it was 15.14 ($F = 13.11; df 2, 58; p = 0.000$).

Faecal incontinence patients who had experienced faecal incontinence symptoms for 2 years or more had a significantly higher ($t = -2.23; df 55; p = 0.03$) mean RFIS score ($M = 10.57$) than those patients who had experienced incontinence for less than 2 years ($M = 7.75$). Patients diagnosed as having mixed faecal incontinence (more than 1 type) had a higher RFIS mean score than patients that were diagnosed as experiencing only one type of faecal incontinence ($t = -3.26; df 59; p = 0.03$).

There was a trend for those having surgical treatment to have higher RFIS scores than those receiving conservative treatment ($t = 1.87; df 56; p = 0.067$) but this did not attain significance at the $p<0.05$ level.

The RFIS does not include any items concerning pad use but pad use should reflect the severity of faecal leakage. Patient scores on the Wexner and St Mark’s items concerning pad use were compared to RFIS scores. Patient scores on the number of pads used (Wexner item) were grouped as ‘never or rarely’; or ‘sometimes, often and always’. The mean RFIS score for the low pad user group was 6.48 and for the higher pad user group it was 11.89. This difference was statistically significant ($t = -5.10; df 56; p = 0.000$). A similar significant finding was found when using the St Mark’s pad item.

RFIS scores were compared with the soiling item from the 7 day bowel diaries. The RFIS mean score was significantly higher for patients who reported soiling 10 or more times per week ($t = -2.31; df 38; p = 0.03$) compared to those that reported soiling less than this.
The study protocol contains a number of other (non RFIS) items about whether particular faecal leakage symptoms were present (e.g. leak mucus, incomplete bowel emptying, faecal urgency, faecal incontinence at night, seepage following a bowel movement, leaking without knowing it and pad use). These symptom scores were added to form an alternate faecal incontinence index of severity and then formed into 3 groups (low/moderate/high scores). There was a significant difference in RFIS pre-treatment scores for these groups ($F = 14.19$, df $2$, $53$; $p = 0.000$) For the mild group the RFIS mean was 5.50; for the moderate group it was 9.96 and for the more severe group the RFIS mean was 13.31.

These analyses reflect that the RFIS discriminates well between different levels of incontinence severity as measured by other clinical and symptom indicators whereas it does not discriminate by unrelated health and demographic variables. When RFIS was examined at baseline by a general item on health status there were no significant differences between those in Excellent/Very Good, Good, or Fair to Poor Health Status. Similarly there were no statistically significant differences in RUIS scores by gender, age group, co-morbidity number, type of treatment or BMI.

At follow up the mean RFIS score was 6.64 (SD = 4.76, N = 39) showing a significant improvement for the patient group. When examined by patient post treatment severity ratings there was a significant difference ($t = 4.33$, df $37$; $p = 0.000$) in RFIS mean scores between the less severe ($M = 4.60$) and the more severe incontinence groups ($M = 10.29$).

The patient’s ratings concerning the success of their treatment was that 7.7% indicated their incontinence had not been helped by treatment, 35% reported some improvement, 41% considered themselves to be partly cured and 15% indicated that their continence was cured. There was a significant difference in RFIS post treatment scores between these groups ($t = 4.01$, df $37$; $p = 0.000$). For those that were not helped or had little improvement the RFIS mean score was 9.59 and for those that were partly cured/cured the mean was 4.36. Table 4.8 shows that the RFIS and the Wexner were the most sensitive instruments for this comparison. There were also significant differences in RFIS post-treatment mean scores ($t = 3.80$, df $37$; $p = 0.001$) between those that considered they still had some/major incontinence problems and those that had slight/no incontinence problems.

Both the RUIS and the RFIS have demonstrated they are sensitive to detecting change in the patient’s incontinence status following treatment. Change scores on the RUIS (pre-post) ranged from an improvement of 15 points to a deterioration of 8 points in the context of a 16 point scale and change scores on the RFIS (pre-post) ranged from an improvement of 18 points to a deterioration of 8 points within a 20 point scale. The average magnitude of change score for the RUIS was 4.83 RUIS scores and for the RFIS it was 4.32 scores. This would suggest both instruments have the capacity to detect both an improvement and deterioration in patient incontinence status.

There was a significant improvement of 4.07 RUIS scores and an improvement of 3.11 mean RFIS scores following treatment. Both instruments are also capable of detecting differences associated with type of treatment. In Section 4 it is shown that both the RUIS and the RFIS were equally or more sensitive to detecting change than most other widely used incontinence instruments.

RFIS and RUIS change groups (improved by 2 RUIS/RFIS scores vs. less than 2 RUIS/RFIS scores) were analysed by the Patient Global Rating of Improvement and these analyses suggested that changes in patient ratings, at the group level, were associated with a change score of greater than 2 points. For the RFIS an improvement of 4 points was associated with a change in pad use (reduced use / stopped using).

Data analyses indicate the RUIS and the RFIS have adequate internal consistency reliability in clinical settings as well as community settings (refer Section 4). The internal consistency alpha of the RUIS at pre-treatment 0.73 which is considered adequate as contrasted with the UDI-6 alpha = 0.64, the ICIQ-SF alpha = 0.65 and the ISI alpha = 0.54 which are considered unacceptable (Streiner and Norman, 2003). The internal consistency reliability of the RUIS for a combined sample of urinary and faecal incontinence patients was 0.84 (N = 255).
At pre-treatment the RFIS alpha was 0.78 which is considered good as compared with an alpha of 0.65 for the Wexner and an alpha of 0.65 for the St Mark’s Incontinence Score which are considered unacceptable. The internal consistency of the RFIS for a combined sample of urinary and faecal incontinence patients was 0.91. This clearly shows the RUIS and the RFIS have better psychometric properties than other widely used incontinence instruments.

RUIS test-retest reliability was assessed at 2 weeks post treatment using the intra class correlation coefficient. For the urinary sample the ICC = 0.77 (N = 60) and for the retest sample of all incontinence patients ICC = 0.80 (N = 78). RFIS test-retest reliability was similar with an ICC = 0.79 for the faecal sample (N = 19) and an ICC = 0.80 for the total retest sample. These test-retest reliabilities were higher than for the other measures of urinary and faecal incontinence.

The principal component analyses undertaken indicated the internal structure of the instruments was appropriate. The 5 items of the RUIS were analysed (N =195) and there was only 1 factor extracted with an eigenvalue of 2.43. This could be described as a general urinary incontinence/leakage factor explaining 49% of the variance. All RUIS items loaded above 0.64 on this factor which indicated a moderate to high level of component saturation (0.69).

The 5 items of the RFIS (N = 61) were analysed and there was only one factor extracted (eigenvalue = 2.71) and this could be described as a general faecal incontinence factor explaining 54% of the variance. All RFIS items had loadings of 0.65 or above on this factor indicating a moderate to high level of component saturation (0.73). These analyses of internal structure were confirmed by re-analyses of the community survey data.

In terms of examining the severity of incontinence conditions the instruments performed as expected, with the clinical sample having higher scores (reflecting greater incontinence) than the community population sample. The RUIS and the RFIS also demonstrated good discriminatory power (validity) by differentiating between patients with different levels of continence severity (e.g. no incontinence, mild, moderate, and severe incontinence).

The continence measures were shown to correlate with other measures of incontinence in the expected directions which is also an indication of their convergent validity. The pre-treatment correlations between RUIS and UDI-6 r = 0.76 (p<0.01); RUIS and ISI r = 0.76 (p<0.01); RUIS and ICIQ-SF r = 0.74; RUIS and WEI (urinary) Incontinence Symptom Index r = 0.72 (p<0.01); RUIS and WEI Bother Index r = 0.66; RUIS and Incontinence Impact Questionnaire r = 0.53 (p<0.01). The RUIS also had a substantial correlation with patient rated severity at pre-treatment rs = 0.62 (p<0.01). These correlations are high, significant and in the expected directions providing good construct validity for the RUIS. As may be expected the RUIS correlations with other urinary incontinence symptom measures were higher than with the Incontinence Impact Questionnaire and the WEI Bother Index which are more concerned with the impact of urinary incontinence on daily activities.

At pre-treatment RFIS correlated with the Wexner Incontinence Scale r = 0.88 (p<0.01); with St Mark’s score r = 0.85 (p<0.01); and with the pre-treatment Patient Incontinence Severity Rating rs = 0.52 (p<0.01). Correlations between RFIS with Wexner Type Specification (impact) items were all significant at the p<0.01 level save for sexual relations (p>0.05) and church attendance (p>0.05). The correlation of RFIS with the Faecal Incontinence Quality of Life (FIQL) coping items was r = 0.57 (p<0.01). These findings provide evidence of construct validity for this measure.

Sansoni et al. (2006) reported that urinary and faecal incontinence measures had negative but significant correlations with measures of health related quality of life reflecting the burden of disease for this condition (e.g. the higher the incontinence score the lower the physical functioning or health status score). At pre-treatment the RUIS correlation with the Physical Function Scale of the SF-36V2 was r = -0.15 (p = 0.039) confirming this finding. The RFIS correlation with the Physical Function Scale of the SF-36V2 was r = -0.27 (p = 0.037) at pre-treatment which is also consistent with findings in the literature showing a negative association between faecal incontinence and generic measures of HRQOL. It is interesting that this is a higher correlation than was found for urinary incontinence suggesting that the impact of faecal incontinence may be greater although it should be noted that the faecal incontinence sample was somewhat older and with a greater level of co-morbidities than the urinary incontinence sample.
The SF-36 profiles for urinary and faecal incontinence patients can be found in Section 4.5. The scores for the SF-36 are transformed scores to provide an average of 50 and a Standard Deviation of 10 based on the norms for the population. For urinary incontinence the Physical Component Summary Scale mean was 45.59 and for the Mental Component Summary Scale it was 43.43. For faecal incontinence most profile means are in the vicinity of 42-43 and thus are about 7 points or 0.7 SD = units below the Australian population norms (50). The Physical Summary Component Score was 43.59 and the Mental Health Component Score was 43.41 which are significantly lower than the Australian norm (Hawthorne et al., 2007). This also confirms that patients with incontinence conditions have poorer physical and mental health status scores compared to the Australian population.

When the RUIS and the RFIS were examined by baseline health status there were no significant differences between those in Excellent/Very Good, Good, or Fair/Poor Health. An interesting finding was that while RUIS change scores did not significantly vary by gender, education or age they did vary by health status with those reporting at baseline that they experienced excellent/very good health obtaining the greatest benefit from treatment. This association was not found for RFIS change scores.

As a number of other urinary and faecal items were contained within the patient protocols for the purposes of cross validation analyses were undertaken to see if the RUIS and the RFIS could be improved by the addition or substitution of other incontinence items. Section 4 provides a thorough discussion of this issue. It was found that no other item consistently improved the psychometric properties of either instrument.

In conclusion the RFIS and the RUIS possess evaluative discrimination by clinician and patient assessed incontinence severity but do not discriminate by unrelated health or socio-demographic variables. Similarly, the RFIS and the RUIS appear to be responsive over time to changes in incontinence status. These two findings suggest they have both content and construct validity; i.e. they assesses the underlying condition of incontinence and this assessment appears to be independent of possible confounding variables. The data analyses confirm that the revised incontinence have demonstrated good psychometric properties in clinical settings and this confirms findings from earlier community sample data (Sansoni et al., 2006).

The findings concerning the Short Assessment of Patient Satisfaction (SAPS) also indicated that it was sensitive to changes in patient status as a result of treatment with those patients reporting the greatest levels of improvement also reporting higher levels of satisfaction with their treatment. It also has good internal consistency (alpha = 0.85) reliability. At post-treatment the RUIS correlation with untransformed patient satisfaction scores (SAPS) was r = -0.44 (p<0.00) which indicates there was an association between higher RUIS scores (indicating greater incontinence) had lower patient satisfaction scores (less satisfaction). The RFIS correlation with patient satisfaction scores was r = -0.28 (p = 0.09).

RUIS and RFIS scores at post-treatment were also analysed by two other satisfaction with outcome items which were concerned with whether patients considered their treatment successful and whether they still had problems with incontinence.

There were significant differences (t = 4.01, df 37; p = 0.000), in RFIS post-treatment mean scores between those that considered themselves cured/partly cured, and those who had little improvement or their incontinence was not helped. There were also significant differences (t = 3.80, df 37; p = 0.001), in RFIS post-treatment mean scores between those that considered that they still had major problems/ some problems or slight problems/ no problems at all.

Similarly there was a significant difference (F Welch = 41.44, df 3, 94; p = 0.000), in RUIS post-treatment mean scores between those that considered themselves cured (M = 1.70), partly cured (M = 5.63), had some improvement (M = 9.35) and those whose incontinence was not helped (M = 12.29). In the case of an item concerning whether the patient still had problems with incontinence there was a significant difference (F = 124.65, df 3, 96; p = 0.000), in RUIS post-treatment mean scores between those that considered that they had major problems (M = 12.85), some problems (M = 8.83) slight problems (M = 5.12) and those that had no problems at all (M = 1.27). All group comparisons were significant.
These findings provided further evidence that the RFIS and the RUIS can discriminate well between levels in related variables, in an ordered way and in the expected directions. The indications are that the RUIS performed well in clinical settings and demonstrated: adequate to good internal consistency reliability; correlations with other measures were in the expected directions indicating construct validity, and it discriminated well between groups varying in incontinence severity. There is good evidence that it was sensitive to changes in continence status. Overall, the RUIS had superior psychometric properties to other commonly used scales such as the UDI-6, the ICIQ-SF and the ISI. It is a short, valid and reliable instrument suitable for use in routine practice in clinical settings and should also be considered by researchers and epidemiologists in population health settings.

The RFIS also performed well in clinical settings demonstrating good internal consistency reliability; correlations with other measures were in the expected directions providing evidence of validity; and there is evidence that it was sensitive to changes in continence status as a result of treatment making it suitable for outcome evaluation. The RFIS has also been shown to have superior psychometric properties to other commonly used faecal incontinence scales such as the Wexner and the St Mark’s Incontinence Score.

A burden of disease analysis was undertaken using population prevalence estimates from the 2004 SAHOS dataset which contained a number of incontinence items and scales (e.g. UDI-6, ISI, Wexner). Based on disutility due to incontinence this study has shown that the excess burden of disease associated with incontinence in Australia is considerable, with the estimate, from a societal perspective, at $25 billion per annum. These findings, showing that society carries considerable incontinence excess burden of disease, contrasts with other estimates. Earlier estimates (Mathers et al., 1999; Goss, 2008) did not include those with slight or moderate urinary incontinence in their analyses yet the current study found that these cases were responsible for three-quarters of all excess costs. These earlier studies also did not consider faecal incontinence.

The Faecal Cost of Illness sub-study is discussed in Section 4.7. Very few studies have conducted face to face direct enquiry of the cost of leakage of faeces and few studies have related this to the severity of faecal incontinence by using a standardised measure of severity. A sub-study of 54 faecal patients was interviewed using the Dowell Bryant Cost of Incontinence Index. Patient ‘out of pocket’ costs amounted to a median of $70.89 per annum and the bulk of Medicare costs (which included medical consultation and rebates for physiology testing) had a median cost of $576.92 per annum. It was found that the Total Costs (which includes personal items and investigations) did not relate directly to severity largely because the costs of investigation are largely fixed in Australia. However, the total Personal Costs of hygiene items (pads, creams) increased with the severity of incontinence.

Research Recommendations

A number of further research activities could be considered by the Department of Health and Ageing.

The Australian Longitudinal Study of Women’s Health has included the RUIS in the latest survey for the cohort of older women (N = 5000 across two time points). An analysis of these data would be extremely useful in helping to confirm the reliability and validity of the RUIS in a large population sample of elderly women as well as providing detailed information on prevalence, severity and the correlates of urinary incontinence (function, health status, social isolation) in this population group.

It should be noted that the focus of the current project was to develop validated continence assessment tools for Australian adults and that it was not proposed to customise these instruments for use by other target groups such as people from Culturally and Linguistically Diverse Backgrounds, Aboriginal and Torres Strait Islander Groups, Children, Proxies or Carers - within the confines of this project. It is recognised that the tools may need to be modified for such groups and this could form a potential follow up project.

Although the RFIS was found to perform well in clinical settings and to have superior measurement properties when compared with other faecal incontinence instruments these findings
are based on a sample of 61 patients (48 females and 13 males at pre-treatment and 31 females and 8 males at post-treatment). It would be desirable if a follow up study could continue to collect more data on faecal incontinence (particularly in males) using the same study protocols to analyse gender and type of treatment aspects for faecal incontinence in more detail. Given the lower prevalence of faecal incontinence, particularly for males, the cooperation of a large number of clinics would be required. However, there are few studies in the research literature addressing male faecal incontinence in particular and this would help to further address this gap. It would also be desirable if further data collection for males with urinary incontinence could be undertaken.

General Recommendations

With only 5 items each the RFIS and the RUIS are short and simple to use and score and continence clinics treating incontinence patients should be encouraged to use them both as assessment measures and as an outcome evaluation measures in routine practice. The use of such measures can provide effective feedback to clinicians concerning the effectiveness of their treatments, can facilitate the systematic review and monitoring of patients, and assist in identifying ways to improve practice. Similarly the SAPS is a short and effective measure of patient satisfaction with treatment which can readily identify patient concerns for the clinic or practice.

It would be desirable if these data were to be collected routinely both prior to and following treatment. These are simple tools with which clinics can report on the effectiveness of their treatments. In the longer term the Continence Outcomes Section could consider the development of a continence outcomes data collaborative as has occurred with rehabilitation and palliative care (Australian Rehabilitation Outcomes Centre, the Palliative Care Outcomes Collaboration) although it is suggested that an online, real time framework is used rather than static warehousing given recent developments in this field. For example the Diabetes Educators Association of Australia and a number of State Asthma groups use an online community to manage patients, share data, share practices, and recruit patients to major conjoint research initiatives. Through such online collaborations clinics can compare their treatment outcomes and patterns of practice with other related clinics. Such organizations are useful change agents in promoting best practice within the field.

Papers about this study have been presented at the Australian Health Outcomes Conference in 2008 (peer reviewed) and the Continence Foundation of Australia Conference in 2009 and in a variety of international forums such as the International Continence Society and the International Society of Quality of Life Research. On project completion materials will be placed on relevant websites such as www.bladderbowel.gov.au and at www.continence.org.au. A presentation on the final outcomes of this study will also be presented at the CFA conference in Melbourne 16-19th November 2011 and poster and presentations are scheduled for the ICS Conference in Scotland in late August. These activities, which are beyond the scope of this project, will assist in providing both local and international recognition of these tools which would also greatly facilitate their adoption in local practice settings.
8 References
Bartlett, M.S. (1954), A note on the multiplying factors for various chi square approximations. Journal of the Royal Statistical Society, 16 (Series B), 296-8
Bjorner, J. B., Kosinski, M. and Ware, J. E., Jr. (2003), Calibration of an item pool for assessing the burden of headaches: an application of item response theory to the headache impact test (HIT), Quality of Life Research, 12, 8, 913-33.


Technical Manual: Revised Incontinence and Patient Satisfaction Tools

Associate Professor Janet Sansoni¹, Associate Professor Graeme Hawthorne², Mr Glenn Fleming¹, Dr Elizabeth Owen¹ and Mr Nick Marosszeky³

1. Centre for Health Service Development, University of Wollongong
2. Mental Health Evaluation Unit, Department of Psychiatry, The University of Melbourne
3. Education Assessment Australia, University of New South Wales Global

* These instruments are copyright to the University of Wollongong with a license to the Commonwealth of Australia and the University of Melbourne. These instruments are available free of charge but permission for use should be sought from the authors.
Introduction and Background: Revised Incontinence Tools

Background

A Continence Outcome Measurement Suite Project (COMS: Thomas et al., 2006) was commissioned by the Australian Government Department of Health and Ageing, National Continence Management Strategy Research Program with the goal of recommending a suite of continence outcome measures to be used by clinicians and researchers in Australia. This project was finalised in early 2006 (Thomas et al., 2006).

Recommendations from this report led to a related project Measuring Incontinence in Australia (Hawthorne, 2006). Measuring Incontinence in Australia (2006) assessed a number of the recommended measures (Urogenital Distress Inventory 6, Incontinence Severity Index, and the Wexner Faecal Continence Grading Scale) by including them in the autumn 2004 South Australian Health Omnibus Survey (SAHOS: Harrison Health Research, 2004), which was a community population survey. The study provided Australian prevalence estimates for both faecal and urinary incontinence based on this community survey.

For urinary incontinence, the results suggested that the preferred urinary incontinence measure was the Incontinence Severity Index (ISI; Sandvik et al., 2000). It was found to possess superior measurement properties in comparison with the Urogenital Distress Inventory (UDI-6; Uebersax et al., 1995). As the UDI-6 contains some items that may be endorsed by those without urinary incontinence (e.g. pain in the lower abdominal region), the UDI-6 may overstate incontinence prevalence and the impact of this on peoples' lives. Given its psychometric properties, there was a case for major revision of the UDI-6. Although the ISI was the preferred measure, because it violated the assumptions of classic psychometric theory relating to scale stability as it contained only two items, further research into its properties was also recommended.

Regarding the measurement of faecal incontinence, as Hawthorne (2006) indicated the current definition of faecal incontinence by the International Continence Society excludes flatus, yet an item on flatus is included in the Wexner Faecal Continence Grading Scale (Jorge and Wexner, 1993). In addition to this definitional inconsistency, the evidence from Hawthorne (2005) and AIHW (2006) suggested that the inclusion of the flatus item led to overestimates of faecal incontinence prevalence. It was recommended that further work on the Wexner be undertaken to remove the flatus item and to improve the measurement properties of this scale.

A follow up project Refining Continence Measurement Tools (Sansoni et al., 2006) was then undertaken to further analyse the SAHOS dataset to develop better instruments for the assessment of incontinence in Australia. This project made use of both Classic Test Theory and Modern Test Theory approaches for the analysis of data. From the analysis of the urinary and faecal incontinence items and scales included in the 2004 SAHOS this study developed some revised scales for urinary and faecal incontinence (Revised Urinary Incontinence Scale (RUIS); Revised Faecal Incontinence Scale (RFIS)) which improved the screening assessment of incontinence when compared with the original measures (Sansoni et al., 2006). Both scales were found to have excellent internal consistency reliability (RUIS 0.91; RFIS 0.89) in a community survey sample. A study using the RUIS in a clinical sample (Hawthorne et al., 2006) indicated that the RUIS could describe more severe cases of incontinence than would be found in a population survey sample and that it was sensitive to change / improvement arising from treatment.

One of the limitations of using community survey data is that, as the data is collected in face to face interviews the data are at the level of subjective reports of incontinence symptoms rather than confirmed diagnoses. These considerations mean that in a community survey there will be a limited range of responses to incontinence items particularly those pertaining to more severe levels of symptoms. Thus it was necessary to trial the revised continence measures in a range of clinical settings in follow-up field trials.

Recently the RUIS and the RFIS were assessed in clinical settings by the Validation and Clinical Translation of the Revised Continence and Patient Satisfaction Tools Project (Sansoni et al., 2011). The findings from this study are described briefly below.
Revised Urinary Incontinence Scale (RUIS): Baseline Data

Urinary incontinence patients (N = 195) were recruited consecutively from 7 urinary continence clinics (specialist and community) across 4 regions in Australia. The Sansoni et al., 2011 study examined clinical and patient definitions of incontinence status, treatment outcomes and success across 4 treatment types (Continence Advising, Physiotherapy, Surgery and Combined/Other Treatments). Study eligibility criteria were attending a clinic to receive treatment for urinary incontinence, age between 18-85, and having sufficient English to complete a self-report questionnaire.

This manual reports on the analysis of 195 urinary incontinence cases with complete baseline data and 100 participants with full data (pre-post) available. In this study (Sansoni et al., 2011) most participants were Australian born (68%), were over 50 years (72%), had achieved completion of high school and/or a post school qualification (70%), were working (PT/FT) (44%) and were female (86%).

Physiotherapy was the most common treatment (57%), followed by continence advising (20%) and surgery (19%). With regard to type of urinary incontinence 80% of patients were identified by the clinicians as having stress incontinence symptoms and 63% had urge incontinence symptoms and many of these patients (49%) had mixed incontinence or a combination of these types. Only 5% of patients were identified by clinicians as having overflow incontinence and 1.5% of patients were identified as having functional incontinence.

Many of the patients had experienced incontinence symptoms for more than 2 years (65%) and only 16% had developed their incontinence symptoms within the last year. Only 10% of the sample had no other illnesses or health conditions, 17% of patients had one other health condition and 73% of the sample had 2 or more co-morbidities. The most common co-morbidities were problems with the neck, back or spine (47%), arthritis (44%), high blood pressure (34%) and migraine headaches (28%).

Figure A1 shows that at baseline for all 195 cases the mean RUIS score was 10.92 (SD = 3.33; N = 195). The mean RUIS score for the sample of females at pre-treatment was 10.90 (SD = 3.16, N = 167) and the mean RUIS score for the sample of males at pre-treatment was 11.07 (SD = 4.18, N = 28) (NB: Similar median results were found for both genders). There was no significant difference between these total scores when analysed by gender (F = 0.07, p>0.05). There were no significant gender differences between the means for any of the individual items (p>0.05). This data may suggest the scale is appropriate for use with both genders. For most RUIS items, at baseline, 62-68% of the clinical sample experienced these symptoms moderately or greatly.

Figure A1: RUIS pre-treatment scores
Table A1 provides a summary of patient health and incontinence status indicators at baseline in comparison to RUIS scores and the statistics relating to these comparisons.

### Table A1: Health and incontinence status at baseline in comparison with RUIS scores

<table>
<thead>
<tr>
<th>Variable</th>
<th>Classifications (N)</th>
<th>RUIS Mean(SD)</th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Male N = 28; 16%</td>
<td>11.07 (4.18)</td>
<td>t = 0.25, df 193; p = 0.80</td>
</tr>
<tr>
<td></td>
<td>Female N = 167; 86%</td>
<td>10.90 (3.81)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>All N = 195</td>
<td>10.92 (3.33)</td>
<td></td>
</tr>
<tr>
<td>Age Groups</td>
<td>Less than 40 yrs N = 18; 9%</td>
<td>11.72 (3.39)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>40-49 N = 36; 19%</td>
<td>10.92 (3.18)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>50-59 N = 47; 24%</td>
<td>10.94 (3.20)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>60-69 N = 62; 32%</td>
<td>10.69 (3.47)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>70 yrs or more N =32; 16%</td>
<td>10.91 (3.52)</td>
<td>F= 0.33, df 4, 190; p = 0.86</td>
</tr>
<tr>
<td>Number of Co-morbidities</td>
<td>0/1 N = 54; 28%</td>
<td>11.30 (3.13)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2/3/4 N = 97; 50%</td>
<td>10.70 (3.60)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5+ N = 44; 23%</td>
<td>10.95 (2.97)</td>
<td>F = 0.55, df 2, 192; p = 0.58</td>
</tr>
<tr>
<td>General Health Status</td>
<td>Excellent/Very good N = 79; 41%</td>
<td>10.71 (3.13)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Good N = 70; 36%</td>
<td>10.94 (3.51)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fair/Poor N = 46; 24%</td>
<td>11.26 (3.44)</td>
<td>F = 0.40, df 2, 192; p = 0.67</td>
</tr>
<tr>
<td>Symptom Duration</td>
<td>Less than 2 years N = 61; 35%</td>
<td>11.05 (3.24)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 years + N = 109; 65%</td>
<td>11.20 (3.10)</td>
<td>t = -0.30, df 168; p = 0.76</td>
</tr>
<tr>
<td>Clinician Rated Incontinence</td>
<td>Mild N = 65; 34%</td>
<td>9.22 (2.97)</td>
<td></td>
</tr>
<tr>
<td>Severity</td>
<td>Moderate N = 99; 51%</td>
<td>11.79 (3.12)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Severe N = 30; 15%</td>
<td>12.13 (2.68)</td>
<td>F = 16.99, df 2, 191; p = 0.000</td>
</tr>
<tr>
<td>Patient Rated Incontinence</td>
<td>Mild N = 66; 34%</td>
<td>8.36 (2.78)</td>
<td></td>
</tr>
<tr>
<td>Severity</td>
<td>Moderate N = 94; 48%</td>
<td>11.60 (2.85)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Severe N = 34; 18%</td>
<td>14.03 (1.66)</td>
<td>FW* = 80.46, df 2, 109.07; p = 0.000</td>
</tr>
<tr>
<td>Pad Use</td>
<td>&lt;1 per day N = 92</td>
<td>8.06 (3.45)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>≥ 1 Pads per day N = 102</td>
<td>11.55 (2.94)</td>
<td>t = -6.22, df 192; p = 0.000</td>
</tr>
<tr>
<td>Pad Size</td>
<td>No Pad/ Thin Pad N = 97</td>
<td>9.21 (3.14)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Medium Pad N = 60</td>
<td>12.00 (2.54)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Large Pad N = 37</td>
<td>13.57 (2.45)</td>
<td>F = 38.56, df 2, 191; p = 0.000</td>
</tr>
<tr>
<td>Double Incontinence</td>
<td>Ul only N = 155</td>
<td>10.60 (3.37)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ul and FI N = 40</td>
<td>12.18 (2.89)</td>
<td>t = -2.71, df 93; p = 0.007</td>
</tr>
<tr>
<td>Type of Treatment</td>
<td>Conservative N = 151</td>
<td>10.71 (3.46)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Surgical N = 33</td>
<td>12.03 (2.23)</td>
<td>t = -2.10, df 70; p = 0.007</td>
</tr>
<tr>
<td>No. of Ul leak symptoms</td>
<td>2 or less N = 70</td>
<td>8.60 (3.25)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3-4 N = 72</td>
<td>11.46 (2.61)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5+ N = 52</td>
<td>13.21 (2.16)</td>
<td>FW* = 43.84, df 2, 125.42; p = 0.000</td>
</tr>
</tbody>
</table>

* Welch Statistic

When examined by clinician ratings of pre-treatment severity there was a significant difference in total RUIS scores between those that had mild urinary incontinence compared with those that had moderate or severe incontinence (F = 16.99, df 2, 191; p = 0.000). The mean RUIS score for the...
mild group was 9.22, for the moderate group it was 11.79 and for the severe group it was 12.13. There were significant differences in RUIS group means between those that had mild incontinence and those with moderate or severe incontinence (p < 0.05).

Similarly, when examined by the pre-treatment patient rating of the severity of urinary incontinence the RUIS mean for those in a mild state was 8.36, for those with moderate incontinence it was 11.60 and for those with severe incontinence it was 14.03; this analysis was statistically significant (F Welch = 80.46, df 2, 109.07; p = 0.000) and all group comparisons were significant (p<0.05).

The mean RUIS scores were significantly higher (t = -2.76, df 70.49; p = 0.007) for those that received surgical vs. conservative treatments – this might be expected as generally surgery is used with patients with more severe incontinence or for whom conservative treatment has failed. The pre-treatment RUIS mean for the surgical group was 12.03 and for the conservative treatment group (physiotherapy and continence advising) it was 10.71.

At pre-treatment there was a significant differences (t = -2.71, df 193; p = 0.007) in RUIS mean scores between those that had both faecal and urinary incontinence (double incontinence) as against urinary incontinence only. The mean for those with urinary incontinence only was 10.60 and for those with double incontinence it was 12.18. Patients with mixed types of incontinence had higher RUIS scores than those with only urge incontinence (F = 9.83, df 2, 176; p = 0.000) and all group comparisons were significant (p<0.05).

The RUIS does not include any items concerning pad use but pad use should reflect the severity of urinary leakage. Patient scores on the Wei items concerning the number and the size of pads were compared to RUIS scores. Patient scores on the number of pads used were grouped as less than 1 per day and 1 per day or more. The mean RUIS score for the low pad use group was 8.06 and for the high pad use group it was 11.55. This difference was statistically significant (t = -6.22, df 192; p = 0.000).

For the pad size item three groups were classified as no pad/thin pad or tissue; medium or regular pad; and large/maxi pad or absorbent disposable undergarment. This analysis was significant (F = 38.56; df 2,191; p = 0.000). The mean RUIS score for those with no pad/thin pad/tissue was 9.21; for those with medium pads it was 12.00 and for those using large pads it was 13.57. These group comparisons were statistically significant (p<0.05).

The ICIQ-SF (Avery et al., 2004) also contains items about whether particular symptoms associated with urinary leakage are present or absent (e.g. leaks before you can get to toilet, when you cough or sneeze, when asleep, when physically active or exercising, when finished urinating and are dressed, leaks for no obvious reason and leaking all the time). The number of these symptoms present (= 1) at baseline, plus an additional item concerning leaking without knowing about it, were added to form an index of leakage severity. From an examination of the data distribution the groups were classified as 2 or less symptoms; 3 and 4 symptoms, and 5 or more symptoms. There was a significant difference in RUIS pre-treatment scores for these groups. For those that had 2 or less symptoms the mean RUIS score was 8.60 for those with three and four leakage symptoms it was 11.46 and for those with 5 or more symptoms it was 13.21 (F Welch = 43.84, df 2, 125.42; p = 0.000). All group comparisons were significant. All the instruments discriminated significantly (p = 0.000) by the alternate index of severity but the RUIS was the most sensitive measure and had the highest F value whereas the ISI was the least sensitive measure for this analysis.

These analyses of the baseline data reflected that the RUIS discriminated well between levels of incontinence severity as measured by other indicators.

When the RUIS was examined by baseline health status there were no significant differences between those in Excellent/Very Good, Good, or Fair/Poor Health (F = 0.40; p>0.05); similarly there were no statistically significant associations by gender, age group, education level, or work status. There was a statistically significant association between BMI and RUIS ($ r_s = 0.20, p<0.02$), but the common explained proportion of variance was very small (~4%). This would suggest the RUIS is assessing the underlying condition of urinary incontinence and this assessment appears to be independent of some common confounders.
### Table A2: RUIS in relation to other incontinence variables at follow-up

<table>
<thead>
<tr>
<th>Variables</th>
<th>Classifications</th>
<th>(N)</th>
<th>Mean (SD)</th>
<th>Statistics</th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>RUIS</td>
<td>Pre-treatment</td>
<td>N = 100</td>
<td>11.02 (3.08)</td>
<td></td>
<td>t paired = 8.56, df 99; p = 0.000</td>
</tr>
<tr>
<td></td>
<td>Post-treatment</td>
<td>N = 100</td>
<td>6.95 (4.77)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean Change Score</td>
<td></td>
<td>4.07 (4.76)</td>
<td>t paired = 8.56, df 99; p = 0.000</td>
<td></td>
</tr>
<tr>
<td>ISI</td>
<td>Pre-treatment</td>
<td>N = 100</td>
<td>6.44 (3.26)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Post-treatment</td>
<td>N = 100</td>
<td>3.74 (3.65)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean Change Score</td>
<td></td>
<td>2.70 (4.06)</td>
<td>t paired = 6.65, df 99; p = 0.000</td>
<td></td>
</tr>
<tr>
<td>UDI-6</td>
<td>Pre-treatment</td>
<td>N = 100</td>
<td>8.82 (3.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Post-treatment</td>
<td>N =100</td>
<td>5.46 (4.06)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean Change Score</td>
<td></td>
<td>3.36 (3.73)</td>
<td>t paired = 9.00, df 99; p = 0.000</td>
<td></td>
</tr>
<tr>
<td>ICIQ-SF</td>
<td>Pre-treatment</td>
<td>N = 99</td>
<td>11.81(4.25)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Post-treatment</td>
<td>N = 99</td>
<td>8.03 (4.98)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean Change Score</td>
<td></td>
<td>3.78 (5.18)</td>
<td>t paired = 7.25, df 98; p = 0.000</td>
<td></td>
</tr>
<tr>
<td>RUIS post scores by</td>
<td>Moderate/Severe</td>
<td>N = 32</td>
<td>12.34 (2.62)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient Rated Severity (post)</td>
<td>Mild</td>
<td>N = 31</td>
<td>6.48 (2.50)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Normal</td>
<td>N = 37</td>
<td>2.68 (2.56)</td>
<td>F = 122.85, df 2.97; p = 0.000</td>
<td></td>
</tr>
<tr>
<td>RUIS post scores by</td>
<td>Severe</td>
<td>N = 14</td>
<td>11.29 (4.01)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinician Rated Severity (post)</td>
<td>Moderate and Mild</td>
<td>N = 54</td>
<td>7.98 (4.25)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Normal</td>
<td>N = 29</td>
<td>2.72 (2.74)</td>
<td>F = 26.86, df 2.94; p = 0.000</td>
<td></td>
</tr>
<tr>
<td>RUIS post scores by</td>
<td>Not Helped</td>
<td>N = 17</td>
<td>12.29 (2.91)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient Rated Outcome</td>
<td>Some Improvement</td>
<td>N = 23</td>
<td>9.35 (3.33)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Partially Cured</td>
<td>N = 38</td>
<td>5.63 (3.19)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cured</td>
<td>N = 20</td>
<td>1.70 (3.06)</td>
<td>F = 41.41, df 3.94; p = 0.000</td>
<td></td>
</tr>
<tr>
<td>RUIS post scores by</td>
<td>Major Problems</td>
<td>N = 26</td>
<td>12.85 (2.34)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current Incontinence Problems (post)</td>
<td>Some Problems</td>
<td>N = 18</td>
<td>8.83 (2.53)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Slight Problems</td>
<td>N = 34</td>
<td>5.12 (1.92)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>No Problems</td>
<td>N = 22</td>
<td>1.27 (2.10)</td>
<td>F = 124.65, df 3.96; p = 0.000</td>
<td></td>
</tr>
<tr>
<td>RUIS Change by Patient Rated</td>
<td>Outcome</td>
<td>N = 17</td>
<td>-0.88 (2.85)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RUIS Change by Patient Rated</td>
<td>Improvement</td>
<td>N = 22</td>
<td>-0.41 (2.84)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A Little Better /Much Better</td>
<td>N = 44</td>
<td>2.80 (3.16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Very Much Better</td>
<td>N = 33</td>
<td>8.82 (3.44)</td>
<td>F = 43.53, df 3.95; p = 0.000</td>
<td></td>
</tr>
<tr>
<td>RUIS Change by Clinician Rated Improvement</td>
<td>Worse/ No Change</td>
<td>N = 12</td>
<td>0.42 (3.26)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Little and Much Better</td>
<td>N = 53</td>
<td>2.83 (3.90)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Very Much Better</td>
<td>N = 31</td>
<td>7.97 (4.34)</td>
<td>F = 22.48, df 2.93; p = 0.000</td>
<td></td>
</tr>
<tr>
<td>Patient Rated Improvement by</td>
<td>2 point RUIS Change</td>
<td>N = 67</td>
<td>5.15 (1.05)</td>
<td>t = 6.44, df 45.92; p = 0.000</td>
<td></td>
</tr>
<tr>
<td>Clinician Rated Improvement by</td>
<td>2 point RUIS Change</td>
<td>N = 66</td>
<td>5.15 (0.92)</td>
<td>t = 4.29, df 95; p = 0.000</td>
<td></td>
</tr>
<tr>
<td>RUIS Change by Baseline health status</td>
<td>Excellent/Very Good</td>
<td>N = 41</td>
<td>5.73 (4.51)</td>
<td>F = 6.87, df 2.97; p = 0.002</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>N = 35</td>
<td>3.91 (4.70)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fair/Poor</td>
<td>N = 24</td>
<td>1.46 (4.17)</td>
<td>F = 6.87, df 2.97; p = 0.002</td>
<td></td>
</tr>
<tr>
<td>RUIS Change by Patient Rated</td>
<td>Severity at Baseline</td>
<td>N = 31</td>
<td>3.32 (3.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Moderate</td>
<td>N = 51</td>
<td>4.22 (5.21)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Severe</td>
<td>N =18</td>
<td>4.94 (5.45)</td>
<td>F Welch = 0.85,df 2.43; p = 0.50</td>
<td></td>
</tr>
</tbody>
</table>
Revised Urinary Incontinence Scale (RUIS) Post-treatment Data

Table A2 shows the changes in RUIS, UDI-6 and ISI mean scores from pre-treatment to post-treatment. It also shows RUIS scores in relation to other incontinence indicators at follow up and the statistics relating to these comparisons. At follow up post-treatment the mean RUIS score was 6.95 (SD = 4.76, N = 100). For females the mean was 6.92 and for males the mean was 7.21. This shows a significant improvement for the patient group.

When RUIS post-treatment scores were examined by the clinical post-treatment rating of incontinence severity there was a significant difference (F = 28.68, df 2, 94; p = 0.000) in mean scores by level of severity. Those that the clinicians considered now to be ‘normal’ had a mean RUIS score of 2.72, those considered ‘mild or moderate’ had a mean score of 7.62, and those still rated as ‘severe’ had a mean score of 11.29.

Similarly, when examined by post-treatment patient rated severity the mean for those in a normal state was 2.68; for those with mild incontinence it was 6.48; and for those with moderate and severe incontinence it was 12.34. This finding was statistically significant (F = 122.65; df 2, 97; p = 0.000) and all group comparisons were significant.

The patient’s ratings concerning their treatment outcome was that 17% indicated their incontinence had not been helped by treatment, 23% reported some improvement, 39% considered themselves to be partly cured and 20% indicated that their continence was cured. There was a significant difference in RUIS post treatment scores between these groups (F = 41.41, df 3, 94; p = 0.000). For those that were ‘not helped’ the mean was 12.29, for those that had ‘some improvement’ the mean was 9.35, the mean for those that were ‘partly cured’ was 5.63 and for those who considered themselves ‘cured’ it was 1.70. All group comparisons were significant (p < 0.05).

In the case of the follow-up item concerning whether the patient still had problems with incontinence there was a significant difference (F = 124.65, df 3, 96; p = 0.000), in RUIS post-treatment mean scores between those that considered that they still had major problems (M = 12.85), some problems (M = 8.83) slight problems (M = 5.12) and those that had no problems at all (M = 1.27) and all group comparisons were significant.

Change scores on the RUIS (pre-post) ranged from an improvement of 15 points to a deterioration of 8 points. Examination of pre-post scores revealed a statistically significant improvement of 4.07 points (SD = 4.76, CI 3.13-5.01, N = 100) (paired t-test, t = 8.56, df 99; p = 0.000). The frequency distributions of RUIS pre-treatment and post-treatment scores are provided in Figure A2.

Figure A2: RUIS pre-treatment and post treatment scores

<table>
<thead>
<tr>
<th>RUIS pre-treatment scores for post treatment sample</th>
<th>RUIS post-treatment scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean = 11.02, SD = 3.08, N = 100</td>
<td>Mean = 6.95, SD = 4.77, N = 100</td>
</tr>
</tbody>
</table>

Mean = 11.02, SD = 3.08, N = 100

Mean = 6.95, SD = 4.77, N = 100
When examined by pre-treatment incontinence severity status there were no statistically significant differences in change scores between the groups. This indicates there was no significant difference in the degree of improvement from treatment based on the initial patient severity rating as all groups improved to a similar level. The mean change scores were 3.32 points for the mild incontinence group, 4.22 points for the moderate group and 4.94 points for the severe group (F Welch = 0.85, df 2, 43.25; p>0.05).

RUIS change scores did not significantly vary by gender, education or age. They did, however, vary by health status with those reporting at baseline that they experienced excellent/very good health obtaining the greater benefit from treatment than those in fair or poor health. The mean change scores were 5.73 points for those in excellent/very good health, 3.91 for those in good health and 1.46 for those in fair/poor health (F = 6.87, df 2, 97; p<0.01).

Reliability: RUIS

The RUIS pre-treatment internal consistency reliability alpha = 0.73, (N = 195 urinary incontinence patients) which is considered adequate (Streiner and Norman, 2003). By comparison, for UDI-6, a scale with 6 items the pre-treatment alpha = 0.64. For the ICIQ-SF alpha = 0.65 and for the ISI it was 0.54. Streiner and Norman (2003) consider alpha levels below 0.70 to be inadequate. For the sample of all urinary and faecal incontinence patients the RUIS pre-treatment alpha was 0.84 (N = 254) which is considered to be very good (Streiner and Norman, 2003). The RUIS post-treatment alpha = 0.90, N = 100 which is similar to the post-treatment alpha (0.85) reported by Hawthorne et al. (2006).

Test-retest reliability was assessed at two weeks post-treatment using the Intra Class Correlation Coefficient. For the urinary sample (N = 60) the ICC = 0.77. For the total retest sample of incontinence patients (N = 78) the ICC = 0.80. The test-retest reliability was superior to that found for the UDI-6 (0.74), the ICIQ-SF (0.67) and the ISI (0.76).

The standard error of measurement (SEM) is related to reliability and estimates the extent, to which a test provides accurate scores (or the observed score reflects the true score), given there will always be some degree of variation in scores in repeated measures using same test. A lower SEM indicates greater precision in measurement. Using Cronbach’s alphas for the urinary sample the SEM for the RUIS = 1.73, for the UDI-6 it was 2.09, for the ISI it was 2.32 and for the ICIQ-SF it was 2.79.

Internal Structure: RUIS

The sample sizes were sufficient for analysis with a sample of 195 urinary patients at pre-treatment (Gaudagnoli and Velicer, 1988; Pallant, 2011). For all analyses presented below the Kaiser-Meyer-Olkin value was >0.6 (Kaiser 1970, 1974) and Bartlett’s Test of Sphericity (Bartlett 1954) reached statistical significance, supporting the factorability of the correlation matrices.

The RUIS appeared to be a uni-dimensional scale with all RUIS items loading above 0.64 on the one urinary incontinence factor extracted (eigenvalue = 2.43; accounting for 49% variance) which was characterised as urinary leakage. The analysis of the community survey data produces a similar structure with a similar pattern of item loadings. The one component extracted explains more variance (74%) but this may be explained by the greater homogeneity of this sample.
Table A3: Principal components analysis of RUIS items

<table>
<thead>
<tr>
<th>RUIS Items</th>
<th>Urinary Incontinence Patients (N = 195)</th>
<th>Community Survey (N = 2915)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Factor 1 (49%)*</td>
<td>Factor 1 (74%)*</td>
</tr>
<tr>
<td>RUIS1 - Urge</td>
<td>0.64</td>
<td>0.80</td>
</tr>
<tr>
<td>RUIS2 - Stress</td>
<td>0.67</td>
<td>0.83</td>
</tr>
<tr>
<td>RUIS3 – leak small amounts</td>
<td>0.80</td>
<td>0.88</td>
</tr>
<tr>
<td>RUIS4 – leak frequency</td>
<td>0.72</td>
<td>0.91</td>
</tr>
<tr>
<td>RUIS5 – leak volume</td>
<td>0.64</td>
<td>0.90</td>
</tr>
</tbody>
</table>

* = proportion of variance explained

In 2006 (Sansoni et al., 2006) the authors analysed a number of urinary incontinence items which included the 5 RUIS items (3 from UDI-6 and 2 from ISI), the 3 other UDI items concerning the frequency of urination, pain, and difficulty emptying the bladder. This analysis was repeated using the clinical sample of urinary patients (N = 195).

Table A4: Rotated factor matrices for urinary incontinence items

<table>
<thead>
<tr>
<th>Scale</th>
<th>Community Survey*</th>
<th>Clinical Survey*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2006 N = 2915</td>
<td>2011 N = 195</td>
</tr>
<tr>
<td>Freq Urination</td>
<td>0.48</td>
<td>0.41</td>
</tr>
<tr>
<td>Urgency Leakage</td>
<td>RUIS1 0.74</td>
<td>0.60</td>
</tr>
<tr>
<td>Stress Leakage</td>
<td>RUIS2 0.82</td>
<td>0.67</td>
</tr>
<tr>
<td>Leak Small Amount</td>
<td>RUIS3 0.85</td>
<td>0.77</td>
</tr>
<tr>
<td>Emptying Bladder</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain Lower Abdominal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leakage Frequency</td>
<td>RUIS4 0.89</td>
<td>0.71</td>
</tr>
<tr>
<td>Leakage Amount</td>
<td>RUIS5 0.89</td>
<td>0.66</td>
</tr>
</tbody>
</table>

* = Principal Components Analyses with Varimax Rotation

The 2006 analysis (N = 2,915) produced a 2 factor/component solution accounting for 67% of the variance with all RUIS items loading highly (loadings of 0.6 or above) on factor/component 1 (eigenvalue = 4.29). This was described as a ‘urinary leakage factor’ and it explained 54% of the variance. The other items concerning lower abdominal pain and emptying bladder from the UDI loaded on factor/component 2 (explaining 13 % of the variance) and this was described as ‘other bladder symptoms’. The frequency of urination item had low to moderate loadings on both factors (0.48 on the urinary leakage factor and 0.49 on the other bladder symptoms factor) which suggested it did not warrant inclusion in the scale. In 2006 (Sansoni et al., 2006) the RUIS items were selected with regard to their loadings on the primary urinary incontinence factor/component. This pattern of factor/component loadings has been replicated in this recent analysis of urinary incontinence patients (N = 195) which confirms the current descriptive system.

Responsiveness: Capacity to Detect Change for the RUIS

Change scores were calculated by subtracting the post-treatment score from the pre-treatment score and thus a reduction in scores on this scale is indicative of improvement.

Most patients improved as a result of their treatment (Figure A3). There was a significant improvement of 3.11 mean RFIS scores (paired t test, t = 3.89, df 37; p = 0.000) following treatment (refer Table A2). In Sansoni et al. (2011) it was shown that the RUIS was equally or
more sensitive to detecting change than other widely used incontinence instruments. The relative efficiency statistic was used to compare the UDI-6, the ICIQ-SF and the RUIS against the least sensitive measure, the ISI (RE = 1.00). The RE = 1.66 for the RUIS, RE = 1.83 for the UDI-6, and the RE was 1.19 for the ICIQ-SF. This indicates that the UDI-6 and the RUIS are far more sensitive to change than the ISI and the ICIQ-SF. The UDI-6 was slightly more sensitive than the RUIS but it should be noted that the RUIS is a shorter instrument than the UDI.

The Kazis effect size (ES) is the mean change score divided by the standard deviation of the baseline score and this was -1.32 for the RUIS. This is classed as a large effect size (Kazis et al., 1989; Cohen, 1988; Crosby et al., 2003). The Kazis effect size for the UDI-6 was -1.02, for the ICIQ-SF it was -0.89 and for the ISI it was -0.82. Note that for all instruments a reduction in scores equated to an improvement in incontinence status. This indicated that all measures were responsive to change over time but the RUIS had the largest effect size of all these measures.

The responsiveness of the instrument is also concerned with the capacity of the instrument to detect change regardless of whether it is improvement or deterioration. Ignoring signs indicating the direction of change the average change score for the sample was 4.83 RUIS points SD = 3.97 N = 100 (as compared with UDI-6 = 4.12 points). Change scores on the RUIS (pre-post) ranged from an improvement of 15 points to a deterioration of 8 points in the context of a 16 point scale. These findings would also suggest that the instrument has the capacity to detect both an improvement and deterioration in patient incontinence status.

RUIS change scores were also examined by the patient’s and the clinician’s global ratings of improvement. These analyses suggested that a change score of > 2 RUIS may be the minimal detectable difference concerning change in the patients’ and the clinicians’ perception of improvement.

The SEM for the RUIS of 1.73 (95% CI SEM = 3.39) would suggest that a score change of 3-4 points is likely to be a more clinically and statistically reliable estimate for the purposes of patient monitoring although it is noted that the SEM reported is for the instrument as a whole and standard errors of measurement vary by score levels (they more be more or less for people scoring at the extremes of the scale) (Spratt, 2009).

Further analyses, using a larger post-treatment sample combined with the use of other clinical indicators as anchor points, will be required to further address this issue.

**Figure A3: RUIS change scores**

![RUIS change scores graph]

Dark Blue = Improved; White = Little or No Change; Light Blue = Worse
Correlations with Other Measures: RUIS

The RUIS has been shown to correlate with other measures of incontinence in the expected directions which is also an indication of validity. The RUIS was correlated with a number of other measures of urinary incontinence. Pre-treatment correlations between RUIS and UDI-6 \( r = 0.76 \) (\( p < 0.01 \)); RUIS and ISI \( r = 0.76 \) (\( p < 0.01 \)); RUIS and ICIQ-SF \( r = 0.74 \) (\( p < 0.01 \)); RUIS and WEI (urinary) Incontinence Symptom Index \( r = 0.72 \) (\( p < 0.01 \)); RUIS and WEI Bothe Index \( r = 0.66 \); RUIS and Incontinence Impact Questionnaire \( r = 0.53 \) (\( p < 0.01 \)). The RUIS also had a substantial correlation with patient rated severity at pre-treatment \( r_s = 0.62 \) (\( p < 0.01 \)). The correlation between RUIS and the pre-treatment clinical rating of severity was significant (\( p < 0.01 \)) but somewhat lower \( r_s = 0.37 \) and this pattern was also found for all other incontinence instruments.

Overall, these correlations were high, significant and in the expected directions providing good construct validity for the RUIS. As may be expected the RUIS correlations with other urinary incontinence symptom measures were higher than with the Incontinence Impact Questionnaire and the WEI Bothe Index which are more concerned with the impact of urinary incontinence on daily activities rather than the extent of the symptom itself.

Sansoni et al. (2006) reported that urinary and faecal incontinence measures had negative but significant correlations with measures of health related quality of life reflecting the burden of disease for this condition (e.g. the higher the incontinence score the lower the physical functioning or health status score). At pre-treatment the RUIS correlation with the Physical Function Scale of the SF-36V2 was \( r = -0.15 \) (\( p = 0.039 \)) confirming this finding. It is interesting that this is a lower correlation than was found for faecal incontinence suggesting that the impact of faecal incontinence on physical function may be greater.

The RUIS correlation with the untransformed patient satisfaction scores was \( r = -0.44 \) (\( p < 0.01 \)) which indicates there was an association between higher RUIS scores (reflecting greater incontinence post -treatment) and lower patient satisfaction scores.

Ceiling and Floor Effects

At pre-treatment scores on the RUIS ranged from 0 – 16 indicating the full range of scores were used. There was 0.5% of the sample on the floor of the instrument (lowest possible score) and 5.6% of the sample at the ceiling of the scale (maximum possible score).

Generally a figure of less than 15% of the sample at the floor or the ceiling is preferred (Streiner and Norman, 2003). It is clear that this instrument does not suffer from significant floor or ceiling effects.

Effect of Treatment

There was a statistically significant difference in the degree of improvement by treatment type (\( F = 9.35 \), df 2, 94; \( p = 0.000 \)) for the RUIS (Refer Table A2). For those receiving continence advising the mean RUIS change score was 2.00, for physiotherapy it was 3.09 and for surgery it was 7.07. There was a greater degree of improvement in post-treatment RUIS scores for those receiving surgical rather than conservative treatments although all treatment groups showed improvement. This is despite the fact one may expect some rate of complications (2% risk of voiding difficulty and 5% risk of urge incontinence) for surgical treatments. However, it should be noted that the change in mean scores for all treatment types from pre-treatment to post-treatment was significant (e.g. CNA mean difference = 2.00; \( t = 2.69 \), df 12; \( p < 0.02 \)).

Suggested Cutpoints for Interpretation: RUIS

It would be advisable to refer people who score 4 or more on the RUIS for further continence assessment. This is two standard deviations below the mean for the total clinical sample but one standard deviation above the mean for the community survey sample (\( M = 1.74; \ SD = 2.93 \)). A score of 8 was the 23rd percentile for the clinical sample and scores of 8 or below might be considered 'mild', a score of 10 was the 50th percentile which might be considered 'moderate' and
a score of 14 represented the 75th percentile which might be considered ‘severe’ (refer Figure A1 and Table A8).

The clinician pre-treatment ratings indicated that a RUIS score of 9 (M = 9.22) or below was considered ‘mild’, a score of 10-12 was considered ‘moderate’ (M = 11.79) and a score above 12 (M = 12.13) was classified as severe. The patient pre-treatment ratings indicated that a score of 8-9 or below was considered ‘mild’ (M = 8.36), a score of 10 through 12 was considered ‘moderate’ (M = 11.60) and a score of 13 or above was considered ‘severe’ (M = 14.03). At post-treatment a score of 3 or less was classified as ‘normal’ by the clinicians and by the patient’s rating of severity.

The ICIQ-SF scale contains 3 items (scale range is from 0-21) concerning the frequency and the volume of urinary leakage and the interference this causes in everyday life. The ICIQ-SF total scores, were recoded to form three groups based on scores +/- 1 standard deviation from the mean (low scores 0-6, N = 37 /moderate scores 7-16, N =125 / high scores 17-21, N = 29). The RUIS means for these three groups were then examined. The low ICIQ-SF score group had a RUIS mean of 7.22 (SD = 3.01, CI 6.21-8.22), the moderate score ICIQ-SF group had a RUIS mean of 11.26 (SD = 2.72, CI 10.74-11.74) and the high score ICIQ-SF group had a RUIS mean of 14.21 (SD = 1.47, CI 13.65-14.77). This would suggest that RUIS scores of 5 through 8 could be considered to reflect ‘mild incontinence’; that scores from 10 through 12 might reflect ‘moderate incontinence’, and that scores of 13 or above might reflect ‘severe’ incontinence.

At pre-treatment the RUIS mean for those wore less than 1 pad per day was 8.06 and the mean for those that wore 1 pad per day or more was 11.55 (t = -6.22, df 192, p = 0.000). At post-treatment the mean for those that did not wear pads was 3.74 and the mean for those that were still wearing pads post-treatment was 9.69 (t = -7.92, df 98; p = 0.000) This suggests that a useful cutpoint based on behavioural anchor points would be a RUIS score of < 9 to distinguish between those patients whose incontinence was not sufficiently severe to wear pads (mild incontinence) and those who needed to wear pads (moderate incontinence).

The following RUIS score ranges are suggested to aide interpretation:
0-3: no urinary incontinence or extremely mild (‘slight’) or occasional (‘rare’) incontinence
4-8: mild urinary incontinence
9-12: moderate urinary incontinence
13-16: severe urinary incontinence (scores of 15 – 16 could be considered very severe).

**Revised Faecal Incontinence Scale (RFIS): Baseline Data**

Most participants were Australian born (74%); spoke English as their primary language (95%); were over 50 years (80%); had achieved completion of high school or a post school qualification (72%); were retired (56%) and were female (84%). The average age of patients was 60.97 years – a somewhat older group than for urinary incontinence.

With regard to type of faecal incontinence, 58% experienced passive incontinence symptoms, 50% experienced urge incontinence, 50% experienced faecal seepage and many of these patients (55%) had mixed incontinence or a combination of these types. Clinicians identified 47% of patients as having flatus incontinence symptoms usually in combination with other forms of faecal incontinence but 4 patients had flatus incontinence only.

Continence Advising was the most common treatment (39%), followed by physiotherapy (28%) and surgery (26%). Many of the patients had experienced incontinence symptoms for more than 1 year (79%), a large number of patients had experienced symptoms for 3 years of more (51%) and only 21% had first received treatment for their faecal incontinence symptoms within the last year. Only 8% of the sample had no other illnesses or health conditions, 16% had one co-morbidity and 76% of the sample had 2 or more co-morbidities. The most common co morbidities were problems with the neck, back or spine (53%), arthritis (57%), high blood pressure (31%) and migraine headaches (25%).
Table A5 shows health and incontinence status at baseline in comparison with RFIS scores and the statistical data relating to these analyses. At baseline for all 61 cases the mean RFIS score was 9.66 (SD = 4.66) and scores ranged from 0-20. For most RFIS items at baseline 50-68% of the sample experienced these symptoms ‘sometimes’ through to ‘always’. The mean RFIS score for the samples at pre-treatment was 9.10 for males and 9.76 for females. This is in contrast to the very low mean scores for the RFIS in the community sample (N = 2915) of 0.43 (females) and 0.25 (males) respectively (SAHOS: Harrison Health Research, 2004; Sansoni et al., 2006; refer Table A9).

There was no significant difference between RFIS total scores when analysed by gender (t = -0.41, df 59; p>0.05) and there were no significant gender differences found for any individual items or in the pattern of responses by gender for any items. This might suggest the scale may be suitable for use with both genders although the sample of males is small so this will need to be confirmed with a larger sample.

When examined by the baseline clinician rating of the severity of faecal incontinence there was a significant difference between RFIS scores for the more severe and less severe faecal incontinence groups (F = 3.17, df 2, 56; p = 0.05). The mean for those with mild incontinence was 8.05, for those with moderate incontinence it was 9.76 and for those with severe incontinence it was 12.08. Similarly, when examined by the baseline patient rating of the severity of faecal incontinence the mean for those in a mild state was 7.12, for those with moderate incontinence it was 10.64 and for those with severe incontinence it was 15.14 (F = 13.11, df 2, 58; p = 0.000).

Faecal incontinence patients who had experienced faecal incontinence symptoms for two years or more had a significantly higher (t = -2.23, df 55; p = 0.03) mean RFIS score (M = 10.57) than those patients who had experienced incontinence for less than 2 years (M = 7.75). Excluding flatus (as anal incontinence) patients diagnosed with mixed faecal incontinence types had a higher RFIS mean score that those diagnosed as having only one type of faecal incontinence (t = -3.26, df 59, p = 0.002). Bowel diary data indicated that those who used pads 10 or more times per week had higher RFIS mean scores than those who used pads 9 or less times per week (t = -2.31, df 38, p = 0.03).

There was a trend for those having surgical treatment to have higher RFIS scores than those receiving conservative treatment (t = -1.87, df 57; p = 0.067) but this did not attain significance at the p<0.05 level.

The RFIS does not include any items concerning pad use but pad use should reflect the severity of faecal leakage. Patient scores on the Wexner and St Mark’s items concerning pad use were compared to RFIS scores. Patient scores on the number of pads used (Wexner item) were grouped as ‘never or rarely’; or ‘sometimes, often and always’. The mean RFIS score for the low pad use group was 6.48 and for the high pad use group it was 11.89. This difference was statistically significant (t = -5.10, df 56; p = 0.000). For the St Mark’s Incontinence Score pad the
mean RUIS score for the 'no pad' group was 7.23 and for the 'wears pad' group it was 11.28 and this difference was significant (t = -3.45, df 56; p = 0.001).

The study protocol contains a number of other (non RFIS) items about whether particular faecal leakage symptoms were present (e.g. leak mucus, incomplete bowel emptying, faecal urgency, faecal incontinence at night, seepage following a bowel movement and leaking without knowing it, and pad use). These symptom scores were added to form an alternate index of severity and then formed into 3 groups (=/- 0.5 SD units from the mean). There was a significant difference in RFIS pre-treatment scores for these groups (F = 14.19, df 2, 53; p = 0.000) For the low score group the RFIS mean was 5.50, for the moderate score group it was 9.96 and for the higher score group the RFIS mean was 13.31. For comparisons with the other measures the pad use item was deleted from the alternate index (as both Wexner and St Mark’s contain pad items) and the groups dichotomized. For all three instruments there was a significant difference in pre-treatment scores in relation to the alternate index of severity (RFIS t = -3.05, df = 54, p = 0.002; Wexner t = -2.75, df = 54, p = 0.008; SMIS t = -3.09, df = 53, p = 0.003). The RFIS and the SMIS were marginally more sensitive measures for this comparison.

These analyses reflect that the RFIS discriminates well between different levels of incontinence severity as measured by other clinical and symptom indicators.

When examined at baseline by a general item on health status there were no significant differences between those in Excellent/Very Good, Good, or Fair/Poor Health (F = 0.80; p>0.05); similarly there were no statistically significant associations by gender, age group, co morbidity number, type of treatment or BMI.
Table A5: Health and incontinence status at baseline in comparison with RFIS scores

<table>
<thead>
<tr>
<th>Variable</th>
<th>Classifications</th>
<th>(N)</th>
<th>RFIS (Mean (SD))</th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Male</td>
<td>N = 10; 16%</td>
<td>9.10 (3.60)</td>
<td>t = -0.41, df 59; p = 0.68</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>N = 51; 84%</td>
<td>9.76 (4.87)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>N = 61</td>
<td>9.66 (4.66)</td>
<td></td>
</tr>
<tr>
<td>Age Groups</td>
<td>Less than 50 yrs</td>
<td>N = 12; 20%</td>
<td>7.42 (5.35)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>50-59</td>
<td>N = 10; 16%</td>
<td>10.00 (3.83)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>60-69</td>
<td>N = 23; 38%</td>
<td>11.17 (4.69)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>70 yrs or more</td>
<td>N = 16; 26%</td>
<td>8.94 (4.09)</td>
<td>F = 1.97, df 3,57; p = 0.13</td>
</tr>
<tr>
<td>Number of Co-morbidities</td>
<td>0/1</td>
<td>N = 15; 25%</td>
<td>9.93 (5.50)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2/3/4</td>
<td>N = 32; 46%</td>
<td>9.28 (4.73)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5+</td>
<td>N = 14; 23%</td>
<td>10.21 (3.68)</td>
<td>F = 0.22, df 2,58; p= 0.80</td>
</tr>
<tr>
<td>Symptom Duration</td>
<td>Less than 2 years</td>
<td>N = 20; 35%</td>
<td>7.75 (4.95)</td>
<td>t = -2.23, df 55; p = 0.03</td>
</tr>
<tr>
<td></td>
<td>2 years +</td>
<td>N = 37; 65%</td>
<td>10.57 (4.34)</td>
<td></td>
</tr>
<tr>
<td>Body Mass Index</td>
<td>Less than 30</td>
<td>N = 41; 79%</td>
<td>9.78 (4.830)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>30 or more</td>
<td>N = 11; 21%</td>
<td>8.00 (4.12)</td>
<td>t = 1.12, df 50; p = 0.27</td>
</tr>
<tr>
<td>General Health Status</td>
<td>Excellent/Very good</td>
<td>N = 18; 30%</td>
<td>8.50 (4.64)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>N = 24; 39%</td>
<td>10.00 (4.44)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fair/Poor</td>
<td>N = 19; 31%</td>
<td>10.32 (4.99)</td>
<td>F = 0.80, df 2,58; p = 0.45</td>
</tr>
<tr>
<td>Clinician Rated Incontinence</td>
<td>Mild</td>
<td>N = 22; 37%</td>
<td>8.05 (4.48)</td>
<td>F = 3.17, df 2,56; p = 0.05</td>
</tr>
<tr>
<td>Severity</td>
<td>Moderate</td>
<td>N = 25; 42%</td>
<td>9.76 (4.31)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Severe</td>
<td>N = 12; 20%</td>
<td>12.08 (4.89)</td>
<td></td>
</tr>
<tr>
<td>Patient Rated Incontinence</td>
<td>Mild</td>
<td>N = 26; 43%</td>
<td>7.12 (3.80)</td>
<td>F = 13.11,df 2,58; p = 0.000</td>
</tr>
<tr>
<td>Severity</td>
<td>Moderate</td>
<td>N = 28; 46%</td>
<td>10.64 (3.97)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Severe</td>
<td>N = 7; 11%</td>
<td>15.14 (4.34)</td>
<td></td>
</tr>
<tr>
<td>Diagnosis Type: Faecal Incontinence</td>
<td>1 Type only</td>
<td>N = 27</td>
<td>7.63 (4.21)</td>
<td>t = -3.26, df 59; p = 0.002</td>
</tr>
<tr>
<td></td>
<td>Mixed Types FI</td>
<td>N = 34</td>
<td>11.26 (4.42)</td>
<td></td>
</tr>
<tr>
<td>Bowel Diary Soiling</td>
<td>&lt; 9 per week</td>
<td>N = 29</td>
<td>8.41 (4.73)</td>
<td>t = -2.31, df 38; p = 0.03</td>
</tr>
<tr>
<td></td>
<td>≥ 10 per week</td>
<td>N = 11</td>
<td>12.09 (3.78)</td>
<td></td>
</tr>
<tr>
<td>Type of Treatment</td>
<td>Conservative</td>
<td>N = 42</td>
<td>8.88 (4.89)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Surgery</td>
<td>N = 17</td>
<td>11.35 (3.79)</td>
<td>t = -1.87, df 57; p = 0.067</td>
</tr>
<tr>
<td>Wexner Pad Use Item</td>
<td>Never/Rarely</td>
<td>N = 23</td>
<td>6.48 (3.47)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sometimes-Always</td>
<td>N = 35</td>
<td>11.89 (4.23)</td>
<td>t = -5.10, df 56; p = 0.000</td>
</tr>
<tr>
<td>Alternate FI Symptoms Index by RFIS</td>
<td>Low</td>
<td>N = 12</td>
<td>5.50 (2.75)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Moderate</td>
<td>N = 28</td>
<td>9.96 (4.23)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>N = 16</td>
<td>13.31 (3.65)</td>
<td>F = 14.19, df 2,53, p= 0.000</td>
</tr>
</tbody>
</table>
RFIS Post-treatment data

Figure A5 shows the distribution of scores at pre-treatment and post–treatment for the RFIS. At follow up the mean RFIS score was 6.64 (SD = 4.76, N = 39) showing a significant improvement for the patient group.

Figure A5: RFIS pre-treatment and post treatment scores

Table A6 shows RFIS scores in relation to other incontinence variables at follow up and the statistics relating to these comparisons.

When examined by patient post treatment severity ratings there was a significant difference (t = 4.43, df 37; p = 0.000) in RFIS mean scores between the less severe (M = 4.60) and the more severe incontinence groups (M = 10.29). Table A6 shows the RFIS and the Wexner were the most sensitive instruments for this comparison (p = 0.000) with the St Mark’s being slightly less sensitive (p = 0.006).

The patient’s ratings concerning the outcomes of their treatment was that 7.7% indicated their incontinence had not been helped by treatment, 35% reported some improvement, 41% considered themselves to be partly cured and 15% indicated that their continence was cured. There was a significant difference in RFIS post treatment scores between these groups (t = 4.01, df 37; p = 0.000). For those that were not helped or had little improvement the RFIS mean score was 9.59 and for those that were partly cured/cured the mean was 4.36 and the RFIS and the Wexner scales were the most sensitive instruments for this comparison (p = 0.000).

There were also significant differences (t = 3.80, df 37; p = 0.001), in RFIS post-treatment mean scores between those that considered that they still had major problems/ some problems following treatment and those that had slight problems/no problems.

RFIS scores discriminated between levels of pad use at post–treatment (p = 0.000). The mean for those that used pads never or rarely at post-treatment was 4.29 whereas for those that used pads sometimes/always the mean was 10.40. RFIS scores also discriminated between levels of the Alternate Index of Severity (comprised of non RFIS other faecal items including pad use) at post-treatment (F = 16.14, df 2, 36; p = 0.000).
Table A6: RFIS and other incontinence variables at follow-up

<table>
<thead>
<tr>
<th>Variable</th>
<th>Classifications</th>
<th>N</th>
<th>Mean (SD)</th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>RFIS Scores</td>
<td>Pre-treatment</td>
<td>38</td>
<td>9.79 (4.68)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Post-treatment</td>
<td>38</td>
<td>6.68 (4.82)</td>
<td></td>
</tr>
<tr>
<td><strong>Mean Change Score</strong></td>
<td></td>
<td></td>
<td>3.11 (4.92)</td>
<td>t paired = 3.89, df 37; p = 0.000</td>
</tr>
<tr>
<td>Wexner Scores</td>
<td>Pre-treatment</td>
<td>38</td>
<td>9.58 (4.74)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Post-treatment</td>
<td>38</td>
<td>6.89 (4.83)</td>
<td></td>
</tr>
<tr>
<td><strong>Mean Change Score</strong></td>
<td></td>
<td></td>
<td>2.68 (5.18)</td>
<td>t paired = 3.19, df 37; p = 0.003</td>
</tr>
<tr>
<td>St Mark’s Scores</td>
<td>Pre-treatment</td>
<td>37</td>
<td>12.46 (5.40)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Post-treatment</td>
<td>37</td>
<td>8.97 (5.56)</td>
<td></td>
</tr>
<tr>
<td><strong>Mean Change Score</strong></td>
<td></td>
<td></td>
<td>3.49 (5.08)</td>
<td>t paired = 4.17, df 36; p = 0.000</td>
</tr>
<tr>
<td>RFIS post scores by Patient</td>
<td>Moderate/Severe</td>
<td>14</td>
<td>10.29 (4.55)</td>
<td></td>
</tr>
<tr>
<td>Rated Severity (post)</td>
<td>Normal/Mild</td>
<td>25</td>
<td>4.60 (3.56)</td>
<td>t = 4.33, df 37; p = 0.000</td>
</tr>
<tr>
<td>Wexner post scores by Patient</td>
<td>Moderate/Severe</td>
<td>14</td>
<td>10.43 (4.99)</td>
<td></td>
</tr>
<tr>
<td>Rated Severity (post)</td>
<td>Normal/Mild</td>
<td>25</td>
<td>4.72 (3.36)</td>
<td>t = 4.27, df 37; p = 0.000</td>
</tr>
<tr>
<td>SMIS post scores by Patient</td>
<td>Moderate/Severe</td>
<td>14</td>
<td>12.07 (5.27)</td>
<td></td>
</tr>
<tr>
<td>Rated Severity (post)</td>
<td>Normal/Mild</td>
<td>25</td>
<td>7.13 (4.82)</td>
<td>t = 2.95, df 36; p = 0.006</td>
</tr>
<tr>
<td>RFIS post scores by Patient</td>
<td>Not Helped/ Little</td>
<td>17</td>
<td>9.59 (4.37)</td>
<td></td>
</tr>
<tr>
<td>Rated Outcome (post)</td>
<td>Improvement</td>
<td>22</td>
<td>4.36 (3.75)</td>
<td>t = 4.01, df 37; p = 0.000</td>
</tr>
<tr>
<td>RFIS post scores by Current)</td>
<td>Major/ Some Problems</td>
<td>18</td>
<td>9.33 (4.60)</td>
<td></td>
</tr>
<tr>
<td>Incontinence Problems (post)</td>
<td>No Problems/ Slight</td>
<td>21</td>
<td>4.33 (3.61)</td>
<td>t = 3.80, df 37; p = 0.001</td>
</tr>
<tr>
<td>RFIS post scores by Pad Use</td>
<td>Sometimes to Always</td>
<td>15</td>
<td>10.40 (4.76)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Never or Rarely</td>
<td>24</td>
<td>4.29 (2.96)</td>
<td>t = 4.96, df 37; p = 0.000</td>
</tr>
<tr>
<td>RFIS post scores by Alt. Index</td>
<td>Low</td>
<td>15</td>
<td>3.53 (2.67)</td>
<td></td>
</tr>
<tr>
<td>Severity</td>
<td>Moderate</td>
<td>16</td>
<td>6.69 (4.22)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>8</td>
<td>12.38 (3.54)</td>
<td>F = 16.14, df 2,36; p = 0.000</td>
</tr>
<tr>
<td>RFIS <strong>Change</strong> by Patient</td>
<td>Not Helped/Little</td>
<td>17</td>
<td>0.59 (4.39)</td>
<td></td>
</tr>
<tr>
<td>Rated Outcome</td>
<td>Improvement</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Partly Cured/Cured</td>
<td>22</td>
<td>5.14 (4.43)</td>
<td>t = 3.16, df 37; p = 0.003</td>
</tr>
<tr>
<td>RFIS <strong>Change</strong> by Patient</td>
<td>Worse</td>
<td>5</td>
<td>-0.80 (4.32)</td>
<td></td>
</tr>
<tr>
<td>Rated Improvement</td>
<td>No Change/A Little</td>
<td>11</td>
<td>0.64 (4.08)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Better</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Much/Very Much Better</td>
<td>21</td>
<td>5.29 (4.53)</td>
<td>F = 6.34, df 2,34; p = 0.005</td>
</tr>
<tr>
<td>Patient Rated Improvement by 2</td>
<td>Improved by ≥ 2</td>
<td>22</td>
<td>5.27 (1.08)</td>
<td></td>
</tr>
<tr>
<td>point RFIS <strong>Change</strong></td>
<td>Improved by &lt; 2</td>
<td>15</td>
<td>4.13 (0.99)</td>
<td>t = 3.26, df 35; p = 0.002</td>
</tr>
<tr>
<td>RFIS <strong>Change</strong> by Patient</td>
<td>Normal/Mild</td>
<td>19</td>
<td>2.84 (6.08)</td>
<td></td>
</tr>
<tr>
<td>Rated Severity at Baseline</td>
<td>Moderate/Severe</td>
<td>19</td>
<td>3.37 (3.56)</td>
<td>t = -0.11, df 36; p = 0.75</td>
</tr>
</tbody>
</table>

* There were 39 patients that submitted post-treatment forms but for 1 subject no pre-treatment data was available

Table A6 shows the changes in the RFIS mean scores from pre-treatment to post-treatment in comparison with other measures of faecal incontinence. Change scores on the RFIS (pre-post)
ranged from an improvement of 18 points to a deterioration of 8 points. Examination of pre-post mean scores revealed a statistically significant improvement of 3.11 RFIS scores (paired t-test, t = 3.89, df 37; p = 0.000) and the RFIS and the SMIS were slightly more sensitive measures for detecting change than the Wexner.

RFIS change scores were examined by Patient Rated Outcome. The RFIS mean change score for those that considered themselves not helped/little improved was 0.59, for those that considered themselves cured/partly cured it was 5.14 (t = 3.16, df 37; p = 0.003).

RFIS change scores were also examined by Patient Rated Improvement. The RFIS mean change score for those that considered themselves worse was -0.80, for those that experienced no change or were a little better it was 0.64 and for those who were much/very much better it was 5.29 (F = 6.34, df 2,34; p = 0.005).

When examined by the pre-treatment patient rating of the severity of faecal incontinence there was no statistically significant difference in RFIS change scores (t = -0.11, df 36; p> 0.05) for those with normal/mild incontinence (M = 2.84) and for those with moderate/severe incontinence (M = 3.37). This indicates there was no significant difference in the degree of improvement during treatment based on the initial patient severity rating – both groups improved. This finding was the same for the other measures of faecal incontinence.

When these scores were dichotomized (improvers vs. no change/deterioration) and examined there were no significant differences by patient severity at baseline, health status, gender, education, age group or BMI.

Reliability: RFIS

The RFIS pre-treatment Cronbach’s alpha was 0.78, N = 61 (faecal incontinence patients). The pre-treatment alphas for the Wexner and the St Mark’s were 0.65 and these are considered marginal as they are less than 0.70. The RFIS pre-treatment alpha for a combined sample of urinary and faecal incontinence patients was 0.91 (N = 254) which is the similar to the reliability estimate for the community population sample (Sansoni et al., 2006). Post-treatment the RFIS alpha was 0.86 (faecal incontinence sample, N = 39) compared with the Wexner = 0.77 and the St Mark’s = 0.66.

Test-retest reliability for the RFIS was also assessed at 2 weeks after completion of treatment and the submission of post-treatment forms using the Intra Class Correlation Coefficient (ICC). For the faecal sample (N = 19) the ICC = 0.79 and for the Wexner test-retest ICC = 0.74 and for the St Mark’s ICC = 0.68. For the total retest sample of incontinence patients (N = 78) the RFIS ICC = 0.80.

The standard error of measurement (SEM) is related to reliability and estimates the extent, to which a test provides accurate scores (or the observed score reflects the true score), given there will always be some degree of variation in individual scores in repeated measures using same test. A lower SEM indicates greater precision in measurement. Using Cronbach’s alphas for the faecal sample (N = 61) the SEM for the RFIS = 2.19, for the Wexner it was 2.70 and for the St Mark’s it was 3.06.

Internal Structure: RFIS

For all analyses presented below the Kaiser-Meyer-Olkin value was >0.6 (Kaiser 1970, 1974) and Bartlett’s Test of Sphericity (Bartlett 1954) reached statistical significance, supporting the factorability of the correlation matrices.

The RFIS appears to be uni-dimensional with all items loading above 0.65 on the one faecal incontinence factor extracted (eigenvalue = 2.71; accounting for 54% of the variance) which was characterised as faecal soiling/leakage (Sansoni et al., 2011). In order to address sample size limitations (N = 61) this analysis was compared with the same analysis conducted on the community survey sample (N = 2915). The analysis of the community survey data produces a similar structure with a similar pattern of item factor loadings. The one factor extracted explains more of the variance (63%) but this may be explained by the greater homogeneity of this sample.
Table A7: Principal components analysis of RFIS items

<table>
<thead>
<tr>
<th>RFIS Items</th>
<th>Factor 1 (54%)*</th>
<th>Factor 1 (63%)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Faecal Incontinence Patients (N = 61)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leak solid</td>
<td>0.65</td>
<td>0.76</td>
</tr>
<tr>
<td>Leak liquid</td>
<td>0.74</td>
<td>0.85</td>
</tr>
<tr>
<td>Leak stool/urgency</td>
<td>0.82</td>
<td>0.84</td>
</tr>
<tr>
<td>Leak/change underwear</td>
<td>0.76</td>
<td>0.81</td>
</tr>
<tr>
<td>Alter Lifestyle</td>
<td>0.69</td>
<td>0.70</td>
</tr>
<tr>
<td>Community Survey (N = 2917)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* = proportion of variance explained

In 2006 the authors used PCA to analyse a number of faecal incontinence items which included the 5 RFIS items, the 2 other Wexner items concerning wearing a pad and flatus, items on bowel patterns/movements and a faecal urge item. The results of this analysis indicated that the RFIS items and the pad item loaded highly on Factor/Component 1 (explaining 40% variance) and urge and flatus items loading on Factor/Component 2 (explaining 11% variance) with only the bowel movement item loading on Factor/Component 3 (explaining 10% variance). The RFIS items were selected with regard to their loadings on the primary faecal incontinence factor (defined as soiling and leakage). The flatus item was not included in the RFIS as it had a low loading on this factor, it is not consistent with the definition of faecal incontinence (Abrams et al., 2008) and an analysis of internal consistency reliability of the Wexner indicated the item should be deleted. Item 4 of the RFIS (Does stool leak so that you have to change your underwear?) had higher loadings on the primary faecal incontinence than the pad item from the Wexner, and given the overlap between these items only the RFIS 4 item was selected for inclusion.

This analysis was repeated with the sample of faecal incontinence patients (N = 61) to re-examine the pattern of item loadings. This produced a 2 component solution accounting for 58% of the variance. The primary faecal incontinence factor accounted for 40% of the variance (eigenvalue = 3.17) and the other bowel symptoms factor/component accounted for 19% of the variance (eigenvalue = 1.48) which is very similar pattern to the 2006 analysis. All RFIS items had loadings of 0.55 or above on the primary faecal incontinence factor which confirms the current descriptive system. One minor difference concerned the lifestyle item which has moderate loadings on both factors in the faecal sample vs. the community sample. However, it is noted that the lifestyle item had quite high loadings on the primary faecal incontinence factor in the analyses of the RFIS items.

Responsiveness: Capacity to Detect Change for the RFIS

Change scores were calculated by subtracting the post-treatment score from the pre-treatment score and on this scale a reduction in score reflects an improvement.

Most patients improved as a result of their treatment (Figure A6). There was a significant improvement of 3.11 mean RFIS scores (paired t test, t = 3.89, df 37; p = 0.000) following treatment (refer Table A5).

The Kazis effect size (ES) for the RFIS was -0.66 (95%CI: -2.15 - +0.82), and this is classed as a moderate effect size as it is over 0.5 (Kazis et al, 1989; Cohen, 1988). All three incontinence measures were responsive over time and expressed as the Kazis’ ES the score changes for the Wexner and the SMIS were ES = -0.57 (95%CI: -2.08 – +0.94) and -0.65 (95%CI: -2.39 - +1.09), respectively.
Centre for Health Service Development

Figure A6: RFIS change scores

The relative efficiency statistic was used to compare the RFIS and the St Mark’s against the least sensitive measure, the Wexner (RE = 1.00). The RE = 1.49 for the RFIS and the RE = 1.71 for the St Mark’s indicating that both these instruments were more sensitive to detecting change than the Wexner. It should be noted that for the SMIS analysis 3 items of missing data were imputed using the last value carried forward (Eliot and Hawthorne, 2005) and if data is not imputed the RE = 1.43. The RFIS is also a shorter instrument than the St Mark’s.

The change scores on the RFIS (pre-post) ranged from an improvement of 18 points to a deterioration of 8 points within a 20 point scale (refer Figure A6). The average magnitude of change score for the RFIS was 4.32 scores. This would suggest the RFIS has the capacity to detect both an improvement and deterioration in patient incontinence status.

RFIS change scores were also examined by the patient’s global ratings of improvement. The data suggested that there is a significant change in the patients’ average ratings of improvement associated with an RFIS change score of 2 points (t = 3.26, df 35; p = 0.002). This might suggest that a change in the vicinity of > 2 RUIS scores may be the minimal detectable difference concerning patients’ perceptions of improvement, at the group level.

The standard error of measurement (SEM) for the RFIS was 2.19 (95% CI SEM = 3.39) suggested that a score change of 3-4 points is more likely to be clinically and statistically reliable estimate for the purposes of patient monitoring although it is noted that the SEM reported is for the scale overall and standard errors of measurement vary by score levels (they may be more or less for people scoring at the extremes of the scales) (Spratt, 2009)

Change scores pre- and post-treatment were also anchored against changes in patient behaviour. Pad use was examined across pre-treatment to post-treatment. Four groups were identified (no pads used at pre or post treatment, reduced pad use from pre to post-treatment, stopped using pads after treatment, and no change/increased use of pads). RFIS post scores were examined by these pad groups with RFIS baseline scores treated as a covariate to control for any pre-existing differences between the groups (F = 5.59, df = 1,3,4,33; p<0.01). This analysis suggested that a change in score of >4.00 points on the RFIS was associated with a reduction in the use of incontinence pads (the mean RFIS reduction for those no longer using pads was 8.20 RFIS points and for those who had reduced their use of pads it was 4.78). A change of ≥4 points appears to represent a clinically important change in that this score change was associated with a change in patient behaviour.

A more detailed discussion of these analyses can be found in Sansoni et al. (2011).
Correlations with Other Measures: RFIS

At pre-treatment the RFIS correlation with Wexner Incontinence Score $r = 0.88$ ($p<0.01$); with St Mark’s Incontinence Score $r = 0.85$ ($p<0.01$); and with the pre-treatment Patient Incontinence Severity Rating $r_s = 0.52$ ($p<0.01$). The correlation between RFIS and the pre-treatment clinical rating of severity was significant ($p<0.01$) but somewhat lower $r_s = 0.29$ and correlations of this magnitude were also found for other faecal incontinence instruments. Correlations between RFIS with Wexner Type Specification (impact) items were all significant at the $p<0.01$ level save for sexual relations ($p>0.05$) and church attendance ($p>0.05$). The correlation of RFIS with the Faecal Incontinence Quality of Life (FIQL) coping items was $r = 0.57$ ($p<0.01$). These findings provide evidence of construct validity for this measure.

Sansoni et al. (2006) reported that urinary and faecal incontinence measures had negative but significant correlations with measures of health related quality of life reflecting the burden of disease for this condition (e.g. the higher the incontinence score the lower the physical functioning or health status score). The RFIS correlation with the Physical Function Scale of the SF-36V2 was $r = -0.27$ ($p = 0.037$) at pre-treatment for the faecal sample. It is interesting that this is a higher correlation than was found for urinary incontinence suggesting that the impact of faecal incontinence may be greater although it should be noted that the faecal incontinence sample was somewhat older and with a greater level of co-morbidities than the urinary incontinence sample.

For the faecal sample ($N = 50$) the Physical Component Summary (PCS) score was 43.59 and for the Mental Component Summary (MCS) score it was 43.41 which was lower than the Australian population norm ($t_{pcs} = -4.54$, $t_{mcs} = -3.75$, df 49; $p < 0.01$) (Hawthorne et al., 2007). These results were consistent with findings in the literature showing a small negative association between faecal incontinence and generic measures of HRQOL (Sansoni et al., 2006).

Ceiling and Floor Effects

At pre-treatment scores on the RFIS ranged from 0 – 20 (two people with faecal incontinence received a score of 0 on the RFIS as it does not include an item on flatus) indicating the full range of scores were used. There was 3.3% of the sample on the floor of the instrument (lowest possible score) and 1.6% of the sample at the ceiling of the scale (maximum possible score).

Generally a figure of less than 15% of the sample at the floor or the ceiling is preferred (Streiner and Norman, 2003). It is clear that this instrument does not suffer from significant floor or ceiling effects.

Effect of Treatment

An analysis of RFIS change scores by type of treatment was undertaken (refer Table A4). For this analysis the four cases receiving medicinal treatments were excluded. As the sample size for physiotherapy was small ($N = 9$) and given the treatment provided was reported to be very similar to that provided by Continence Nurse Advisors these groups were combined for further analyses and classed as ‘conservative therapy’. The mean change score for conservative treatment was 2.52 RFIS scores ($t = 2.26$, df 24; $p = 0.033$) and for surgery it was 4.50 RFIS scores ($t = 4.86$, df 11; $p = 0.000$). For this sample size, this difference in improvement by treatment type was not significant – both treatment groups improved ($t = 1.14$; df 35; $p>0.05$).

Suggested Cutpoints for Interpretation: RFIS

It would be advisable to refer people who score 4 or more on this RFIS for further continence assessment. This is more than one standard deviation below the mean for the total clinical sample and represents the 10th percentile of the clinical sample (refer Figure A4 and Table A9). It is more than two standard deviation units above the mean for the community sample. To obtain this score one would need to endorse ‘slightly’ or ‘rarely’ on most faecal incontinence items.

An analysis of the distribution of the clinical sample scores indicated that a score of 6 was the 25th percentile for the clinical sample which might be considered ‘mild’, a score of 10 was the 50th
percentile which might be considered ‘moderate’ and a score of 14 represented the 80th percentile which might be considered ‘severe’.

Other faecal leakage items that are not contained in the RFIS (leak mucus, incomplete bowel empty, faecal urge, leak when asleep/at night, seepage following bowel movement, leak not knowing and pad use) were summed to form an alternate index of the severity of faecal leakage. This index was used to assist in determining cutpoints for the RFIS. Table A5 shows there was a significant difference in RFIS pre-treatment scores for these groups. For the mild/low score group the RFIS mean was 5.50, for the moderate group it was 9.96 and for the more severe/ high scores group the RFIS mean was 13.31.

Clinician ratings of ‘mild’ equated to a RFIS mean score of 8.05 (N = 22), for ‘moderate’ this equated to an RFIS mean score of 9.76 (N = 25) and for severe this equated to a mean score of 12.08 (N = 12). Patient pre-treatment severity ratings indicated that the mean for a rating of ‘mild’ was 7.12 (N = 26), the mean for a rating of moderate was 10.64 (N = 28), and the mean for a rating of severe was 15.14 (N = 7). At post treatment the clinical ratings suggested a score of 3 or less would be considered ‘normal’ as did the patient severity ratings.

Pad use from pre-treatment to post treatment was examined and four groups were identified. Four groups were identified (no pads used at pre or post treatment, reduced pad use from pre to post-treatment, stopped using pads after treatment, and no change/increased use of pads). Figure A7 suggested that a critical cutpoint score on the RFIS was at 7-points.

**Figure A7: RFIS by changes in pad use, pre-post treatment**

![Figure A7: RFIS by changes in pad use, pre-post treatment](image)

Statistics: ANOVA, $F = 5.59$, $df = 1,3,4,33$, $p < 0.01$

Most participants at both baseline and follow up with scores <7-points had faecal incontinence insufficiently severe to warrant the use of incontinence pads. In contrast, most of those with scores ≥7-points reported using pads. Based on a cutpoint of 7 RFIS points, at baseline the relative risk of a participant using an incontinence pad sometimes/often/always was almost three times that of a
person with a RFIS score of ≤6 (RR = 2.79; 95%CI: 1.26 – 5.81). At follow up they were 8-times more likely to be using incontinence pads based on the same cutpoint (RR = 8.41; 95%CI: 2.19 – 32.76).

Given these considerations the following RFIS score ranges are suggested to aide interpretation:

0-3: no faecal incontinence or extremely mild incontinence symptoms
4-6: mild faecal incontinence
7-12: moderate faecal incontinence
13-20: severe faecal incontinence (RFIS scores above 17 could be considered very severe)

Conclusion

The RFIS and the RUIS possess evaluative discrimination by patient assessed incontinence severity but do not discriminate by unrelated health or socio-demographic variables. Similarly, the RFIS and the RUIS appear to be responsive over time to changes in incontinence status. These two findings suggest they have both content and construct validity; i.e. they assess the underlying condition of incontinence and this assessment appears to be independent of possible confounding variables. The data analyses confirmed that the revised incontinence tools have superior psychometric properties to comparable instruments in the clinical setting and this confirms findings from earlier community sample data (Sansonni et al., 2006).

With only 5 items each the RFIS and the RUIS are short and simple to use and score. Continence clinics treating incontinence patients or aged care assessors should find them easy to use both as assessment measures and as an outcome evaluation measures in routine practice. The use of such measures can provide effective feedback to clinicians concerning the effectiveness of their treatments, can facilitate the systematic review and monitoring of patients, and can assist in identifying ways to improve practice. The Short Assessment of Patient Satisfaction Scale is also a useful tool for monitoring patient satisfaction with the outcomes of treatment and further information on this scale is provided later in this manual.

References


Revised Urinary Incontinence Scale (RUIS)

Do you experience and if so how much are you bothered by:

1. Urine leakage related to the feeling of urgency
   - Not at all
   - Slightly
   - Moderately
   - Greatly

2. Urine leakage related to physical activity, coughing or sneezing
   - Not at all
   - Slightly
   - Moderately
   - Greatly

3. Small amounts of urine leakage (drops)
   - Not at all
   - Slightly
   - Moderately
   - Greatly

4. How often do you experience urine leakage?
   - Never
   - Less than once a month
   - A few times a month
   - A few times a week
   - Every day and/or night

5. How much urine do you lose each time?
   - None
   - Drops
   - Small splashes
   - More
Revised Urinary Incontinence Scale (RUIS): Scoring Guide

Do you experience and if so how much are you bothered by:

1. Urine leakage related to the feeling of urgency
   - Not at all: 0
   - Slightly: 1
   - Moderately: 2
   - Greatly: 3

2. Urine leakage related to physical activity, coughing or sneezing
   - Not at all: 0
   - Slightly: 1
   - Moderately: 2
   - Greatly: 3

3. Small amounts of urine leakage (drops)
   - Not at all: 0
   - Slightly: 1
   - Moderately: 2
   - Greatly: 3

4. How often do you experience urine leakage?
   - Never: 0
   - Less than once a month: 1
   - A few times a month: 2
   - A few times a week: 3
   - Every day and/or night: 4

5. How much urine do you lose each time?
   - None: 0
   - Drops: 1
   - Small splashes: 2
   - More: 3

The RUIS total score is calculated by adding a person’s score for each question. Adding the score for the five questions results in a possible score range from 0 – 16.
Revised Urinary Incontinence Scale Scoring Instructions

People respond to the Revised Urinary Incontinence Scale (RUIS) questions by selecting one particular response option from the set of standard response options for each question. These response options can then be scored by using the numbers presented in brackets to the right of each response option. The RUIS total score is then calculated by adding up a person’s score for each question. Adding the score for each of the five questions results in a possible score range of 0 - 16.

In a pre-treatment clinical sample and a community population survey the following descriptive statistics for the RUIS (adults over 18 years) were derived:

Table A8: Clinical and community sample statistics for the RUIS

<table>
<thead>
<tr>
<th></th>
<th>RUIS: Clinical Sample</th>
<th>RUIS: Community Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Male</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>11.07</td>
<td>0.70</td>
</tr>
<tr>
<td>Std Deviation</td>
<td>4.18</td>
<td>1.97</td>
</tr>
<tr>
<td>N</td>
<td>28</td>
<td>1203</td>
</tr>
<tr>
<td>Range</td>
<td>0-16</td>
<td>0-14</td>
</tr>
<tr>
<td><strong>Female</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>10.90</td>
<td>2.47</td>
</tr>
<tr>
<td>Std Deviation</td>
<td>3.33</td>
<td>3.31</td>
</tr>
<tr>
<td>N</td>
<td>167</td>
<td>1712</td>
</tr>
<tr>
<td>Range</td>
<td>3-16</td>
<td>0-16</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>10.92</td>
<td>1.74</td>
</tr>
<tr>
<td>Std Deviation</td>
<td>3.33</td>
<td>2.93</td>
</tr>
<tr>
<td>N</td>
<td>195</td>
<td>2915</td>
</tr>
</tbody>
</table>

Given the data presented above it would be advisable to refer people who score 4 or more on the RUIS for further continence assessment. This is two standard deviations below the mean for the total clinical sample but 1 standard deviation above the average score (1.74: SD = 2.93) found in the community survey sample. To obtain these scores one would need to endorse at least ‘slightly’ or ‘rarely’ on most incontinence items. A score of 4 represents the 10th percentile of the clinical sample.

The RUIS score distributions, clinician and patient severity ratings and other indicators of severity were examined to provide the following RUIS score ranges as an aide to interpretation:

- 0-3: no urinary incontinence or extremely mild (‘slight’) or occasional (‘rare’) incontinence
- 4-8: mild urinary incontinence
- 9-12: moderate urinary incontinence
- 13-16: severe urinary incontinence (scores of 15 – 16 could be considered very severe).

Finally, users should check that each question has a response option selected in order to avoid any missing data.
Revised Faecal Incontinence Scale (RFIS)

1. Do you leak, have accidents or lose control with solid stool?
   - **Never**
   - **Rarely** i.e. less than once in the past four weeks
   - **Sometimes** i.e. less than once a week, but once or more in the past four weeks
   - **Often or usually** i.e. less than once a day but once a week or more
   - **Always** i.e. once or more per day or whenever you have a bowel movement

2. Do you leak, have accidents or lose control with liquid stool?
   - **Never**
   - **Rarely** i.e. less than once in the past four weeks
   - **Sometimes** i.e. less than once a week, but once or more in the past four weeks
   - **Often or usually** i.e. less than once a day but once a week or more
   - **Always** i.e. once or more per day or whenever you have a bowel movement

3. Do you leak stool if you don’t get to the toilet in time?
   - **Never**
   - **Rarely** i.e. less than once in the past four weeks
   - **Sometimes** i.e. less than once a week, but once or more in the past four weeks
   - **Often or usually** i.e. less than once a day but once a week or more
   - **Always** i.e. once or more per day or whenever you have a bowel movement

4. Does stool leak so that you have to change your underwear?
   - **Never**
   - **Rarely** i.e. less than once in the past four weeks
   - **Sometimes** i.e. less than once a week, but once or more in the past four weeks
   - **Often or usually** i.e. less than once a day but once a week or more
   - **Always** i.e. once or more per day or whenever you have a bowel movement

5. Does bowel or stool leakage cause you to alter your lifestyle?
   - **Never**
   - **Rarely** i.e. less than once in the past four weeks
   - **Sometimes** i.e. less than once a week, but once or more in the past four weeks
   - **Often or usually** i.e. less than once a day but once a week or more
   - **Always** i.e. once or more per day or whenever you have a bowel movement
Revised Faecal Incontinence Scale (RFIS) Scoring Guide

1. **Do you leak, have accidents or lose control with solid stool?**
   - **Never** 0
   - **Rarely** i.e. less than once in the past four weeks 1
   - **Sometimes** i.e. less than once a week, but once or more in the past four weeks 2
   - **Often or usually** i.e. less than once a day but once a week or more 3
   - **Always** i.e. once or more per day or whenever you have a bowel movement 4

2. **Do you leak, have accidents or lose control with liquid stool?**
   - **Never** 0
   - **Rarely** i.e. less than once in the past four weeks 1
   - **Sometimes** i.e. less than once a week, but once or more in the past four weeks 2
   - **Often or usually** i.e. less than once a day but once a week or more 3
   - **Always** i.e. once or more per day or whenever you have a bowel movement 4

3. **Do you leak stool if you don’t get to the toilet in time?**
   - **Never** 0
   - **Rarely** i.e. less than once in the past four weeks 1
   - **Sometimes** i.e. less than once a week, but once or more in the past four weeks 2
   - **Often or usually** i.e. less than once a day but once a week or more 3
   - **Always** i.e. once or more per day or whenever you have a bowel movement 4

4. **Does stool leak so that you have to change your underwear?**
   - **Never** 0
   - **Rarely** i.e. less than once in the past four weeks 1
   - **Sometimes** i.e. less than once a week, but once or more in the past four weeks 2
   - **Often or usually** i.e. less than once a day but once a week or more 3
   - **Always** i.e. once or more per day or whenever you have a bowel movement 4

5. **Does bowel or stool leakage cause you to alter your lifestyle?**
   - **Never** 0
   - **Rarely** i.e. less than once in the past four weeks 1
   - **Sometimes** i.e. less than once a week, but once or more in the past four weeks 2
   - **Often or usually** i.e. less than once a day but once a week or more 3
   - **Always** i.e. once or more per day or whenever you have a bowel movement 4

The RFIS total score is calculated by adding up a person’s score for each question. Adding the score for each of the five questions results in a possible score range of 0 – 20.
Revised Faecal Incontinence Scale Scoring Instructions

People respond to the Revised Faecal Incontinence Scale (RFIS) questions by selecting one particular response option from the set of standard response options for each question. These response options can then be scored by using the numbers presented in brackets to the right of each response option. The RFIS total score is then calculated by adding up a person’s score for each question. Adding the score for each of the five questions results in a possible score range of 0 - 20.

In a pre-treatment clinical sample and a community population survey the following descriptive statistics for the RFIS (adults over 18 years) were derived:

**Table A9: Clinical and community sample statistics for the RFIS**

<table>
<thead>
<tr>
<th></th>
<th>RFIS: Clinical Sample</th>
<th>RFIS: Community Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Male</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>9.10</td>
<td>0.25</td>
</tr>
<tr>
<td>Std Deviation</td>
<td>3.6</td>
<td>1.04</td>
</tr>
<tr>
<td>N</td>
<td>10</td>
<td>1201</td>
</tr>
<tr>
<td>Range</td>
<td>3-16</td>
<td>0-12</td>
</tr>
<tr>
<td><strong>Female</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>9.76</td>
<td>0.43</td>
</tr>
<tr>
<td>Std Deviation</td>
<td>4.86</td>
<td>1.56</td>
</tr>
<tr>
<td>N</td>
<td>51</td>
<td>1714</td>
</tr>
<tr>
<td>Range</td>
<td>0-20</td>
<td>0-18</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>9.66</td>
<td>0.36</td>
</tr>
<tr>
<td>Std Deviation</td>
<td>4.66</td>
<td>1.37</td>
</tr>
<tr>
<td>N</td>
<td>61</td>
<td>2915</td>
</tr>
<tr>
<td>Range</td>
<td>0-20</td>
<td>0-18</td>
</tr>
</tbody>
</table>

It would be advisable to refer people who score 4 or more on this scale for further continence assessment. This is more than one standard deviation below the mean for the total clinical sample and represents the 10th percentile of the clinical sample. It is more than two standard deviation units above the mean for the community sample. To obtain this score one would need to endorse ‘slightly’ or ‘rarely’ on most faecal incontinence items.

The RFIS score distributions, clinician and patient severity ratings and other indicators of severity were examined to provide the following RFIS score ranges as an aide to interpretation:

0-3: no faecal incontinence or extremely mild incontinence symptoms
4-8: mild faecal incontinence
9-12: moderate faecal incontinence
13-20: severe faecal incontinence (RFIS scores above 17 could be considered very severe)

The RFIS is a measure of faecal (vs. anal) incontinence and does not contain an item on flatus. If flatus is an issue the following items from the Wexner Scale (Jorge and Wexner, 1993) or the Colorectal Anal Distress Scale (CRADI–8; Barber et. al., 2005) could be used as a separate exercise:

**Wexner: Do you leak, have accidents or lose control with gas (flatus or wind)?**

- **Never**
- **Rarely** i.e. less than once in the past four weeks
- **Sometimes** i.e. less than once a week, but once or more in the past four weeks
- **Often or usually** i.e. less than once a day but once a week or more
- **Always** i.e. once or more per day or whenever you have a bowel movement

0
1
2
3
4
For patients that identify themselves as experiencing flatus often or usually (3) or always (4) referral for continence assessment could also be considered.

**CRADI-8: Do you lose gas from your rectum beyond your control? (CRADI- 8)**

- **Never**
- **Rarely** *i.e. less than once in the past four weeks* 1
- **Sometimes** *i.e. less than once a week, but once or more in the past four weeks* 2
- **Often** *i.e. less than once a day but once a week or more* 3
- **Usually** *i.e. once per day* 4
- **Always** *i.e. several times per day* 5

**CRADI-8: If so, how much does this bother you?**

- **Not at all**
- **Somewhat**
- **Moderately** 2
- **Quite a bit** 3

For the CRADI-8 items it is suggested that the scores from the 2 items are summed. If the patient scores 6 or above it is recommended they are referred for further continence assessment. The Wexner item is simpler to use but the CRADI-8 items are slightly more sensitive to flatus status.

Finally, users should check that each question has a response option selected in order to avoid any missing data.
Introduction and Background: Short Assessment of Patient Satisfaction (SAPS)

Background
Assessing patient satisfaction with health care services serves three purposes: it enables monitoring of the patient and the identification of areas of health care that are of concern to the patient; it can assist with involving the patient in medical decision-making about possible future treatments; and it can be used to audit health care for monitoring health care standards and the legitimizing of health care policy (Hawthorne, 2006).

In the field of incontinence, only one condition-specific patient satisfaction instrument had been published at commencement of the Australian National Continence Management Strategy (NCMS), the Genito-Urinary Treatment Satisfaction Scale (Hawthorne and Harmer, 2000). One of the recommendations of the first NCMS report was that there should be a separate report on patient satisfaction measures that could be considered for inclusion in the national suite of outcome measures (Thomas et al., 2006). The subsequent review of patient satisfaction measures reported that no one generic instrument was to be preferred and recommended the development of a short generic instrument (Hawthorne, 2006). Following development of the draft SAPS (Hawthorne et al., 2006) review by international colleagues suggested changes to the wording of some items and the final SAPS is presented in this manual.

The SAPS descriptive system
The SAPS is a participant reported outcome measure (PRO) which assesses overall patient satisfaction. Based on Donabedian’s (1988, 1996) definition that patient satisfaction represents the patient’s judgement on care quality, the SAPS comprises 7 items each of which is representative of the key domains of patient satisfaction that are widely agreed to contribute to the overall concept of patient satisfaction: access (SAPS item #6), information (#2), relationship with the clinician (#5), participation in decision-making (#4), technical skill (#3), effectiveness of treatment (#1) and a general satisfaction domain (#7) (Hawthorne, 2006).

Each SAPS item response scale consists of five possible response levels. To avoid response bias, these response scales vary; there are three positive items (SAPS #2, #4, #6) and four negative items (#1, #3, #5, #7). A copy of the SAPS descriptive system can be found at Attachment 1.

Psychometric properties of the SAPS
The measurement properties of the SAPS have been examined in a clinical sample of patients with faecal and urinary incontinence (N = 136). The validation was based on three extensive tests: an examination of the data at the item and instrument levels, an assessment of the internal structure of the SAPS using both traditional and Rasch analyses, and tests of its discriminatory function by known groups selected to demonstrate convergent and divergent scores.

Descriptive statistics showed that scores were well distributed with just 13% of cases reporting ceiling (compared to ~20% among other leading patient satisfaction instruments (4)) scores and no cases with floor scores. Importantly, response bias was present in just 2% of cases. The overall mean score was 22 points (standard deviation 5-points) – although those using the SAPS should take note that these scores should not be used as norms because participants were not a random sample of those with incontinence. The internal structure of the SAPS showed that it was unidimensional on the principal component analysis; the SAPS explained 53% of the overall variance. The Rasch analysis supported this; no items were reported to be misfitting, no items exhibited disordered thresholds, and there was no differential item functioning (DIF) by gender or age group (two items did report DIF by continence status (faecal/urinary), but this may have been due to non-SAPS situational conditions). The discriminatory function of the SAPS appears excellent. The SAPS discriminated by current health status, clinician and patient estimates of current incontinence severity, treatment type and patient reported change in incontinence severity over time. Finally, the reliability of the SAPS was Cronbach’s alpha α = 0.85, suggesting it is a reliable measure.
Administrating the SAPS

The SAPS is copyright to the University of Wollongong with a license to the Commonwealth of Australia and the University of Melbourne. This copyright is not aimed at restricting use of the SAPS, but is to ensure that SAPS users report scores that are comparable across different projects. The following website provides information on free registration for using the SAPS: www.psychiatry.unimelb.edu.au/centres-units/cpro/index.html

The SAPS was designed for administration after completion of treatment. Although the SAPS can be administered immediately after treatment this is only appropriate where the SAPS is used as a discussion stimulus for an open and frank discussion between the patient and treating clinician.

When the SAPS is used as an outcome measure, it is recommended that between 2-weeks and 3-months elapses post-treatment prior to administration. The reason for this is to enable the patient to make a considered judgement regarding satisfaction with their health care (where it is administered immediately after treatment, the patient may be in a dependent relationship with the clinician, the outcomes from treatment may still be uncertain, and the patient may be influenced by extraneous events (e.g. difficulty with public transport or parking).

Administration of the SAPS should not be undertaken by the treating clinician, but by an independent third party, such as a practice nurse. Where possible, administration should be separate from a clinic visit (during such visits patients may be upset when waiting to be seen or upset by the results of their visit). Administration can be either by letter or over the telephone.

Patients should be provided with a copy of the SAPS which they mark, without instruction from the clinician or interviewer. Completion of the SAPS takes about one minute.

Scoring the SAPS

There are three steps in scoring the SAPS:

1. Prior to scoring, the negative items need to be reversed. These are items #1, #3, #5, and #7. After reversal check that all items are scored 0, 1, 2, 3, 4, where a value of 0 always represents the worst possible outcome (the patient expressing extreme dissatisfaction) and 4 the best possible outcome (extreme satisfaction). All other values are illegal and indicate that there has been a scoring error. Any other values must be checked and corrected.

2. Check for any missing data on any SAPS items. If data are missing, try to find out why the data are missing. If possible, go back to the patient and ask them to complete the missing datum item. Where this is not possible and it is reasonable to assume that any missing data are random (e.g. a person missed answering a question) then missing data can be imputed.
   - Imputation of missing data is preferably carried out using horizontal mean imputation. This is where the missing datum is given the average value of present data for the case. For example, suppose Case A provided the following responses (after reversal): 3, 4, 2, 3, -, 4, 4. Item #5 is missing and would be assigned the rounded up average score across the other items (3, 4, 2, 3, 4, 4 = 20/6 = 3.33 = 3).
   - Missing data imputation is recommended where up to 2 items are missing from the SAPS. Where three or more items are missing, the case should be dropped for data analysis (listwise deletion).
   - The computerised scoring algorithm for the SAPS includes a section that automatically calculates missing data.

3. Once items have reversed, checked and missing values imputed, then the SAPS can be scored by summing item scores e.g. for Case A above, after imputation, the score would be 3+4+2+3+3+4+4 = 23.
   - Although the SAPS is easy to score by hand, this is much slower and more likely to contain errors when compared with available computer-scoring algorithms; these are compatible with Excel, Access and SPSS and available from the Melbourne University Mental Health Evaluation Unit website at: http://www.psychiatry.unimelb.edu.au/centres-units/cpro/index.html.
Interpretation of SAPS scores

SAPS scores can be presented either as continuous scores or as classification scores.

1. Continuous scores. This is the summed score for individuals and the mean score for groups of individuals. When reporting group scores, the average or mean score, the number in the group and the standard deviation should be reported. Where inferential statistical tests using parametric procedures are undertaken, where the scores are statistically significantly skewed they should be suitably transformed prior to data analysis. If the data are skewed and cannot be suitably transformed, then median scores and interquartile ranges should be reported and non-parametric statistical tests used. In the validation sample (N = 136 cases) the overall mean score was 22 points (standard deviation 5-points) – although these scores should not be used as norms because participants were not a random sample of those with incontinence.

2. Categorical scores. The literature on patient satisfaction shows that between 70-90% of patients are satisfied with their health care. This should be kept in mind when interpreting SAPS scores, and where continuous scores are insufficient for interpreting SAPS scores, categorical scores can be used instead. In general, SAPS scores can be interpreted as follows:

- **0 to 10 = Very dissatisfied.** To obtain a score in this range, a person must have indicated that they are dissatisfied or very dissatisfied on four or more items (i.e. over half of all items). Any patient obtaining scores in this range is indicating that their health care has failed them badly and that they are in need of urgent help.

- **11 to 18 = Dissatisfied.** To obtain a score in this range, a person must have indicated that they are dissatisfied or very dissatisfied on at least two items (i.e. two aspects of their health care), or that they have refused to endorse being very satisfied on any item. Patients obtaining scores in this range are indicating health care failure in several areas of their health care and are in need of help in these areas.

- **19 to 26 = Satisfied.** To obtain a score in this range, a person must have indicated that they are very satisfied or satisfied on over half SAPS items (4/7). These patients should be asked about those areas of health care they found unsatisfactory and efforts made to improve such areas.

- **27 to 28 = Very satisfied.** To obtain a score in this range, a person must have indicated they are very satisfied or satisfied on all seven SAPS items. These patients are indicating that all aspects of their health care have met or exceeded their expectations.

References


Attachment 1: Copy of the SAPS

Instructions: After reading each question, circle the answer that best describes you. The order of the answers varies between the questions, so take a moment to read each question carefully. We know that sometimes answers may not describe you exactly, so please pick the answer that most closely describes you. When you have finished, please check that you have answered all questions.

1. How satisfied are you with the effect of your {treatment/care}?  
   - Very satisfied 0  
   - Satisfied. 1  
   - Neither satisfied nor dissatisfied. 2  
   - Dissatisfied. 3  
   - Very dissatisfied. 4

2. How satisfied are you with the explanations the {doctor/other health professional} has given you about the results of your {treatment/care}?  
   - Very dissatisfied. 0  
   - Dissatisfied. 1  
   - Neither satisfied nor dissatisfied. 2  
   - Satisfied. 3  
   - Very satisfied. 4

3. The {doctor/other health professional} was very careful to check everything when examining you.  
   - Strongly agree 0  
   - Agree 1  
   - Not sure 2  
   - Disagree 3  
   - Strongly disagree 4

4. How satisfied were you with the choices you had in decisions affecting your health care?  
   - Very dissatisfied 0  
   - Dissatisfied. 1  
   - Neither satisfied nor dissatisfied. 2  
   - Satisfied. 3  
   - Very satisfied. 4

5. How much of the time did you feel respected by the {doctor/other health professional}?  
   - All of the time 0  
   - Most of the time 1  
   - About half the time 2  
   - Some of the time 3  
   - None of the time 4

6. The time you had with the {doctor/other health professional} was too short.  
   - Strongly agree 0  
   - Agree 1  
   - Not sure 2  
   - Disagree 3  
   - Strongly disagree 4

7. Are you satisfied with the care you received in the {hospital/clinic}?  
   - Very satisfied 0  
   - Satisfied. 1  
   - Neither satisfied nor dissatisfied. 2  
   - Dissatisfied. 3  
   - Very dissatisfied. 4

Scoring  
1. Reverse the scores for items #1, #3, #5, #7  
2. Sum all scores. The score range is from 0 (extremely dissatisfied) to 28 (extremely satisfied)
Attachment B: Brochures for the Incontinence and Patient Satisfaction Tools
Tools for Assessing and Monitoring Urinary Incontinence: The Revised Urinary Incontinence Scale (RUIS)

Background

The RUIS is a short, reliable and valid five item scale that can be used to assess urinary incontinence and to monitor patient outcomes following treatment. It was originally developed by selecting the best performing urinary incontinence items (selected from standardised measures such as the Urogenital Distress Inventory 6 and the Incontinence Severity Index) which were included in a large community survey of 2,915 Australians in 2006. The RUIS has recently been validated in clinical settings (Sansoni et al., 2006; 2011) with support from the Australian Government Department of Health and Ageing. These studies have shown that the RUIS is a valid and reliable measure of urinary incontinence. Internal consistency reliability is Cronbach’s alpha $\alpha = 0.73$ (urinary incontinence sample, N = 195), alpha = 0.84 (all incontinence patients N = 254) and alpha = 0.91 (community sample N = 2,915). It has high and statistically significant correlations with other measures of urinary incontinence and other clinical indicators of incontinence severity and has better measurement properties than comparable measures (Sansoni et al., 2011). With only 5 items the RUIS is short and simple to use and score. Most patients will only take a minute to complete it.

Why Use a Standardised Measure of Urinary Incontinence?

This means you are using the same yardstick to assess all patients. This combined with your clinical judgement will help to inform the best treatment for the patient. The use of such measures can also provide effective feedback to clinicians concerning the effectiveness of their treatments, can facilitate the systematic review and monitoring of patients, and can assist in identifying ways to improve practice. It is also useful information to demonstrate the effectiveness of your service.

Continence clinics treating incontinence patients or aged care assessors should find it easy to use it both as assessment measure and as an outcome evaluation measure for routine practice.

The RUIS contains the following items:

Do you experience and if so how much are you bothered by:

1. Urine leakage related to the feeling of urgency
   - Not at all 0
   - Slightly 1
   - Moderately 2
   - Greatly 3

2. Urine leakage related to physical activity, coughing or sneezing
   - Not at all 0
   - Slightly 1
   - Moderately 2
   - Greatly 3

3. Small amounts of urine leakage (drops)
   - Not at all 0
   - Slightly 1
   - Moderately 2
   - Greatly 3

4. How often do you experience urine leakage?
   - Never 0
   - Less than once a month 1
   - A few times a month 2
   - A few times a week 3
   - Every day and/or night 4

5. How much urine do you lose each time?
   - None 0
   - Drops 1
   - Small splashes 2
   - More 3

Scoring

The RUIS total score is then calculated by adding up a person’s score for each question. Adding the score for each of the five questions results in a possible score range of 0 - 16.
Interpreting Scores

The average score for patients receiving treatment for urinary incontinence is 10.92 (N = 195). The mean RUIS scores for female urinary incontinence patients was 10.90 and for males it was 11.07. By contrast the average RUIS score in a large community survey was 1.74 (N = 2,915); for females the mean was 2.47 and for males it was 0.70.

A score of less than 4 indicates that the patient has no urinary incontinence or very mild incontinence symptoms.

Patients with a score of 4 in screening surveys may require further assessment by a continence practitioner. To obtain these scores one would need to endorse ‘slightly’ or ‘rarely’ on most incontinence items.

Based on the distribution of scores in the clinical sample and comparisons with other indicators such as pad use, a score of 4-8 is considered mild, a score of 9-12 is considered moderate and a score of 13 or above is considered severe.

The cut points are supported by clinician and patient ratings of incontinence severity. The clinician pre-treatment ratings indicated that a RUIS score of 9 was considered ‘mild’, a score of 11 was considered ‘moderate’ and a score above 12 was classified as severe which provides some clinical confirmation for the suggested cut points. At post-treatment a score of 3 or less was classified as ‘normal’ by clinicians and patients.

Sensitivity to Detecting Improvement and Change in Patient Incontinence

The RUIS is sensitive to change as a result of treatment and is equally or more sensitive than comparable measures. In the clinical study (Sansoni et al., 2011) it was shown that there was a significant improvement (p<0.01) of an average of 4 RUIS scores following treatment across all types of treatment (continence advising, physiotherapy and surgery). You can easily demonstrate that you have made a difference to patient outcomes. You can also easily identify those patients that have not improved or are deteriorating and this can be very useful for patient review and referral.

Further Information

The above is a very brief summary concerning the RUIS. Further Information can be found at www.bladderbowel.gov.au where copies of the Validation Report and the Technical Manual can be found. This instrument is copyright to the University of Wollongong with a license to the Commonwealth of Australia and the University of Melbourne. This instrument is available free of charge but permission for use should be sought from the authors by contacting Associate Professor Jan Sansoni at janet.sansoni@grapevine.com.au.

Relevant Reports


Study funded by the Australian Government Department of Health and Ageing as part of the National Continence Management Strategy
Tools for Assessing and Monitoring Faecal Incontinence: 
The Revised Faecal Incontinence Scale (RFIS)

Background
The RFIS is a short, reliable and valid five item scale that can be used to assess faecal incontinence and to monitor patient outcomes following treatment. It was developed by selecting the best performing faecal incontinence items (selected from standardised measures such as the Wexner Faecal Continence Grading Scale and faecal incontinence items developed by specialist clinicians) which were tested in a large community survey of 2,915 Australians adults in 2006. The RFIS has recently been validated in clinical settings (Sansoni et al., 2006; 2011) with the support of the Australian Government Department of Health and Ageing. This has shown that the RFIS is a valid and reliable measure of faecal incontinence. The internal consistency reliability is Cronbach’s alpha $\alpha = 0.78$ (faecal incontinence sample; $N = 61$), $\alpha = 0.91$ (all incontinence patients; $N = 254$) and $\alpha = 0.89$ (adults, community sample $N = 2,915$). It has high and statistically significant correlations with other measures of faecal incontinence and other indicators of incontinence severity and had better measurement properties than comparable scales (Sansoni et al., 2011).

The RFIS has 5 items and is short and simple to use and score. Most patients will only take a minute to complete it.

Why Use a Standardised Measure of Faecal Incontinence?
This means you are using the same yardstick to assess all patients. This combined with your clinical judgement will help to inform the best treatment for the patient. The use of such measures can also provide effective feedback to clinicians concerning the effectiveness of their treatments, can facilitate the systematic review and monitoring of patients, and can assist in identifying ways to improve practice. It is also useful information to demonstrate the effectiveness of your service. Continence clinics treating incontinence patients or aged care assessors should find it easy to use both as assessment measure and as an outcome evaluation measures for routine practice. The RFIS contains the following items:

1. Do you leak, have accidents or lose control with solid stool?

   Never
   Rarely i.e. less than once in the past four weeks
   Sometimes i.e. less than once a week, but once or more in the past four weeks
   Often or usually i.e. less than once a day but once a week or more
   Always i.e. once or more per day or whenever you have a bowel movement

   Scoring
   The RFIS total score is then calculated by adding up a person’s score for each question. Adding the score for each of the five questions results in a possible score range of 0 - 20.

2. Do you leak, have accidents or lose control with liquid stool?

   Never
   Rarely i.e. less than once in the past four weeks
   Sometimes i.e. less than once a week, but once or more in the past four weeks
   Often or usually i.e. less than once a day but once a week or more
   Always i.e. once or more per day or whenever you have a bowel movement

3. Do you leak stool if you don’t get to the toilet in time?

   Never
   Rarely i.e. less than once in the past four weeks
   Sometimes i.e. less than once a week, but once or more in the past four weeks
   Often or usually i.e. less than once a day but once a week or more
   Always i.e. once or more per day or whenever you have a bowel movement

4. Does stool leak so that you have to change your underwear?

   Never
   Rarely i.e. less than once in the past four weeks
   Sometimes i.e. less than once a week, but once or more in the past four weeks
   Often or usually i.e. less than once a day but once a week or more
   Always i.e. once or more per day or whenever you have a bowel movement

5. Does bowel or stool leakage cause you to alter your lifestyle?

   Never
   Rarely i.e. less than once in the past four weeks
   Sometimes i.e. less than once a week, but once or more in the past four weeks
   Often or usually i.e. less than once a day but once a week or more
   Always i.e. once or more per day or whenever you have a bowel movement

Scoring
The RFIS total score is then calculated by adding up a person’s score for each question. Adding the score for each of the five questions results in a possible score range of 0 - 20.
Interpreting Scores

The average score for patients receiving treatment for faecal incontinence is 9.66. A score of less than 4 indicates that the patient has no faecal incontinence or very mild symptoms. Patients with a score of 4 in screening surveys require further assessment by a continence practitioner. Scores from 4-6 are considered mild, a score of 7-12 is moderate and a score of 13 or above is considered severe and these suggested cut points are supported by comparison with other clinical indicators and clinician and patient ratings of incontinence severity.

Flatus: As it is a measure of faecal (vs. anal) incontinence the RFIS does not contain an item on flatus. If flatus is an issue of concern the following items from the Wexner Continence Scale (Jorge and Wexner, 1993) or the Colorectal Anal Distress Scale (CRADI-8; Barber et al., 2005) could be used as a separate exercise:

**Wexner**

Do you leak, have accidents or lose control with gas (flatus or wind)?

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never</td>
<td>0</td>
</tr>
<tr>
<td>Rarely i.e. less than once in the past four weeks</td>
<td>1</td>
</tr>
<tr>
<td>Sometimes i.e. less than once a week, but once or more in the past four weeks</td>
<td>2</td>
</tr>
<tr>
<td>Often or usually i.e. less than once a day but once a week or more</td>
<td>3</td>
</tr>
<tr>
<td>Always i.e. once or more per day or whenever you have a bowel movement</td>
<td>4</td>
</tr>
</tbody>
</table>

**CRADI-8**

Do you lose gas from your rectum beyond your control?

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never</td>
<td>0</td>
</tr>
<tr>
<td>Rarely i.e. less than once in the past four weeks</td>
<td>1</td>
</tr>
<tr>
<td>Sometimes i.e. less than once a week, but once or more in the past four weeks</td>
<td>2</td>
</tr>
<tr>
<td>Often i.e. less than once a day but once a week or more</td>
<td>3</td>
</tr>
<tr>
<td>Usually i.e. once per day</td>
<td>4</td>
</tr>
<tr>
<td>Always i.e. several times per day</td>
<td>5</td>
</tr>
</tbody>
</table>

If so, how much does this bother you?

<table>
<thead>
<tr>
<th>Intensity</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at all</td>
<td>0</td>
</tr>
<tr>
<td>Somewhat</td>
<td>1</td>
</tr>
<tr>
<td>Moderately</td>
<td>2</td>
</tr>
<tr>
<td>Quite a bit</td>
<td>3</td>
</tr>
</tbody>
</table>

For patients that identify themselves as experiencing flatus often or usually (3) or always (4) on the Wexner item referral for continence assessment could be considered. For the CRADI-8 items it is suggested that the scores from the 2 items are summed. If the patient scores 6 or above it is recommended they are referred for further continence assessment. The Wexner item is simpler to use but the CRADI-8 items are slightly more sensitive to flatus status.

Sensitivity to Detecting Improvement and Change in Patient Incontinence

The RFIS is sensitive to change as a result of treatment. In Sansoni et al., (2011) it was shown that there was an average improvement of 3 RFIS scores following treatment and that all types of treatment (conservative treatment and surgery) were effective. It was equally or more sensitive to change than other faecal incontinence scales. You can demonstrate that you have made a difference to patient outcomes. You can also easily identify those patients that have not improved or have deteriorated and this can be very useful for patient review and referral.

Further Information

The above is a very brief summary concerning the RFIS. Further Information can be found at [www.bladderbowel.gov.au](http://www.bladderbowel.gov.au) where copies of the Validation Report and the Technical Manual for the RFIS can be found. This instrument is copyright to the University of Wollongong with a license to the Commonwealth of Australia and the University of Melbourne. This instrument is available free of charge but permission for use should be sought from Associate Professor Jan Sansoni at janet.sansoni@grapevine.com.au.

Relevant Reports


Study funded by the Australian Government Department of Health and Ageing as part of the National Continence Management Strategy
The Short Assessment of Patient Satisfaction (SAPS)

Background

The Short Assessment of Patient Satisfaction (SAPS) is a short, reliable and valid seven item scale that can be used to assess patient satisfaction with their treatment. In 2006 (Hawthorne 2006, Hawthorne et al., 2006) a study was undertaken to examine a number of the leading patient satisfaction measures with urinary incontinence patients. The items from all these patient satisfaction scales were pooled and the SAPS was developed by selecting the items with best measurement properties and the most comprehensive coverage of the domains of patient satisfaction. The SAPS consists seven items assessing the core domains of patient satisfaction which include treatment satisfaction, explanation of treatment results, clinician care, participation in medical decision-making, respect by the clinician, time with the clinician, and satisfaction with hospital/clinic care. Responses scales are 5-point scales (see below).

The SAPS has been validated in clinical settings (Hawthorne et al., 2006; Sansoni et al., 2011) with support from the Australian Government Department of Health and Ageing. These studies have shown that the SAPS is a valid and reliable measure of patient satisfaction. Reliability is Cronbach’s alpha $\alpha = 0.85$; it correlates highly with other measures of patient satisfaction, and correlates well with other indicators of treatment outcomes.

The SAPS is a generic measure of patient satisfaction. Although it was developed and validated in continence settings it can be used in any service settings with any treatment group. This means patient satisfaction scores in different treatment settings can be compared. With only 7 items the SAPS is short and simple to use and score. Most patients will only take a minute to complete it.

Why Use a Standardised Measure of Patient Satisfaction?

This means you are using the same yardstick to assess all patients. The use of such measures can provide effective feedback to clinicians concerning the patient's view of the effectiveness of their treatments, and can assist in identifying ways to improve practice and to address patient concerns. It is also useful information to demonstrate the effectiveness of your service. Continence clinics treating incontinence patients should find it easy to use as an outcome evaluation measure in routine practice.

Instructions: After reading each question, circle the answer that best describes you. The order of the answers varies between the questions, so take a moment to read each question carefully.

We know that sometimes answers may not describe you exactly, so please pick the answer that most closely describes you.

When you have finished, please check that you have answered all questions.

1. How satisfied are you with the effect of your {treatment/care}?
   - Very satisfied: 0
   - Satisfied: 1
   - Neither satisfied nor dissatisfied: 2
   - Dissatisfied: 3
   - Very dissatisfied: 4

2. How satisfied are you with the explanations the {doctor/other health professional} has given you about the results of your {treatment/care}?
   - Very dissatisfied: 0
   - Dissatisfied: 1
   - Neither satisfied nor dissatisfied: 2
   - Satisfied: 3
   - Very satisfied: 4

3. The {doctor/other health professional} was very careful to check everything when examining you.
   - Strongly agree: 0
   - Agree: 1
   - Not sure: 2
   - Disagree: 3
   - Strongly disagree: 4

4. How satisfied were you with the choices you had in decisions affecting your health care?
   - Very dissatisfied: 0
   - Dissatisfied: 1
   - Neither satisfied nor dissatisfied: 2
   - Satisfied: 3
   - Very satisfied: 4
5. How much of the time did you feel respected by the {doctor/other health professional}?

- All of the time                      0
- Most of the time                    1
- About half the time                 2
- Some of the time                    3
- None of the time                    4

6. The time you had with the {doctor/other health professional} was too short.

- Strongly agree                      0
- Agree                               1
- Not sure                            2
- Disagree                            3
- Strongly disagree                   4

7. Are you satisfied with the care you received in the {hospital/clinic}?

- Very satisfied                      0
- Satisfied                           1
- Neither satisfied nor dissatisfied  2
- Dissatisfied                        3
- Very dissatisfied                   4

Scoring
1. Reverse the scores for items #1, #3, #5, #7
2. Sum all scores. The score range is from 0 (extremely dissatisfied) to 28 (extremely satisfied)

Interpreting Scores
The literature on patient satisfaction shows that between 70-90% of patients are satisfied with their health care. This should be kept in mind when interpreting SAPS scores. In general, SAPS scores can be interpreted as follows:

- 0 to 10 = Very dissatisfied. To obtain a score in this range, a person must have indicated that they are dissatisfied or very dissatisfied on four or more items (i.e. over half of all items). Any patient obtaining scores in this range is indicating that their health care has failed them badly and that they are in need of urgent help.

- 11 to 18 = Dissatisfied. To obtain a score in this range, a person must have indicated that they are dissatisfied or very dissatisfied on at least two items (i.e. two aspects of their health care), or that they have refused to endorse being very satisfied on any item. Patients obtaining scores in this range are indicating health care failure in several areas of their health care and are in need of help in these areas.

- 19 to 26 = Satisfied. To obtain a score in this range, a person must have indicated that they are very satisfied or satisfied on over half SAPS items (4/7). These patients should be asked about those areas of health care they found unsatisfactory and efforts made to improve such areas.

- 27 to 28 = Very satisfied. To obtain a score in this range, a person must have indicated they are very satisfied or satisfied on all seven SAPS items. These patients are indicating that all aspects of their health care have met or exceeded their expectations.

In a recent study (Sansoni et al., 2011) the average score for all patients receiving incontinence treatment (N = 139) was 21.96 (SD 4.85); for females it was 21.75 and for males it was 23.09.

Further Information
Further Information can be found at www.bladderbowel.gov.au or from the Mental Health Evaluation Unit (MHEU), Department of Psychiatry, University of Melbourne at www.psychiatry.unimelb.edu.au/centres-units/cpro/index.html. These websites have downloadable copies of the Patient Administration Form, the Registration Form and the Validation Report. The SAPS is available free of charge but permission for use should be sought from the MHEU at the web address above. Additional information can also be obtained from Associate Professor Graeme Hawthorne at graemeeh@unimelb.edu.au.

Relevant Reports
