UNIVERSITY OF BIRMINGHAM

Accuracy of bladder ultrasound in the diagnosis of detrusor overactivity (BUS study)

A Thesis submitted to the Faculty of Medicine and Dentistry of the University of Birmingham for the degree of Doctor of Medicine
by

Suneetha Rachaneni
7/18/2015

Academic department of Obstetrics and Gynaecology
School of Clinical and Experimental Medicine
College of Medical and Dental Sciences
University of Birmingham

Contents

ACCURACY OF BLADDER ULTRASOUND IN THE DIAGNOSIS OF DETRUSOR OVERACTIVITY (BUS STUDY) **1**

LIST OF FIGURES **6**

SYNOPSIS **9**

DEDICATION **10**

ACKNOWLEDGEMENTS **11**

EXECUTIVE ABSTRACT **12**

OBJECTIVES 12

METHODS 12

RESULTS 13

CONCLUSION 15

LIST OF ABBREVIATIONS **16**

PUBLICATIONS FROM THIS THESIS **18**

CHAPTER 1: INTRODUCTION **19**

1.1 BACKGROUND 19

1.2 OVERACTIVE BLADDER (OAB) 19

1.3 MIXED URINARY INCONTINENCE 20

1.4 UNDERLYING PATHOLOGY 21

1.5 EVALUATION OF SYMPTOMS OF OAB 23

1.6 URODYNAMIC EVALUATION OF OAB 26

1.7 BLADDER WALL THICKNESS 26

1.71 Technique and the route of scanning *27*

1.72 Variations in the technique of BWT *29*

1.73 Route of BWT scanning *29*

1.74 Challenges in BWT measurement *29*

1.75 Bladder wall thickening in LUTS-Possible clinical applications *30*

1.8 Acceptability of diagnostic tests-BWT and Urodynamics 32

1.81 STAI scores *33*

1.9 Rationale for the study 33

1.91 Aims and Objectives of the study 34

CHAPTER 2: UPDATE OF THE SYSTEMATIC REVIEW ON THE ACCURACY OF BWT AS A BIOMARKER IN THE DIAGNOSIS OF DETRUSOR OVERACTIVITY 36

Introduction 36

2.1Evidence for the Accuracy of Bladder Ultrasound as a test for diagnosing DO 36

2.2 Methods 36

2.3 Search strategy 37

2.31 Identification of studies and data extraction 37

2.32 Methodological quality assessment 38

2.33 Data Synthesis 38

2.4 Results 39

2.5 Discussion 45

2.51Main findings *45*

2.52 Strengths & weaknesses *46*

2.53 Interpretation of findings and conclusion *47*

CHAPTER 3 55

ACCURACY OF BLADDER WALL THICKNESS ON TRANSVAGINAL ULTRASOUND IN DIAGNOSING DETRUSOR OVERACTIVITY: MATERIALS AND METHODS 55

Introduction 55

3.1 Study design 56

3.2 Eligibility 56

3.3 Sample size 57

3.4 Recruitment of participants and consent 57

3.5 Setting of tests 58

3.6 The index test: Bladder ultrasound 59

3.7 The Reference Standard-Urodynamics 60

3.71 Urodynamics in Overactive bladder (OAB) 62
3.8 Statistical methodology for diagnostic accuracy of BWT 63
3.81Primary analysis *63*
3.811Diagnostic accuracy of BWT scan *63*
3.82 Secondary analysis *64*
3.83 Subgroup and other analysis *64*
3.84 Sensitivity analysis *65*
3.85 Missing outcome data *65*
CHAPTER 4: ACCURACY OF TRANSVAGINAL BLADDER WALL THICKNESS IN DIAGNOSING DETRUSOR OVERACTIVITY: RESULTS 67
Introduction 67
4.1 Index test - bladder wall thickness via ultrasound 67
4.2 Reference standard – Urodynamics 68
4.3 Data analysis 68
4.4 Results 68
4.41 Recruitment of participants in the BUS study *68*
4.42 Demographics of study population: *70*
4.43 Index Test and the Reference Standard *72*
4.431 Bladder wall thickness *73*
4.432 Urodynamic findings *73*
4.7 Estimates of test accuracy 76
4.8 Discussion 80
4.81 Summary of main findings *80*
4.82 Strengths and limitations *80*
4.83 Interpretation of findings *85*
4.9 Conclusion 86
CHAPTER 5: INTEROBSERVER VARIABILITY IN THE SONOGRAPHIC ASSESSMENT OF BLADDER WALL THICKNESS 87
Introduction 87

5.1 OBJECTIVES 87
5.2 METHODS 88
5.3 ANALYSES 90
5.4 RESULTS 91
5.41 SUB-STUDY A - INTRA-OBSERVER REPEATABILITY OF THE SAME SCANS 91
5.42 SUB-STUDY B INTER-OBSERVER REPEATABILITY OF THE SAME SCANS 93
5.43 SUB-STUDY C INTER-OBSERVER REPEATABILITY OF DIFFERENT SCANS 94
5.5 DISCUSSION 96
5.51 Findings in the context of existing evidence 97
5.52 Strengths and limitations of the study 99
5.53 Interpretation 100
5.6 CONCLUSION 101
CHAPTER 6: EVALUATION OF PATIENT ACCEPTABILITY OF BLADDER ULTRASOUND (BUS) AND URODYNAMICS 104
INTRODUCTION 104
6.1 METHODS 104
6.2 RESULTS 105
6.3 DISCUSSION 106
6.4 CONCLUSION 110
CHAPTER 7: DISCUSSION 111
7.1 INTRODUCTION 111
7.2 EVALUATION OF DIAGNOSTIC TESTS 111
7.3 STRENGTHS AND LIMITATIONS 113
7.4 PERFORMANCE OF OTHER BIOMARKERS FOR DO 117
7.5 IMPLICATIONS FOR PRACTICE 117
7.6 CONCLUSION 118
7.7 FUNDING 118

List of figures

Figure 1: Transvaginal scan of the normal bladder wall showing measurements of the trigone and parts of the dome 28

Figure 2: Flow chart for the selection of studies included in the update of systematic review of the accuracy of bladder wall thickness in detrusor overactivity 39

Figure 3: Quality of studies included in the update of systematic review of the accuracy BWT in the diagnosis of DO 40

Figure 4: Risk of bias assessment 43

Figure 5: Forest plot of sensitivities and specificities for BWT in diagnosing DO 44

Figure 6: Hierarchical SROC curve of test accuracy data from meta-analysis on diagnostic accuracy of BWT in diagnosing DO 45

Figure 7: Thickened bladder wall as seen on transvaginal scan 60

Figure 8: Study flow chart 66

Figure 9: Progress of recruitment of patients in the BUS study 69

Figure 10: Participant flow diagram 72

Figure 11: Histogram of average bladder wall thickness measurements 73

Figure 12: ROC curve analysis for BWT in the study population 77

Figure 13: Box and whisker plot comparing BWT with DO diagnosis 78

Figure 14: Scatter plot for sub-study A- Intra-observer repeatability of the same scans 92

Figure 15: Bland-Altman analysis for sub-study A 92

Figure 16: Scatter plot for sub-study B-Inter-observer repeatability of the same scans 93

Figure 17: Bland Altman analysis for sub-study B 94

Figure 18: Scatter plot for sub-study C- Inter-observer repeatability of different scans 94

Figure 19: Bland-Altman analysis for sub-study C 95

Appendix 129

Appendix 1: STARD Checklist for reporting of studies of diagnostic accuracy 129

Appendix 2: Study accrual 131

Appendix 3: Sensitivity analyses 132
Figure 3.1: ROC curve excluding those results where the urodynamics test was not blind to the results of the ultrasound test 132
Figure 3.2: ROC curve excluding those results where there was more than four weeks between the tests 132
Figure 3.3: ROC curve incorporating incomplete ultrasound measurements 133
Figure 3.4: ROC curve using ambulatory urodynamic diagnosis where available 133
Appendix 4: Exploratory analyses 134
Figure 4.1: ROC curve including the urgency alone group (as per clinical history; excluding mixed stress/urge incontinence group: 134
Figure 4.2: ROC curve including the 'pure' DO group only (diagnosis of DO/low compliance/DO plus low compliance, excluding 'mixed' DO - DO with another diagnosis of USI or voiding dysfunction; 134
Figure 4.3: ROC curve including the 'wet' DO group only (excluding 'dry' DO); 135
Figure 4.4: Box and whisker plot comparing BWT for 'wet' and 'dry' DO 135
Figure 4.5: ROC curve from exploratory analysis using the trigone measurement alone for BWT ... 136
Figure 4.6: ROC curve excluding those who had a detrusor pressure rise upon provocation testing 'provoked DO' 136
Figure 4.7: ROC curve excluding those who had PVR>30ml upon testing 137
Figure 4.8: ROC curve using the average dome, 1cm left of dome, 1cm right of dome 137
Appendix 5 138
Table 5.1: Results of ROC curve analysis in pre-specified subgroupings 138
Table 5.2: Results of univariate analysis exploring factors possibly associated with DO diagnosis .. 138
Table 5.3: Results of multivariable analysis exploring factors possibly associated with DO diagnosis 139
Appendix 6 140
6a) BUS Standard Operating Procedure for Scanning 140
6b) Standard Operating Procedure for Urodynamics 147

List of tables

1. Table 1: Conditions that cause overactive bladder 23
2. Table 2: Characteristics of the included studies in the update of the systematic review of accuracy of transvaginal/ translabial bladder / detrusor wall thickness in detrusor overactivity……48
3. Table 3: Characteristics of the included studies for the systematic review on BWT by transabdominal BWT/DWT ……..51
4. Table 4: Results of studies included in the systematic review of the accuracy of ultrasound measurement of BWT in diagnosing DO----------52
5. Table 5: Recruitment by centre 70
6. Table 6: Characteristics of patients in the BUS study 71
7. Table 7: Summary of all urodynamic diagnosis in the cohort of women recruited in this study 74
8. Table 8: Urodynamic parameters in the BUS study 75
9. Table 9: Comparison of index and reference standard results – dichotomised at 5mm 76
10. Table 10: Estimates of BWT test accuracy – dichotomised at 5mm 76
11. Table 11: Estimates of BWT test accuracy – trichotomised at 3, 3-5, >=5mm 77
12. Table 12: Comparison of distribution of BWT measures between the full BUS study and the sub-studies 95
13. Table 13: Details the experience of the observers in the sub studies---------- -96
14. Table 14: Estimates of measures of analytical and individual variability---------- 96
15. Table 15: Pre-existing evidence on inter and intra-observer variation of BWT ----------102
16. Table 16: Pain during after the urodynamics and bladder ultrasound tests----------105
17. Table 17: Acceptability of urodynamics and transvaginal bladder ultrasound tests---------106
18. Table 18: STAI-SF 106

Synopsis

This thesis sheds light on the following areas in the investigation of overactive bladder:

1. Update of the systematic review of the existing evidence on the role of bladder wall thickness (BWT) in the diagnosis of detrusor overactivity. 2. Accuracy of BWT in the diagnosis of detrusor overactivity. 3. Reproducibility of transvaginal BWT on ultrasound scan.4. A comparison of patient acceptability of both the diagnostic techniques-transvaginal BWT scan vs. urodynamics in women with overactive bladder.

Based on the findings of this study, I recommend that BWT scan is not an accurate diagnostic test in diagnosing detrusor overactivity in women with overactive bladder.

Dedication

I dedicate this thesis to my husband Muralidhar and my children Vaishnavi and Abhishikth for giving me time to pursue my research and academic interests and to my parents who enthused me through difficult times and taught me to persevere with patience.

Acknowledgements

This work was undertaken while I was a research fellow coordinating the BUS trial and was based in the School of Clinical and Experimental Medicine, University of Birmingham from September 2011 to July 2013.The trial was funded by the NIHR/HTA. My sincere thanks to NIHR/HTA without whose support this project would not have been possible.

I am indebted to my research supervisor Dr Pallavi Latthe, Consultant in Obstetrics and Gynaecology and Subspecialist in Urogynaecology, for her expert guidance and continued encouragement and support through the years and many hours spent on research supervision leading to this thesis. I would also like to thank my other supervisor, Prof Arri Coomarasamy for his continued help and mentoring. I sincerely thank Dr Anu Sajja for being my co-observer in the reproducibility of BWT study. I am extremely grateful to Joan and Lorna, the healthcare assistants on the Urogynaecology unit for their valuable help during BWT scanning. This work would not have been possible without the help of Mr Lee Middleton, Dr Jane Daniels, Prof Jon Deeks and Mrs Laura Gennard. Finally, I would like to emphasise that this work would not be complete without the co-investigators, sonographers and research nurses in the multicentre BUS study.

Executive abstract

Objectives

This thesis has the following objectives:

Primary:

To estimate the accuracy of bladder wall thickness (BWT), measurement by transvaginal bladder ultrasound (BUS) in diagnosing detrusor overactivity (DO)

Secondary:

To generate pooled evidence on the role of BWT scan in diagnosing DO by an update of systematic review

To assess whether measurements of BWT have adequate reliability and reproducibility

To investigate the acceptability of BWT scan compared to urodynamics

Methods

A cross-sectional test accuracy study was undertaken to the estimate the accuracy of bladder wall thickness (BWT) in diagnosing DO in 687 women with overactive bladder (OAB) symptoms or urgency predominant mixed incontinence. Eligible women from 22 UK hospitals were recruited and BWT was measured using transvaginal ultrasound scans (index test). All women had the reference test of multichannel urodynamics, which was undertaken blind to the findings of the transvaginal ultrasound. The primary analysis involved calculations of sensitivity, specificity, predictive values and likelihood ratios using a BWT of 5mm as a cut-off (>=5mm indicating presence of DO, <5mm indicating absence of DO). A receiver operator characteristics (ROC) curve was constructed and the area under the curve

computed (with 95% CI) to give an overall estimate of BWT accuracy across all thresholds of BWT.

The intra- and inter-observer reproducibility of measuring BWT was assessed in three subsets of the transvaginal scans. For each subset, BWT measurements were analysed using one-way analysis of variance and the intra-class correlation and repeatability coefficients were derived.

The acceptability of transvaginal BWT and urodynamics from the patient's perspective was evaluated through the completion of self-reported questionnaires containing visual analogue scales for pain, levels of embarrassment, ordinal scales for acceptability and a generic state anxiety measure. Mean differences and 95% confidence intervals were determined by a paired t-test for pain and anxiety scores. Wilcoxon signed-rank test was used for acceptability responses and McNemar's test for binary responses.

Results

Test accuracy & reproducibility studies

The mean age of the 687 women consenting to the study was 52.7 years (SD 13.9) and the average BMI was 30.6 (SD 12.2). Fifty five percent (387/687) of the women were postmenopausal. According to the clinical history, 61% (419/687) had urgency predominant mixed incontinence and 33% (226/687) reported only urinary urgency along with increased frequency. The median duration of symptoms was 3.0 years (IQR: 1.6, 7.0). Six hundred and forty four participants had both the index and reference standard carried out.

Estimation of the accuracy of BWT showed poor sensitivity, specificity and likelihood ratios at all pre-specified cut-offs. The ROC curve showed no evidence of discrimination at any threshold between those with and without DO (p=0.25): the AUC was 0.53, 95%CI: (0.48, 0.57). Furthermore, there was no evidence that the mean BWT measurements were any higher in the DO positive group compared with the DO negative group: 4.85mm (SD: 1.36) versus 4.70mm (SD: 1.29); p=0.19. Extensive sensitivity analyses and subgroup analyses did not alter these findings.

For the intra-observer variation study, the individual variability in standard deviation (SD) was 1.04mm with an analytical variation of 0.42mm.In the inter-observer variation study on interpretation of stored images, we found an SD of 1.23mm with an analytical variation of 0.35mm. In the prospective interobserver variation study of repeated scans performed by two independent observers, the SD was 0.95mm with an analytical variability of 0.76mm.

Acceptability of tests

Six hundred and forty-six (94%) participants in the study completed the acceptability questionnaire following both tests. Pain levels following both tests appeared relatively low, with scores during and shortly after urodynamics slightly higher than the corresponding scores during and after a BWT scan. There was a trend towards greater acceptability of BWT scan compared to urodynamics (p<0.001), (81% versus 56%). Fewer women felt that they would recommend urodynamics to a friend compared to a BWT scan (86% v 96%; p<0.001) and have the same test again (88% v 97%, p<0.001). Nearly 20% of women reported moderate levels of embarrassment with urodynamics compared to 10% with BWT scan (p<0.001). Both the tests appeared to provoke moderate levels of anxiety levels (12.6 for

BWT scan and 12.9 for urodynamics), although the scores were only slightly higher with urodynamics (0.3 points difference on a 4-24 scale, 95%CI: 0.1 to 0.5; p=0.02).

Conclusion

There was no evidence that transvaginal BWT had any relationship with DO, regardless of the cut-off point. Bladder wall thickness measurement did not discriminate women with DO versus those without DO and hence is not an accurate test for diagnosing DO. In the presence of high levels of analytical variation for the measurement of BWT, it is unlikely that BWT measurement made by transvaginal ultrasound has sufficient reliability and reproducibility to be a precise diagnostic test. Women experienced higher levels of embarrassment and a lower rate of acceptability with urodynamics compared to the BWT scan procedure.

List of Abbreviations

AUC	Area Under Curve
BCTU	Birmingham Clinical Trials Unit
BTX-A,	Botulinum toxin serotype A, Onabotulinum A
BOO	Bladder outflow obstruction
BUS	Bladder Ultrasound
BWT	Bladder Wall Thickness
CI	Confidence interval
CISC	Clean Intermittent self-catheterisation
DO	Detrusor Overactivity
DWT	Detrusor Wall Thickness
EQ-5D	EuroQol EQ-5D
GUP	Good Urodynamic Practice
ICC	Interstitial cells of Cajal
ICECAP	Investigating Choice Experiments CAPability measure
ICIQ	International Consultation on Incontinence Questionnaire
ICS	International Continence Society
I-QoL	Incontinence-specific Quality of Life Questionnaire
IUGA	International Urogynaecology Association
LC	Low Compliance
LR	Likelihood ratio
LUTS	Lower Urinary Tract Symptoms
MCC	Maximum cystometric capacity
MESA	Medical, Epidemiological, and Social Aspects of Ageing
MUI	Mixed Urinary Incontinence
NICE	National Institute for health and care Excellence
NPV	Negative predictive value
OAB	Overactive Bladder
PFMT	Pelvic floor muscle training
POP	Pelvic Organ Prolapse
PPV	Positive predictive value
PSA	Probabilistic Sensitivity Analysis
PTNS	Percutaneous Tibial Nerve Stimulation

PVR	post void residual
QC	Quality Control
QoL	Quality of Life
ROC	Receiver operating characteristic
SNS	Sacral Nerve Stimulation
SOP	Standard Operating Procedure
STAI-SF	State-Trait Anxiety Inventory short form
STARD	Standards for Reporting of Diagnostic Accuracy
SUI	Stress Urinary Incontinence
UDI-UI	Urogenital Distress Inventory-Urinary Incontinence
UDS	Urodynamics
UI	Urinary incontinence
UUI	Urgency urinary Incontinence
UK	United Kingdom
UPP	Urethral Pressure Profile
USI	Urodynamic Stress Incontinence
USS	Urgency Severity Scores
UTI	Urinary Tract Infection
VAS	Visual Analogue Scale
VD	Voiding Dysfunction

Publications from this thesis

1. Rachaneni S, Balogun M, Latthe P. Bladder-wall-thickness ultrasound scan in the investigation of LUTS in women: challenges and limitations. Int Urogynecol J. 2013 May; 24(5):725-8

Manuscripts submitted for publication

1. Rachaneni S; McCooty S; Middleton LJ; Brookes VL; Daniels JP; Coomarasamy A ; Verghese T; Balogun M; Gorantis I; Barton P; Roberts T; Deeks JJ; Latthe P Accuracy of bladder wall thickness in the diagnosis of detrusor overactivity: A study to evaluate if ultrasound can reduce the need for urodynamics -Monograph submitted to the Health Technology Assessment(HTA) currently undergoing revision.

Presentations from this thesis

1. Rachaneni S; McCooty S; Parsons M; Toozs-Hobson P; Middleton LJ; Latthe P. LC vs DO Podium presentation at the UKCS meeting held in Bradford, UK on 18/4/15.
2. Rachaneni S; Sajja A; Latthe P. Reproducibility of bladder wall thickness scan-A prospective study. Non discussed poster presented at the ICS 2013 conference held in Barcelona from 26/8/2013 to 30/8/2013.

Chapter 1: Introduction

1.1 Background

Urinary incontinence (UI) is defined by the International Continence Society as 'the complaint of any involuntary leakage of urine'. Urinary incontinence has a significant negative impact on the quality of life. Women affected by this condition find it psychologically distressing and socially restrictive. Urgency UI (involuntary leakage during or immediately preceeded by urgency), stress UI (involuntary urine leakage with physical activity or on sneezing or coughing), and mixed UI (a combination of the above two types) are the common types on incontinence (Haylen et al. 2010).

1.2 Overactive bladder (OAB)

International Continence Society (ICS)/International Urogynaecology Association defines overactive bladder(OAB) as a symptom complex of urinary urgency (intense, sudden desire to void) with or without incontinence, increased urinary frequency, or nocturia in the absence of infection or other proven pathology(Haylen, de, Freeman, Swift, Berghmans, Lee, Monga, Petri, Rizk, Sand, & Schaer 2010). Increased urinary frequency and urgency symptoms are more prevalent in patients with OAB than incontinence. Urgency incontinence which is the most distressing symptom of OAB, affects only a third of patients (Abrams 2003).

Millions of people worldwide suffer from OAB. In the epiLUTS study, 12.8% of the population were found to suffer from OAB (Coyne et al. 2009). Prevalence and severity of OAB are known to increase with age. The prevalence seems to increase from 14.9% in the 18-29yrs group, to 21.3% in the 30-39yrs group, 32.9% in the 40-49yrs group, 35.8% in the 50-59yrs group and 39.8% in the 60-69yrs group (Coyne et al. 2013). It is anticipated that the

burden of OAB will increase with the increase in longevity, over the next few decades from 500 million in 2013 to 546 million by 2018 (20.1% increase) (Irwin et al. 2011).

OAB can have an enormous impact on the overall wellbeing of the affected women (Irwin et al. 2006). They tend to severely restrict their fluid intake to control urinary frequency and carry out 'toilet mapping' to cope with the feeling of urgency and urgency incontinence (Irwin, Milsom, Kopp, Abrams, & Cardozo 2006). The fear of coital incontinence may have a serious impact on their relationships (Jha et al. 2012). Low mood and depression due to social restriction and fear of embarrassment are reported more frequently in women with OAB. These symptoms may also result in significant financial implications (e.g. cost of pads, prescriptions, time off work, job losses, effects on the family etc.) (Freeman and Adekanmi 2005). In the elderly population, nocturia secondary to urgency and urgency incontinence predisposes to falls and fractures (Mobley and Baum 2014).

1.3 Mixed urinary incontinence

The standardisation report ICS/IUGA describes mixed urinary incontinence (MUI) as a complaint of involuntary loss of urine associated with urgency and also with effort or physical exertion or on sneezing or coughing (Haylen, de, Freeman, Swift, Berghmans, Lee, Monga, Petri, Rizk, Sand, & Schaer 2010). MUI is the commonest form of UI.

In a prevalence study on women complaining of lower urinary tract symptoms(LUTS), MUI with equal severity of urgency and stress was the commonest condition in 56%, stress predominant MUI in 29% and urgency predominant MUI in 15%. These prevalence rates on urodynamics changed to MUI with both stress incontinence (SUI) and detrusor overactivity (DO) in 18%, SUI in 42%, DO in 25% and normal findings in 15%. Urodynamic stress

incontinence was diagnosed in 82% of stress predominant MUI. In the urgency predominant MUI group, only 64% had DO on urodynamics (Digesu et al. 2003).

Appropriate categorisation of women into urgency predominant or stress predominant MUI based on clinical history has been a matter of great debate. Medical, Epidemiological, and Social Aspects of Ageing (MESA) Questionnaire, urogenital distress inventory-urinary incontinence (UDI-UI) and visual analogue scale(VAS) scores for urgency component and stress component evaluation along with bladder dairies may be used to categorise women with mixed incontinence based on the predominant subcomponent (stress or urgency)(Brubaker et al. 2011). In busy clinical practice, however, women are categorised based on the symptoms they think are the most bothersome.

1.4 Underlying pathology

Detrusor overactivity (DO) is defined as the occurrence of involuntary detrusor contractions during filling cystometry. These contractions, which may be spontaneous or provoked, produce a wave form on the cystometrogram of variable duration and amplitude (Haylen, 2010). Neurogenic detrusor overactivity is a condition where DO is demonstrated with evidence of a relevant neurological disorder (Haylen, 2010). DO may be the pathology behind OAB symptoms in 54-58% of people. The remaining 42-46% of the patients may have other pathologies causing OAB symptoms (Hashim and Abrams 2006)(Table 1).

The pathophysiology of the OAB and DO are not yet understood completely. The following theories have been proposed:

a)The smooth muscle of the detrusor and the urothelium might generate increased afferent activity resulting in the symptoms of OAB (Andersson 2002).

b)Spinal parasympathetic outflow is absent during the filling phase of the micturition cycle (de Groat 2006). Despite the absence of parasympathetic outflow, non-synchronized local contraction and relaxation might give rise to OAB symptoms (Andersson 2006).

c) Cajal-like interstitial cells (ICC) may generate electrical potentials and stimulate detrusor contractions. The alterations in the transduction of the signals of ICCs between nerves and detrusor smooth muscle cells, along with alteration between afferent nerve endings and the urothelium (via suburothelial ICCs) may lead to a disturbance in spontaneous contractility resulting in DO (Juszczak 2014)(Kubota 2011).In vitro studies have shown that muscle strips from overactive bladder have increased spontaneous contractile activity than those from normal bladder (Kinder and Mundy 1987).

d) Neuronal and non-neuronal sources release mediators to stimulate the detrusor myogenic contractile activity (Andersson and Wein 2004). Detrusor smooth muscle cells may become hyperexcitable, start reacting to minor stimuli resulting in untimely bladder contractions and give rise to the symptom of urgency (myogenic theory). The pathology may be purely intrinsic to the detrusor muscle (Darblade et al. 2006).

e) In normal bladder, there is a lack of coordination among various detrusor muscle units with some units contracting and others relaxing to stabilise intravesical pressure. Changes in intercellular communication may lead to an increased coordination between various units resulting in high amplitude contractions in patients with DO (Brading 1997). Other aetiological theories include failure of pelvic floor inhibitory reflexes during the filling phase and abnormalities in neurotransmission(Steers 2000).

Table 1: Conditions that cause overactive bladder

Lower urinary tract conditions	Mechanism of effect
Detrusor overactivity	Involuntary detrusor contractions during filling/storage phase of the bladder
Recurrent urinary tract infections	Alterations in bladder mucosa may activate some pathways leading to irritative bladder symptoms
Bladder outflow obstruction	Detrusor hypertrophy secondary to obstruction may sometimes be iatrogenic(after a midurethral sling)
Oestrogen deficiency	Bladder mucosal atrophy, atrophic vaginitis and urethritis
Neurological conditions: Multiple sclerosis, Cerebrovascular accidents, multi-infarct dementia, Parkinson's disease, Space occupying lesions compressing the spine-benign or malignant	Impairment of central inhibition of bladder causing causing neurogenic DO

1.5 Evaluation of symptoms of OAB

A detailed clinical history of urinary symptoms alone may not help in the identification of the underlying pathology (Jackson 1997;James et al. 1999). Clinical evaluation should include a 3 day bladder diary of urinary habits (Smith et al. 2013) and a thorough clinical examination to rule out urogenital atrophy, any pelvic organ prolapse or any other urogenital abnormalities. The National Institute of Health and Care Excellence (NICE) guideline on urinary incontinence in women recommends conservative management as a first step in treatment. In women unresponsive to conservative management, the recommendation is to perform urodynamics to diagnose DO. NICE also recommends the incontinence-specific quality-of-life (QoL) scales in assessing the impact of urinary symptoms eg. ICIQ, UISS, ISI,BFLUTS, I-QOL, KHQ5, SEAPI-QMM, and SUIQQ (Smith, Bevan, Douglas, & James 2013).

1.51 Clinical history

Clinical history-taking for UI includes the type of incontinence, duration and severity of UI, impact on the quality of life, exacerbating factors like physical activity, quantity and quality of fluid intake and any medications, coexisting medical, surgical or gynaecological conditions. During the clinical assessment, the NICE recommends identification of other conditions where specialist input may be needed(Smith, Bevan, Douglas, & James 2013). In a literature review, sensitivity and specificity of clinical history compared to urodynamics was reported to be 0.69 and 0.60 for urgency incontinence/OAB, and 0.51 and 0.66 for women with MUI. Sensitivity for predicting DO in women with clinical symptoms of OAB was 0.76, but the specificity was only 0.57 (Colli et al. 2003). In a systematic review, sensitivity of clinical history for diagnosing DO was found to be 0.61 (0.57-0.65) and the specificity was 0.87 (0.85-0.89). Approximately, 3-15% of women complaining of UI may have normal findings on urodynamics (Martin et al. 2006). Hence clinical history alone may not help in the identification of the underlying pathology.

1.52 Bladder diaries in the assessment of OAB

Bladder diaries may have diagnostic and prognostic value in the investigation and treatment of LUTS (Coyne, Sexton, Thompson, Milsom, Irwin, Kopp, Chapple, Kaplan, Tubaro, Aiyer, & Wein 2009). When the accuracy of bladder diaries in OAB patients against the urodynamic diagnosis of DO was evaluated, the sensitivity was found to be 0.88 and specificity 0.83(Parsons et al. 2007).

Data from the three day bladder diaries was found to be as accurate as those from the seven day diaries. Diaries of urinary habits collected over a three day period may improve the

accuracy of data collection by reducing the possibility of incorrect information from women who might complete some of the missing data from memory (Dmochowski et al. 2005).

Severity of urinary urgency may be measured by urgency severity scores. Episodes of urgency, timing, activities predisposing to the sensation, the severity, the volume of leakage and the number of pads used are recorded (Chung et al. 2011).

Women need to understand which predisposing factors make them leak and provide an accurate recording of diaries for their UI. Some patients might fill in only a part of the diary resulting in missing data. Other sources of inaccuracies/incomplete data may happen when women sometimes attempt to reconstruct their voiding habits from memory rather than fill their bladder diaries contemporaneously.

1.53 Questionnaires

1.531 ICIQ-OAB SF

The International consultation on incontinence modular questionnaire-short form (ICIQ OAB-SF) allows a subjective measure of the severity of each component of OAB and its effect on QoL. The four scored items of the ICIQ-OAB are: Urinary frequency and its ‘bother’ on visual analogue scale (VAS) ranging from 0 ‘‘Not at all’’ to 10 ‘‘A great deal’’, [2] nocturia and its ‘bother’ [3] urgency and its ‘bother’ and [4] urgency incontinence and its ‘bother’. Shorter questionnaires are advantageous in settings, where the availability of time is limited. High ICIQ scores have a positive correlation with maximal detrusor pressure levels and low ICIQ scores have a negative correlation with the volume at the first sensation

(Seckiner et al. 2007). ICIQ-OAB SF questionnaire was used in the study as it was validated, short, simple and more likely to be completed by the patients.

1.6 Urodynamic evaluation of OAB

Urodynamics is a test used to assess the storage and voiding function of the bladder (Digesu et al. 2004). At present, laboratory urodynamics remains the gold standard test for assessment of LUTS (Coyne, Sexton, Thompson, Milsom, Irwin, Kopp, Chapple, Kaplan, Tubaro, Aiyer, & Wein 2009).Urodynamics consists of uroflowmetry and cystometry. Urodynamics done concomitantly with urethral pressure profile (UPP) may provide information about the detrusor contractility and urethral function along with the coordination between both (Digesu, Hutchings, Salvatore, Selvaggi, Milani, & Khullar 2004).

In a Cochrane review, the role of urodynamics in the management of urinary incontinence has been questioned (Clement, 2013). Multichannel urodynamics have a 50–80% false negative rate compared to ambulatory urodynamics for diagnosing DO(Dokmeci et al. 2010). However, in asymptomatic women volunteers, ambulatory urodynamics seem to have a high false positive rate for DO (van Waalwijk van Doorn ES, 1992). Due to the poor reproducibility of the findings during urodynamics, the bladder has been blamed to be an 'unreliable witness'(Hashim, 2006). In a study on women with a clinical history of pure SUI, a discordant finding of DO was found in 25% of the patients on urodynamics (Serati et al. 2013). Inspite of all the above mentioned weaknesses, in the absence of a better diagnostic test, urodynamics was taken as the reference standard for the study.

1.7 Bladder wall thickness

Ultrasound scanning has been used in diagnosing bladder or urethral diverticula, the assessment of bladder neck in SUI and voiding dysfunction. Sonographic BWT has been

studied in various populations—women and children, and adult men with and without prostate hypertrophy. The technique was first described by Singer et al. in 1981(Singer D, 1981).

Normal BWT in adult women may range between 3–6 mm (Yang, 2002). The average BWT increased from 1.5mm to 2.76 mm in the paediatric age group to 3.0±1.0mm in adult women without LUTS. Bladder wall thickness had a small positive correlation with increasing age (Elbadawi et al. 1998). Increased interstitial collagen deposition may explain the small increase in BWT with age. With very small age related increase in the BWT, correction with respect to age may not be required (Hakenberg et al. 2000).

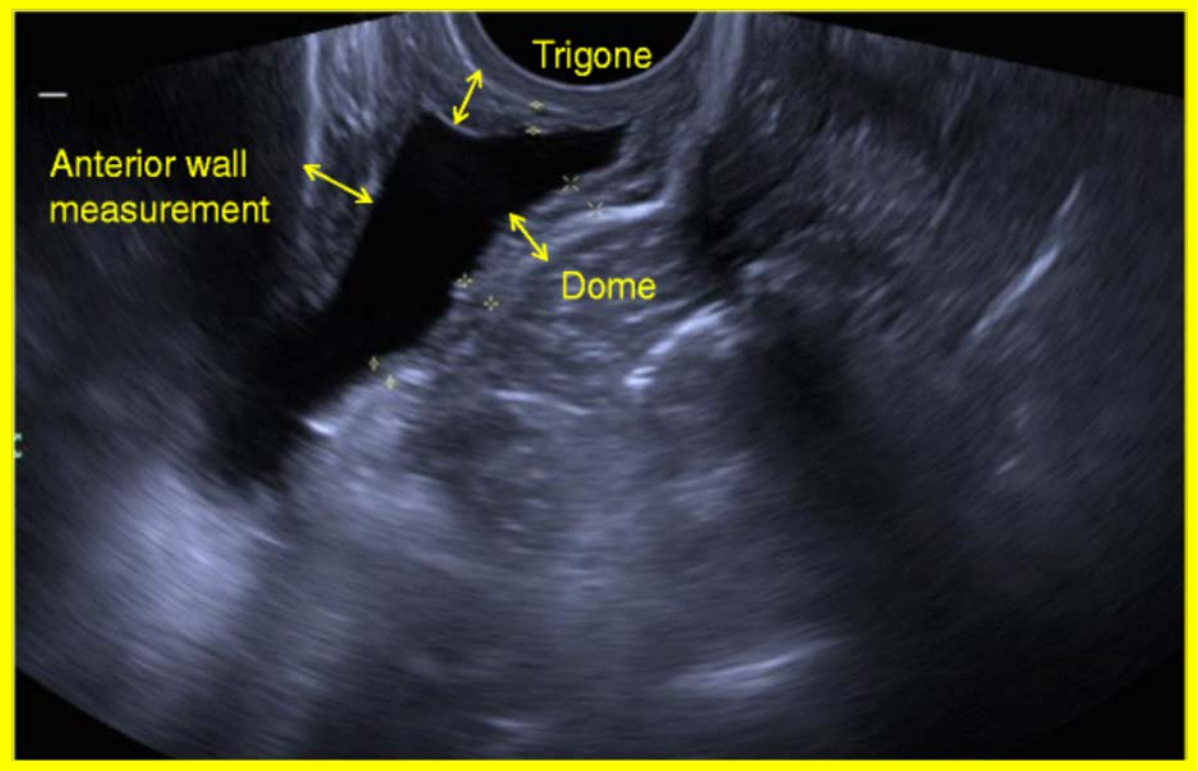
Measurement of BWT may be useful in differentiating women with DO from those with SUI (Kuhn et al. 2011). In women with LUTS who show equivocal findings on urodynamics, BWT may prove to be a useful adjunctive test (Robinson, 2002).

1.71 Technique and the route of scanning

The bladder wall consists of the outer serosa, the detrusor muscle and the inner urothelium. On ultrasound scan of the bladder wall, the serosa and the urothelium have a hyperechoic appearance and the detrusor muscle, a hypoechoic appearance (Fig. 1). Bladder wall thickness is measured with the transvaginal probe in the sagittal plane introduced 1 cm beyond the vaginal introitus in the midline and the entire thickness of all the three layers measured. To measure the detrusor wall thickness (DWT) only, the hyperechoic urothelium and serosa are excluded from the measurement (Farag, 2011). The echo-poor central area of the urethra is a useful landmark. Tangential measurements of BWT might erroneously suggest bladder-wall thickening, perpendicular imaging is recommended. When the probe is

perpendicular to the bladder wall, the hyperechoic layers of serosa and the urothelium appear as thin, distinct lines (Oelke, 2010). Measurements are to be made at maximum magnification. The BWT measurement is carried out in three sites-anterior wall, trigone and dome and an average BWT calculated. Bladder wall appearance/thickness is known to be fairly constant at small volumes of 0–50 ml (Khullar, 1994). Figure 1 depicts the measurement of a normal bladder wall.

Figure 1: Transvaginal scan of the normal bladder wall showing measurements of the trigone and parts of the dome



With the transvaginal probe closest to the bladder, the margins of the bladder wall are more clearly visualised (Panayi, 2010). With transabdominal, introital (translabial) or perineal sonography, the probe is further away from the bladder and the resulting attenuation of ultrasound waves reduce the image quality.

1.72 Variations in the technique of BWT

Transvaginal BWT measurement was first described by Khullar et al in the parasagittal plane. He measured the BWT at the trigone, dome and the anterior wall and calculated the average of the three measurements(Khullar, 1994). Some investigators measured only the dome thickness with the probe in the midline. The measurements were made in the sagittal plane at three different sites on the dome—midline and 1 cm on either side—and calculated mean BWT(Lekskulchai, 2008).

1.73 Route of BWT scanning

There has been much debate on the ideal route of bladder scanning-transvaginal, transperineal or transabdominal. We have chosen the transvaginal route of scanning for BWT due to the low interobserver difference, good quality images and the ease of the technique compared to transabdominal or translabial scanning(Panayi, 2010).

1.74 Challenges in BWT measurement

Transvaginal probes with high-frequency (e.g.>7 MHz) have a resolution of <0.13 mm and are ideal for measuring thin-walled structures like BWT. Appropriate gain settings and the focal zone adjustments are needed to ensure good visualisation of bladder wall. The uterovesical peritoneal reflection on to the dome of the bladder may add 1 mm to the BWT measurement (Jequier, 1987). Care should be taken not to include the vaginal wall while measuring BWT at the trigone. Visualising and measuring anterior wall thickness can be challenging in patients with cystocele and those who have undergone bladder-neck surgery. The bladder wall may be irregular and undulating and the outline of the margins may not be seen clearly. Sometimes the BWT may not be visualised perpendicularly depending on how the surrounding structures are indenting the bladder erroneously increasing the BWT(Rachaneni, 2013). Thickened urothelium (secondary to infection, catheterisation) may give a false positive thickened bladder wall measurement.

The demarcation of the different layers of bladder wall may be seen only when the probe is exactly perpendicular to the bladder wall. With the bladder being a curvilinear organ, the differentiation of all the three layers with transvaginal probe may be possible only at the trigone and not on the anterior wall and dome. On transabdominal scan, the demarcation of three layers may be possible for the anterior wall and not from the dome or trigone.

1.75 Bladder wall thickening in LUTS-Possible clinical applications

Overactive and obstructed bladders may be distinguishable from other LUTS by measuring BWT. In OAB patients, repeated detrusor contractions may give rise to the perception of urgency during the filling phase of micturition cycle. In response to the symptom of urgency and the fear of leakage, women may be voluntary activate the urethral sphincter and the pelvic floor musculature to increase the urethral closure pressure. These detrusor contractions against a closed urethral sphincter may result in the bladder wall thickening or hypertrophy (Khullar et al. 1996) (Serati et al. 2011).Bladder outflow obstruction may also result in detrusor hypertrophy(Harrison et al. 1987). Women who have an over corrected of bladder neck angle following SUI surgery may also show thickened bladder walls(Martan et al. 2001).

The study by Panayi et al. found that women with a mean BWT >5 mm had a significant increase in urinary frequency and higher visual analogue score for urgency compared with those with a mean BWT of <5 mm (Panayi, 2010). In a systematic review on the diagnostic accuracy of BWT scanning, for a 5-mm cut off, the sensitivity in diagnosing DO ranged from 37 % to 84 % and specificity from 78 % to 89 % (Latthe, 2010). Using 6 mm as cut off, Robinson et al. calculated a sensitivity of 22 % and specificity of 89 % for BWT as a marker for DO on ambulatory urodynamics (Robinson, 2002). As the cut-off increased to 6.5 mm in

another study, sensitivity was 13.3 %, specificity increased to 97.7 %, with a positive predictive value of 71.4 % and a negative predictive value of 72.1 % (Serati, 2010).

Women with a previous history of paediatric nocturnal enuresis (enuresis after school entry) had thicker bladder walls with an average detrusor wall thickness (thickness of bladder wall after excluding the measurement of mucosa and the serosa) of 4.7(SD 2.1, range 2.1–10.6) mm compared to 4.2 (SD 1.7, range 1.5–14.2) mm in the non-nocturnal enuresis group (Lekskulchai O and Dietz 2006). In a pediatric study on the treatment with Desmopressin in children with nocturnal enuresis, response was poorer in children who had thickened bladder walls and reduced bladder capacity (Yeung, Eschenbacher, & Pauls 2014).

In women complaining of DO and coital incontinence at orgasm, the bladder wall was found to be thicker compared to women with only DO (5.8 +/- 0.6 mm vs. 5.2 +/- 1.2 mm [P=0.007]) (Serati, Cattoni, Braga, Siesto, & Salvatore 2011).

A significant correlation between BWT and detrusor pressure at maximum flow rate (p det/Q max values) was seen in a small study(Kuhn, 2011). Bladder wall thickness measurement may be a promising technique to diagnose bladder outflow obstruction (BOO) that have the potential to replace urodynamics needing further evaluation(Belal, 2006).

Women with higher grades of uterovaginal prolapse may have a significant distortion of anatomy of anterior compartment and a relative urethral kink. The resulting infravesical obstruction may thicken the bladder wall(Farag, 2011) . However, in a retrospective study in patients with voiding dysfunction, there was no increase in mean BWT (Lekskulchai, 2009).

So far the role of thickened bladder wall in identifying the above mentioned conditions has been evaluated in studies which were lacking in methodological quality and hence no firm conclusions can be drawn.

1.8 Acceptability of diagnostic tests-BWT and Urodynamics

A diagnostic test is of no use in clinical practice if it is not generally acceptable to the population it is being offered to. Our index test was BWT measured on transvaginal scan. Transvaginal ultrasound scan is often used in gynaecological investigations. Transvaginal ultrasound scan has become an acceptable investigation among pregnant as well as non-pregnant women across various ethnicities (Atalabi et al. 2012;Rosati and Guariglia 2000).

In a study on patient experience of invasive diagnostic procedures, a third of pregnant women undergoing transvaginal scan experience mild pain or discomfort. Very few women described the experience as ‘excruciating’, ‘horrible’ or ‘distressing’. On a scale from 0 to 5, the measured perception of difficulty was rated as 1.3. When compared to other procedures which are perceived as uncomfortable or painful, women felt that transvaginal scan fared better than having a dental filling, a cervical smear or having a blood sample taken. However, they felt it was more discomforting than abdominal scan(Clement et al. 2003).

Urodynamic test involves inserting a catheter into the bladder and another into the rectum/vagina. Pain, embarrassment and distress are reported by women undergoing urodynamics because of the invasive nature of the intervention(Yeung et al. 2014).Women with younger age, history of anxiety or depression, and a diagnosis of OAB and painful bladder syndrome have been reported to have more negative experiences during urodynamics (Yeung, Eschenbacher, & Pauls 2014). Currently, there are no published validated

questionnaires on the acceptability of urodynamics. To test the acceptability of urodynamics and BWT testing, a questionnaire was constructed using validated instruments like visual analogue scale(VAS) measurements on pain during and after the test (Hjermstad et al. 2011), levels of embarrassment a person experienced, recommendation of the test to a friend and whether they were willing to go through the same test again (NHS Friends and Family test 2012).

1.81 STAI scores

The State-Trait Anxiety Inventory (STAI) is a tool designed to differentiate and measure anxiety as a state and as a trait. The stable traits of an individual like calmness, confidence and anxiety proneness are measured on Trait anxiety scale (T-Anxiety). State Anxiety Scale is designed to measure the anxiety a person is experiencing at the time of evaluation. It also measures temporary changes in anxiety in response to a particular situation like worry apprehension, tension and nervousness. People who are more anxiety prone (higher T-Anxiety scores) have a lower threshold of perception of danger or a threat. State anxiety scores are directly proportional to their trait anxiety scores (Spielberger 1983). State-Trait Anxiety Inventory-Short form (STAI-SF) questions were included in the acceptability questionnaire to study and compare the anxiety provoked by the index test and the reference standard.

1.9 Rationale for the study

The only investigation extensively used for investigation of LUTS is urodynamics. Urodynamics is an invasive, expensive investigation which is poorly tolerated and has a higher risk of urinary tract infections (Foon et al. 2012). A search for an alternative diagnostic test without the above mentioned weaknesses of urodynamics has identified thickened bladder wall as a possible marker for DO in a few observational studies(Khullar, 1994) (Kuhn, 2011) . Bladder wall thickness scan by transvaginal route is less invasive, less

embarrassing, less uncomfortable and better tolerated test and it does not carry any risk of urinary infection. A systematic review of studies on the diagnostic accuracy of BWT has shown that there is absence of good quality evidence and there exists an equipoise on the clinical utility of BWT scan in diagnosing DO(Latthe et al. 2010). The technique and sites of BWT measurement has not been standardised with some studies measuring DWT (Lekskulchai, 2008) and some BWT (Khullar, 1994). The site of measurement of BWT was variable with some taking multiple measurements of the bladder dome (Lekskulchai, 2008) and some measuring it in three different sites-dome, anterior wall and trigone (Robinson, 2002). There is a lack of consensus on the appropriate cut off required to make a diagnosis of DO (Latthe, 2010).

Determining the accuracy of BWT is an essential first step in establishing its clinical effectiveness. If BWT scan proves to be an accurate test for diagnosing DO compared to urodynamics, it can replace urodynamics and improve patient experience in diagnostic workup, reduce the risk of urinary infection and hence my question of interest. Optimum BWT cut-off value for diagnosing various LUTS may then be standardised. A subsequent health economic evaluation may show the cost savings to the health care systems by replacing urodynamics with the less expensive BWT scan. Whether BWT scanning will help in the understanding of disease development, progression and prediction of response to treatment needs to be explored. The present study was designed to elicit the diagnostic accuracy of BWT in diagnosing DO.

1.91 Aims and Objectives of the study

The primary aim of this study was to evaluate whether BWT measurement has sufficient accuracy to reduce the need for urodynamics in the diagnosis of DO.

The secondary aims were to investigate the relative acceptability of urodynamics and BWT scan and to examine the reproducibility of BWT.

Chapter 2: Update of the systematic review on the accuracy of BWT as a biomarker in the diagnosis of detrusor overactivity

Introduction

The search for more reliable, less invasive simple and cost-effective tests as alternatives to urodynamics has led to the exploration of bladder wall thickness (BWT) scan as an alternative diagnostic technique. Detrusor muscle of the bladder may hypertrophy in response to increased workload of frequent involuntary contractions which are the hallmark of DO (Khullar et al. 1994). The hypertrophied detrusor may be visualized as thickened bladder wall on ultrasound scan. A systematic review was carried out in 2010 on the accuracy of BWT on transvaginal scan in the diagnosis of DO showed a trend towards increased BWT in women with DO(Latthe, Champaneria, & Khan 2010). In order to know whether there is any recent robust evidence on the diagnostic accuracy of BWT and to inform whether BWT can be used as an alternative to urodynamic diagnosis of DO, a comprehensive update of systematic review was carried out.

2.1Evidence for the Accuracy of Bladder Ultrasound as a test for diagnosing DO

An update of the systematic review of test accuracy studies on women who had an ultrasound scan to measure bladder wall thickness (BWT) against the reference standard of urodynamics to confirm DO was performed.

2.2 Methods

After formulating a protocol a priori using recommended methods (Khan et al. 2001), we carried out a systematic review by independent observers (SR and PL). A meta-analysis was performed using Rev Man 5.3 software and pooled estimates of test accuracy were calculated.

2.3 Search strategy

Electronic searches were conducted in Medline, MEDION, EMBASE, LILACS, CINAHL and websites for abstracts from the annual meetings of ICS, IUGA, UKCS, EUA, AUGS from database inception to March 2015. We also carried out hand searches from the reference list of published articles. We used combinations of MeSH subheadings and key words. Our key words included ‘ultrasound’, ‘bladder’, ‘bladder wall thickness’, ‘detrusor wall thickness’, ‘overactive bladder’ and ‘detrusor overactivity’. We adapted a search strategy for each database and restricted our searches to studies on ‘women’. We did not apply any language restrictions or methodological filters, as this might lead to the omission of relevant papers.

2.31 Identification of studies and data extraction

Study selection was undertaken by electronic searches without any language restrictions. Selected studies were thoroughly scrutinized and the manuscripts of all the citations were obtained for a critical review. The four part question was:

1. Population: Adults with urinary urgency, with or without frequency, nocturia and urgency incontinence

2. Index Test: Bladder or detrusor wall thickness

3. Reference standard: Urodynamic (laboratory and/or ambulatory) observation of DO

4. Study design: All studies in which patients had undergone both the index test (ultrasound scan for bladder wall or detrusor wall thickness and reference standard, urodynamics

SR and PL independently scrutinised all of the manuscripts to determine if they met the above criteria. Data on study characteristics, quality and results was extracted onto a data collection proforma. Data were also extracted on patient characteristics, methods of

recruitment and the procedural details of the two diagnostic tests- ultrasound and urodynamics. Any disagreements between the two reviewers on the inclusion/exclusion of a manuscript or on the extracted data were resolved by consensus. We have tried to contact the primary authors for missing data or additional data but there was no response.

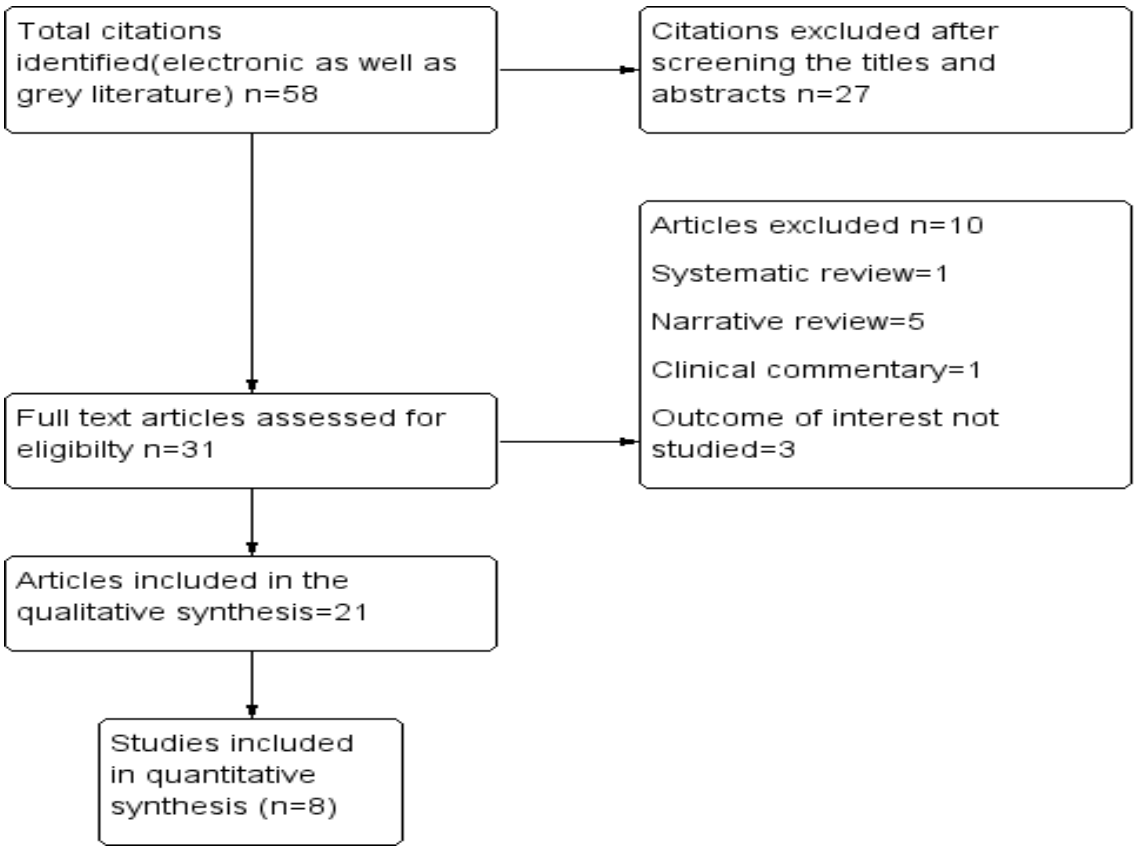
2.32 Methodological quality assessment

We assessed the robustness of the methodology of the included manuscripts using the QUADAS checklist (Whiting et al. 2003). We looked at the appropriateness of the study design, relevant features of the study population, conduct of the index test and the reference standard of each study. We considered the study to be of good quality if it satisfied the QUADAS criteria. We have assessed the risk of bias in the included studies based on the criteria from Cochrane handbook for diagnostic test accuracy reviews.

2.33 Data Synthesis

To summarize sensitivity and specificity data for BWT, we constructed a Forest plot from the 2x2 contingency tables cross-classifying index (BWT) test results and the reference standard (urodynamic diagnoses) using the Rev Man 5.3 software(Figure 5). For the meta-analysis, we have used the hierarchical summary receiver operating characteristic (HSROC) model to accommodate the limitations (paucity of robust data). As we were unsure whether the diagnostic accuracy was related to BWT threshold, we used allowances for asymmetry of the HSROC curve where the $\beta \neq 0$.We have calculated the pooled sensitivity, the specificity of the collated results of the studies included in the systematic review from the summary point of the SROC curve.

Figure 2: Flow chart for the selection of studies included in the update of systematic review of the accuracy of bladder wall thickness in detrusor overactivity



2.4 Results

Figure 2 summarises the process of systematic search and of study selection. A total of 54 relevant studies were identified. Out of the 54 articles, 21 studies (18 prospective and 3 retrospective) satisfied our inclusion criteria for the systematic review. The studies on pooling had 4151 women. Ten of the these 21 studies were case control in design((Abou-Gamrah et al. 2014;Chung et al. 2010;Ibrahim S and Najdy M 2011;Kuo 2009;Minardi et al. 2007;Otsuki E N et al. 2014;Ozturk H et al. 2011;Parsons, Amundsen, Cardozo, Vella, Webster, & Coats 2007;Soligo M et al. 2002;Yang and Huang 2002)

Figure 3: Quality of studies included in the update of systematic review of the accuracy BWT in the diagnosis of DO

(stacked bar chart used with numbers inside bars indicate the number of studies)

Transabdominal ultrasound was used in six studies at different bladder volumes to measure detrusor wall thickness (DWT)(Blatt et al. 2008;Chan L et al. 2005;Chung, Chiu, Kuo, Chuang, Wang, Guan, & Chancellor 2010;Kuo 2009;Ozturk H, Aydur E, Irkilata H, Seckin B, & Dayanc M 2011;Silva et al. 2014)(Tables 3 and 4) There was increased DWT(0.75mm) on transabdominal scan when measured at maximum capacity in women with DO with a sensitivity (73%) specificity (67%).and an AUC(0.776). However, women with DO had a lower bladder capacity compared to women without DO and no correction was employed to accommodate this variation (Kuo 2009). In other studies on transabdominal BWT, no such

increase was found at bladder capacity and at the first sensation of void (Chan L, The S, Tse V, & Titus J 2005).

Twelve studies suggested that BWT may be useful to diagnose DO, although the reported sensitivity/specificity/area under the curve(AUC) varied amongst the studies (Tables 3 and 4) (Ibrahim S & Najdy M 2011;Khullar, Salvatore, Cardozo, Bourne, Abbott, & Kelleher 1994;Khullar, Cardozo, Salvatore, & Hill 1996;Kuhn, Genoud, Robinson, Herrmann, Gunthert, Brandner, & Raio 2011;Minardi, Piloni, Amadi, El, Milanese, & Muzzonigro 2007;Otsuki E N, Júnior E A, Oliveira E, Castelo Girão M J B, & Jármy-Di Bella Z I K 2014;Ozturk H, Aydur E, Irkilata H, Seckin B, & Dayanc M 2011;Parsons, Amundsen, Cardozo, Vella, Webster, & Coats 2007;Robinson et al. 2002;Serati et al. 2010). Of these, five were prospective cohort studies and seven were case control studies. BWT was found to be increased in women with DO when compared to BWT with in patients with other types of incontinence in the above mentioned studies.

In the study by Khullar et al, 42 women who had BWT of > 5mm and 4 women who had BWT of <3.5mm and did not have DO on urodynamics were further subjected to a secondary diagnostic test-the ambulatory urodynamics. Of these 36 women were diagnosed to have DO. So, 25% of the study participants had two reference standard tests leading to differential verification bias (Khullar, Cardozo, Salvatore, & Hill 1996)(Table 2).

In another study by Robinson et al, the patients included were 128 women with OAB who did not have DO on conventional urodynamics and were subjected to ambulatory urodynamics which is a different reference standard. Hence this group is not representative of the standard

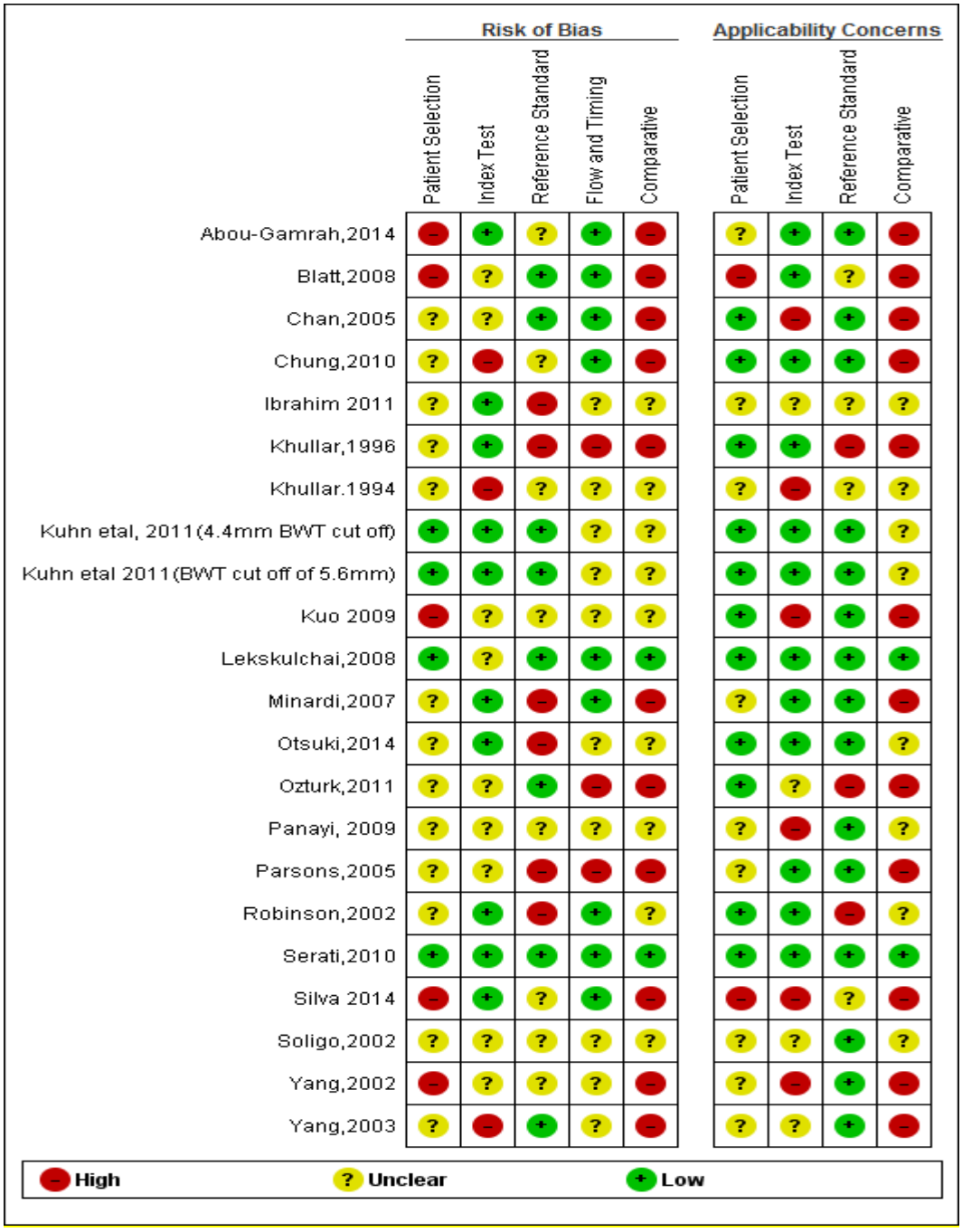
OAB population. The BWT was studied as a secondary diagnostic test rather than a primary index test and so the results of this study may not be applicable in routine clinical practice(Robinson, Anders, Cardozo, Bidmead, Toozs-Hobson, & Khullar 2002)(Table 2).

On review of test accuracy data, the sensitivity ranged from 40-90.6% and specificity from 78-96.6% for a mean BWT of 5 mm (Abou-Gamrah, Fawzy, Sammour, & Tadros 2014;Khullar, Cardozo, Salvatore, & Hill 1996;Kuhn, Genoud, Robinson, Herrmann, Gunthert, Brandner, & Raio 2011)(Table 4).

Only 7/21 studies reported on construction of the ROC curves (Abou-Gamrah, Fawzy, Sammour, & Tadros 2014;Ibrahim S & Najdy M 2011;Kuhn, Genoud, Robinson, Herrmann, Gunthert, Brandner, & Raio 2011;Kuo 2009;Lekskulchai & Dietz 2008;Serati, Salvatore, Cattoni, Soligo, Cromi, & Ghezzi 2010;Serati, Cattoni, Siesto, Braga, Sorice, Cantaluppi, Cromi, Ghezzi, Vitobello, Bolis, & Salvatore 2013). In a study by Kuhn et al, the ROC curve was used to predict forms of urinary incontinence and to demonstrate the relationship between SUI, OAB and obstructed outflow rather than to predict the diagnostic accuracy in each of the different urodynamic diagnoses. In the other six studies with the ROC curves, the AUC ranged from 0.606 to 0.776 (Table 4).

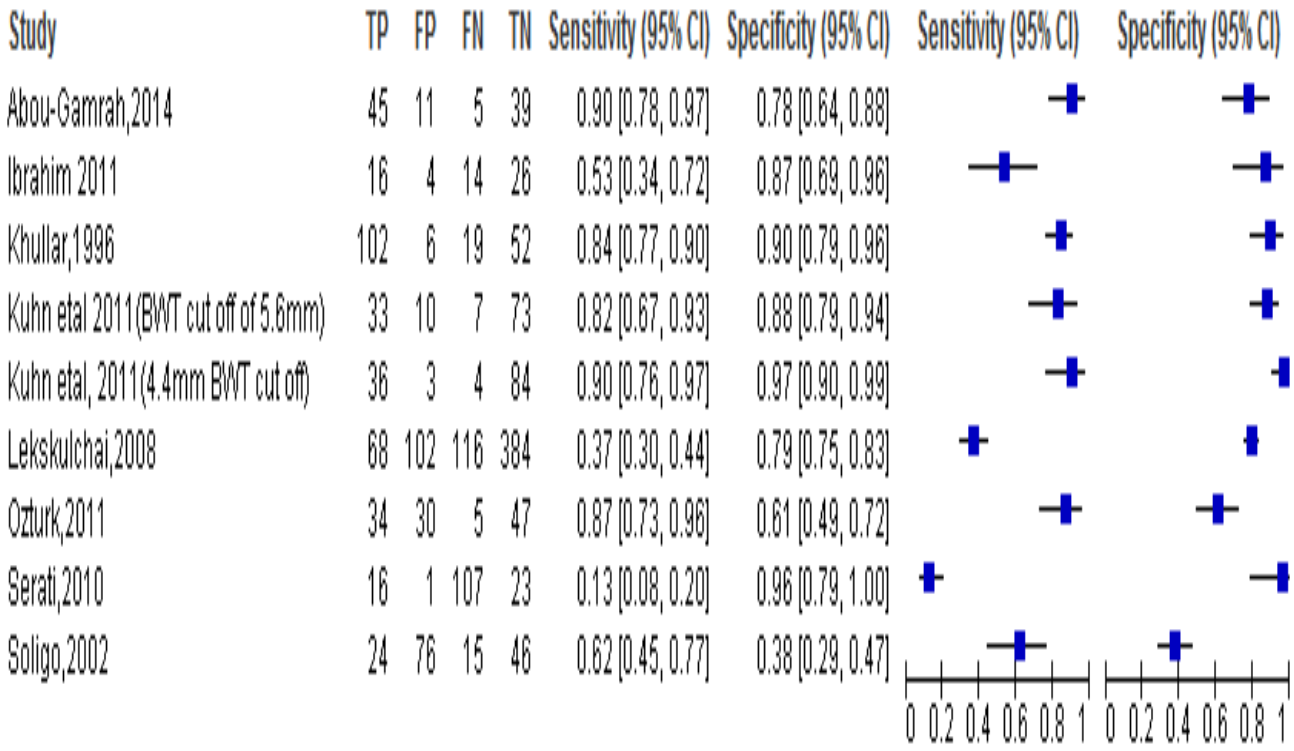
On quality assessment based on QUADAS checklist, there was a significant variation across the studies (see figure 3). The route and the technique of performing the BWT scan, the urodynamic tests, (video-urodynamics, laboratory urodynamics and a combination of laboratory and ambulatory urodynamics), were highly variable across the studies. On risk of bias assessment, 18/21 studies were found to be ‘high in concern’ (Figure 4).

Figure 4: Risk of bias assessment



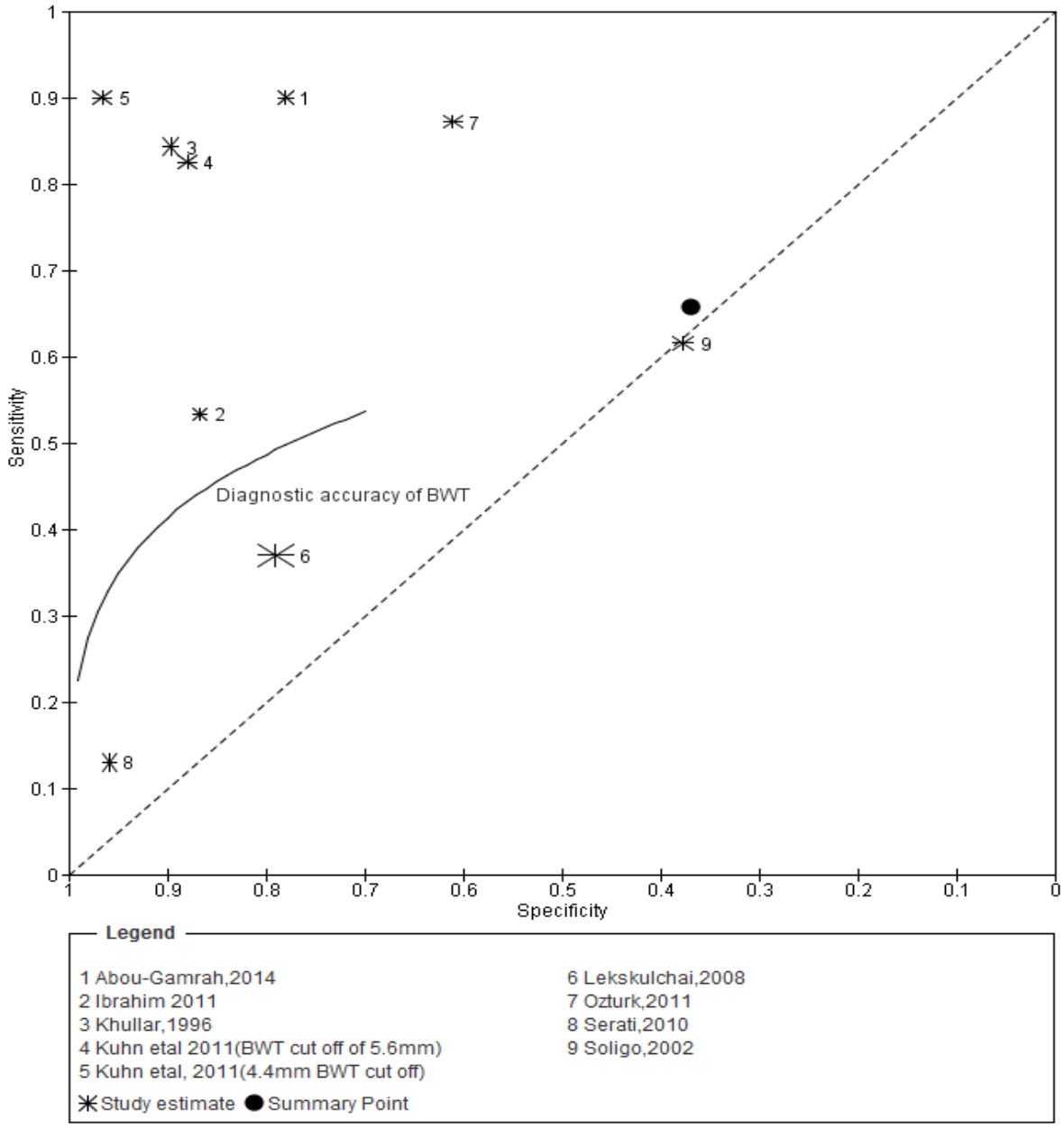
Visual inspection of forest plot has shown a significant variation of sensitivities and specificities was found for individual studies (Figure 5).

Figure 5: Forest plot of sensitivities and specificities for BWT in diagnosing DO



Using the Rev Man 5.3 software, I have constructed a HSROC curve. Each of the included studies in the meta-analysis is represented by a star on the HSROC curve.The size of the each star has varied according to the sample size of the study.The solid dot on the HSROC graph is the summary point estimate of sensitivity and specificity of studies included in the meta-analyses. The summary point has given a pooled estimate of sensitivity of 66% and specificity of 38% on SROC curve (Figure 6)

Figure 6: Hierarchical SROC curve of test accuracy data from meta-analysis on diagnostic accuracy of BWT in diagnosing DO



2.5 Discussion

2.51Main findings

The pooled estimates of sensitivity and specificity suggest questionable the reliability of BWT as a diagnostic test for OAB patients. The studies suffer from significant design flaws like underpowered and insufficient or inaccurate statistical evaluation.

2.52 Strengths & weaknesses

The strengths of this review include a thorough and systematic search using the recommended methods and a predesigned protocol. The study selection and extraction of the data was carried out by two reviewers independently. We attempted to include the grey literature to reduce the risk of publication bias. Of the 21 studies which fulfilled our inclusion criteria, six of them have been conference abstracts and 15 were published articles. We have assessed the quality of the included studies using the QUADAS checklist (Figure 3) and the risk of bias (Figure 4) based on the criteria laid out by the Cochrane handbook for systematic reviews for diagnostic test accuracy.

Small case-control/cohort studies on BWT in diagnosing DO seem to have erroneously amplified the effectiveness of BWT as a diagnostic marker of DO (Lijmer et al. 1999).Nearly half of the studies(10/21) included were case control in design(Tables 2 and 3). None of the studies were adequately powered to study the diagnostic accuracy of BWT in DO. The inclusion/exclusion criteria were not stringent enough to exclude the confounding factors. One study included a significant number of women with previous stress incontinence surgery who developed OAB symptoms. Inclusion of patients with iatrogenic partial bladder outflow obstruction which is an independent risk factor for bladder wall thickening might have altered the results of this study (Kuhn, Genoud, Robinson, Herrmann, Gunthert, Brandner, & Raio 2011).

Included studies showed significant heterogeneity in the site and route of BWT and also in the reference standards used. The techniques of measurement-BWT(full thickness of the bladder wall) or DWT(thickness of only the detrusor excluding the serosa and mucosa of the

bladder), sites of measurement(tables 2 and 3) (only dome thickness compared to that of an average BWT measurement were variable across different studies (Khullar, Salvatore, Cardozo, Bourne, Abbott, & Kelleher 1994;Robinson, Anders, Cardozo, Bidmead, Toozs-Hobson, & Khullar 2002) (Lekskulchai and Dietz 2008). Each of these techniques have their own weaknesses and so far have not been standardised (Fransisco Cruz et al. 2009).

The statistical evaluation in the studies ranged from point estimates of sensitivity, specificity, positive and negative predictive values to construction of ROC curves and the estimation of AUC. Thirteen of the included studies estimated mean BWT/DWT in the group with DO and compared it with those without DO. Their inaccurate study design and the methodology flaws were reflected in poor data presentation. Consequently 13/21 studies could not be included in the meta-analysis. Only 8 studies were included in the meta-analysis and these studies have exhibited significant inter-study heterogeneity with regards to study design, technique and route of BWT, cut off of BWT used and their sample sizes. Hence the results of meta-analysis should be interpreted with caution.

2.53 Interpretation of findings and conclusion

All of the studies have significant methodological flaws which might have had an impact on the results questioning the usefulness of BWT as a diagnostic marker of DO. The pooled estimates indicative of low sensitivity and specificity from the SROC curve on the diagnostic accuracy of BWT could be due to the poor performance of the test itself or due to other factors like inappropriate study design, inadequate sample size and variable techniques and cut-offs of BWT used. No firm conclusion could be drawn from the systematic review due to the poor quality of the included studies. An appropriately designed, sufficiently powered

diagnostic accuracy study is needed to estimate the role of BWT and hence the BUS study was planned.

Table 2: Characteristics of the included studies in the update of the systematic review of accuracy of transvaginal/ translabial bladder / detrusor wall thickness in detrusor overactivity

Study, date, country, design	**Population**	**Test**	**Reference Standard**
Otsuki et al, 2014,Brazil, prospective case control study	n=91 women Cases:30 with DO Controls:31 with no incontinence and 30 with SUI	Transvaginal BWT measurements in the parasagittal plane at anterior wall, trigone and dome	Urodynamics only in cases –DO and SUI. Controls did not undergo urodynamics.
Abou-Gamrah et al,2014, Egypt, Prospective ?Case control study	Sample size calculation: 41 women in each of the two groups to detect a difference of 1 mm in mean BWT between both groups with a standard deviation (SD) of 1.6 with power of 80 % and alpha error 0.05. DO=50 women USI=50 women	Transvaginal ultrasound at a postvoid residual (PVR) of <50ml by an operator who was blind to urodynamics result. BWT at three sites: thickest part of the dome, the trigone, and the anterior wall.	Urodynamics performed as per ICS guidelines
Kuhn et al, 2011,UK, Prospective Cohort study	122 women: Previous incontinence surgery=39, SUI=59, DO=40, Obstruction =24	Transvaginal BWT was performed in all women at PVR<50ml. Technique not described. Clinician measuring BWT was blinded to urodynamics.	Urodynamics performed in all women
Ibrahim et.al,2011.UK, Egypt Prospective case control study(Conference abstract)	60 women Detrusor instability=30 Healthy controls=30	Transvaginal BWT was performed at <50ml at the trigone, dome and anterior wall. Clinician measuring BWT was blinded to the urodynamics result	Urodynamics was performed only in cases. The authors do not seem to make a distinction between detrusor instability and OAB.
Panayi. et al,2010,UK Prospective (Conference abstract)	182 women underwent both urodynamics and transvaginal BWT scan	BWT was measured at trigone, dome and anterior wall	All the participants underwent urodynamics
Serati. et al 2010	247 women	All the study participants had urodynamics and transvaginal	Urodynamics in all women according to

Italy, Prospective Cohort		BWT ultrasound scans by the technique described by Khullar et al. BWT scan operator was blinded to urodynamics result.	Good Urodynamics Practice guidelines of the ICS
Lekskulchai et al 2008 Australia Retrospective cohort	686 women underwent multichannel urodynamics and translabial ultrasound	Translabial ultrasound for DWT measurement(the iso-to hypo-echogenic layer at the bladder dome opposite the internal urethral meatus and two additional measurements within 2 cm of the mid-sagittal plane)	Multi-channel urodynamics confirmed to ICS standards
Minardi et al 2007 Italy Prospective Case-Control	80 women- 66 cases with SUI=36 Urgency incontinence=30 Controls=14	DWT was measured at bladder dome Unsure whether translabial or introital ultrasound was used for DWT measurement.	Urodynamics according to ICS criteria
Parsons et al, 2005,UK Prospective case control (Conference abstract)	250 women: Cases=194, Women with DO=31 Controls=61 Withdrawals=26	Transvaginal BWT at trigone, dome and anterior wall.	Video urodynamics in all women 31 women (18.5%) had DO
Yang et al 2003 Taiwan Retrospective	492 women with LUTS who had undergone BWT and urodynamics and who had normal urinalysis findings, negative urine culture results, or both. DO=38, SUI=248, MUI=39 Hypersensitive bladder=35	BWT at dome or trigone using transvaginal ultrasound	Urodynamics at a filling rate of 80 ml/min with patient sitting upright in a birthing chair.
Yang et al 2002, Taiwan, Retrospective case control	1049 women- SUI=764, Detrusor instability=190, Hypersensitive bladder=95. Controls=36	BWT at dome or trigone using transvaginal ultrasound	Urodynamics= uroflowmetry, filling and voiding phase cystometry, a urethral pressure profile at both resting and at valsalva. Detrusor instability was diagnosed if there was a detrusor contraction in association with urgency, leakage, or both
Soligo et al, 2002, Italy and UK Prospective case control study (Conference	161 women- OAB=70 Normal controls=91. Detrusor instability=24	Sites of BWT measurement not elaborated	Urodynamics performed in all women

abstract)			
Robinson et al 2002 UK, Prospective	128 women with OAB with normal or equivocal urodynamics referred for ambulatory urodynamics. Clinician performing urodynamics were blinded to BWT	Measurement of transvaginal BWT (Khullar's technique)	Ambulatory urodynamics- DO diagnosed if detrusor pressure rise was recorded in association with symptoms of urgency and/ or urge incontinence
Khullar et al 1996 UK, Prospective	184 patients attending clinic for laboratory urodynamics DO=107 women Normal urodynamics but with a mean BWT >5mm who underwent ambulatory urodynamics=42	BWT after emptying bladder post micturition residual was checked to ensure <50 mls. The measurements were made at maximum magnification in 3 places: Perpendicular to the luminal surface at the thickest part of the trigone, dome and the anterior wall .Mean BWT = dome+ anterior wall+ trigone/3	Videourodynamics per ICS standard BWT <3.5mm or with >5mm but no DO went onto have ambulatory urodynamics on a separate day Urodynamics observer was blinded to BWT results
Khullar et al 1994 UK, Prospective	45 women Detrusor instability =19 SUI =20.	Technique as described above. Operator performing BWT was blinded to urodynamic diagnosis	Video urodynamics in supine position

Table 3: Characteristics of the included studies for the systematic review on BWT by transabdominal BWT/DWT

Study, date, country, design	**Population**	**Test**	**Reference Standard**
Silva et al, 2014,Brazil, Cohort study-Prospective	272 women with spinal cord injury	BWT measured transabdominally midway between anterior wall of the bladder and the lateral bladder wall	Multichannel urodynamics along with urethral external sphincter myography
Ozturk et al 2011,Turkey, Case-control Study Conference abstract	82 women: DO=39, SUI=43, Controls=31	Transabdominal DWT at 200ml volume at three different places-anterior wall, right and left lateral wall.	Women were categorised into SUI and DO based on urodynamics
Chung et al, 2010, Taiwan, Prospective Case control	122 women-83 cases Normal urodynamics=28 Increased bladder sensation=30 DO=30 Normal controls=39	Transabdominal DWT was measured at natural bladder filling and during catheter filling.	88 Women underwent Video-urodynamics
Kuo 2009,Taiwan Prospective Case control	92 women OAB=81 Controls=11 Transvaginal detrusor wall thickness(DWT) scan on empty bladder and transabdominal DWT at bladder capacity	Transvaginal DWT was measured at bladder neck, bladder base, anterior and posterior wall Trans-abdominal DWT was measured on the anterior wall at three sites	All women underwent video-urodynamics
Blatt et al 2008, Australia, Prospective cohort	180 patients including 107 women. Normal urodynamics=69, BOO=39, Increased bladder sensation=38 DO=34	Mean BWT of two transabdominal BWT measurements on the anterior bladder wall, 1 cm apart in the midline at 200 ml bladder volume	Video urodynamics as per the ICS standards
Chan et al,2005,Australia, Prospective cohort study,(Conference abstract)	86 women with OAB: Normal urodynamics=42, Sensory urgency=22 DO=22	Transabdominal BWT. Two measurements of anterior BWT (1 cm apart in midline at 200 mls filling and at capacity	Urodynamics method not specified

Table 4: Results of studies included in the systematic review of the accuracy of ultrasound measurement of BWT in diagnosing DO

Study, date, country, design	**ROC/AUC**	**Sensitivity**	**Specificity**	**Mean among those with DO(+/- 2SD or 95% confidence interval)**	**Mean among those without DO(+/- 2SD or 95% confidence interval)**	**Comments**
Otsuki et al, 2014,Brazil, prospective case control study	The ROC revealed an area under the curve of 0.962 (95% CI, 0.90–1.01) for BWT as a diagnostic marker.			6.2mm	4.5mm	BWT and vesical pressure at involuntary detrusor contraction (r = 0.39, P = 0.017) were directly correlated
Silva et al, 2014,Brazil,	AUC = 0.624, 95 % CI (0.530, 0.718), p = 0.011			4.2mm in neurogenic DO/DSD	3.6 mm in reduced bladder compliance	
Abou-Gamrah et al,2014, Egypt	AUC =0.905	90%	78 %			BWT cut-off used=4.78
Kuhn et al, 2011,UK, Prospective Cohort study	A single ROC curve was constructed for all the three conditions-SUI, DO and obstruction. AUC =0.87 (95% CI 0.78–0.97; P < 0.0001) for predicting different kinds of UI!	90.6% (4.4 mm BWT cut-off) 83.3% (5.6mm cut-off)	96.6% (4.4 mm BWT cut-off) 87.5% (5.6mm cut-off)	4.97±0.63 mm	3.78 ±0 .39mm	Outflow obstruction: 6.01± 0.73mm (P < 0.0001)
Ozturk et al 2011,Turkey		87.1%	60.8%			At a cut off of 4.88mm PPV =53.9% NPV=90%
Ibrahim et.al,2011.UK, Egypt	AUC =0.73.	70%	74%	7.79mm	6.60 mm	BWT cut off of 4.48mm, PPV =73% and NPV = 71% for OAB
Panayi. et al,2010,UK						BWT > 5mm in 65% of women with DO BWT > 5mm in 85% of women with OAB
Chung et al, 2010,						No difference of DWT in women with DO compared to

Taiwan						OAB without DO or from the asymptomatic control group at a bladder volume of 250–300ml. No difference in the DWT between natural filled and catheter filled bladders at 250ml volume
Serati. et al 2010 Italy,	AUC for urgency incontinence=0.645, pure DO= 0.702 and all DO (provoked as well as spontaneous DO) = 0.704.					Positive predictive value=100% for all DO at a cut-off of 6.5mm
Kuo 2009,Taiwan	*Natural filling*: On ROC curves, transabdominal DWT cut off of 0.75 mm at bladder capacity(AUC of 0.776, standard error 0.068, 95% confidence interval 0.643, 0.909) *Catheter filling*: The AUC was 0.648, standard error of 0.078, 95% confidence interval 0.495, 0.802)	73%	67%			No difference in transvaginal DWT at the bladder neck, anterior wall, posterior wall, and bladder base among women with DO, OAB without DO or normal controls. Transabdominal DWT was greater in DO group at maximum capacity.
Lekskulchai et al 2008 Australia	AUC= 0.606 (0.56-0.65)	37%	79%	DWT in DO=4.7+/-1.9mm	DWT in normal women=4.1 ± 1.6mm (P<0.001)	For 5 mm DWT cut-off
Blatt et al 2008, Australia,				1.9mm	Normal urodynamics=2.0 BOO=2.1, Increased bladder sensation =1.8 mm	No difference in the mean BWT between the groups (ANOVA p = 0.064)
Minardi et al 2007 Italy				Urgency incontinence= 7.1 +/1.6mm	USI=4.1 +/-1.1mm Controls=3.9 +/-1.9mm (P=0.019)	
Parsons et				4.862mm	4.085mm in OAB without DO	

al,2005,UK					Controls=3.92	
Chan et al,2005,Australia				1.7mm	Normal urodynamics=1.7mm, Sensory urgency=1.6	No difference between BWT at 200 mls and at cystometric capacity. No difference between BWT in normal urodynamics vs BWT in DO and sensory urgency. Mean BWT in patients and DO = 1.7 mm, (p = 0.18).
Yang et al 2003 Taiwan	DO= Thickened bladder wall found in all LUTS except hypersensitive bladder-a nonspecific marker for several LUTS			5.5 (5-6.6)mm	USI= 5.6-6.4 mm MUI= 5-6.2mm	
Yang et al 2002, Taiwan, Retrospective case control	Mean BWT of Detrusor instability (DI)-,			5.8 ± 1.9	SUI-6.0 ± 2.4, Hypersensitive bladder=5.3 ± 1.9 Controls= 4.9 ± 2.1	Negative correlation between BWT at trigone and dome to that of resting bladder neck angle (P =0.006 and 0.019, respectively)
Soligo et al, 2002, Italy and UK				5mm with 95% CI of 4.6 and 5.3	3.6mm CI of 3.4 and 3.9	PPV of BWT=83.3% NPV= 83.2%
Robinson et al 2002 UK, Prospective				6.7 (95% CI 6-7.4)mm	Normal=5.1mm (95% CI 4.6-5.6) USI=4.8 mm (95% CI 4.4-5.3) MUI=5.8mm (95% CI 5.1–6.5)	
Khullar et al 1996 UK	For a BWT cut off of 5mm, Sensitivity= 84 Specificity=89	84 (75.8 - 89.7)%	89(78.8 - 96.11) %			
Khullar et al 1994 UK	Mean BWT			6.7	GSI =3.5	

Chapter 3

Accuracy of bladder wall thickness on transvaginal ultrasound in diagnosing Detrusor Overactivity: Materials and Methods

Introduction

From the systematic review in chapter 2 it was obvious that there was a lack of consensus on the standardisation of the technique, route of BWT, optimisation of cut off in diagnosing DO. A prospective, appropriately designed, adequately powered, multicentre diagnostic accuracy study was planned [Accuracy of bladder wall thickness ultrasound in the diagnosis of Detrusor Overactivity (BUS study)]. It was approved by Nottingham Research Ethics Committee (ethics no10/H0408/57) and funded by the National Institute of Health Research/ Health Technology Assessment Board. The study has been carried out in 22 centres in the UK. The units also represented the spectrum of patients, from a busy district general hospital to tertiary referral centres. Using recommended methods for diagnostic accuracy evaluation (Irwig et al. 2002), a prospective study protocol was developed with a classic test accuracy design and reported according to the STARD standards (Bossuyt, Reitsma, Bruns, Gatsonis, Glasziou, Irwig, Lijmer, Moher, Rennie, & de Vet 2003). Nottingham Research Ethics Committee and NHS trust research governance approval was obtained for 22 recruiting hospitals in the UK with the Birmingham Women's Hospital and University of Birmingham acting as the study sponsors. Recruiting hospitals included both specialised tertiary referral centres and district general hospitals. Our study population was drawn from a large socioeconomically and ethnically diverse group of women. The study was monitored by an independent data monitoring committee.

3.1 Study design

The study evaluated the accuracy of bladder wall thickness ultrasound (BWT) in making a diagnosis of DO using laboratory multichannel urodynamics as the reference standard. For the index test, BWT was measured on transvaginal ultrasound scans, which is a continuous variable reported in millimetres.

3.2 Eligibility

Inclusion criteria

- Frequency of 9 or more voids in 24 hours on at least 1 out of the 3 days on bladder diary
- Mild – severe urgency on at least one occasion in 3 day bladder diary
- Post void residual (PVR) volume <100 mls on screening
- Written informed consent
- If patient has had previous stress urinary incontinence (SUI) surgery &/or Botox, it was >6 months ago

Exclusion criteria

- Pregnancy and up to 6 weeks postpartum.
- Symptoms of SUI or stress predominant mixed incontinence
- Evidence of cystitis (dipstick positive for leucocytes/nitrites)
- Voiding difficulties (PVR >100 ml)
- Prolapse > grade II (any compartment)
- Urodynamics assessment in the past 6 months
- Use of antimuscarinics for more than 6 months continuously
- Current use of anti-muscarinics (e.g. Tolterodine, solifenacin and oxybutynin)

- If the woman is taking anti-muscarinics for <6 months at the point of consent, she will be eligible if the medication is ceased immediately and there is a delay of at least 2 weeks until the index and reference tests are carried out

3.3 Sample size

A minimum target sample size of 600 women was pre-specified in order to obtain estimates of sensitivities and specificities with 95% confidence intervals of width 10% or less anticipating sensitivity and specificity values between 70% and 95%. The computation was based on a prevalence of 50% for DO (Hashim & Abrams 2006), providing 300 women each for the estimate of sensitivity and specificity.

3.4 Recruitment of participants and consent

The potential participants referred with frequency and urgency to the urogynaecology, urology or urodynamics clinics were identified and study information leaflets were sent along with their appointment letters. Study information leaflets were also given to women who presented primarily with prolapse (grade 2 or less) but had coexisting urinary frequency and urgency symptoms.

Patients who fulfilled the entry criteria were approached and consent was sought from the willing in a two stage process. The study information leaflets along with a sample consent form and bladder diaries were posted to all prospective participants with their clinic invitation letter. Research nurses and principal investigators in the recruiting hospitals were trained to consent participants by reinforcing the information provided and answering any questions that the women may have had. Women were provided enough time for consideration and opportunity to ask questions and then approached for consent and study recruitment.

Wherever necessary, appropriate interpreters were used to aid discussion relating to study participation.

All of the recruiting nurses and the doctors were trained regarding the introduction of information about the study and instruction on their roles from the local coordinating clinicians and the by the Clinical Research Fellow. Team meetings were organised and newsletters periodically sent to reinforce the study recruitment and procedures from the Study Office. A screening log was maintained for the women screened including those who agreed to participate and also those who declined participation.

Recruitment was organised and supported by a dedicated team under the supervision of clinical principal investigator (PI). Documentation was provided by the BUS Study Office and the clinical research fellow supported clinics in some centres local to Birmingham. The Clinical Research Fellow (SR), liaised with the local PI at each centre, provided ultrasound scan (USS) training, recruited patients at the main centre, dealt with any problems with recruitment and conducted quality assurance of the scans.

3.5 Setting of tests

Urodynamics was carried out by health care professionals (doctors or nurses) who were already carrying out the procedure in routine clinical practice. For BWT scanning, hands on training was delivered on site for each of the recruiting sites for the clinicians (doctors or sonographers) and two scan training workshops were carried out at the Birmingham Women's Hospital.

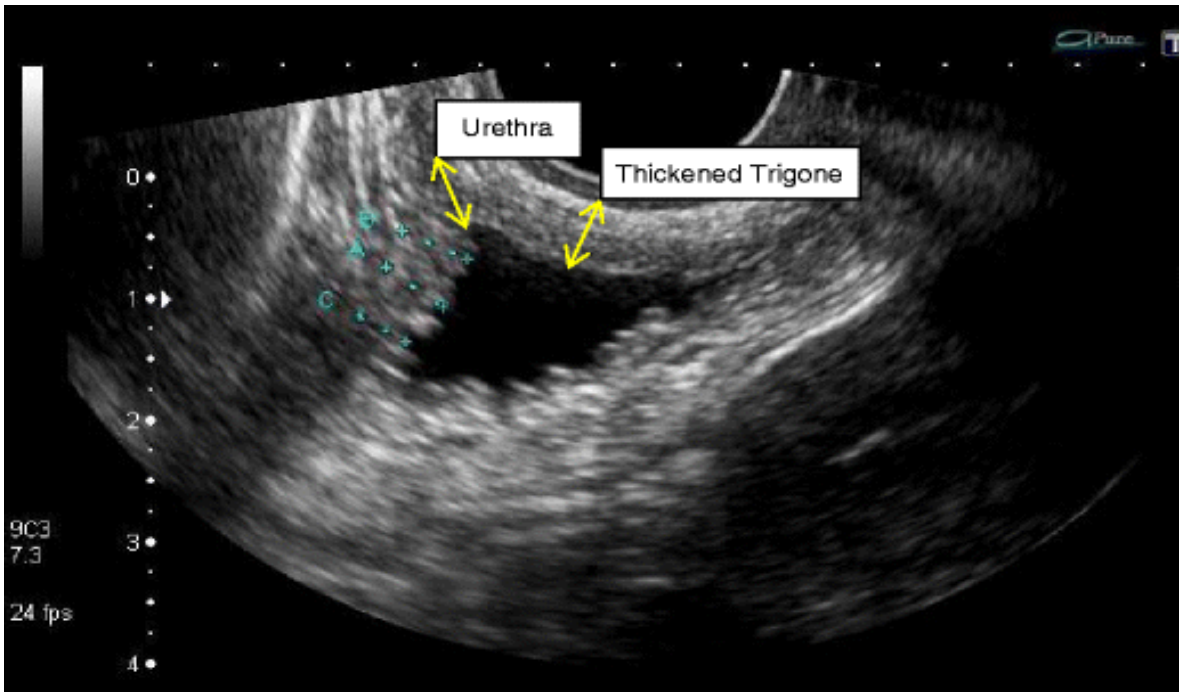
At centres, where the BWT scan was done in the radiology department, the urodynamics was carried out blinded to the ultrasound result, within 4 weeks of the BWT scan. If this 4 week cut-off period was breached, there was a further 4 week window in which the second test data was collected, but classed as a protocol violation. Both the urodynamics and BWT were performed by independent observers blinded to each other.

3.6 The index test: Bladder ultrasound

BWT was measured with the transvaginal ultrasound probe in the sagittal plane, introduced 1cm beyond the vaginal introitus in the midline. The participants were advised to pass urine and the PVR urine was measured using the technique described by Haylen (Haylen 2007). The BWT was measured at a PVR volume of less than 30ml as the thickness was found to be fairly constant when measured in the range of bladder volumes of 0-50ml (Khullar, Salvatore, Cardozo, Bourne, Abbott, & Kelleher 1994). If the bladder had >30 mls PVR, participant was asked to double void and the PVR was measured again. If the PVR was </=100mls then the patient was included in the study.

The thickness of all the three layers of the bladder wall- serosa, the detrusor muscle and the urothelium was measured on transvaginal ultrasound and a mean BWT calculated. The sites of BWT measurement on the anterior wall, trigone and dome was chosen based on the previous evidence (Khullar, 1994) (Robinson, 2002) .The average BWT calculated from three site measurement of anterior wall, dome and trigone might be a better representative of BWT better than by measuring dome thickness alone. The technique and settings for BWT has been described in chapter 1 pages 27 and 28. Figure 7 depicts the measurement of a thickened bladder wall.

Figure 7: Thickened bladder wall as seen on transvaginal scan



The margins of the bladder wall were found to be more clearly visualised by the transvaginal route, as the more distant the probe was from the visualised structure, the higher the attenuation of ultrasound waves resulting in poor-quality images. The interobserver variability in BWT measurement was found to the lowest in transvaginal scan compared to transabdominal and transperineal scanning(Panayi, 2010). Hence we have chosen transvaginal route in measuring BWT. For the standard operating procedure followed

for BWT, please see Appendix 6 a.

3.7 The Reference Standard-Urodynamics

Urodynamics is considered ‘gold standard’ for investigation of LUTS and has been universally adapted (Abrams 2006). Urodynamics is believed to provide an explanation to the pathology behind LUTS by reproduction of patients symptoms(Abrams et al. 2009). The test is accompanied by imaging during video urodynamics.

Urodynamics are performed in three stages (Abrams 2006):

1. Uroflowmetry

2. Filling cystomtery

3. Voiding cystometry

Uroflowmetry is a noninvasive study of assess voiding pattern of the bladder. Uroflowmetry is carried out by asking the patient with a reasonably full bladder to void in a flowmeter in privacy. Uroflowmetry is used to diagnose bladder outflow obstruction and also as a measure of voiding function (Abrams 2006).

Cystometry is an interactive test of the storage (filling cystometry) and the voiding function (voiding flowmetry) of the bladder. The filling phase is recorded on cystometrogram as variations in detrusor pressure with incremental filling. Detrusor pressure is calculated by the subtraction of abdominal pressure from the vesical pressure. The rate of filling the bladder should be based on patient's symptoms. Bladder sensation, bladder pain, detrusor activity, compliance of the bladder wall and cystometric capacity may be studied during filling cystometry (Abrams 2006;Abrams, Artibani, Cardozo, Dmochowski, van, & Sand 2009).

The coordination between detrusor activity and urethral function may be studied during voiding cystometry. Detrusor underactivity, detrusor sphincter dyssynergia and dysfunctional voiding may be diagnosed during this phase of urodynamics (Abrams 2006). Provocation tests like coughing and running taps are carried out to detect provoked DO (Abrams 2006).For the standard operating procedure followed for urodynamics, please see

Appendix 6 b.

Laboratory urodynamics is chosen as the reference standard as it is the most common urodynamic procedure used in investigating LUTS around the world. Laboratory urodynamics has been criticized to be not a very physiological test due to the short duration involved, retrograde filling using a urethral catheter and the rapid filling of the bladder which might be much faster than the physiological rate of urine production.

Ambulatory urodynamics may be a used when patients with troublesome urinary incontinence have normal urodynamics (Swithinbank et al. 1999). During ambulatory urodynamics, bladder is allowed to fill naturally, bladder pressure is measured by a micro-transducer and the patient is allowed to move around mimicking a physiologic state. Ambulatory urodynamics may have a higher pickup rate of DO and have shown better correlation to clinical symptoms (Radley et al. 2001). However, in studies of healthy asymptomatic volunteers, 16% on laboratory urodynamics and 48% of women on ambulatory urodynamics were diagnosed to have DO indicating a high false positive rate (Heslington and Hilton 1996). Moreover, the technique of ambulatory urodynamics is more time consuming, has not been standardised and not universally available. In the present study, ambulatory urodynamics was offered to women with OAB symptoms if they had normal urodynamics but only at the Birmingham Women's Hospital.

3.71 Urodynamics in Overactive bladder (OAB)

Detrusor Overactivity (DO) is defined as the urodynamic observation of involuntary detrusor contractions during the filling phase of cystometry which may be spontaneous or provoked (Abrams 2003). DO was observed in 44% of OAB dry and 58% of OAB wet patients in a study by Hashim and Abrams (Hashim & Abrams 2006). DO may be classified into idiopathic (absence of a defined cause) and neurogenic DO (presence of a neurological

condition). Also, DO may be phasic where the individual experiences urgency but may or may not have incontinence and terminal DO where there seems to be a single involuntary uninhibited contraction emptying the bladder (Abrams 2003). Bladder compliance is a mathematical calculation obtained by dividing the change in bladder volume (DV) by the detrusor pressure change (D Pdet) during that change in bladder volume (DV/DP det). It gives an indication on the elasticity of the bladder wall.

Presence of artefacts may mislead and interfere with the diagnosis. Abdominal pressure artefacts appear when patients strain during voiding or with rectal peristalsis, making it difficult to see whether detrusor activity was present (Hogan et al. 2012) Evidence on the reproducibility of urodynamic findings is conflicting (Homma et al. 2000).

3.8 Statistical methodology for diagnostic accuracy of BWT

A statistical analysis plan with proposed methods to analyse study data on the accuracy of BWT in the diagnosis of DO was drawn a priori.

3.81Primary analysis

3.811Diagnostic accuracy of BWT scan

Calculations of test accuracy-sensitivity, specificity, predictive values and likelihood ratios(Altman, 1994 112 /id) (Altman, 1994 111 /id) (Altman, 1994 113 /id) using a BWT of 5mm as a cut-off (>=5mm indicating presence of DO, <5mm indicating absence of DO) were planned. Using binomial exact methods, 95% CIs will be calculated for all estimates (Clopper C, 1934) . The cut off of 5mm mean BWT in diagnosing DO was based on previous evidence from a systematic review on the diagnostic accuracy of BWT (Latthe, 2010).

3.82 Secondary analysis

Likelihood ratios for BWT will be calculated (along with 95% CI) for the three BWT cut-offs: <3mm / >=3mm to <5mm / >=5mm. A ROC curve (Altman, 1994) will be constructed (plot of sensitivity versus 1-specificity) (e.g. every 0.25mm) between the highest and lowest measurements obtained for exploration of an 'optimum' cut-off of BWT. Accuracy estimates will be reported along with estimates at the Q* threshold (point where sensitivity and specificity are maximised). The area under the ROC curve (probability that a random person with DO has a higher measurement of BWT than a random person without DO, with 0.5 the lowest possible value and 1.0 the highest) will also be calculated, along with a 95% CI, as an overall estimate of BWT accuracy. Statistical significance will be tested by comparing against the uninformative model (i.e. where AUC=0.5) using a non-parametric approach(Xiao-Hua Zhou NAO and Donna K.McClish 2011). A t-test on BWT comparing those with a positive and negative diagnosis of DO will be performed and box-and-whisker plots generated.

3.83 Subgroup and other analysis

Estimates of accuracy (as per the primary and secondary outcomes) within subgroups will be calculated for the following variables:

previous treatment with antimuscarinics

clinical history suggesting mixed incontinence

presence of urinary tract infections in previous 12 months

patients with voiding difficulties

patients who also have ambulatory urodynamics for DO verification

past history of incontinence surgery

The importance of these subgroups will be explored using a logistic regression model with sensitivity or specificity as the outcome variable and a variable representing the subgroup as the explanatory variable. Odds ratio, 95% CI and related p-values will be calculated for these analyses.

Other exploratory analysis will include a multivariable logistic regression analysis to explore possible predictors of DO. Combinations of history (which includes the subgroup variables listed above along with, but not limited to: age, menopausal status, parity, body mass index) and BWT will be used as the explanatory variables with the urodynamic diagnoses (DO: yes/no) as the dependent variable.

3.84 Sensitivity analysis

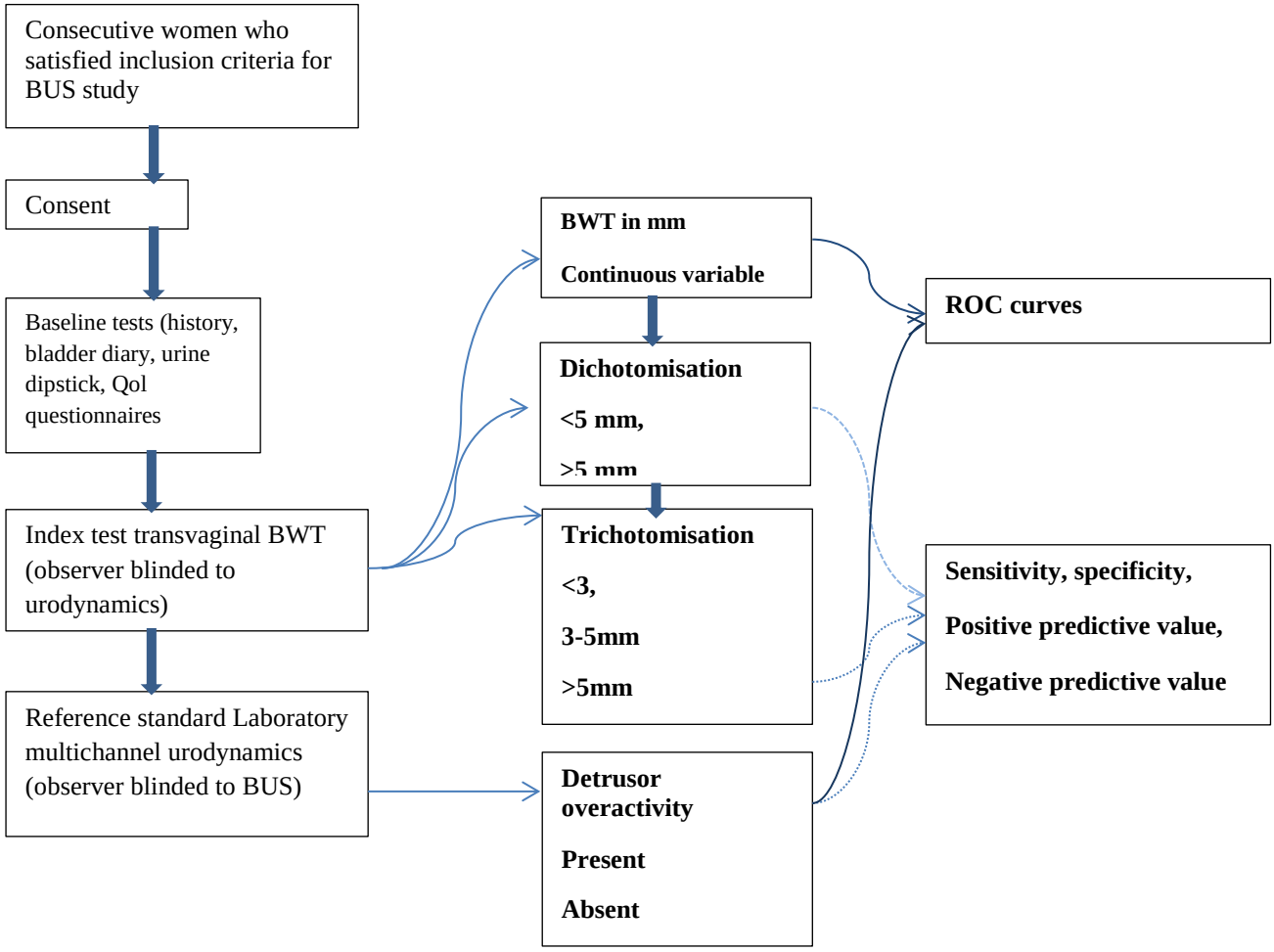
Sensitivity analysis including missing explanatory variable data inputted though multiple imputation techniques will be performed. Where urodynamics were performed without blinding to the results of the ultrasound, exploratory sensitivity analysis will be performed without these tests to see how this affects the primary and secondary outcomes. Sensitivity analysis will also be completed for those patients with more than four week interval between the ultrasound and urodynamics.

3.85 Missing outcome data

The study will carry on recruiting until we have reached at least 600 patients who have been through the ultrasound and urodynamics. Missing data items will be flagged to the data manager through the data entry system at the entry stage.

If all three(dome, trigone and anterior wall) of the ultrasound measurements used to calculate mean BWT are unavailable (following a review by the data manager and statistician) then these results will not count towards the final recruitment total. If one or two of the three measurements are unavailable then the average BWT will be calculated from the remaining data. A sensitivity analysis will be performed without these particular results to see how this affects the primary and secondary analysis. The flow chart of patient pathway including the statistical methods is depicted in figure 8 below.

Figure 8: Study flow chart



Chapter 4: Accuracy of transvaginal bladder wall thickness in diagnosing detrusor overactivity: Results

Introduction

A diagnostic test is used to predict the post-test probability of a disease, i. e, discriminate a person with the disease from a person without the disease. The threshold at which a test result is significant needs to be determined if the result is reported in continuous variables. Accuracy of the test in discriminating disease positive and disease negative states at that threshold needs to be ascertained to interpret the test result (Mallett et al. 2012).With a new diagnostic test is considered, rigorous evaluation of its accuracy is necessary to prevent or minimise its inappropriate use and interpretation, thus minimising any harmful or unwanted clinical consequences .

This chapter describes a prospective test accuracy study to determine the accuracy of bladder wall thickness (BWT) in diagnosing detrusor overactivity (DO) in women with overactive bladder (OAB) symptoms.

Population: Women with OAB or urgency predominant mixed incontinence

Index test: BWT on transvaginal ultrasound scan

Reference standard: Multichannel urodynamics

Outcome: Diagnosis of DO.

4.1 Index test - bladder wall thickness via ultrasound

For the technique of bladder wall thickness, please see chapter 3 and appendix 6a for the standard operating procedure (SOP).

4.2 Reference standard – Urodynamics

Please refer to appendix 6b for the SOP for urodynamics. The aim of urodynamics was to reproduce patient's symptoms and correlate them with the changes in detrusor pressure. The effect of bladder filling on the patient's perception of symptoms was assessed. Table 8 depicts the various urodynamic parameters in the study population.

4.3 Data analysis

The detailed statistical methods for primary and secondary analyses have been elaborated in chapter 3 under the sections 3.81 and 3.82 in pages 64 and 65. The sensitivity analyses were performed as previously planned and documented in the appendix section.

4.4 Results

4.41 Recruitment of participants in the BUS study

Recruitment of participants started in January 2011 and closed in March 2013. A total of 1310 women were approached and the course of recruitment from March 2011 to march 2013 is depicted in figure 9. Our study recruitment was faster than the projected course as seen in figure 9. Six hundred and eighty seven women, consenting to participate were recruited into the study from 22 centres. Details of recruitment of study participants from each centre is given in table 5. The study over-recruited to compensate for withdrawals and also to accommodate the participants who did not undergo either the index or the reference standard tests.

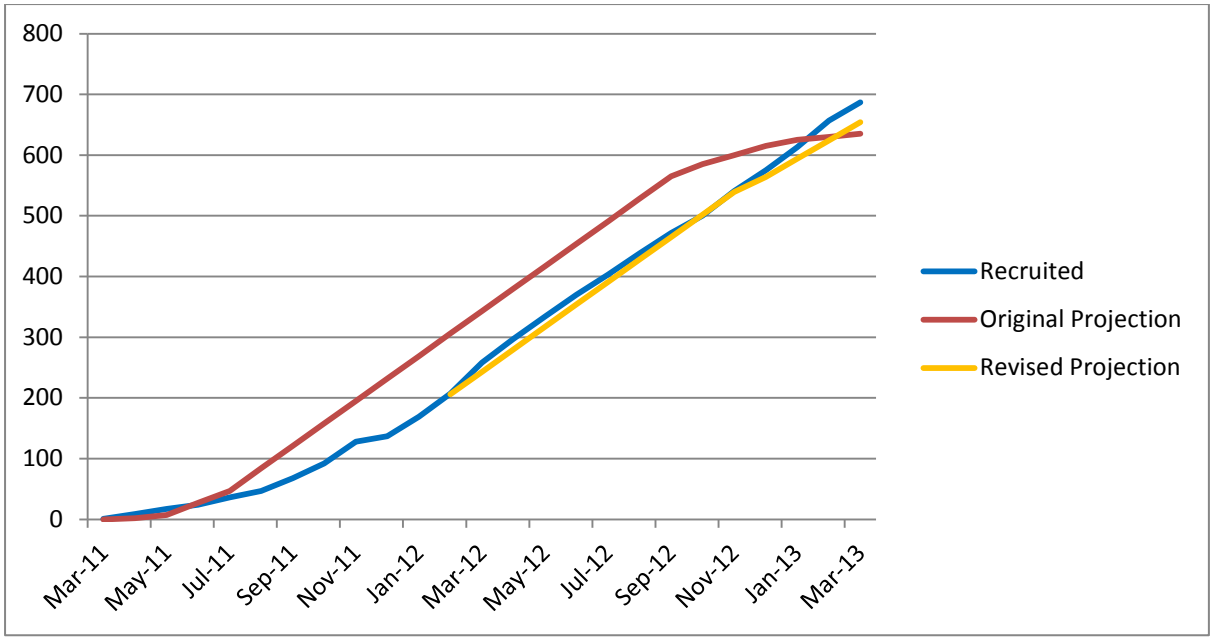


Figure 9: Progress of recruitment of patients in the BUS study

Majority of the patients were recruited from Birmingham Women's Hospital (n=254) followed by Medway Maritime Hospital (n=109) (Table 5).

Table 5: Recruitment by centre

Recruiting Centre	Number
Birmingham Women's Hospital	254
Medway Maritime Hospital, Kent	109
Mayday University Hospital, Croydon	92
Basingstoke and North Hampshire Hospital	30
St Mary's Hospital, Manchester	26
Staffordshire General Hospital	26
Stepping Hill Hospital	25
Ormskirk & District General Hospital	19
New Cross Hospital, Wolverhampton	16
Royal Bournemouth General Hospital	16
The Alexandra Hospital, Redditch	14
City General Hospital (University Hospital of North Staffordshire)	10
Crosshouse Hospital, Ayrshire	9
Manor Hospital, Walsall	8
Northampton General Hospital	6
Royal Hallamshire Hospital, Sheffield	6
Derriford Hospital, Plymouth	5
Pinderfields General Hospital	4
St Mary's Hospital, Paddington	4
Southern General Hospital, Glasgow	3
The Royal London Hospital	3
Sandwell General Hospital, Birmingham	2
Total	687

4.42 Demographics of study population:

The study population represented the typical OAB spectrum of patients. The mean age of women was 52.7 years (SD 13.9) with an average BMI was 30.6 (SD 12.2). Fifty five percent (387/687) of the women were postmenopausal. According to the clinical history 61% (419/687) of the women had urgency predominant mixed incontinence and 33% (226/687)

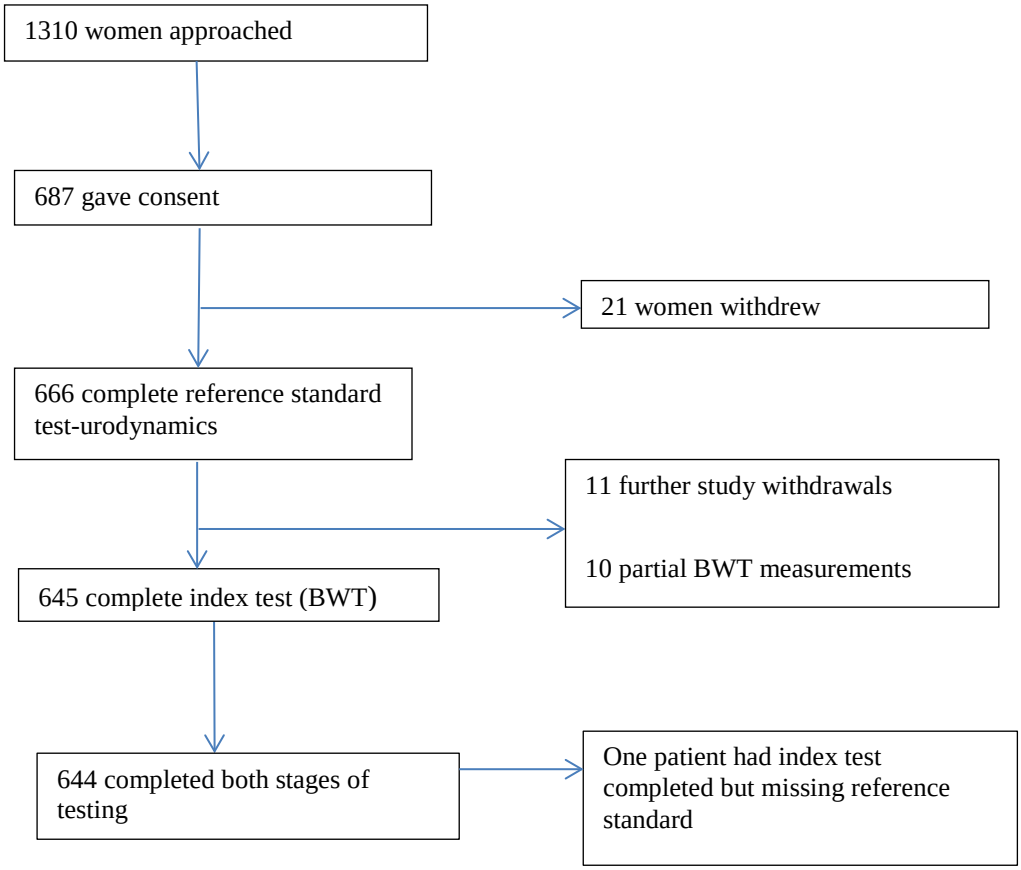
reported only urinary urgency along with increased frequency. The median duration of symptoms was 3.0 years (IQR: 1.6, 7.0) (Table 6).

Table 6: Characteristics of patients in the BUS study

Age (years)	Mean (SD)	52.7 (13.9)
	Missing	0 (-)
Ethnicity	White British/Irish/Other	538 (78%)
	Asian Pak/Ind/Bang/other	72 (10%)
	Black Carrib/African/other	49 (7%)
	Mixed/other	18 (3%)
	Not given/missing	10 (1%)
Parity	0	69 (10%)
	1	90 (13%)
	2	241 (35%)
	3	152 (22%)
	4	56 (8%)
	>4	63 (9%)
	Missing	16 (2%)
Post-menopausal (lmp>1 year)	Yes	378 (55%)
	No	293 (43%)
	Missing	16 (2%)
BMI (kg/m^2)	Mean (SD)	30.6 (12.2)
	Missing	28
Incontinence type	Mixed urinary incontinence	419 (61%)
	Urgency incontinence alone	226 (33%)
	Stress incontinence alone	4 (1%)
	Neither	19 (3%)
	Missing	19 (3%)
If mixed, what started first (n=419)?	Urgency	226 (54%)
	Stress	107 (26%)
	Unsure	54 (13%)
	Missing	32 (8%)
Current or previous treatment with anti-muscarinics	Yes	226 (33%)
	No	444 (65%)
	Missing	17 (2%)
Recurrent cystitis (3 or more in last 12 months)	Yes	50 (7%)
	No	606 (88%)
	Missing	31 (5%)
Voiding difficulties	Yes	286 (42%)
	No	374 (54%)
	Missing	27 (4%)
Vaginal birth	Yes	561 (82%)
	No	95 (14%)
	Missing	31 (5%)
Previous incontinence surgery	Yes	36 (5%)
	No	623 (91%)
	Missing	28 (4%)
Previous POP/UI surgery	Yes	56 (8%)
	No	603 (88%)
	Missing	28 (4%)

Following consent, the flow of patients through the diagnostic pathway (index test and reference standard is shown in figure 10.

Figure 10: Participant flow diagram



The number of participants with a complete reference standard diagnosis was 644/687 (97%). The other 21 (3%) decided to withdraw from the study before any testing could take place (Figure 10).

4.43 Index Test and the Reference Standard

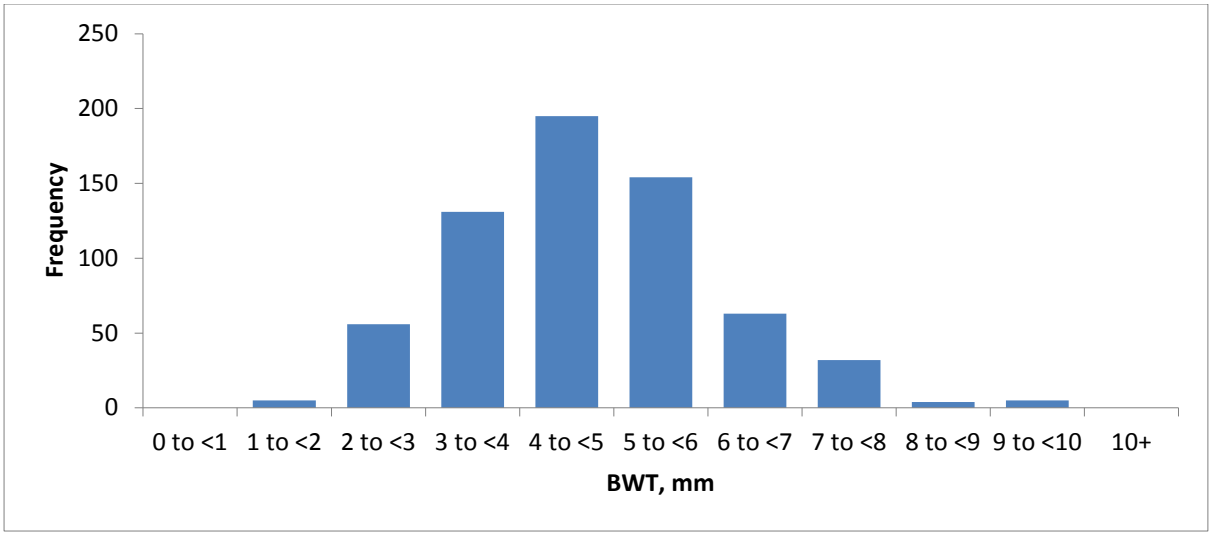
The majority of the participants 68% (446/660) had both BWT scan and urodynamics performed on the same day. Only a small proportion (26/660, 4%) were performed more than four weeks apart. No serious adverse events were reported following either test, although 49/479 (10%) of the respondents reported having a UTI within two weeks of testing at six

month follow-up; 36/48 (75%,) of these were diagnosed by a General Practitioner or in a hospital and resulted in antibiotic use in the 83% of cases (39/47 83%).

4.431 Bladder wall thickness

The number of participants with all three BWT measurements (trigone, dome midline, anterior wall midline) available was 645 (94%). Of the remainder, 32 (5%) withdrew from the study or were lost to follow-up and did not have any recorded measurements and 10 (1%) had partial measurements recorded (9 with 2/3 of the measurements and 1 with 1/3 of the measurements). Summary statistics and distribution of BWT are provided in Figure 11.

Figure 11: Histogram of average bladder wall thickness measurements



4.432 Urodynamic findings

The number of participants with a complete reference standard diagnosis was 666 (97%), the other 21 (3%) were withdrawn from study or lost to follow-up. Of these, 399 (60%) were diagnosed with DO (95%CI: 56% to 64%) (Table 7). Of the 399, two hundred and forty-five were given further sub-diagnosis of ‘wet’ DO (DO with urinary incontinence) (61%) and 154 as ‘dry’ DO (DO without urinary incontinence) (39%). The participants also had their DO diagnosis sub-categorised as systolic ‘spontaneous’ DO (detrusor contraction during the filling phase: 182/369, 49%, 30 observations missing),provoked DO (if the detrusor

contraction occurred during or after provocation tests like cough, star jumps, running water or immersion of hands in cold water: 56/369). Details of the measurements made in this testing are given in Table 8. Of these, 399 (60%) were diagnosed with DO (95%CI: 56% to 64%). Of the 399, two hundred and forty-five were given a further sub-diagnosis of 'wet' DO (61%) and 154 as 'dry' DO (39%). The participants also had their DO diagnosis sub-categorised as systolic or 'spontaneous' DO (detrusor contraction during the filling phase: 182/369, 49%, 30 observations missing), provoked DO (if the detrusor contraction occurred during or after provocative measures like cough, running water or immersion of hands in cold water: 56/369, 15%), or both (131/369, 36%)(Table 8).

Table 7: Summary of all urodynamic diagnosis in the cohort of women recruited in this study (n=666)

Including DO (n=399)	Numbers (%)
DO only	258 (39)
DO+USI	97 (15)
DO+ voiding dysfunction (VD)	18 (3)
DO+USI+VD	12 (2)
DO+ Low compliance (LC)	8 (1)
DO+USI+LC	5 (1)
DO+USI+LC+VD	1 (<1)
Not including DO (n=267)	
Normal	124 (19)
USI only	78 (12)
LC only	36 (5)
VD only	14 (2)
USI+VD	8 (1)
USI+LC	6 (1)
LC+VD	1 (<1)

Table 8: Urodynamic parameters in the BUS study (n=666)

Uroflowmetry	
	Frequency (%) for binary data, median [IQR] for continuous data, n=number of values recorded
Patient had comfortably full bladder=yes	502/655 (77%)
Volume voided, ml	129 [58, 245], n=627
Post void residual volume, ml	10 [2, 40], n=619
Maximum flow rate, ml/sec	16 [9, 25], n=596
Filling cystometry	
Patient in recommended sitting position for test=yes	457/664 (69%)
Fill rate, ml/min	100 [70, 100], n=660
First desire, ml	135 [84, 197], n=644
Normal desire, ml	200 [140, 268], n=587
Strong desire, ml	272 [199, 357], n=562
Pain (if reported), ml	300 [203, 395], n=154
Leakage (if applicable), ml	10 [0, 100], n=229
Total volume in bladder at the end of filling, ml	421 [314, 498], n=639
Rise in detrusor pressure upon filling=yes	350/598 (59%)
Detrusor pressure at start, cm/H_2O	0 [-1, 1], n=638
Detrusor pressure rise on filling to 500 ml, cm/H_2O	12 [6, 21], n=576
Detrusor pressure rise when complaint of urgency, cm/H2O	12 [5, 21], n=557
Provocation test (where performed)	
Detrusor pressure rise with cough=yes	101/517 (20%)
Detrusor pressure rise with running tap=yes	119/367 (32%)
Detrusor pressure rise with exercise	39/124 (31%)
Flow cystometry	
Peak flow rate, ml/sec	20 [15, 28], n=624
Maximum voiding pressure, cm/H_2O	41 [29, 60], n=577
Residual volume, ml	0 [0, 20], n=540

4.7 Estimates of test accuracy

Six hundred and forty four participants had both complete index and reference standard results. Estimation of the accuracy of BWT showed poor sensitivity, specificity and likelihood ratios at the pre-specified cut-offs of 5mm (Tables 9 and 10), <3/3-5/>=5mm (Table 11) and overall (Figure 12). The AUC was 0.53, 95%CI: (0.48, 0.57) indicating that there was no evidence that BWT had any ability to discriminate between those with and without DO (p=0.25) (Figure: 12). Furthermore, there was no evidence that the mean BWT measurements were any higher in the DO positive group compared to the DO negative group: 4.85mm (SD: 1.36) versus 4.70mm (SD: 1.29); p=0.19 (Figure 13) or that it had any relationship with ICIQ-OAB symptoms score when measured at presentation (*r*=-0.01; p=0.88) (Appendix: 5.3).

Table 9: Comparison of index and reference standard results – dichotomised at 5mm

		Reference standard (Urodynamics)		
		DO	No DO	Total
Index test: BWT by ultrasound	Positive Result (>=5 mm)	165	98	263 (41%)
	Negative Result (<5 mm)	223	158	381 (59%)
	Total	388 (60%)	256 (40%)	644

Table 10: Estimates of BWT test accuracy – dichotomised at 5mm

	Value	95% CI
Sensitivity	43%	38 to 48%
Specificity	62%	55 to 68%
PPV	63%	57 to 69%
NPV	41%	36 to 47%
LR+	1.11	0.92 to 1.35
LR-	0.93	0.82 to 1.06

Table 11: Estimates of BWT test accuracy – trichotomised at 3, 3-5, >=5mm

		Reference standard (Urodynamics)			LR+	LR-	95% CI
		DO	**No DO**	**Total**			
Index test: BWT by ultrasound	++Result (>=5 mm)	165	98	263	1.11	0.93	0.92 to 1.35
	+ Results (3-5 mm)	193	132	325	0.96	1.04	0.83 to 1.13
	Result (<3 mm)	30	26	56	0.76	1.03	0.46 to 1.26
	Total	388 (60%)	256 (40%)	644			

Figure 12: ROC curve analysis for BWT in the study population

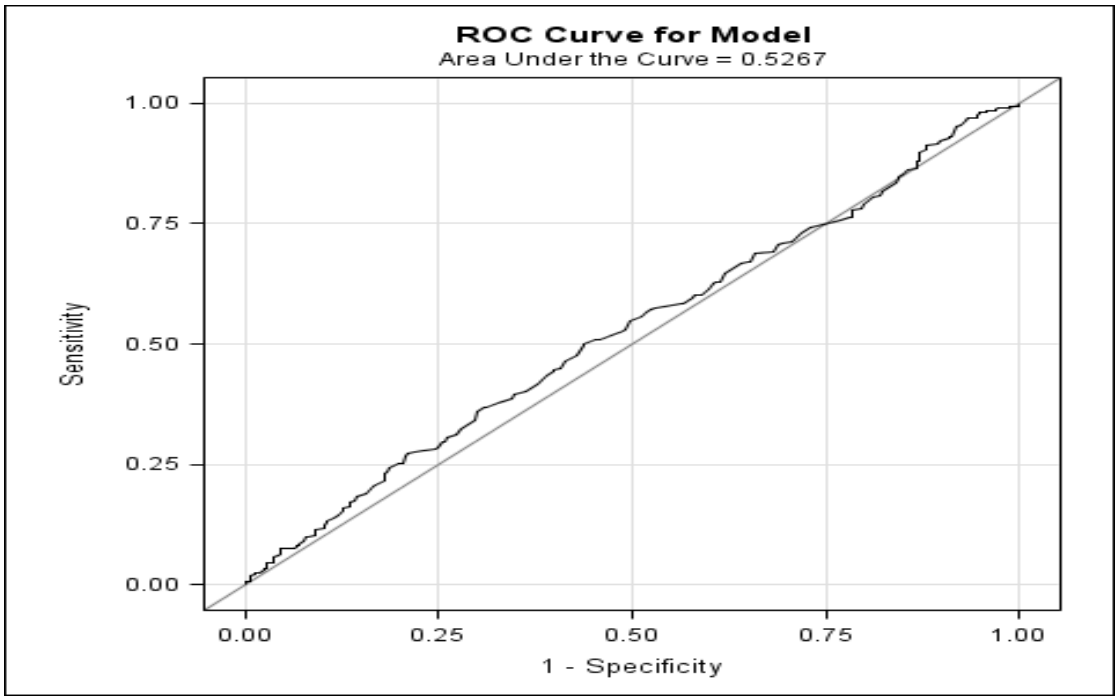
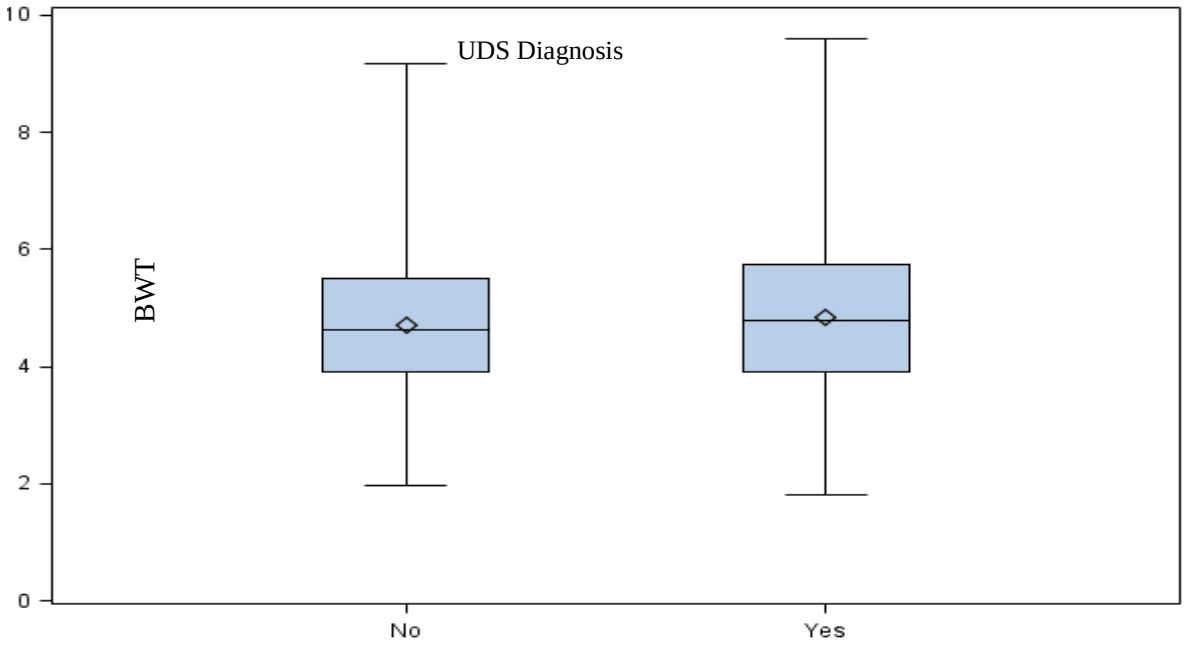


Figure 13: Box and whisker plot comparing BWT with DO diagnosis



Sensitivity analyses included construction of ROC curves and calculation of AUC and were limited to excluding those patients where it was revealed that the urodynamics test result was not blinded to the results of the ultrasound(Appendix:3 Figure 3.1); excluding those patients where there was more than four weeks between index and reference standard tests(Appendix:3 Figure:3.2); including results of incomplete ultrasound measurements, i.e. where not all three components of BWT were recorded (in these cases if one or two measurements were missing the average of the remaining values was taken to be BWT)(Appendix:3 Figure 3.3); replacing the original urodynamics diagnosis with that from the additional ambulatory urodynamics test where available (this only happened in fourteen instances all from one centre) (Appendix:3 Figure 3.4); using the trigone measurement alone as BWT(Appendix:4 Figure:4.5); excluding those who had 'provoked DO' (detrusor pressure rise upon provocation testing – 187 cases)(Appendix:4 Figure:4.6) & excluding those who had post void residual>30ml upon BWT testing (Appendix:4 Figure 4.7);.

The populations looked at in the exploratory analyses were: a) Women with urgency alone on clinical history (i.e. excluding those with mixed stress/urge incontinence) (Appendix: 4 Figure 4.1); b) Women with 'pure' DO only (i.e. not alongside another diagnosis from urodynamics) (Appendix: 4 Figure: 4.2); and c) Women with 'wet' DO only (i.e. not including those with 'dry' DO) (Appendix: 4 Figure: 4.3).

Subgroup analyses were performed by the generation of ROC curves for each subgroup.(Appendix 5,Table:5.1) The AUC for each subgroup was compared using a large sample chi-squared test for independent curves(Pepe 2004). The subgroups used here to dichotomise patient groups were: previous treatment with antimuscarinics, a clinical history suggesting mixed incontinence; presence of a urinary tract infection in the previous twelve months; voiding difficulties; previous incontinence surgery and the BMI (<25, >=25).

An exploratory logistic regression analysis was undertaken to assess which variables may be associated with a diagnosis of DO. They included pre-test ICIQ score, BWT, age, duration of symptoms, ethnicity, vaginal birth, menopausal status, parity and previous POP surgery Covariates were considered individually and then in a multivariable analysis (Appendix: 5 Table: 5.3). Three multivariable models were constructed: one using all possible explanatory variables; another using all possible explanatory variables but using a multiple imputation approach to generate missing responses (Schafer JL 1997) and another using a backward-step process to eliminate unimportant variables (a level of p=0.1 was used here as criteria for staying in the model). Whether BWT had any relationship with baseline ICIQ-OAB score using a simple linear regression model was also examined.

Sensitivity analysis or the unplanned exploratory analysis as described above did not change the interpretation of these findings. There was some evidence that those diagnosed with 'wet' DO had higher BWT than those with 'dry' DO (4.94mm vs. dry 4.69mm; p=0.08) (Appendix 4.4). However, when the BWT for the wet DO group was analysed alone, the AUC was only 0.55, 95%CI: 0.50, 0.59) (Appendix: 4 Figure: 4.3). There was no evidence that BWT performed any differently in any of the pre-specified subgroups (Appendix 5.1).

In the multivariable exploration of factors possibly associated with DO diagnosis, only higher pre-test ICIQ score (i.e. worse symptoms) was associated with DO (OR: 1.21, 95%CI: 1.13, 1.29; p<0.0001 from the model including all possible variables), i.e. the odds of DO diagnosis were increased by 21% for every point increase in ICIQ score (Appendix 5 Table: 5.3). Previous treatment with antimuscarinics and previous urine infection in the last twelve months also showed some relationship but these were of borderline statistical significance.

4.8 Discussion

4.81 Summary of main findings

To date, BUS is the largest diagnostic accuracy study to estimate the test accuracy of BWT in diagnosing DO. We could not find any evidence that BWT had any relationship with DO. It appeared to be no more accurate than chance at diagnosing DO with an AUC of 0.53(95%CI: 0.48, 0.57) (Figure: 12). Extensive sensitivity analyses and subgroup analysis did not alter the interpretation of these findings. Based on this evidence, we conclude that BWT is not a useful test in diagnosing DO and should not be used clinical practice.

4.82 Strengths and limitations

The main strength of this study is its methodology. The protocol was drawn up 'a priori', peer reviewed and given robust oversight with six monthly meetings of independent Data

Monitoring and Study Steering Committees. Blinding of operators performing BWT scan and urodynamics was ensured for 610/629 (97%) women recruited into the study. Verification bias was minimised by incorporating a complete verification design. Disease progression bias was minimised by conducting BWT and urodynamics within a short time span of each other, often within the same day.

A sample size calculation was performed to ensure that the study is powered to estimate diagnostic accuracy of BWT, and the study recruited beyond the target. A number of sensitivity analyses were performed on the primary population to test the robustness of the results to protocol deviations and missing data. Participants in the study were recruited from several centres (tertiary referral centres and district general hospitals) (Table 5).Women were of different ages, ethnicities and social background, recruited from various parts of UK, providing high level of generalisability across the NHS (Table 6). The spectrum variation (recruiting women from various ethnicities, varying degree of severity of OAB or urgency predominant mixed incontinence) in the study was a strength of the design, improving the generalizability of the findings. Exploratory analyses were performed to gauge the effect of changing the population of interest and also to see which parameters were associated with DO diagnosis and there was only some relation between ICIQ scores and DO. Furthermore, BWT had no relationship to ICIQ score upon presentation, indicating that it has no relationship with symptom severity. ICIQ score.

It is believed that spontaneous DO is secondary to pathology in detrusor muscle whereas provoked DO is due to pathology in bladder neck. Provoked DO which was previously called urethral instability could be due to primary urethral aetiology (Ulmsten 1997;Ulmsten and

Falconer 1999). Based on the theory that bladder wall thickens with frequent detrusor contractions against a closed sphincter, a sensitivity analyses to evaluate diagnostic accuracy of BWT excluding those with provoked DO was carried out. However, the ROC for women with only spontaneous DO has shown an AUC of 0.54.indicating absence of any significance (Appendix: 4.6).

The technique of transvaginal BWT measurement is easy to perform, with the urinary bladder being an anterior and relatively superficial midline structure. To standardize the performance of BWT scan, a standard operating procedure (SOP) for carrying out the ultrasound was developed (Appendix 6a). Hands on training at individual recruitment sites to the investigators was provided and two workshops were organised on BWT measurements. Quality assurance tests on both urodynamics and BWT were carried out. For the study findings to be reliable, uniformity in the conduct of recruitment/diagnostic testing is essential for multicenter clinical trials using urodynamics. With the use of regular training updates on standardized urodynamic testing procedures, interpretation guidelines and the quality assurance audits, the technical quality of urodynamics was improved. To improve reliability in the way BWT and urodynamics were carried out, several proactive measures were undertaken eg. standard urodynamic test protocol, guidelines for interpretation, audits of the traces centrally every six months, training and assessing competence in measurements of BWT, checking reproducibility of scans, independent data monitoring committee etc.

In an update of the systematic review on the diagnostic accuracy of BWT in diagnosing DO (Chapter: 2); ten of the twenty one studies were of case-control design (Tables: 2 and 3). Seven of the ten case-control studies have shown a significantly increased BWT in DO

compared with patients without DO (Abou-Gamrah, Fawzy, Sammour, & Tadros 2014;Ibrahim S & Najdy M 2011;Minardi, Piloni, Amadi, El, Milanese, & Muzzonigro 2007;Otsuki E N, Júnior E A, Oliveira E, Castelo Girão M J B, & Jármy-Di Bella Z I K 2014;Ozturk H, Aydur E, Irkilata H, Seckin B, & Dayanc M 2011;Parsons, Amundsen, Cardozo, Vella, Webster, & Coats 2007;Soligo M, Salvatore, Luppino G., Arcari V, Milani, & Khullar 2002) The results from these studies need to be interpreted with caution as the comparison with the 'non disease controls' in ten studies might have introduced significant bias by over-estimating the diagnostic accuracy of the test (Lijmer, Mol, Heisterkamp, Bonsel, Prins, van der Meulen, & Bossuyt 1999). All of the prospective and retrospective cohort studies on diagnostic accuracy of BWT also suffered from multiple weaknesses like absence of power calculation, being single centre (mostly tertiary) studies with limited generalizability, use of different routes of ultrasound like the transabdominal, translabial with varying bladder volumes during measurement and using additional reference standards like ambulatory urodynamics (Khullar, Cardozo, Salvatore, & Hill 1996;Robinson, Anders, Cardozo, Bidmead, Toozs-Hobson, & Khullar 2002). The methodology of these studies was not robust to draw meaningful comparisons with our study (Figures 3 and 4).

Of the remaining 11 prospective and retrospective cohort studies, only five report test accuracy data (sensitivity and specificity). Except one study (Yang & Huang 2002) all others report significant differences in mean BWT between groups, but the difference was not significant enough to be used as a diagnostic test. Kuhn et al, recruited consecutive women undergoing urodynamics for any lower urinary tract symptoms. They compared vaginal (5MHz probe), perineal and abdominal ultrasound measurement of BWT after filling the bladder to 50 mls with a catheter. Out of 125 women, 21 were excluded as they had voiding dysfunction. This study had a different objective to that of our study, was from a single centre

with overlapping patient population, had slightly different method of transvaginal scan, was not powered for diagnostic accuracy and BWT was not blinded to urodynamic diagnoses with a potential to introduce bias (Kuhn, Genoud, Robinson, Herrmann, Gunthert, Brandner, & Raio 2011).

In a prospective study consecutive women with OAB symptoms, SUI and all types of MUI who had urodynamics and the transvaginal scan for BWT were recruited. The mean BMI of women in their study was lower compared to our study. The technique was slightly different- 5 m HZ probe and a parasagittal view (as opposed to 7-9 MHz and sagittal view) was used for BWT measurement. They concluded that although there is a relationship between BWT and DO at higher BWT values beyond 6.5 mm, only 6.5% of their study population would have theoretically avoided urodynamics (Serati, Salvatore, Cattoni, Soligo, Cromi, & Ghezzi 2010). Our study is superior to that of Serati et al, with a more selective patient population (only the ones with OAB/urgency predominant MUI), larger sample size based on power calculation, assessment of inter-observer variation, multicentre recruitment, and prospectively determined sensitivity analysis.

Estimates of diagnostic accuracy of the index test are directly influenced by the quality, the reliability and the reproducibility of the reference standard and the level of agreement between the index test and the reference standard(Lijmer, Mol, Heisterkamp, Bonsel, Prins, van der Meulen, & Bossuyt 1999). Estimates of the sensitivity and specificity or area under the curve (AUC) for new diagnostic tests are difficult to produce when the accuracy of the reference standard is unknown, or known to be imperfect. Many studies have cast doubt on the reproducibility and accuracy of urodynamics, which is our reference standard. During

evaluation of reproducibilty of urodynamics in healthy volunteers and patients with OAB, inconsistencies between serial urodynamic procedures were a common finding (Bellucci et al. 2012;Broekhuis et al. 2010;Brostrom et al. 2002;Digesu et al. 2003b). In a multicentre study with six serial urodynamics in patients with OAB, there was increased variability in pressure measurements than volume measurements (Frenkl et al. 2011).

When estimating the test accuracy of BWT against an imperfect reference standard urodynamics, the accuracy of BWT may have been biased to an unknown degree, or submerged in the 'noise' from the imperfect reference standard. However, the poor accuracy for BWT elicited in our study is likely to be related to inherent inaccuracy of BWT since there was no significant relationship between bladder wall measurements and grades of DO severity. When the test (BWT) values do not differ amongst those with varying grades of the target condition, it can be inferred that the lack of accuracy may be an inherent feature of the index test.

4.83 Interpretation of findings

A key difference between the BUS study and all others is the focus on the accuracy of ultrasound in a group of women with OAB/Urgency predominant mixed incontinence and who do not have signs of pure stress urinary incontinence (SUI). The prevalence of DO in women with OAB, urgency predominant mixed incontinence was found to be 60%, which was similar to other studies. When exploratory analyses was performed comparing women with isolated DO to those of mixed urodynamic incontinence of DO and USI, there was no change in our findings. Even in women with an incompetent sphincter (DO co-existing with USI), the BWT was very similar to those with a competent sphincter. This finding disproves the theory that bladder wall hypertrophies and thickens when it contracts against a closed

sphincter (Khullar, Salvatore, Cardozo, Bourne, Abbott, & Kelleher 1994;Khullar, Cardozo, Salvatore, & Hill 1996).

There is some emerging evidence from observational studies that the response to invasive therapies might be similar in patients with frequency and urgency +/- urgency incontinence, with or without the observation of DO on urodynamics. An adequately powered multicentre RCT of urodynamics vs. no urodynamics in patients with OAB receiving invasive therapies like PTNS, Botox and SNS would be helpful in discerning its role in OAB. There might be a potential in further research into the diagnostic accuracy of the ICIQ-OAB questionnaire.

4.9 Conclusion

Given the poor performance for BWT as tool for diagnosing DO, the recommendation for clinicians is to not use it as an alternative to urodynamics. There is no further role of BWT in women with OAB.

Chapter 5: Interobserver variability in the sonographic assessment of bladder wall thickness

Introduction

Validation of a test will involve the scientific community define a threshold, and then gather sufficient information to allow test to be used with confidence. A minimum standard of reliability and reproducibility need to be met for any diagnostic test as limited or poor reproducibility affects the precision of a test (Bossuyt et al. 2003). Agreement between different clinicians performing the test depends on the ease of interpretation. If the people who actually perform and interpret the test cannot agree on the interpretation, the test results will be of little clinical use.

Inter-observer reproducibility of BWT can be demonstrated by studying the difference between blinded observers when exposing the same patient to the technique independently at different points of time. Reproducibility of BWT is of particular importance given the fact that bladder is distensible organ and its thickness is known to change with the amount of urine present in the bladder. The aim of this study was to assess whether measurement of BWT using transvaginal ultrasound have adequate reliability and reproducibility to detect difference in BWT potentially indicative of DO.

5.1 Objectives

The three key objectives were:

1) To estimate the intra-observer measurement error in interpreting images by comparing blinded duplicate assessments of images by a single observer

2) To estimate the inter-observer measurement error in interpreting images by comparing blinded duplicate assessments of images by different observers

3) To estimate the inter-observer measurement error in the complete scanning and interpretation process by comparing measurements made by different observers on different scans made on women at two separate occasions.

5.2 Methods

The reproducibility study was carried out as a part of the multicentre Accuracy of bladder wall thickness ultrasound in the diagnosis of Detrusor Overactivity (BUS study) which was approved by Nottingham Research Ethics Committee (ethics no10/H0408/57) and funded by the National Institute of Health Research/ Health Technology Assessment Board. BUS study was aimed at evaluating whether BWT measurement could reduce the need for urodynamics in women with OAB.

When a second observer was available, women who agreed to have two transvaginal scans by different operators were recruited into the study evaluating the prospective reproducibility of BWT scans. Recruitment into interobserver studies was opportunistic based on the availability of a second observer and patients who were willing to have a second scan. For the studies evaluating the reproducibility of the interpretation of scans, random selections of images were sent by the recruiting centres at the request of the trial coordinator.

Measurements of BWT at the trigone, dome midline and anterior wall midline were made as per a standardised operating procedure using 2-D transvaginal end firing probe, as shown in Figures 1 and 7. The BWT measurement was defined as the mean of three measurements

made at the locations ([trigone+dome midline+anterior wall midline]/3). The process of measurement required placing a calliper reference point on the image, using a mouse operated cursor on the electronic image, at the interface between the bladder wall and the adjacent tissue or lumen (Figures 1 and 7). Images were saved with and without the calliper placement.

Three sub-studies were undertaken to address the three above mentioned objectives;

In sub-study A, BWT was measured on 37 ultrasound images from individual participants. Repeat measurements were made by the same observer on the same images 6-12 months later. All second measurements were made blind to the original measurement, using images without calliper marks. All images were from the Birmingham Women's Hospital, and were measured on the scan machine. The measurement process was the same from the beginning to the end of the study and included repeat measurements. All measurements complied with our standard operating procedure.

In sub-study B, BWT was measured on ultrasound images from 57 individual participants by a single observer. Repeated measurements were made by one of 8 different observers on the same images, such that there were duplicate measures for each image (Table 12). All second measurements were made blind to the original measurement, using images without calliper marks. Images from the Birmingham Women's Hospital were measured on the scan machine; those performed outside BWH were measured using Digital Imaging and Communications in Medicine (DICOM) software.

In sub-study C, 27 women underwent two separate ultrasound scans at different points of time undertaken by different observers. The second scans and measurements of BWT were made blind to measurement from the first observer. Three observers were used in total, all women received scans from observer one, followed by either observer two or observer three (Table 12). Details of experience of the operators in the reproducibility study have been elucidated in table 13. The previous evidence on the reproducibility of BWT has been summarised in table 15.

5.3 Analyses

For each study BWT measurements were analysed using one-way analysis of variance. One-way ANOVA decomposes the total variation observed (SD_T^2) into that originating from differences between women (SD_I^2 - individual variability) and that caused by the measurement process (SD_A^2 - analytical variability). The estimates are linked as $SD_T = \sqrt{SD_A^2 + SD_I^2}$. The standard deviation for analytical variability, SD_A estimates the measurement error.

Two further statistics were computed from these values. The intraclass correlation coefficient (ICC) described the fraction of the total variance in BWT measurements due to individual rather than analytical variation (SD_I^2/SD_T^2). ICC values lie between zero and one: measurements that are reliable have ICCs approaching one, as the signals (the individual variation) dominate the noise (the analytical variation).

The repeatability coefficient described the *smallest real difference* (SRD) that can be detected with a specified degree of certainty between two measurements, and was computed as

$\sqrt{2}Z\sqrt{SD_A^2}$ (where Z takes the value of 1.96 for a difference which has 95% certainty of being a real effect and not measurement error). SRD values were given in the units of the original measurement. The above analyses were all undertaken assuming exchangeability of observers, i.e. that the ordering of the measurements has no relevance. Generalizability of these findings relies on the observers being presumed to be representative of those who would make the measurements in practice.

The data in scatterplots and Bland-Altman plots was to demonstrate the distribution of measurements and differences between measurements. For these analyses assignment of measurements to particular observers is important. In sub-study A there is a logical choice for the first and second measurements, and the distribution of these differences is of interest. In sub-studies B and C measurements made by observer 1 were arbitrarily taken as the first measurement and remaining observers 2 through 6 were taken as the second measurement.

5.4 Results

A total of 121 women were invited and all of them took part in the sub-studies. The distribution of BWT measures are shown in Table 12. The mean and standard deviation of the BWT measures in each sub-study were similar to that of the BUS cohort. Ranges in the sub-studies were lower, which is expected as ranges increase with sample size.

5.41 Sub-study A - Intra-observer repeatability of the same scans

Paired measurements were available for 37 women. The scatter of measurements is shown in Figure 14, and the distribution of differences in measurements in Figure 15.The later measurements were on average higher than the earlier measurements by 0.35mm, 95%CI: 0.19 to 0.51 ($p<0.0001$) but without any evidence of a relationship between error and the mean BWT value. Differences in measurements of up to 1.5mm were observed.

The standard deviation for the analytical variation for intra-observer variability was estimated as 0.42mm (Table 14).This level of variability compares to a standard deviation of 1.04mm between individual differences, thus 86% of the total variability observed is attributed to individual variability and 14% to measurement error. With this level of measurement error, differences of over 1.16 mm are 95% likely to be real for this single assessor.

Figure 14: Scatter plot for sub-study A- Intra-observer repeatability of the same scans

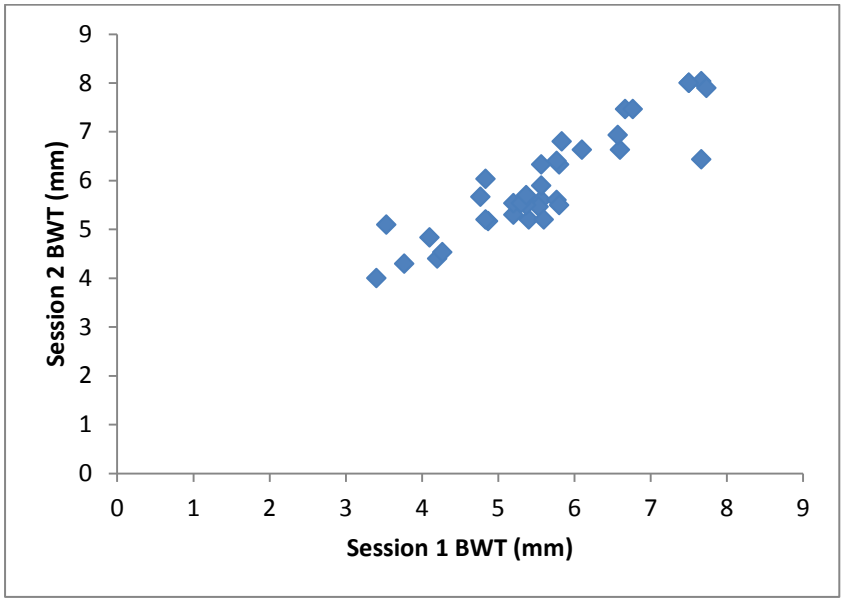
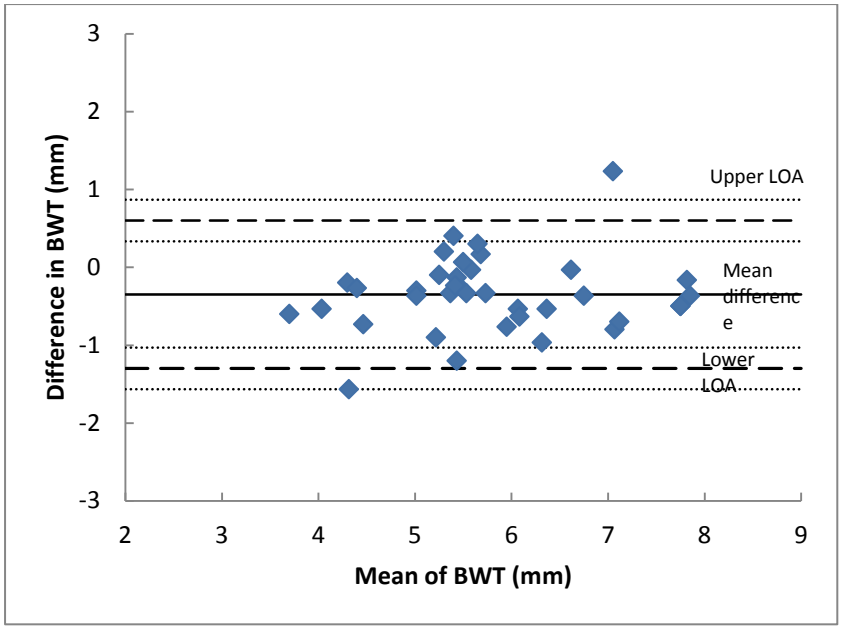


Figure 15: Bland-Altman analysis for sub-study A



5.42 Sub-study B Inter-observer repeatability of the same scans

Paired assessments were available for 57 women made by one of the 8 different observers. The distribution of measurements and differences are shown in (Figures 16 & 17). Differences as large as 2mm were observed. The standard deviation for the analytical variation for inter-observer variability was estimated as 0.35mm (Table 14).This level of variability compares to a standard deviation of 1.23mm between individuals, thus 93% of the total variability observed is attributed to individual variability and 7% to measurement error. With this level of measurement error, differences made by assessors similar to these would need to be at least 0.97 mm to be 95% likely to be real.

Figure 16: Scatter plot for sub-study B-Inter-observer repeatability of the same scans

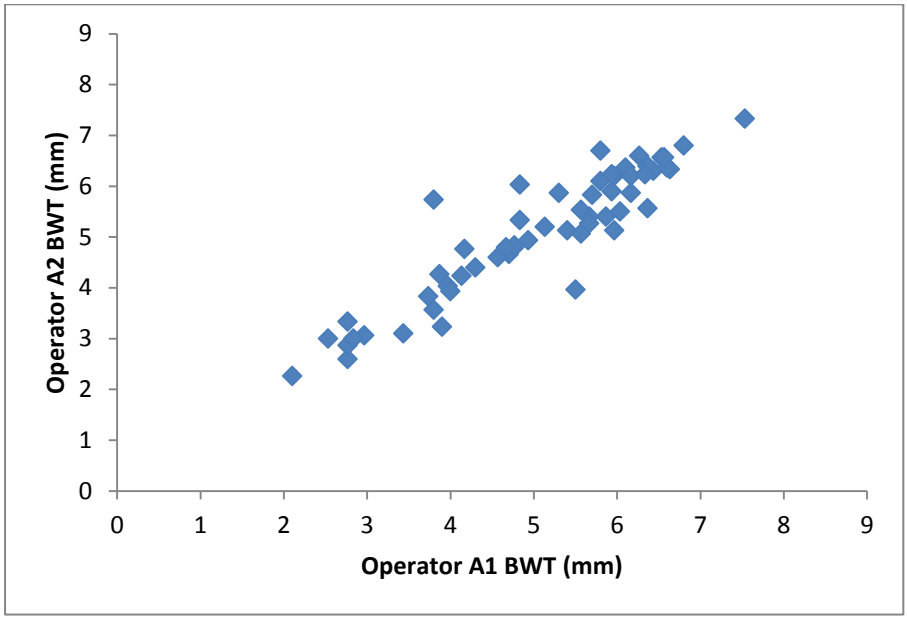
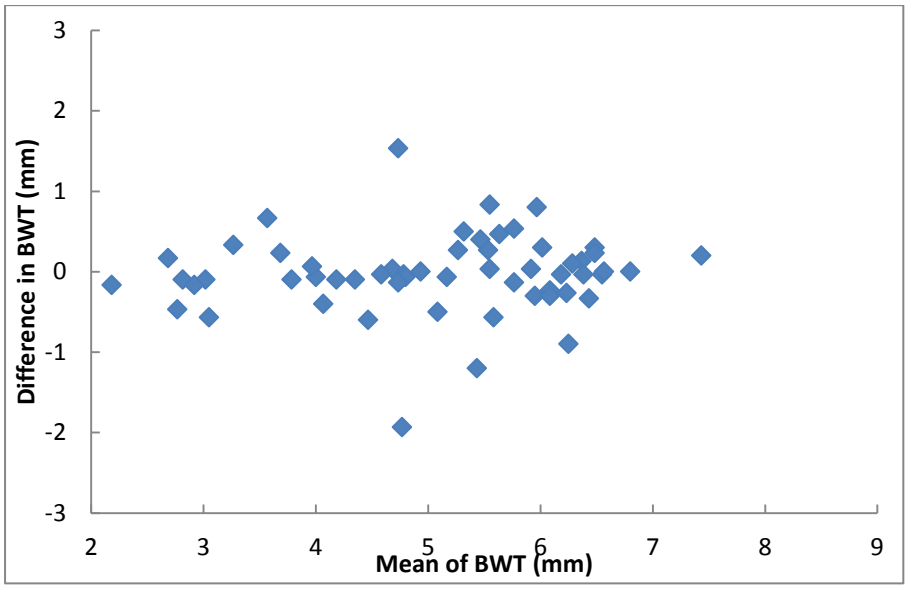


Figure 17: Bland Altman analysis for sub-study B



5.43 Sub-study C Inter-observer repeatability of different scans

Paired measurements were made for 27 women prospectively using 3 different observers. The design of Study C included estimation of variation occurring from repeated scans together with the interpretation of scans. The distribution of measurements is shown in (Figure 18), and of differences in (Figure 19).Maximum differences were again around 2mm, but they were more common in this sub-study than in previous sub-studies A and B.

Figure 18: Scatter plot for sub-study C- Inter-observer repeatability of different scans

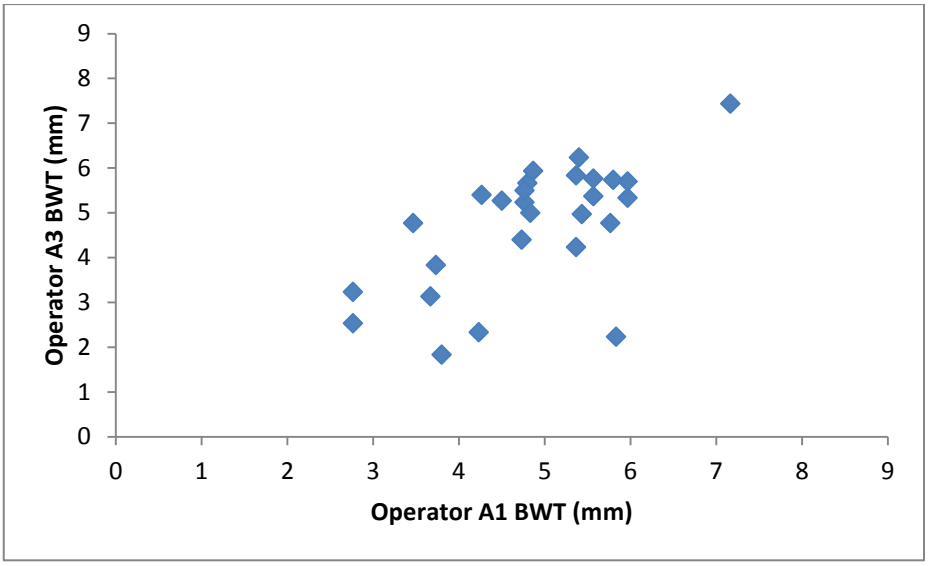
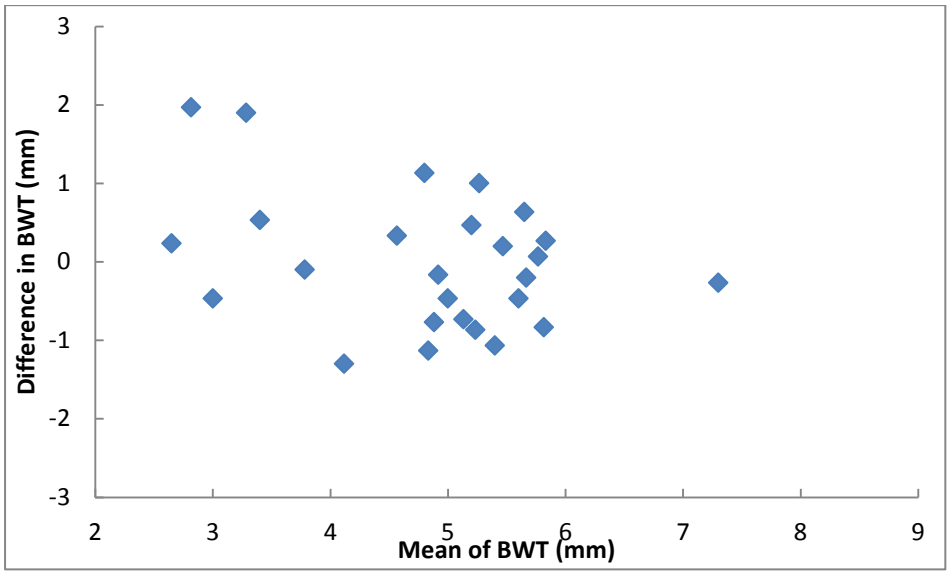


Figure 19: Bland-Altman analysis for sub-study C



The standard deviation for the analytical variation for intra-observer variability of repeated scans was estimated as 0.76mm (Table 14). This level of variability compares to a standard deviation of 0.95mm between individuals, thus 61% of the total variability observed is attributed to individual variability and 39% to measurement error. With this level of measurement error, differences made by assessors similar to these would need to be at least 2.11 mm to be 95% likely to be real.

Table 12: Comparison of distribution of BWT measures between the full BUS study and the sub-studies

Study	**Observer**	**Bladder wall thickness (mm)**			
		N	Mean	SD	range
BUS	Multi-centre study	645	4.78	1.34	1.07-9.60
Sub-study A	Observer 1	37	5.60	1.14	3.40-7.73
	Observer 2	37	5.95	1.08	4.00-8.03
Sub-study B	Observer 1	57	5.05	1.30	2.10-7.53
	Observers 2-6	57	5.07	1.26	2.27-7.33
Sub-study C	Observer 1	27	4.86	1.04	2.77-7.17
	Observers 2-3	27	4.73	1.38	1.83-7.43

Table 13: Details the experience of the observers in the sub studies

Study	Observer and location	'Skill level' (years of experience in O & G)
Sub-study A	Observer 1: Birmingham (n=37 x 2);	15
Sub-study B	Observer 1: Birmingham (n=57)	15
	Observer 2: Birmingham (n=34)	10
	Observer 3: Birmingham (n=4)	5
	Observers 4-6 Bournemouth (n=7); St Mary's (n=4); Medway (n=8);	7-20
Sub-study C	Observer 1: Birmingham (n=27)	15
	Observer 2: Birmingham (n=16)	15
	Observer 3: Birmingham (n=11)	10

Table 14: Estimates of measures of analytical and individual variability

Sub-study (interpretation)	Individual variability SD (mm)	Analytical variability SD (mm)	ICC, 95%CI	Smallest real difference (mm)
Sub-study A (Intra-observer of same scans)	1.04	0.42	0.86 (0.75, 0.92)	1.16
Sub-study B (Inter-observer of same scans)	1.23	0.35	0.93 (0.88, 0.96)	0.97
Sub-study C (Inter-observer from repeat scans)	0.95	0.76	0.61 (0.32, 0.80)	2.11

5.5 Discussion

We undertook three separate studies to investigate the reliability and repeatability of measurements of BWT using transvaginal ultrasound. For the intra-observer variation study, the individual variability in standard deviation (SD) was 1.04mm with an analytical variation of 0.42mm.In the inter-observer variation study on interpretation of stored images, we found an SD of 1.23mm with an analytical variation of 0.35mm and for the inter-observer variation

of repeated scans performed prospectively, the SD was 0.95mm with an analytical variability of 0.76mm (Table 14).

Our analyses have found that differences of less than 2mm in BWT cannot be safely interpreted as indicating real differences in BWT, as such differences are in the realms of those attributable to analytical variability (measurement error).We observed that the process of interpreting scans introduces measurement error of around 1mm, suggesting that the remaining 1mm is attributable to a combination of the scanning process and biological variability. The sub-studies were also designed to identify the magnitude of the possible sources of the analytical variability.

We failed to assess whether the differences in interpreting the scans arise because of within or between observer variability. Surprisingly our estimate of intra-observer variation is greater than that of inter observer variation. As the study samples are not large, and different scans were assessed for these repeat measurements, this observation potentially could be explicable by the play of chance or confounding. However, problems were also experienced with the quality of stored images used in sub-study A and B from the Birmingham Women's hospital which were not saved using DICOM software. The deteriorating quality of stored images may have resulted in measurements on stored images slightly greater than real time measurements and explain the finding that second reads of scans in sub-study A on average gave BWT measures 0.6mm greater than the original measurement.

5.51 Findings in the context of existing evidence

Six previous studies of reliability and reproducibility of BWT have used a variety of ultrasound techniques including transabdominal (Kuo 2009;Pannek, Bartel, Gocking, &

Frotzler 2012) and translabial (Lekskulchai & Dietz 2008) as well as transvaginal scanning (Chung, Liao, Chen, & Kuo 2011;Khullar, Salvatore, Cardozo, Bourne, Abbott, & Kelleher 1994;Panayi et al. 2010b;Tubaro A et al. 2013) as summarised in (Table 15). Several of these studies investigated women with different or a mixture of LUTS, or used varying levels of bladder filling. Of the transvaginal ultrasound studies only two (Kuo 2009;Panayi, Khullar, Fernando, & Tekkis 2010a) included assessment of repeated scans carried out at different points of time as in our third sub-study, evaluating 10 and 25 women respectively. The other two studies (Khullar, Salvatore, Cardozo, Bourne, Abbott, & Kelleher 1994;Tubaro A et al. 2013) evaluated the reproducibility of image interpretation (as in our first and second sub-studies) based on repeated assessment of 10 and 1544 images respectively.

Comparison of findings between these previous studies and the BUS study is complicated due to inappropriate use of Bland-Altman analyses and Pearson's correlation coefficients. Neither of these methods directly estimated the degree of analytical variability allowing assessment of the reproducibility of the measurement and the signal to noise ratio, although a pseudo estimate of analytical variability can be computed from the standard deviation of the differences. Reporting of the study design and statistical analysis was often incomplete or ambiguous. (Table 15) (Panayi, Khullar, Fernando, & Tekkis 2010a;Tubaro A, Khullar V, Oelke, Wijkstra H, Tretter R, Stow B, Huang M, Compion G, & Robinson D 2013). Estimates of the standard deviation of analytical variability vary between 0.3mm and 1.3mm for inter- and intra-observer variation for image interpretation (compared to 0.3-0.4mm for the BUS study), and a standard deviation of 0.4mm for inter-observer variation for repeated scans (compared to 0.8 mm for the BUS study). Kuo et al only reported Pearson's correlation coefficients from which no useful measures of reproducibility can be obtained (Kuo 2009).

5.52 Strengths and limitations of the study

The women reported to our study were recruited prospectively as part of the BUS test accuracy study, which involved good characterisation of symptoms and disease state. They were recruited from standard NHS incontinence clinics, and are thus highly likely to be representative of women in whom BWT would be measured as part of the diagnosis of DO. The sub-samples in each study appear to be representative of the larger cohort. BWT measurements were made according to a standardised protocol implemented following a programme of rigorous investigator training implemented in the larger study, which will have minimised variability due to differences in technique.

Both the number of women and the number of assessors limit the precision of the estimates made, although the study was larger than many previously undertaken. The assessors who partook in the intra observer studies generally had high levels of experience and expertise with the techniques, such that the estimates of operator dependent analytical variability may be lower than those in standard practice (Table 13).

In sub-studies A and B, the use of stored images of the original bladder transvaginal scan was problematic because those not stored using DICOM software were of poorer quality than the original images, as reported in previous inter-observer variation studies using ultrasound (Amer et al. 2002). We found that the brightness of echogenic serosa and mucosa were reduced in the stored ultrasound images, making the bladder wall less distinct and reducing the ability to place the calipers accurately. This may have led to over-estimation in analytical variation in sub-studies A and B, particularly in A, where no images were stored using DICOM.

The site of measurement is important. At the dome of the bladder, the peritoneal reflection may add 1 mm to the apparent wall thickness (Jequier and Rousseau 1987).Visualizing and measuring anterior part of the bladder wall can be challenging in patients with cystocele and those who have undergone bladder-neck surgery. Even in patients with normal anatomy, it may be difficult to see the anterior bladder wall because of the ring-down artefact of the anterior abdominal wall and the shading by the acoustic shadow from the pubic symphysis (Yang and Huang 2003). Sometimes, the bladder outline may be irregular and undulating and the margins may not be seen clearly resulting in significant variation in thickness measurements of stored scan images. The normal bladder is indented by surrounding organs (bowel, uterus).

Depending on how the surrounding structures were indenting the bladder, sometimes the image may show the cross-section of the bladder wall perpendicularly and sometimes tangentially. In addition, the variable shape of the bladder with different degrees of filling may result in errors in calculating bladder volume (Rachaneni et al. 2013a). These variables in BWT measurement may explain the difference in prospective measurements of BWT (sub-study C).

5.53 Interpretation

Ultrasound measurements of BWT have a high level of analytical variation arising from the scanning technique, underlying biological variability and interpretation of images, such that only differences greater than 2mm should be interpreted as indicating real changes in BWT. The range of BWT measurements observed in the full cohort ranged from 1mm to 10mm.To illustrate the potential impact of measurement error of this magnitude, if a threshold of 5mm is used to define test positives, those between 3 and 5mm could be mis-classified as test

negatives through measurement error, and those between 5 and 7mm be misclassified as test positives. In the BUS cohort of 645 women, 326 had values between 3mm and <5mm, and 217 had values between 5mm and <7mm. These groups constitute 84% of the complete sample. Only 41 women (6%) had BWT measures of 7mm and over, and 61 (10%) have measurements less than 3mm.Thus for the majority of women included in the study there is a possibility that a transvaginal ultrasound measurement of BWT would misclassify them using a 5mm threshold due to analytical variation. Rates of potential misclassification for higher or lower thresholds would be lower, but still substantial.

5.6 Conclusion

In the presence of high levels of analytical variation (measurement error) for a relatively small measurement of BWT, it is unlikely that BWT measurement by transvaginal ultrasound has sufficient reliability and reproducibility to be an accurate diagnostic test.

Table 15: Pre-existing evidence on inter and intra-observer variation of BWT

Study	Patients and study design	Technique and route of scan	Results presented	Comment on results
Khullar 1994 Ultrasound Journal of Obstetrics and Gynaecology	10 women each received one scan which was interpreted twice by each of two readers	Transvaginal BWT	Intraobserver difference -0.02mm 95%CI* (-0.22,0.18) Interobserver difference 0.02mm 95%CI* (-0.32, 0.35) *paper states that the CI is computed as 2 standard deviations(SD) either side of the mean, not 2SE.	Not possible to ascertain the analytical variability, the smallest real difference or the intra-class correlation coefficient from the data presented. An approximate (under) estimate of the analytical variability can be obtained by dividing the SD of the differences by √2, i.e. 0.3mm for intra and 0.5mm for inter observers
Lekskulchai 2008 Ultrasound journal of Obstetrics and gynaecology	67 women each had one scan read once by two different readers	Translabial DWT at dome	Intraclass correlation coefficient estimate of ICC=0.82; 95%CI (0.63–0.91)	No estimates of analytical variability or smallest real difference can be computed.
Kuo 2009 IUJ	10 women each received two scans two weeks apart	Transvaginal and transabdominal DWT measurement	Pearson's correlation coefficients are reported for transvaginal measures: Bladder base- 0.833 (p=0.020), Anterior wall- 0.759 (p=0.05) Posterior wall-0.599 (p=0.155) Bladder neck-0.768 (p=0.044)	No estimates of analytical variability, smallest real difference or intraclass correlation coefficients can be computed.
Panayi 2010 BJUI	25 women each had two scans by two different operators on the same day.	Transvaginal BWT at Dome, Anterior wall, and Trigone	Mean difference and 95% confidence interval for the three locations are: 0.13mm (0.08–0.33) 0.10mm (−0.12–0.31) −0.22mm (−0.41–0.01)	Not possible to ascertain the analytical variability, the smallest real difference or the intraclass correlation coefficient from the data presented. An approximate (under) estimate of the analytical variability can be obtained by dividing the sd of differences by √2, i.e. 0.4mm for all three measures.
Pannek 2013 World Journal	10 women had two measurements made by	Transabdominal DWT at three	States that interobserver coefficient of variability	The mean DWT is not reported in the

of Urology	the same observer (and implies these were from different scans which were repeated immediately)	different sites of the bladder	was +14.78%, and the correlation (not stated whether Pearson's or ICC) was 0.984.	paper, thus it is not possible to deduce the analytical variability, the smallest real difference and the ICC.
Tubaro 2013 ICS abstract 136	40 women each had one scan which was interpreted twice by each of three readers A further 1504 images were assessed twice by different readers	Transvaginal BWT	Data were analyzed using the Bland-Altman method, and mean differences and confidence intervals within and between readers presented Standard deviations of differences between pairs of readers were 1.1mm, 1.7mm and 1.8mm.	Not possible to ascertain the analytical variability, the smallest real difference or the intraclass correlation coefficient from the data presented. An approximate (under) estimate of the analytical variability can be obtained by dividing the sd of differences by √2, i.e. 0.8 to 1.3mm

*DWT- detrusor wall thickness; sd- standard deviation

Chapter 6: Evaluation of patient acceptability of Bladder Ultrasound (BUS) and Urodynamics

Introduction

In this chapter a comparative evaluation of the acceptability of performing both transvaginal bladder ultrasound scanning to measure BWT and urodynamics from the patient's perspective was undertaken.

6.1 Methods

Participants (n=687) were the patients who took part in the BWT diagnostic accuracy study (details given in Chapters on methods and Results).They underwent BWT scan and urodynamics in the participating centres, where possible, carried out on the same day. If it was not possible for both the tests to be performed on the same day, they were completed within a four week period. Immediately after each test, acceptability questionnaires were given to the participant for completion. Items included in the questionnaire were as follows:

Pain measured using a Visual Analogue Scale (VAS) on a 0 (no recorded pain) to 100 (worst pain imaginable) scale during and shortly after testing (Jensen et al. 2003;Price et al. 1983).

Acceptability of testing is measured by STAI-SF (six item) to measure generic state anxiety. This is a validated and widely accepted instrument used to assess the intensity of current feelings in relation to how you "generally feel today" (Marteau and Bekker 1992). Scores ranged from 4 (most positive) to 24 (most negative). The short form was used to improve patient compliance as opposed to the long form to accommodate time constraints in busy clinics.

All responses were compared using paired (urodynamics versus BWT scan) methods for dependant data (Agresti A 2011). For VAS and STAI-SF scores, mean differences and 95% confidence intervals were calculated with statistical significance determined by a paired t-test. Wilcoxon signed-rank test was used for ordinal responses and McNemar's test for binary responses. Analysis was performed using SAS software, version 9.2 (SAS Institute).

6.2 Results

6.21 Pain during and after the tests

Pain scores during urodynamics testing (28.1 points) and shortly after (21.1) were higher than the corresponding scores during BWT scan (15.3 points) and after BWT scan (13.3 points) respectively (Table 16).

Table 16: Pain during after the urodynamics and bladder ultrasound tests

	Urodynamics Mean(SD, n)	**Ultrasound Mean (SD, n)**	**Difference 95% CI**	**p-value**
During	28.1 (28.4, 653)	12.8 (19.5, 646)	15.3 (13.1 to 17.6)	<0.001
After	21.1 (26.7, 648)	7.9 (16.1, 646)	13.3 (11.2 to 15.4)	<0.001

6.22 Acceptability

The proportion of women who found the test totally acceptable was higher with BWT scan compared with urodynamics (81% versus 56%; p<0.001), although the number reporting an unacceptable test was still relatively low following urodynamics (2%). (Table 17) More women found the exposure required for the test embarrassing with urodynamics compared with BWT scan (proportion reporting some embarrassment 64% v 48%; p<0.001).Fewer women felt that they would recommend urodynamics to a friend compared to BWT scan (86% v 96%; p<0.001) and have the same test again (88% v 97%, p<0.001).

Table 17: Acceptability of urodynamics and transvaginal bladder ultrasound tests

		Urodynamics (n= 653) **Frequency (%)**	**Ultrasound (n= 648)** **Frequency (%)**	**p-value**
Procedure acceptability	Totally	368 (56%)	521 (81%)	<0.001
	Generally	273 (42%)	124 (19%)	
	Unacceptable	12 (2%)	1 (<1%)	
Exposure for test embarrassing?	Extremely	55 (8%)	21 (3%)	<0.001
	Moderately	128 (20%)	63 (10%)	
	A little	237 (36%)	230 (35%)	
	No	233 (36%)	334 (52%)	
Recommend test to a friend?	Yes	559 (86%)	625 (96%)	<0.001
	No	93 (14%)	23 (4%)	
Have same test again?	Yes	572 (88%)	627 (97%)	<0.001
	No	78 (12%)	20 (3%)	

6.23 Anxiety

Anxiety levels associated with both tests appeared quite moderate (12.6 for BWT scan and 12.9 for urodynamics), although the scores were only slightly higher with urodynamics (0.3 points difference on a 4-24 scale, 95%CI: 0.1 to 0.5; p=0.02). (Table 18)

Table 18: **STAI-SF**

Urodynamics **Mean (SD, n)**	**Ultrasound** **Mean (SD, n)**	**Difference (95%CI)**	**p-value**
12.9 (3.8, 616)	12.6 (3.8, 602)	0.3 (0.1 to 0.5)	0.02

6.3 Discussion

Results show that BWT scan was more acceptable and less embarrassing and painful than

urodynamics. Despite this, a high proportion of women said that would recommend the urodynamics test to a friend (88%) and also have it repeated (86%). Anxiety scores were also higher for- urodynamics compared to BWT scan, though the mean difference was small. To our knowledge this is the first formal evaluation of the comparison of tolerability and acceptability of various diagnostic procedures to evaluate bladder function.

The question is whether urodynamics are likely to be deemed acceptable in terms of having them repeated in view of its clinical importance. Women were aware that their treatment plans were based on the urodynamics diagnoses and not that of BWT scan. This awareness of importance may have contributed to the improved acceptability of having this test repeated again if necessary. I have enclosed a comment made by one of our patients with regards to acceptability of urodynamics:

'Although BWT scan was a generally more acceptable test, I felt that because urodynamics physically replicated my symptoms, it allowed me a better understanding of my condition.'

Pain was higher with urodynamics although the average score was 28 on a scale of 0-100. Origin of pain may be multifactorial. Pain could be due to physical components like the urethral and rectal catheterisation and the artificial filling of the bladder. Following urodynamics, participants often expressed urethral pain. The exact mechanism of pain perception after the urodynamics is not known though slight trauma to the urethra may be considered to be an aetiological factor (Gorton and Stanton 1999).

Elevated anxiety levels during urodynamics may also have contributed to patient's perception of pain. Women are worried about going through invasive procedures and the perception of pain or discomfort associated with them(Marteau & Bekker 1992). Fear of perceiving pain

may increase the anxiety levels in patients going through dental and minor surgical procedures(Terry et al. 2007). The perception of pain and discomfort from invasive procedures may be worsened by pre-existing fear of pain (Bradley and Kennedy 2008).Fear of pain may increase the background anxiety levels (McNeil and Berryman 1989) and may influence their willingness towards invasive testing and follow-ups (Denberg et al. 2005). The level of apprehension and embarrassment during urodynamics had a positive correlation to the level of pain perception. Younger age and the fear of undergoing an invasive procedure were found to be the risk factors for the heightened perception of pain (Yiou et al. 2013).

Results on pain/anxiety and embarrassment provoked are similar to those mentioned in the literature. In a prospective study of 208 patients, although urodynamics was only associated with minor complications, it was perceived to be painful, (pain score of 3.1/10) worrying and traumatic (Ku et al. 2004). In a study of pre-test and post-test evaluation of anxiety with urodynamics, severe pre-test anxiety was reported in only a small fraction of women (4.6%) undergoing urodynamics. Following completion of urodynamics, women reported minimal or no anxiety(77.5%), minimal or no embarrassment (84.1%) and minimal or no physical discomfort (75.5%) (Neustaedter et al. 2011). Younger age, history of anxiety or depression, a diagnosis of OAB and painful bladder syndrome may lead to more negative experiences during urodynamics (Yeung, Eschenbacher, & Pauls 2014).

6.31 Strengths and Limitations

A strong component of the study was that a large number of patients were recruited (n=687) and comprised of a population derived from several geographical areas and various ethnicities within the UK. Data collected from multicentre studies may be more applicable and generalizable than that collected from a single centre. The study provides information to

assist in counselling women who may be apprehensive or anxious regarding an invasive test such as urodynamics.

The instrument to measure anxiety was administered only after each test procedure. The difference between pre and post-test questionnaires to know the anxiety provoked by each procedure should have been measured. This would have yielded valid data about fluctuations in anxiety state before and after each test and a comparison of the difference would have been ideal.

Participants were aware of the fact that the diagnoses and the plan of management were made on the basis the information gained during urodynamics. The awareness of the role of the test in clinical decision making may have introduced bias and influenced the participant's decision to recommend the test or have it repeated again if required. Participants were aware that the BWT scan was used for assessment only and did not aid in the management of a bladder diagnosis.

Compared to patients with low trait anxiety scores, patients with higher trait anxiety will experience increased anxiety with invasive procedures(Spielberger,1983). Anxious women have increased episodes of urgency incontinence and vice versa. High anxiety score on the STAI was a predisposing factor for UI (Marteau & Bekker 1992). UI may increase anxiety of possible social exclusion in the geriatric population(Agresti A 2011).Logically, women with higher trait anxiety levels should experience greater state anxiety to urodynamics compared to BWT scan. Unfortunately, the Trait anxiety was not studied due to time constraints in our busy clinics.

The impact of the information given to the women before the procedure and how well we prepared the women for each diagnostic test was not studied. The information leaflets about both urodynamics and BWT scan procedures were sent out along with the appointment letters. During the clinic visit, the patient's understanding of each procedure was further reinforced. This was our routine practice based on the previous evidence that women were likely to find the test less distressing when they knew what to expect during the course of the investigation(Gorton & Stanton 1999).

6.32 Interpretation of findings

Urodynamics procedure had statistically significant higher levels of pain and a lower rate of acceptability compared to the ultrasound. Inspite of this, majority of women would have repeat urodynamics if needed. The results of this study will help in the preparation of information leaflets for women on urodynamics and/or transvaginal ultrasound, to get a realistic picture of patient experience of these invasive investigations.

6.4 Conclusion

The elevated anxiety levels elicited by invasive diagnostic testing on a background of increased trait anxiety due to lower urinary tract symptoms need careful evaluation and interpretation. This evidence may improve the support women receive during invasive diagnostic testing (urodynamic testing) and their subsequent satisfaction.

Chapter 7: Discussion

7.1 Introduction

Sometimes diagnostic tests are applied in clinical practice without proper evaluation of their test accuracy. Revision of formal probability of a diagnosis with the availability of a test result is essential to decision making. Clinicians seem to use the pre-test probability of DO and formally revise the diagnosis after urodynamics despite the absence of evidence of quantitative accuracy of urodynamic test.

Three distinct studies have been completed as part of the BUS trial:

1. To determine the accuracy (sensitivity, specificity, predictive values, ROC curves, AUC) of transvaginal BWT as the index test and urodynamics as the reference standard in women with OAB

2. To determine the reproducibility of the index test BWT

3. To determine the acceptability of BWT and compare it with urodynamics acceptability

This chapter attempts to focus on the key findings and limitations emerging from the BUS study.

7.2 Evaluation of diagnostic tests

There was no evidence that BWT had any relationship with DO, regardless of the cut-off point (AUC: 0.53, 95% CI: 0.48, 0.57) (Figure: 12). On univariate analyses of the effect of various risk factors for predicting DO(Appendix:5.2), age, BMI, ethnicity, history of mixed incontinence and history of onset of the type of incontinence (SUI first or urgency incontinence first),parity, menopausal status, voiding difficulties, previous history of POP or

incontinence surgery and previous treatment with antimuscarinics were analysed. Only one variable-previous history of treatment with antimuscarinics seems to have attained statistical significance with a p-value of 0.0005, with an OR and 95% CI of 1.84 (1.30, 2.59). This finding could be interpreted as women were appropriately referred to secondary care (NICE CG 171) after a failed trial of antimuscarinics and are probably in the more severe spectrum of the disease.

The effect of different variables in the diagnostic accuracy of BWT in DO were also studied as follows: When a history of urgency alone was considered excluding a history of mixed incontinence, the AUC was 0.528(Appendix: 4.1);When the urodynamic diagnosis of isolated DO was considered after excluding women with urodynamic diagnoses of mixed DO and USI, AUC was only 0.521, 95%CI: (0.476, 0.566) (p=0.37 compared to AUC=0.50) (Appendix: 4.2) Similar results were obtained in women with wet DO on urodynamics. The AUC was only 0.548, 95%CI: (0.502, 0.594) (p=0.04 compared to AUC=0.50) (Appendix: 4.3). None of these variables seemed to have had any effect on the diagnostic accuracy of BWT.

I undertook three separate studies to investigate the intra- and inter-observer variation of BWT using transvaginal ultrasound and concluded that it was unlikely that this measurement would be sufficiently reliable or reproducible to be an accurate diagnostic test. Only differences greater than 2mm could be safely interpreted as real change in BWT meaning that for the vast majority of women (84%) there could be some possibility of misclassification when using a cut-off of 5mm.

Transvaginal ultrasound was more acceptable as well as less painful than urodynamics. Surprisingly, a high proportion of women said that they would recommend the urodynamics test to a friend (88%) and also have it repeated (86%).

7.3 Strengths and limitations

The methodologically robust test accuracy study design and appropriate conduct of the study increase the confidence in the estimates of diagnostic accuracy of BWT. All the criteria for a high-quality test accuracy evaluation, the STARD checklist (Bossuyt, Reitsma, Bruns, Gatsonis, Glasziou, Irwig, Lijmer, Moher, Rennie, & de Vet 2003) have been met (Appendix:1) The study population included only a specific group of women with urgency predominant mixed incontinence/OAB compared to all other studies(Kuhn, Genoud, Robinson, Herrmann, Gunthert, Brandner, & Raio 2011;Lekskulchai & Dietz 2008) where BWT was studied in women with all types of LUTS. There was a very high proportion of index test verification by reference standard; over 93% (644/687) ensuring a near complete verification design. The strengths also include recruitment of a large sample of 687 women in excess of the target sample size of 600, multicentre setting, prospective and consecutive recruitment. Study population were generalizable across the NHS with varying degrees of disease severity and a good representation of ethnic mix (22% non-caucasians) (Table: 6). The large sample also provided the opportunity to analyse the acceptability of both the index test and the reference standard.

Bias in the study methodology was minimised by ensuring that two independent operators performed the index tests and reference standard blinded to each other. Standard protocols for index and reference standard testing were drawn up and quality assurance checks were made

and a pre-specified statistical plan was followed. Comprehensive oversight via an independent Data Monitoring/Trial Steering Committee was employed.

The fact that DO was picked up in 60% of the OAB/urgency predominant MUI population which was similar to other studies on DO prevalence in OAB (Hashim & Abrams 2006) indicate that the study sampling was of the appropriate quality. On evaluation of the reproducibility of the index test, BWT, the results were similar to the SHRINK study (Tubaro A, Khullar V, Oelke, Wijkstra H, Tretter R, Stow B, Huang M, Compion G, & Robinson D 2013), which is a retrospective study of centralised measurement of BWT in 1544 patients with OAB on antimuscarinics. Though their conclusion was that of good reproducibility of BWT, we found significant disagreement between the independent observers on Bland Altman's analysis. There has been no consensus of the acceptable limits of interobserver difference in the published literature for BWT measurement. Contrary to the previous studies on the reproducibility of BWT measurements, BUS study is the first to analyse prospectively, the data on the agreement between the two observers (Bland and Altman 1986). Hence my conclusion of lack of reproducibility of BWT may be considered robust.

Studies have utilised Medical, epidemiological and Social aspects of aging questionnaire (MESA) sub-scale scores for urgency and stress urinary incontinence. They have categorised women who had higher MESA scores for urgency incontinence (compared to their MESA subscale scores for SUI) as urgency predominant mixed incontinence along with urogenital distress inventory (UDI) scores, bladder diaries and urodynamics (Brubaker, Lukacz, Burgio, Zimmern, Norton, Leng, Johnson, Kraus, & Stoddard 2011). We may be criticised for not objectively assessing the severity of each type of incontinence before deciding whether the

patient suffered from urgency predominant mixed incontinence. In this study, direct patient questioning of which is the most bothersome type of incontinence on history taking was employed which is a true reflection of routine clinical practice.

The reference standard, urodynamics is known for its uncertain reproducibility. Previous studies on the reproducibility of urodynamics in OAB reported inconsistencies between serial urodynamic procedures on repeat urodynamics in the same patient (Broekhuis, Kluivers, Hendriks, Massolt, Groen, & Vierhout 2010;Homma, Kondo, Takahashi, Kitamura, & Kawabe 2000;Mortensen et al. 2002;Sorensen et al. 1984). Repeat urodynamics in the same session was poorly reproducible and showed an improved bladder function after the first urodynamics. Nearly 50% of the patients had more than ≥25% change in one or more variables on repeat urodynamics (Homma, Kondo, Takahashi, Kitamura, & Kawabe 2000). Similar findings were reported in another study on healthy volunteers which evaluated repeat urodynamics in the same session with an increase in the volume of first and normal desire during the second urodynamic procedure(Brostrom, Jennum, & Lose 2002).In a multicentre study with six serial urodynamics in patients with OAB, there was increased variability in pressure measurements than volume measurements (Frenkl, Railkar, Palcza, Scott, Alon, Green, & Schaefer 2011). When repeated in the same session in patients with neurogenic LUTS, difference against mean detrusor pressure at maximum flow (*Pdet Qmax*), voided volume and post-void residual urine (*PVR*) showed wide confidence limits of agreement reflecting poor reproducibility and unacceptable discrepancy. However there was excellent reproducibility of neurogenic DO in both the tests (κ=0.87, 95% CI 0.80-0.94) (Bellucci, Wollner, Gregorini, Birnbock, Kozomara, Mehnert, & Kessler 2012). In a videourodynamic study on the reproducibility of voiding flowmetry parameters, high interobserver and

intraobserver agreement of Pdet, Qmax , opening detrusor pressure and closing detrusor pressure was observed (Digesu, Hutchings, Salvatore, Selvaggi, & Khullar 2003b). The poor reproducibility of urodynamics could be due to a combination of variation in the physiology of bladder function and the poor sensitivity and specificity of the technology used (Gupta et al. 2004).

Having an imperfect gold standard (poor reproducibility of urodynamics) for reference may have affected the test accuracy of our index test, the BWT. However, in the absence of another reliable diagnostic test for lower urinary tract dysfunction, we have to utilise urodynamic test which has been the only available diagnostic test so far.

The study design selection of test accuracy for evaluating the accuracy of BWT in diagnosing DO may be questioned. In asymptomatic conditions like cervical premalignant conditions, test accuracy statistics of screening and diagnostic test results may have a major role in clinical decision making. There is some emerging evidence that making a diagnosis of DO does not alter treatment outcomes for interventions including use of antimuscarinics (Malone-Lee and Al-Buheissi 2009), Botulinum toxin, percutaneous tibial nerve stimulation (PTNS) (Vandoninck et al. 2003) and sacral neuromodulation (SNS) (South et al. 2007). This is why one may question the role of diagnostic tests in this condition as it might not alter patient management.

A randomised controlled trial (RCT) comparing treatment outcomes of women who undergo BWT along with office evaluation compared to office evaluation and urodynamics may have been a stronger study design. This design could have shed light on the influence of evaluated

test on the management and prognosis of the OAB/urgency incontinence patients. Nevertheless, the study designs selected depend on the preexisting evidence.

7.4 Performance of other biomarkers for DO

Ongoing research to identify diagnostic and prognostic biomarkers for OAB like Nerve growth factor (NGF), Brain derived nerve factor(BDNF), urinary neurotropins, urinary adenosine triphosphate, urinary prostaglandins and cytokines has not yielded any conclusive result (Antunes-Lopes et al. 2014;Fry et al. 2014).The studies on the role of these biomarkers suffer from serious flaws in design and are under powered (Rachaneni et al. 2013b).The prognostic value of these biomarkers has not been demonstrated (Antunes-Lopes, Cruz, Cruz, & Sievert 2014). Large scale research into the pathophysiology of normal and abnormal bladder function is needed before attempts to identify biomarkers for OAB for non-invasive diagnoses. Ultimately, such knowledge will contribute to an improved understanding of the bladder function and will pave the way towards new treatments for LUTS.

7.5 Implications for practice

Based on the results of the BUS study, I conclude that there is no place for BWT measurement in managing women with OAB. History, clinical examination, incontinence questionnaires, bladder diaries and urodynamic diagnoses are all part of management algorithm of patients with OAB. Hence, the decision making for treatment options may not be based on the diagnoses of DO on urodynamics alone.Evidence from clinical studies indicate that in women with OAB, an urodynamic observation of DO alone does not impact on the outcome of invasive interventions (Groenendijk et al. 2008;Rovner et al. 2011). The question arises whether we need to subject our women to invasive diagnostic tests like urodynamics and/or BWT scan to diagnose DO if it not going to have an impact on the treatment outcome.

The pathology behind the symptoms of OAB need to be studied before any attempts to identify diagnostic markers. In women with OAB/urgency predominant MUI, randomised controlled trials comparing treatment based on urodynamics diagnoses compared to treatment based on clinical assessment (history and examination alone) and related health economic evaluation for these diagnostic interventions are required to consolidate the role of urodynamics in the management of OAB/MUI women. The composite role of patient reported outcome measure like ICIQ-OAB SF questionnaires, bladder diaries and changes in the quality of life measures need to be established in measuring treatment responses to various interventions.Diagnostic accuracy of individual components of office evaluation of OAB may be different compared to the composite test accuracy of all the components of office evaluation. Studying composite test accuracy of various components in this context may be a highly complicated exercise. Further studies need to be planned to look into composite test accuracy of office evaluation with or without urodynamics in OAB.

7.6 Conclusion

Bladder wall thickness measurement did not discriminate women with DO versus those without DO and hence is not an accurate test for diagnosing DO. In the presence of high levels of analytical variation for the measurement of BWT, it is unlikely that BWT measurement made by transvaginal ultrasound has sufficient reliability and reproducibility to be a precise diagnostic test. Women experienced higher levels of embarrassment and a lower rate of acceptability with urodynamics compared to the BWT scan procedure.

7.7 Funding

The research costs of the study were funded by a grant from the National Institute for Health Research (NIHR) Health Technology Assessment Unit awarded to the University of Birmingham.

References

Abou-Gamrah, A., Fawzy, M., Sammour, H., & Tadros, S. 2014. Ultrasound assessment of bladder wall thickness as a screening test for detrusor instability. *Arch.Gynecol.Obstet.*, 289, (5) 1023-1028 available from: PM:24276424

Abrams, P. Urodynamics. [Third], 17-67. 2006. Springer-Verlag.
Ref Type: Online Source

Abrams, P., Artibani, W., Cardozo, L., Dmochowski, R., van, K.P., & Sand, P. 2009. Reviewing the ICS 2002 terminology report: the ongoing debate. *Neurourol.Urodyn.*, 28, (4) 287 available from: PM:19350662

Abrams, P. 2003. Describing bladder storage function: overactive bladder syndrome and detrusor overactivity. *Urology*, 62, (5, Supplement 2) 28-37 available from: http://www.sciencedirect.com/science/article/pii/S0090429503010501

Agresti A. Score and Pseudo-Score Confidence Intervals for Categorical Data Analysis. Stat Biopharm Res 3[2], 163-172. 2011.
Ref Type: Generic

Amer, S.A., Li, T.C., Bygrave, C., Sprigg, A., Saravelos, H., & Cooke, I.D. 2002. An evaluation of the inter-observer and intra-observer variability of the ultrasound diagnosis of polycystic ovaries. *Hum.Reprod.*, 17, (6) 1616-1622 available from: PM:12042287

Andersson, K.E. 2002. Bladder activation: afferent mechanisms. *Urology*, 59, (5 Suppl 1) 43-50 available from: PM:12007522

Andersson, K.E. 2006. Treatment-resistant detrusor overactivity--underlying pharmacology and potential mechanisms. *Int J Clin.Pract.Suppl* (151) 8-16 available from: PM:17169005

Andersson, K.E. & Wein, A.J. 2004. Pharmacology of the lower urinary tract: basis for current and future treatments of urinary incontinence. *Pharmacol.Rev.*, 56, (4) 581-631 available from: PM:15602011

Antunes-Lopes, T., Cruz, C.D., Cruz, F., & Sievert, K.D. 2014. Biomarkers in lower urinary tract symptoms/overactive bladder: a critical overview. *Curr.Opin.Urol.*, 24, (4) 352-357 available from: PM:24841379

Atalabi, O.M., Morhason-Bello, I.O., Adekanmi, A.J., Marinho, A.O., Adedokun, B.O., Kalejaiye, A.O., Sogo, K., & Gbadamosi, S.A. 2012. Transvaginal ultrasonography: a survey of the acceptability and its predictors among a native African women population. *Int J Womens Health*, 4, 1-6 available from: PM:22312194

Awad, S.A. & McGinnis, R.H. 1983. Factors that influence the incidence of detrusor instability in women. *J Urol*, 130, (1) 114-115 available from: PM:6683325

Bellucci, C.H., Wollner, J., Gregorini, F., Birnbock, D., Kozomara, M., Mehnert, U., & Kessler, T.M. 2012. Neurogenic Lower Urinary Tract Dysfunction-Do We Need Same Session Repeat Urodynamic Investigations? *The Journal of Urology*, 187, (4) 1318-1323 available from: http://www.sciencedirect.com/science/article/pii/S0022534711057855

Bland, J.M. & Altman, D.G. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, 1, (8476) 307-310 available from: PM:2868172

Blatt, A.H., Titus, J., & Chan, L. 2008. Ultrasound measurement of bladder wall thickness in the assessment of voiding dysfunction
199. *J Urol*, 179, (6) 2275-2278 available from: PM:18423703

Bossuyt, P.M., Reitsma, J.B., Bruns, D.E., Gatsonis, C.A., Glasziou, P.P., Irwig, L.M., Lijmer, J.G., Moher, D., Rennie, D., & de Vet, H.C. 2003. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ*, 326, (7379) 41-44 available from: PM:12511463

Brading, A.F. 1997. A myogenic basis for the overactive bladder. *Urology*, 50, (6A Suppl) 57-67 available from: PM:9426752

Bradley, P. & Kennedy, Y. 2008. Confronting pain face-to-face at NSU. *Todays.FDA.*, 20, (10) 37-40 available from: PM:19266691

Broekhuis, S.R., Kluivers, K.B., Hendriks, J.C., Massolt, E.T., Groen, J., & Vierhout, M.E. 2010. Reproducibility of same session repeated cystometry and pressure-flow studies in women with symptoms of urinary incontinence. *Neurourol.Urodyn.*, 29, (3) 428-431 available from: PM:19618451

Brostrom, S., Jennum, P., & Lose, G. 2002. Short-term reproducibility of cystometry and pressure-flow micturition studies in healthy women. *Neurourology and Urodynamics*, 21, (5) 457-460 available from: http://dx.doi.org/10.1002/nau.10019

Brubaker, L., Lukacz, E.S., Burgio, K., Zimmern, P., Norton, P., Leng, W., Johnson, H., Kraus, S., & Stoddard, A. 2011. Mixed incontinence: comparing definitions in non-surgical patients. *Neurourol.Urodyn.*, 30, (1) 47-51 available from: PM:21181960

Chan L, The S, Tse V, & Titus J. The value of bladder wall thickness measurement in the assessment of overactive bladder syndrome. Ultrasound in Obstetrics & Gynecology ; **26**, 376-471. 2005. *15th World Congress on Ultrasound in Obstetrics and Gynecology*.
Ref Type: Generic

Chung, S.D., Liao, C.H., Chen, Y.C., & Kuo, H.C. 2011. Urgency severity scale could predict urodynamic detrusor overactivity in patients with overactive bladder syndrome. *Neurourol.Urodyn.*, 30, (7) 1300-1304 available from: PM:21560153

Chung, S.D., Chiu, B., Kuo, H.C., Chuang, Y.C., Wang, C.C., Guan, Z., & Chancellor, M.B. 2010. Transabdominal ultrasonography of detrusor wall thickness in women with overactive bladder
177. *BJU Int*, 105, (5) 668-672 available from: PM:19793377

Clement, S., Candy, B., Heath, V., To, M., & Nicolaides, K.H. 2003. Transvaginal ultrasound in pregnancy: its acceptability to women and maternal psychological morbidity. *Ultrasound Obstet Gynecol*, 22, (5) 508-514 available from: PM:14618665

Clement, K. D., Lapitan, M. C., Omar, M. I., and Glazener, C. M. 2013 Urodynamic studies for management of urinary incontinence in children and adults. Cochrane.Database.Syst.Rev. 10, CD003195.

Colli, E., Artibani, W., Goka, J., Parazzini, F., & Wein, A.J. 2003. Are urodynamic tests useful tools for the initial conservative management of non-neurogenic urinary incontinence? A review of the literature. *Eur.Urol*, 43, (1) 63-69 available from: PM:12507545

Coyne, K.S., Sexton, C.C., Bell, J.A., Thompson, C.L., Dmochowski, R., Bavendam, T., Chen, C.I., & Quentin, C.J. 2013. The prevalence of lower urinary tract symptoms (LUTS) and overactive bladder (OAB) by racial/ethnic group and age: results from OAB-POLL. *Neurourol.Urodyn.*, 32, (3) 230-237 available from: PM:22847394

Coyne, K.S., Sexton, C.C., Thompson, C.L., Milsom, I., Irwin, D., Kopp, Z.S., Chapple, C.R., Kaplan, S., Tubaro, A., Aiyer, L.P., & Wein, A.J. 2009. The prevalence of lower urinary tract symptoms (LUTS) in the USA, the UK and Sweden: results from the Epidemiology of LUTS (EpiLUTS) study. *BJU Int*, 104, (3) 352-360 available from: PM:19281467

Darblade, B., Behr-Roussel, D., Oger, S., Hieble, J.P., Lebret, T., Gorny, D., Benoit, G., Alexandre, L., & Giuliano, F. 2006. Effects of potassium channel modulators on human detrusor smooth muscle myogenic phasic contractile activity: potential therapeutic targets for overactive bladder. *Urology*, 68, (2) 442-448 available from: PM:16904481

de Groat, W.C. 2006. Integrative control of the lower urinary tract: preclinical perspective. *Br J Pharmacol.*, 147 Suppl 2, S25-S40 available from: PM:16465182

Deeks, J.J. & Altman, D.G. 2004. Diagnostic tests 4: likelihood ratios. *BMJ*, 329, (7458) 168-169 available from: PM:15258077

Denberg, T.D., Melhado, T.V., Coombes, J.M., Beaty, B.L., Berman, K., Byers, T.E., Marcus, A.C., Steiner, J.F., & Ahnen, D.J. 2005. Predictors of nonadherence to screening colonoscopy. *J Gen.Intern.Med.*, 20, (11) 989-995 available from: PM:16307622

Digesu, G.A., Khullar, V., Cardozo, L., & Salvatore, S. 2003a. Overactive bladder symptoms: do we need urodynamics? *Neurourol.Urodyn.*, 22, (2) 105-108 available from: PM:12579626

Digesu, G.A., Hutchings, A., Salvatore, S., Selvaggi, L., & Khullar, V. 2003b. Reproducibility and reliability of pressure flow parameters in women. *BJOG.*, 110, (8) 774-776 available from: PM:12892691

Digesu, G.A., Hutchings, A., Salvatore, S., Selvaggi, L., Milani, R., & Khullar, V. 2004. Pressure flow study: a useful diagnostic test of female lower urinary tract symptoms. *Neurourol.Urodyn.*, 23, (2) 104-108 available from: PM:14983419

Dmochowski, R., Sanders S W, Rodney A A, Victor W N, & Willy Davila G. Bladder-health diaries: an assessment of 3-day vs 7-day entries. BJUI 96[Issue 7], 1049-1054. 2005.
Ref Type: Online Source

Dokmeci, F., Seval, M., & Gok, H. 2010. Comparison of ambulatory versus conventional urodynamics in females with urinary incontinence. *Neurourol.Urodyn.*, 29, (4) 518-521 available from: PM:19731314

Elbadawi, A., Diokno, A.C., & Millard, R.J. 1998. The aging bladder: morphology and urodynamics. *World J Urol*, 16 Suppl 1, S10-S34 available from: PM:9775413

Farag, F.F. & Heesakkers, J.P. 2011. Non-invasive techniques in the diagnosis of bladder storage disorders. *Neurourol.Urodyn.*, 30, (8) 1422-1428 available from: PM:21780168

Foon, R., Toozs-Hobson, P., & Latthe, P. 2012. Prophylactic antibiotics to reduce the risk of urinary tract infections after urodynamic studies. *Cochrane Database.Syst.Rev.*, 10, CD008224 available from: PM:23076941

Fransisco Cruz, Heesakkers, J. P., Khullar V, & Tubaro A. Bladder Wall Thickness in Overactive Bladder: Does It
Have a Role? EUROPEAN UROLOGY SUPPLEMENT S 8 8, 769-771. 2009.
Ref Type: Online Source

Freeman, R.M. & Adekanmi, O.A. 2005. Overactive bladder. *Best Pract.Res.Clin.Obstet Gynaecol.*, 19, (6) 829-841 available from: PM:16176887

Frenkl, T.L., Railkar, R., Palcza, J., Scott, B.B., Alon, A., Green, S., & Schaefer, W. 2011. Variability of urodynamic parameters in patients with overactive bladder. *Neurourol.Urodyn.*, 30, (8) 1565-1569 available from: PM:21674594

Fry, C.H., Sahai, A., Vahabi, B., Kanai, A.J., & Birder, L.A. 2014. What is the role for biomarkers for lower urinary tract disorders? ICI-RS 2013. *Neurourol.Urodyn.*, 33, (5) 602-605 available from: PM:24436105

Gorton, E. & Stanton, S. 1999. Women's attitudes to urodynamics: a questionnaire survey. *Br J Obstet Gynaecol.*, 106, (8) 851-856 available from: PM:10453837

Groenendijk, P.M., Nyeholt, A.A., Heesakkers, J.P., van Kerrebroeck, P.E., Hassouna, M.M., Gajewski, J.B., Cappellano, F., Siegel, S.W., Fall, M., Dijkema, H.E., Jonas, U., van den Hombergh, U., & Sacral Nerve Stimulation Study Group 2008. Urodynamic evaluation of sacral neuromodulation for urge urinary incontinence. *BJU International*, 101, (3) 325-329

Gupta, A., Defreitas, G., & Lemack, G.E. 2004. The reproducibility of urodynamic findings in healthy female volunteers: results of repeated studies in the same setting and after short-term follow-up. *Neurourol.Urodyn.*, 23, (4) 311-316 available from: PM:15227647

Hakenberg, O.W., Linne, C., Manseck, A., & Wirth, M.P. 2000. Bladder wall thickness in normal adults and men with mild lower urinary tract symptoms and benign prostatic enlargement. *Neurourology and Urodynamics*, 19, (5) 585-593 available from: http://dx.doi.org/10.1002/1520-6777(2000)19:5<585::AID-NAU5>3.0.CO;2-U

Harrison, S.C., Hunnam, G.R., Farman, P., Ferguson, D.R., & Doyle, P.T. 1987. Bladder instability and denervation in patients with bladder outflow obstruction. *Br J Urol*, 60, (6) 519-522 available from: PM:3427336

Hashim, H. & Abrams, P. 2006. Is the bladder a reliable witness for predicting detrusor overactivity? *J Urol*, 175, (1) 191-194 available from: PM:16406907

Haylen, B.T. 2007. The empty bladder. *Int Urogynecol J Pelvic Floor Dysfunct*, 18, (3) 237-239 available from: PM:16791705

Haylen, B.T., de, R.D., Freeman, R.M., Swift, S.E., Berghmans, B., Lee, J., Monga, A., Petri, E., Rizk, D.E., Sand, P.K., & Schaer, G.N. 2010. An International Urogynecological Association (IUGA)/International Continence Society (ICS) joint report on the terminology for female pelvic floor dysfunction. *Neurourol.Urodyn.*, 29, (1) 4-20 available from: PM:19941278

Heslington, K. & Hilton, P. 1996. Ambulatory monitoring and conventional cystometry in asymptomatic female volunteers. *Br J Obstet.Gynaecol.*, 103, (5) 434-441 available from: PM:8624316

Hjermstad, M.J., Fayers, P.M., Haugen, D.F., Caraceni, A., Hanks, G.W., Loge, J.H., Fainsinger, R., Aass, N., & Kaasa, S. 2011. Studies comparing Numerical Rating Scales, Verbal Rating Scales, and

Visual Analogue Scales for assessment of pain intensity in adults: a systematic literature review. *J Pain Symptom.Manage.*, 41, (6) 1073-1093 available from: PM:21621130

Hogan, S., Gammie, A., & Abrams, P. 2012. Urodynamic features and artefacts. *Neurourol.Urodyn.*, 31, (7) 1104-1117 available from: PM:22473568

Homma, Y., Kondo, Y., Takahashi, S., Kitamura, T., & Kawabe, K. 2000. Reproducibility of cystometry in overactive detrusor. *Eur.Urol*, 38, (6) 681-685 available from: PM:11111184

Ibrahim S & Najdy M. Measuring bladder wall thickness in cases of detrusor instability using transvaginal ultrasound. ICS conference abstract . 2011.
Ref Type: Online Source

Irwig, L., Bossuyt, P., Glasziou, P., Gatsonis, C., & Lijmer, J. 2002. Designing studies to ensure that estimates of test accuracy are transferable. *BMJ*, 324, (7338) 669-671 available from: PM:11895830

Irwin, D.E., Milsom, I., Kopp, Z., Abrams, P., & Cardozo, L. 2006. Impact of overactive bladder symptoms on employment, social interactions and emotional well-being in six European countries. *BJU Int*, 97, (1) 96-100 available from: PM:16336336

Irwin, D.E., Kopp, Z.S., Agatep, B., Milsom, I., & Abrams, P. 2011. Worldwide prevalence estimates of lower urinary tract symptoms, overactive bladder, urinary incontinence and bladder outlet obstruction. *BJU International*, 108, (7) 1132-1138 available from: http://dx.doi.org/10.1111/j.1464-410X.2010.09993.x

Jackson, S. 1997. The patient with an overactive bladder--symptoms and quality-of-life issues. *Urology*, 50, (6A Suppl) 18-22 available from: PM:9426747

James, M., Jackson, S., Shepherd, A., & Abrams, P. 1999. Pure stress leakage symptomatology: is it safe to discount detrusor instability? *Br J Obstet.Gynaecol.*, 106, (12) 1255-1258 available from: PM:10609718

Jensen, M.P., Chen, C., & Brugger, A.M. 2003. Interpretation of visual analog scale ratings and change scores: a reanalysis of two clinical trials of postoperative pain. *J Pain*, 4, (7) 407-414 available from: PM:14622683

Jequier, S. & Rousseau, O. 1987. Sonographic measurements of the normal bladder wall in children. *AJR Am.J Roentgenol.*, 149, (3) 563-566 available from: PM:3303881

Jha, S., Strelley, K., & Radley, S. 2012. Incontinence during intercourse: myths unravelled. *Int Urogynecol J*, 23, (5) 633-637 available from: PM:22237785

Juszczak, K., Maciukiewicz, P., Drewa, T., & Thor, P.J. 2014. Cajal-like interstitial cells as a novel target in detrusor overactivity treatment: true or myth? *Cent.European J Urol*, 66, (4) 413-417 available from: PM:24757530

Khan, K.S., Dinnes, J., & Kleijnen, J. 2001. Systematic reviews to evaluate diagnostic tests. *Eur.J Obstet.Gynecol.Reprod.Biol.*, 95, (1) 6-11 available from: PM:11267714

Khullar, V., Cardozo, L.D., Salvatore, S., & Hill, S. 1996. Ultrasound: a noninvasive screening test for detrusor instability. *Br J Obstet.Gynaecol.*, 103, (9) 904-908 available from: PM:8813311

Khullar, V., Salvatore, S., Cardozo, L., Bourne, T.H., Abbott, D., & Kelleher, C. 1994. A novel technique for measuring bladder wall thickness in women using transvaginal ultrasound. *Ultrasound Obstet.Gynecol.*, 4, (3) 220-223 available from: PM:12797185

Kinder, R.B. & Mundy, A.R. 1987. Pathophysiology of idiopathic detrusor instability and detrusor hyper-reflexia. An in vitro study of human detrusor muscle. *Br J Urol*, 60, (6) 509-515 available from: PM:3427334

Ku, J.H., Kim, S.W., Kim, H.H., Paick, J.S., Son, H., & Oh, S.J. 2004. Patient experience with a urodynamic study: a prospective study in 208 patients. *J Urol*, 171, (6 Pt 1) 2307-2310 available from: PM:15126810

Kubota, Y.,Kojima, Y., Shibata, Y., Imura, M., Sasaki, S., Kohri, K. Role of KIT-Positive Interstitial Cells of Cajal in the Urinary Bladder and Possible Therapeutic Target for Overactive Bladder
1. Adv.Urol 2011, 816342. 2011

Kuhn, A., Genoud, S., Robinson, D., Herrmann, G., Gunthert, A., Brandner, S., & Raio, L. 2011. Sonographic transvaginal bladder wall thickness: does the measurement discriminate between urodynamic diagnoses? *Neurourol.Urodyn.*, 30, (3) 325-328 available from: PM:20949598

Kuo, H.C. 2009. Measurement of detrusor wall thickness in women with overactive bladder by transvaginal and transabdominal sonography. *Int Urogynecol J Pelvic Floor Dysfunct*, 20, (11) 1293-1299 available from: PM:19557300

Latthe, P.M., Champaneria, R., & Khan, K.S. 2010. Systematic review of the accuracy of ultrasound as the method of measuring bladder wall thickness in the diagnosis of detrusor overactivity. *Int Urogynecol J*, 21, (8) 1019-1024 available from: PM:20424825

Lekskulchai O & Dietz, H. P. Detrusor wall thickness in women with a history of childhood nocturnal enuresis. ICS conference abstract [330]. 2006.
Ref Type: Online Source

Lekskulchai, O. & Dietz, H.P. 2008. Detrusor wall thickness as a test for detrusor overactivity in women. *Ultrasound Obstet.Gynecol.*, 32, (4) 535-539 available from: PM:18634113

Lekskulchai, O. & Dietz, H.P. 2009. Is detrusor hypertrophy in women associated with voiding dysfunction? *Aust.N.Z.J Obstet.Gynaecol.*, 49, (6) 653-656 available from: PM:20070717

Lijmer, J.G., Mol, B.W., Heisterkamp, S., Bonsel, G.J., Prins, M.H., van der Meulen, J.H., & Bossuyt, P.M. 1999. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA*, 282, (11) 1061-1066 available from: PM:10493205

Mallett, S., Halligan, S., Thompson, M., Collins, G.S., & Altman, D.G. 2012. Interpreting diagnostic accuracy studies for patient care. *BMJ*, 345, e3999 available from: PM:22750423

Malone-Lee, J.G. & Al-Buheissi, S. 2009. Does urodynamic verification of overactive bladder determine treatment success? Results from a randomized placebo-controlled study. *BJU Int*, 103, (7) 931-937 available from: PM:19281469

Martan, A., Masata, J., Halaska, M., & Voigt, R. 2001. Ultrasound imaging of the lower urinary system in women after Burch colposuspension. *Ultrasound Obstet Gynecol*, 17, (1) 58-64 available from: PM:11244658

Marteau, T.M. & Bekker, H. 1992. The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). *Br J Clin.Psychol.*, 31 (Pt 3), 301-306 available from: PM:1393159

Martin, J.L., Williams, K.S., Sutton, A.J., Abrams, K.R., & Assassa, R.P. 2006. Systematic review and meta-analysis of methods of diagnostic assessment for urinary incontinence. *Neurourol.Urodyn.*, 25, (7) 674-683 available from: PM:17016795

McNeil, D.W. & Berryman, M.L. 1989. Components of dental fear in adults? *Behav.Res.Ther.*, 27, (3) 233-236 available from: PM:2730504

Minardi, D., Piloni, V., Amadi, A., El, A.Z., Milanese, G., & Muzzonigro, G. 2007. Correlation between urodynamics and perineal ultrasound in female patients with urinary incontinence. *Neurourol.Urodyn.*, 26, (2) 176-182 available from: PM:17016799

Mobley, D.F. & Baum, N. 2014. Etiology, evaluation, and management of nocturia in elderly men and women. *Postgrad.Med.*, 126, (2) 147-153 available from: PM:24685979

Mortensen, S., Lose, G., & Thyssen, H. 2002. Repeatability of cystometry and pressure-flow parameters in female patients. *Int Urogynecol J Pelvic Floor Dysfunct*, 13, (2) 72-75 available from: PM:12054185

Neustaedter, E.G., Milne, J., Shorten, K., Weckman, B., Tse, A., & Tange, S. 2011. How well informed are women who undergo urodynamic testing? *Neurourol.Urodyn.*, 30, (4) 572-577 available from: PM:21268100

NHS Friends and Family test. The NHS Friends and Family Test. 2012.
Ref Type: Online Source

Oelke, M. 2010. International Consultation on Incontinence-Research Society (ICI-RS) report on non-invasive urodynamics: the need of standardization of ultrasound bladder and detrusor wall thickness measurements to quantify bladder wall hypertrophy. *Neurourol.Urodyn.*, 29, (4) 634-639 available from: PM:20432327

Otsuki E N, Júnior E A, Oliveira E, Castelo Girão M J B, & Jármy-Di Bella Z I K. Ultrasound Thickness of Bladder Wall in Continent
and Incontinent Women and Its Correlation with Cystometry. Scientific World Journal . 2014. Hindawi Publishing Corporation.
Ref Type: Generic

Ozturk H, Aydur E, Irkilata H, Seckin B, & Dayanc M. The diagnostic role of transabdominal ultrasonographic bladder wall thickness in female urinary incontinence. ICS conference abstract . 2011.
Ref Type: Online Source

Panayi, D.C., Khullar, V., Fernando, R., & Tekkis, P. 2010a. Transvaginal ultrasound measurement of bladder wall thickness: a more reliable approach than transperineal and transabdominal approaches. *BJU Int*, 106, (10) 1519-1522 available from: PM:20438565

Panayi, D.C., Tekkis, P., Fernando, R., Hendricken, C., & Khullar, V. 2010b. Ultrasound measurement of bladder wall thickness is associated with the overactive bladder syndrome. *Neurourol.Urodyn.*, 29, (7) 1295-1298 available from: PM:20127835

Pannek, J., Bartel, P., Gocking, K., & Frotzler, A. 2012. Clinical usefulness of ultrasound assessment of detrusor wall thickness in patients with neurogenic lower urinary tract dysfunction due to spinal cord injury: urodynamics made easy? *World J Urol* 1-6 available from: http://dx.doi.org/10.1007/s00345-012-0970-6

Parsons, M., Amundsen, C.L., Cardozo, L., Vella, M., Webster, G.D., & Coats, A.C. 2007. Bladder diary patterns in detrusor overactivity and urodynamic stress incontinence. *Neurourol.Urodyn.*, 26, (6) 800-806 available from: PM:17335054

Pepe, M. S. The statistical evaluation of medical tests for classification and prediction. 2004. Oxford University Press.
Ref Type: Online Source

Price, D.D., McGrath, P.A., Rafii, A., & Buckingham, B. 1983. The validation of visual analogue scales as ratio scale measures for chronic and experimental pain. *Pain*, 17, (1) 45-56 available from: PM:6226917

Rachaneni, S., Balogun, M., & Latthe, P. 2013a. Bladder-wall-thickness ultrasound scan in the investigation of LUTS in women: challenges and limitations. *Int Urogynecol J*, 24, (5) 725-728 available from: PM:23376906

Rachaneni, S., Arya, P., & Latthe, P. 2013b. Urinary nerve growth factor: a biomarker of detrusor overactivity? A systematic review. *Int.Urogynecol.J.*, 24, (10) 1603-1609 available from: PM:23649686

Radley, S.C., Rosario, D.J., Chapple, C.R., & Farkas, AG. 2001. Conventional and ambulatory urodynamic findings in women with symptoms suggestive of bladder overactivity. *The Journal of Urology*, 166, (6) 2253-2258 available from: http://www.sciencedirect.com/science/article/pii/S0022534705655450

Robinson, D., Anders, K., Cardozo, L., Bidmead, J., Toozs-Hobson, P., & Khullar, V. 2002. Can ultrasound replace ambulatory urodynamics when investigating women with irritative urinary symptoms? *BJOG.*, 109, (2) 145-148 available from: PM:11888096

Rosati, P. & Guariglia, L. 2000. Acceptability of early transvaginal or abdominal sonography in the first half of pregnancy. *Arch.Gynecol Obstet*, 264, (2) 80-83 available from: PM:11045328

Rovner, E., Kennelly, M., Schulte-Baukloh, H., Zhou, J., Haag-Molkenteller, C., & Dasgupta, P. 2011. Urodynamic results and clinical outcomes with intradetrusor injections of onabotulinumtoxinA in a randomized, placebo-controlled dose-finding study in idiopathic overactive bladder. *Neurourol.Urodyn.*, 30, (4) 556-562 available from: PM:21351127

Schafer JL. Analysis of incomplete multivariate data. 1997. Chapman & Hall/CRC press.
Ref Type: Online Source

Seckiner, I., Yesilli, C., Mungan, N.A., Aykanat, A., & Akduman, B. 2007. Correlations between the ICIQ-SF score and urodynamic findings. *Neurourol.Urodyn.*, 26, (4) 492-494 available from: PM:17304520

Serati, M., Cattoni, E., Braga, A., Siesto, G., & Salvatore, S. 2011. Coital incontinence: relation to detrusor overactivity and stress incontinence. A controversial topic. *Neurourol.Urodyn.*, 30, (8) 1415 available from: PM:21661041

Serati, M., Salvatore, S., Cattoni, E., Soligo, M., Cromi, A., & Ghezzi, F. 2010. Ultrasound measurement of bladder wall thickness in different forms of detrusor overactivity. *Int.Urogynecol.J.*, 21, (11) 1405-1411 available from: PM:20535449

Serati, M., Cattoni, E., Siesto, G., Braga, A., Sorice, P., Cantaluppi, S., Cromi, A., Ghezzi, F., Vitobello, D., Bolis, P., & Salvatore, S. 2013. Urodynamic evaluation: can it prevent the need for surgical intervention in women with apparent pure stress urinary incontinence? *BJU Int* available from: PM:23421421

Silva, J.A., Gonsalves, M.D., de Melo, R.T., Carrerette, F.B., & Damiao, R. 2014. Association between the bladder wall thickness and urodynamic findings in patients with spinal cord injury. *World J.Urol.* available from: PM:24573904

Smith, A., Bevan, D., Douglas, H.R., & James, D. 2013. Management of urinary incontinence in women: summary of updated NICE guidance. *BMJ*, 347, f5170 available from: PM:24021756

Soligo M, Salvatore, S., Luppino G., Arcari V, Milani, R., & Khullar, V. Overactive bladder definition and ultrasound measurement of bladder wall thickness:The right way without urodynamics. ICS conference abstract 1. 2002.
Ref Type: Online Source

Sorensen, S.S., Nielsen, J.B., Norgaard, J.P., Knudsen, L.M., & Djurhuus, J.C. 1984. Changes in bladder volumes with repetition of water cystometry. *Urol Res.*, 12, (4) 205-208 available from: PM:6495446

South, M.M., Romero, A.A., Jamison, M.G., Webster, G.D., & Amundsen, C.L. 2007. Detrusor overactivity does not predict outcome of sacral neuromodulation test stimulation. *Int Urogynecol J Pelvic Floor Dysfunct*, 18, (12) 1395-1398 available from: PM:17364132

Spielberger, C. D. C. Manual for the State-Trait Anxiety Inventory(Form Y). 1983. Palo Alto, CA: Mind Garden.
Ref Type: Serial (Book,Monograph)

Steers, W. 2000. Pathogenesis of the overactive bladder and its attendant risk factors. *BJU.Int.*, 85 Suppl 3, 69-1 available from: PM:11954201

Swithinbank, L.V., James, M., Shepherd, A., & Abrams, P. 1999. Role of ambulatory urodynamic monitoring in clinical urological practice. *Neurourology and Urodynamics*, 18, (3) 215-222 available from: http://dx.doi.org/10.1002/(SICI)1520-6777(1999)18:3<215::AID-NAU8>3.0.CO;2-4

Terry, R., Brodie, E.E., & Niven, C.A. 2007. Exploring the phenomenology of memory for pain: is previously experienced acute pain consciously remembered or simply known? *J Pain*, 8, (6) 467-475 available from: PM:17306627

Tubaro A, Khullar V, Oelke, M., Wijkstra H, Tretter R, Stow B, Huang M, Compion G, & Robinson D. Intra- and inter-reader variability of transvaginal ultrasound bladder wall thickness measurements:Results from the SHRINK study . ICS conference abstract . 2013.
Ref Type: Online Source

Tubaro A , Khullar V, Oelke, M., Wijkstra H, Tretter R, Stow B, Huang M, Compion G, & Robinson D. Intra- and inter-reader variability of transvaginal ultrasound bladder wall thickness measurements:Results from the SHRINK study . ICS conference abstract . 2013.
Ref Type: Online Source

Ulmsten, U. 1997. Some reflections and hypotheses on the pathophysiology of female urinary incontinence. *Acta Obstet Gynecol Scand.Suppl*, 166, 3-8 available from: PM:9253370

Ulmsten, U. & Falconer, C. 1999. Connective tissue in female urinary incontinence. *Curr.Opin.Obstet Gynecol*, 11, (5) 509-515 available from: PM:10526931

Vandoninck, V., van Balken, M.R., Finazzi, A.E., Petta, F., Micali, F., Heesakkers, J.P., Debruyne, F.M., Kiemeney, L.A., & Bemelmans, B.L. 2003. Percutaneous tibial nerve stimulation in the treatment of overactive bladder: urodynamic data. *Neurourol.Urodyn.*, 22, (3) 227-232 available from: PM:12707873

Van Waalwijk van Doorn, E.S., Remmers, A, Janknegt, R.A. 1992. Conventional and extramural ambulatory urodynamic testing of the lower urinary tract in female volunteers J. Urol., 147, 1319

Whiting, P., Rutjes, A., Reitsma, J., Bossuyt, P., & Kleijnen, J. 2003. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Medical Research Methodology*, 3, (1) 25 available from: http://www.biomedcentral.com/1471-2288/3/25

Xiao-Hua Zhou NAO & Donna K.McClish. Statistical Methods in Diagnostic Medicine. 2011. Ref Type: Generic

Yang, J.M. & Huang, W.C. 2002. Discrimination of bladder disorders in female lower urinary tract symptoms on ultrasonographic cystourethrography. *J Ultrasound Med.*, 21, (11) 1249-1255 available from: PM:12418766

Yang, J.M. & Huang, W.C. 2003. Bladder wall thickness on ultrasonographic cystourethrography: affecting factors and their implications. *J Ultrasound Med.*, 22, (8) 777-782 available from: PM:12901404

Yeung, J.Y., Eschenbacher, M.A., & Pauls, R.N. 2014. Pain and embarrassment associated with urodynamic testing in women. *Int Urogynecol J*, 25, (5) 645-650 available from: PM:24280994

Yiou, R., Audureau, E., Loche, C.M., Dussaud, M., Lingombet, O., & Binhas, M. 2013. Comprehensive evaluation of embarrassment and pain associated with invasive urodynamics. *Neurourol.Urodyn.* available from: PM:24259301

Appendix

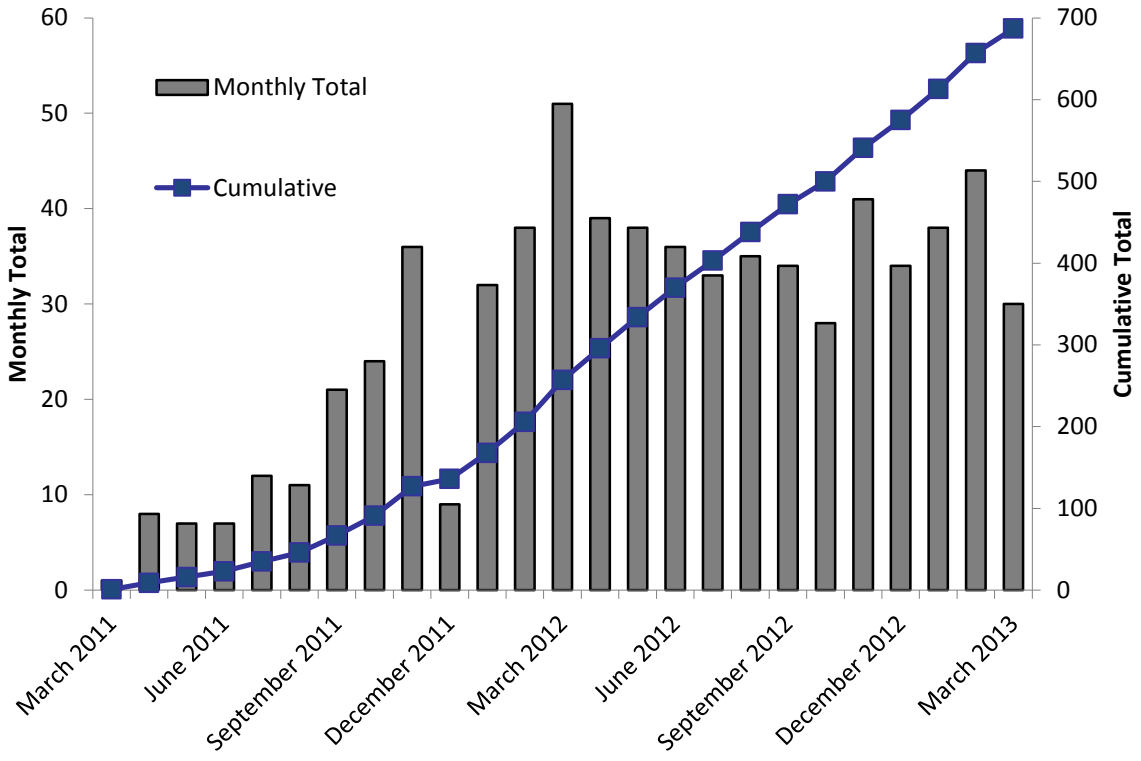
Appendix 1: STARD Checklist for reporting of studies of diagnostic accuracy

Section and Topic	Item #		On page #
TITLE/ABSTRACT/ KEYWORDS	1	Identify the article as a study of diagnostic accuracy (recommend MeSH heading 'sensitivity and specificity').	Title page
INTRODUCTION	2	State the research questions or study aims, such as estimating diagnostic accuracy or comparing accuracy between tests or across participant groups.	29
METHODS			
Participants	3	The study population: The inclusion and exclusion criteria, setting and locations where data were collected.	48-49
	4	Participant recruitment: Was recruitment based on presenting symptoms, results from previous tests, or the fact that the participants had received the index tests or the reference standard?	49-50
	5	Participant sampling: Was the study population a consecutive series of participants defined by the selection criteria in item 3 and 4? If not, specify how participants were further selected.	49
	6	Data collection: Was data collection planned before the index test and reference standard were performed (prospective study) or after (retrospective study)?	Prospective study Page 48
Test methods	7	The reference standard and its rationale.	54
	8	Technical specifications of material and methods involved including how and when measurements were taken, and/or cite references for index tests and reference standard.	Appendix 6 a(UDS) Appendix 6 b (BUS)
	9	Definition of and rationale for the units, cut-offs and/or categories of the results of the index tests and the reference standard.	51-54 Appendix 6
	10	The number, training and expertise of the persons executing and reading the index tests and the reference standard.	50,72 Table 6 for reproducibility study

	11	Whether or not the readers of the index tests and reference standard were blind (masked) to the results of the other test and describe any other clinical information available to the readers.	51
Statistical methods	12	Methods for calculating or comparing measures of diagnostic accuracy, and the statistical methods used to quantify uncertainty (e.g. 95% confidence intervals).	88-89
	13	Methods for calculating test reproducibility, if done.	66-67
RESULTS			
Participants	14	When study was performed, including beginning and end dates of recruitment.	58
	15	Clinical and demographic characteristics of the study population (at least information on age, gender, spectrum of presenting symptoms).	60, Table 4
	16	The number of participants satisfying the criteria for inclusion who did or did not undergo the index tests and/or the reference standard; describe why participants failed to undergo either test (a flow diagram is strongly recommended).	61,Figure 8
Test results	17	Time-interval between the index tests and the reference standard, and any treatment administered in between.	4 wks Page 97-100
	18	Distribution of severity of disease (define criteria) in those with the target condition; other diagnoses in participants without the target condition.	Tables 12 and 13 (UDS), Table 6, Figure 15 (BWT)
	19	A cross tabulation of the results of the index tests (including indeterminate and missing results) by the results of the reference standard; for continuous results, the distribution of the test results by the results of the reference standard.	Table 14
	20	Any adverse events from performing the index tests or the reference standard.	UTI in a few patients following Urodynamics
Estimates	21	Estimates of diagnostic accuracy and measures of statistical uncertainty (e.g. 95% confidence intervals).	104, Tables 15.16
	22	How indeterminate results, missing data and outliers of the index tests were handled.	Appendices 3 and 4

	23	Estimates of variability of diagnostic accuracy between subgroups of participants, readers or centers, if done.	Appendix 5
	24	Estimates of test reproducibility, if done.	Table 7
DISCUSSION	25	Discuss the clinical applicability of the study findings.	111

Appendix 2: Study accrual



Appendix 3: Sensitivity analyses

Figure 3.1: ROC curve excluding those results where the urodynamics test was not blind to the results of the ultrasound test

(16/632 women (3%);

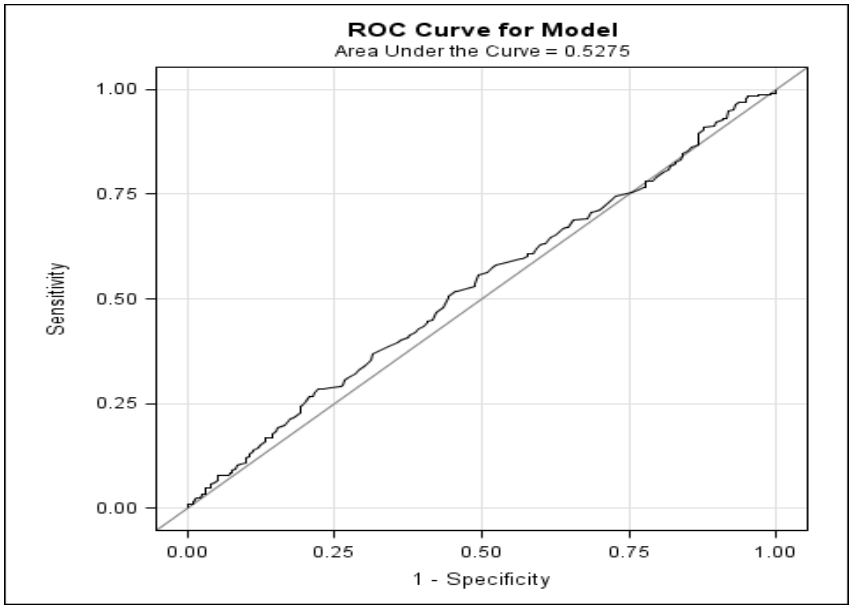


Figure 3.2: ROC curve excluding those results where there was more than four weeks between the tests

(26/660 women (4%);

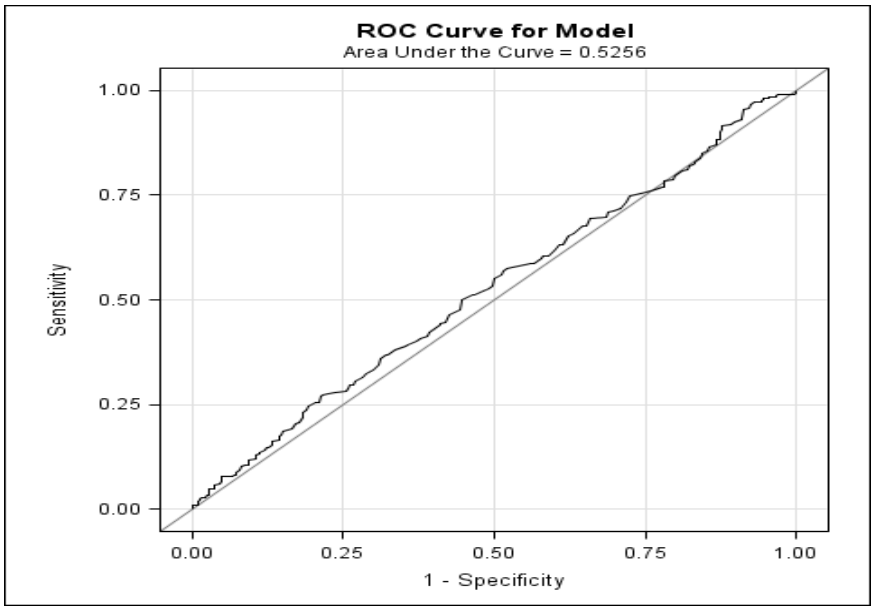


Figure 3.3: ROC curve incorporating incomplete ultrasound measurements

(10 observations – average of remaining one or two measurements used);

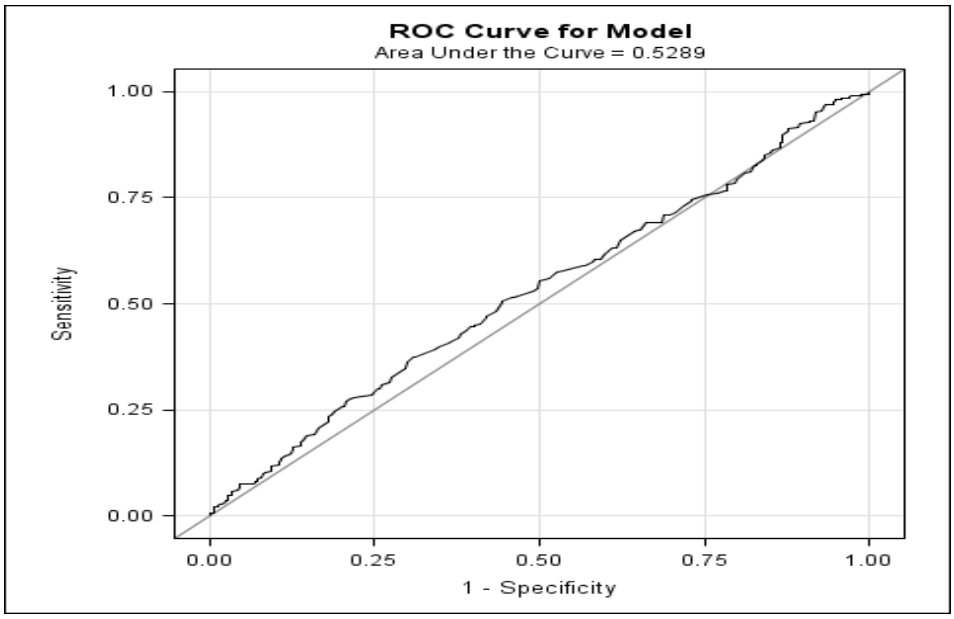
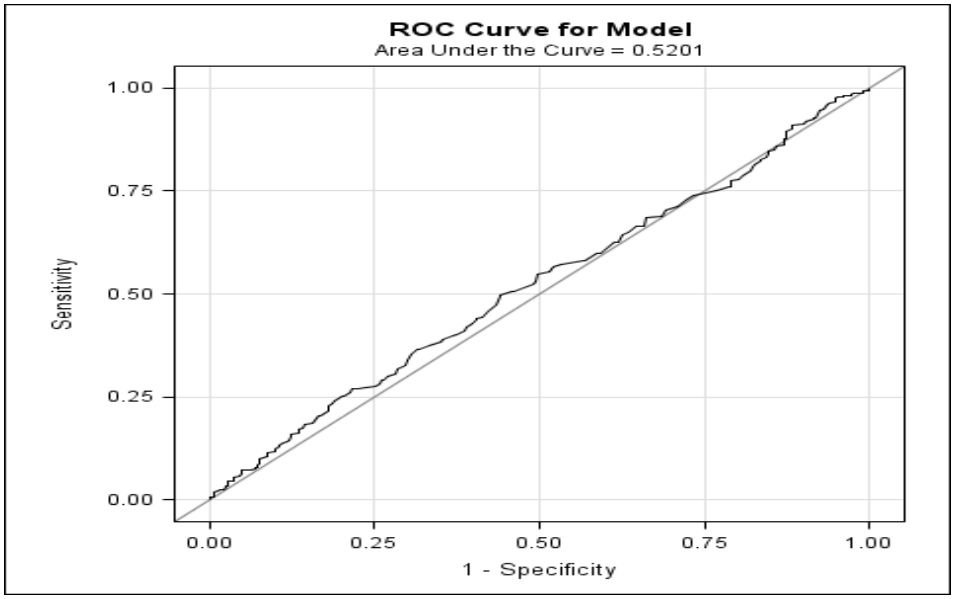


Figure 3.4: ROC curve using ambulatory urodynamic diagnosis where available

(14 participants);



Appendix 4: Exploratory analyses

Figure 4.1: ROC curve including the urgency alone group (as per clinical history; excluding mixed stress/urge incontinence group:

(217 patients)

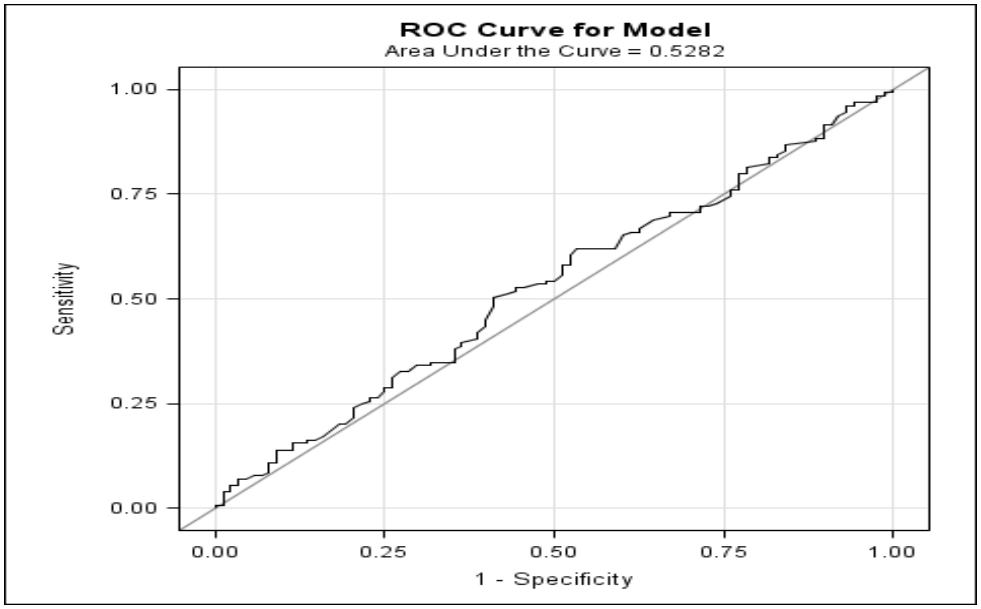


Figure 4.2: ROC curve including the 'pure' DO group only (diagnosis of DO/low compliance/DO plus low compliance, excluding 'mixed' DO - DO with another diagnosis of USI or voiding dysfunction;

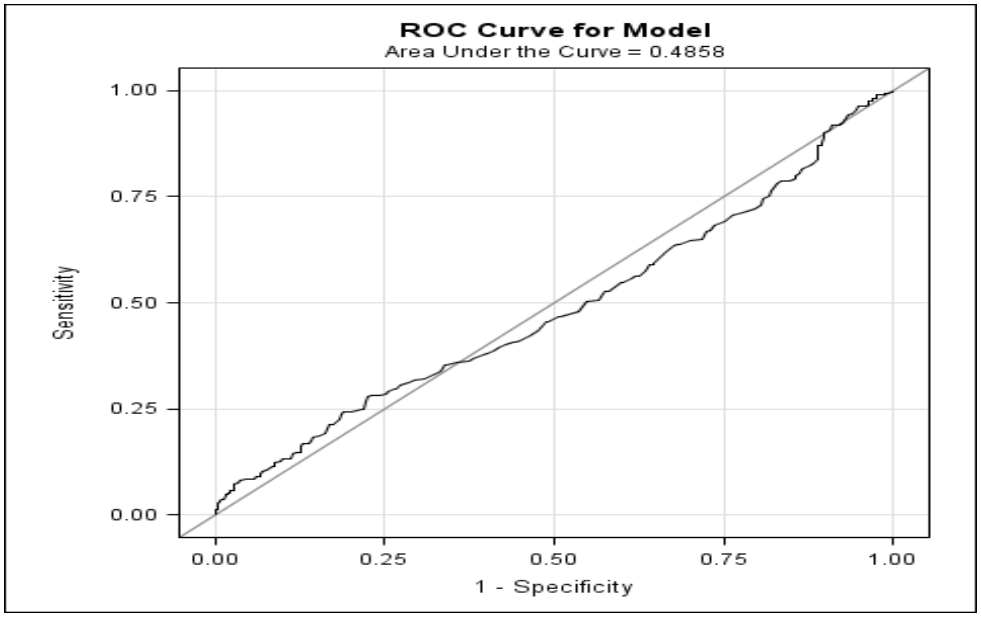


Figure 4.3: ROC curve including the 'wet' DO group only (excluding 'dry' DO);

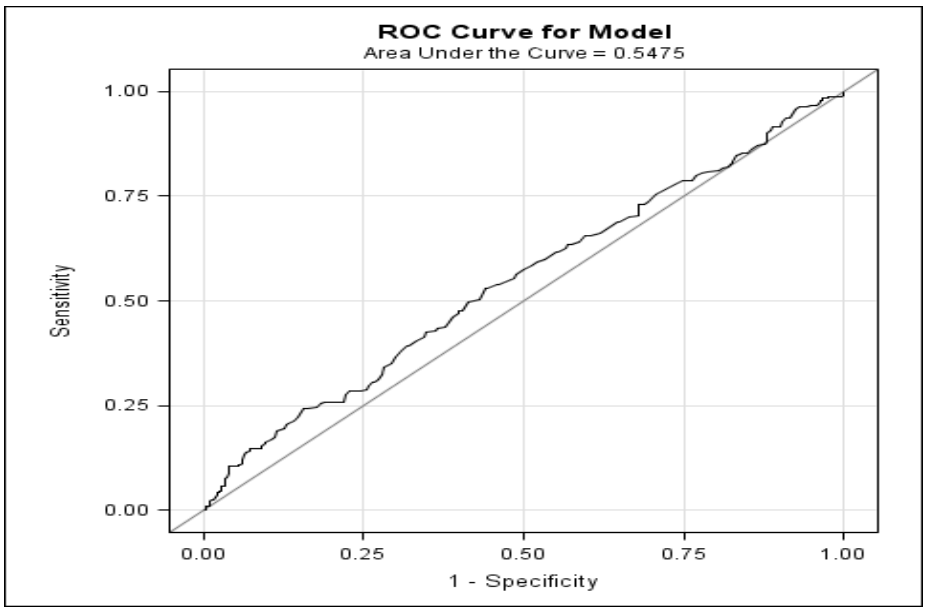


Figure 4.4: Box and whisker plot comparing BWT for 'wet' and 'dry' DO

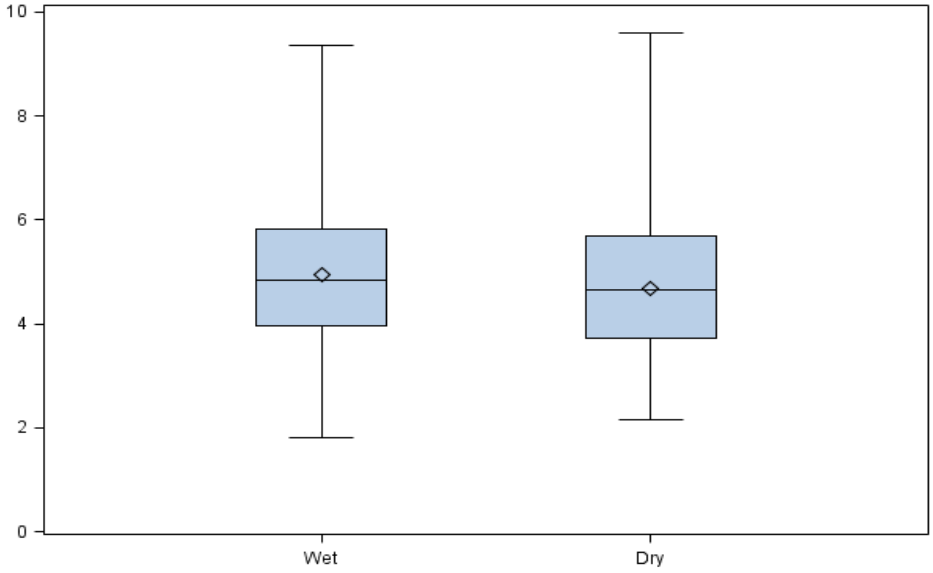


Figure 4.5: ROC curve from exploratory analysis using the trigone measurement alone for BWT

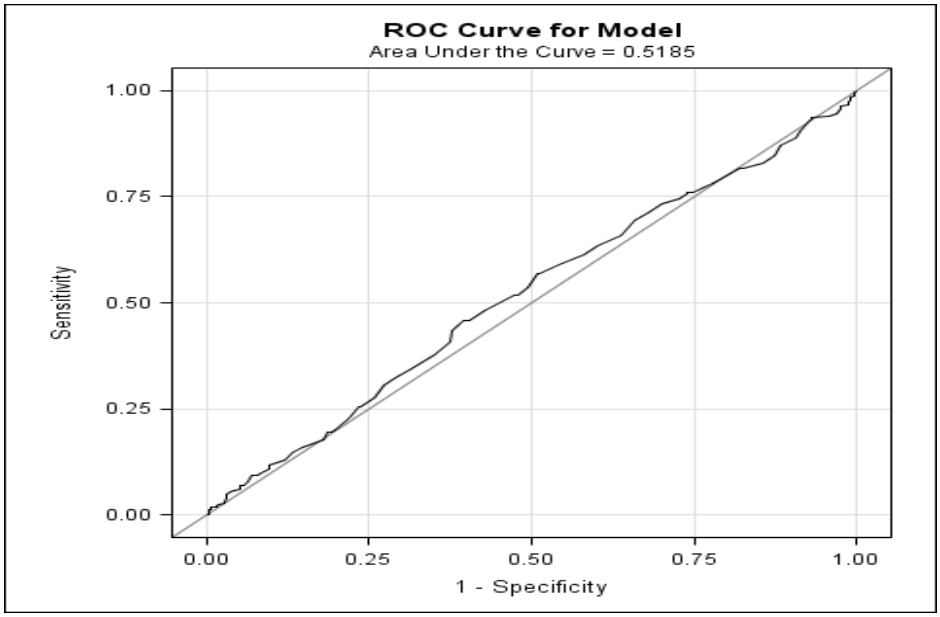


Figure 4.6: ROC curve excluding those who had a detrusor pressure rise upon provocation testing 'provoked DO'

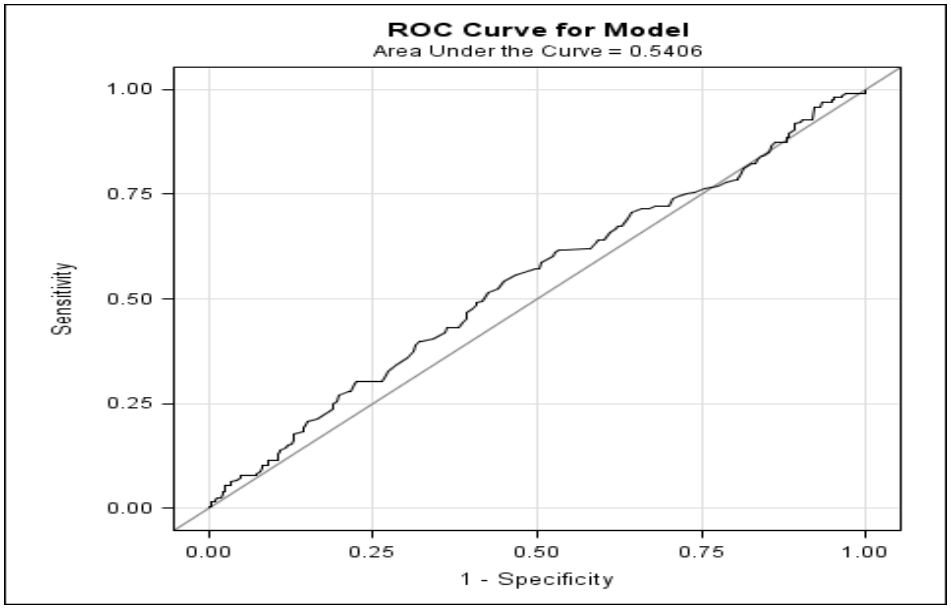


Figure 4.7: ROC curve excluding those who had PVR>30ml upon testing

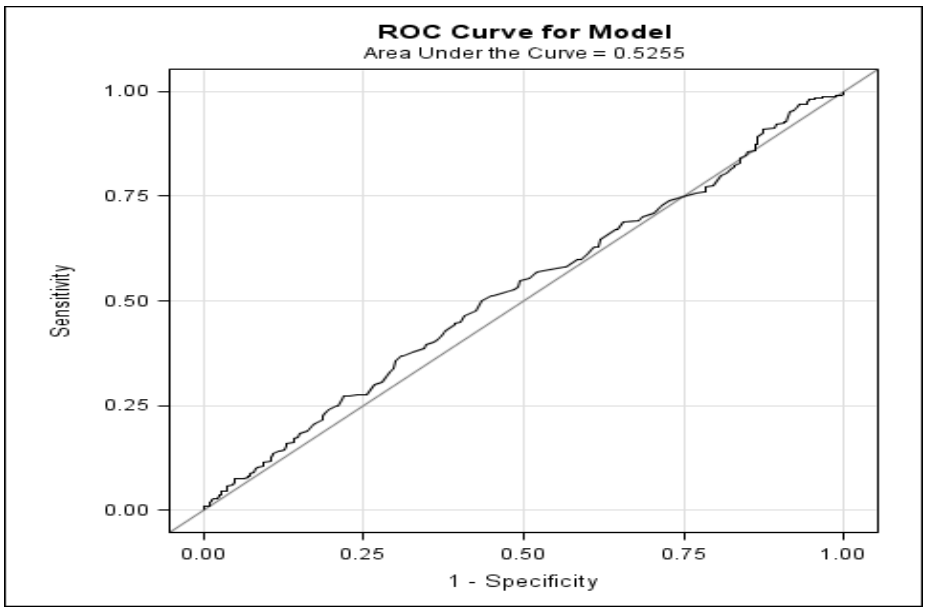
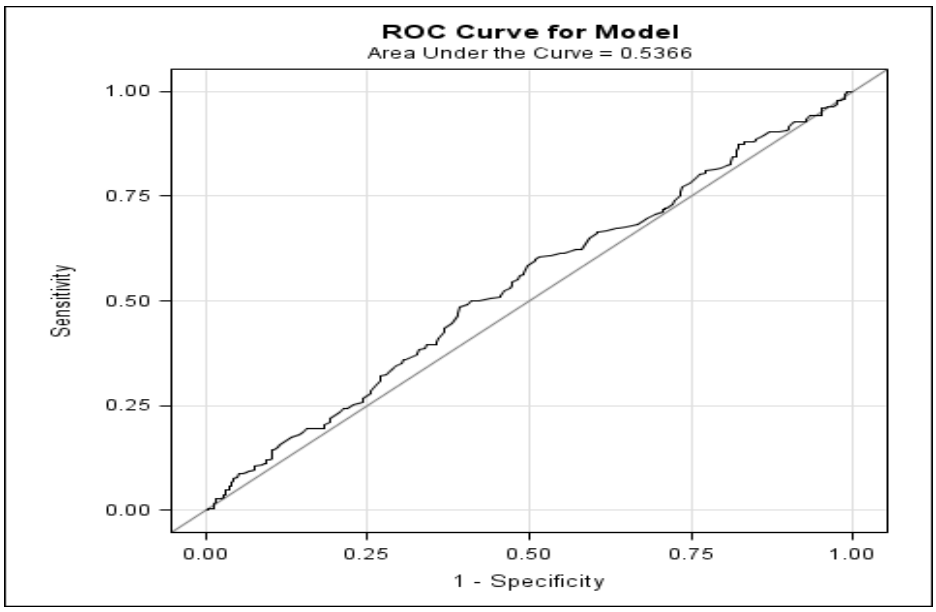


Figure 4.8: ROC curve using the average dome, 1cm left of dome, 1cm right of dome



Appendix 5

Table 5.1: Results of ROC curve analysis in pre-specified subgroupings

Variable		AUC	95% CI	p-value for difference between AUCs
Previous treatment with antimuscarinics	=No	0.536	(0.481, 0.592)	0.48
	=Yes	0.501	(0.420, 0.582)	
Clinical history suggested mixed incontinence	=No	0.534	(0.460, 0.608)	0.73
	=Yes	0.518	(0.460, 0.575)	
Presence of UTI in the last 12 months	=No	0.530	(0.482, 0.578)	0.53
	=Yes	0.586	(0.417, 0.755)	
Patients with voiding difficulties	=No	0.533	(0.472, 0.594)	0.84
	=Yes	0.524	(0.454, 0.593)	
Previous incontinence surgery	=No	0.526	(0.479, 0.573)	0.76
	=Yes	0.493	(0.294, 0.693)	
BMI	<25	0.519	(0.424, 0.614)	0.95
	>=25	0.523	(0.471, 0.575)	

Table 5.2: Results of univariate analysis exploring factors possibly associated with DO diagnosis

Variable	Data type	p-value	OR (95%CI) if statistically important	Frequencies (binary/categorical data)
ICIQ score (best=0, worst=16)	Continuous	<0.0001	1.23 (1.15, 1.31)	
BWT, mm	Continuous	0.19		
Age, years	Continuous	0.66		
Duration of symptoms, years	Continuous	0.45		
BMI, kg/m^2	Continuous	0.38		
Ethnicity (white/black/Asian/other)	Categorical	0.59		
Vaginal birth=yes	Binary	0.64		
Clinical history suggests mixed incontinence=yes	Binary	0.40		
If clinical history suggests mixed incontinence, which came first (stress/urge/unsure/na)	Categorical	0.66		
Previous treatment with antimuscarinics=yes	Binary	0.001	1.74 (1.24, 2.44)	68% (152/222) DO when=yes 56% (245/441) DO when=no
Previous UTI in last 12 months=yes	Binary	0.08	0.60 (0.34, 1.07)	48% (24/50) DO when=yes 61% (363/599) DO when=no

History of voiding difficulties=yes	Binary	0.16		
Post-menopausal=yes	Binary	0.67		
Parity (0/1/2/3/4+)	Categorical	0.27		
Previous incontinence surgery=yes	Binary	0.59		
Previous POP surgery=yes	Binary	0.32		

Table 5.3: Results of multivariable analysis exploring factors possibly associated with DO diagnosis

Model	Significant variables	p-value	OR (95%CI) if significant
Backward selection (p=0.1 to stay in model)	ICIQ score Previous UTI in last 12 months	<0.0001 0.04	1.21 (1.13, 1.29) 0.51 (0.27, 0.97)
All variables included	ICIQ score Previous UTI in last 12 months	<0.0001 0.06	1.21 (1.13, 1.29) 0.53 (0.27, 1.03)
All variables included, multiple imputation used for missing data	ICIQ score Previous treatment with antimuscarinics Previous UTI in last 12 months	<0.0001 0.02 0.07	1.23 (1.15, 1.31) 1.57 (1.09, 2.28) 0.57 (0.31, 1.06)

Appendix 6

6a) BUS Standard Operating Procedure for Scanning

Clinician Preparation

The clinician performing the BUS should be different to the clinician performing the UDS, to ensure blinding between the two tests.

If, for any reason, this is not possible the BUS should be performed BEFORE the UDS.

Patient Preparation

The patient may be seen in various settings:

Clinic (patient may need other assessment)

Scan department (may require renal tract assessment and attends with full bladder)

Ensure patient empties bladder before assessment of Bladder Wall Thickness (BWT) and Post Void Residual (PVR)

Important to stress need to void as completely as possible

Machine & Probe Preparation

Ensure scanner is capable of measurement in millimetres (mm)

Set to the scanner to the Gynae preset

Use a Trans-Vaginal probe (not a rectal probe)

Multifrequency – use between 7 & 9 MHz for optimal image (no lower than 5 MHz)

Prepare the probe:

Clean

Put gel into probe cover, excluding air

Put gel onto tip of probe

Timescales

The ultrasound scan should ideally be completed in a one-stop clinic with the urodynamics test.

The two tests should be undertaken by different clinicians to ensure blinding of results.

If it is not possible to hold one-stop clinics, the ultrasound and urodynamics tests should be undertaken no more than 4 weeks apart.

The second test can be undertaken up to a further 4 weeks after this cut-off (8 weeks in total from the first test), but the data collected will be classed as a protocol violation.

If the patient is happy to have one or both tests re-taken (effectively constituting a second 'set' of tests), they can do so, as long as they haven't become ineligible in the interim (i.e. begun medication).

An interval of more than 4 weeks between the two tests will be designated as breach of protocol and no per patient payment will be provided for these patients.

Ultrasound Assessment

The patient should be in the supine position (stirrups or pad under pelvis as appropriate)

The TV probe should be inserted into introitus in longitudinal plane

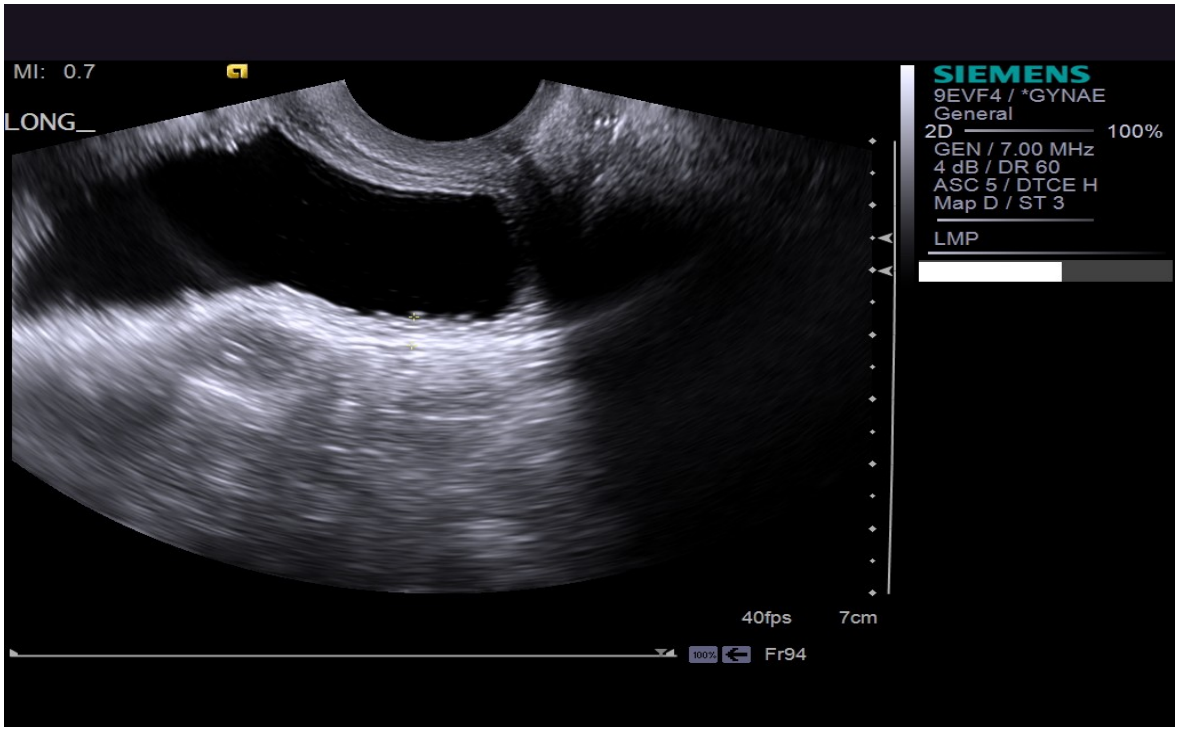
Position should then be assessed on screen

Identify the urethra in sagittal orientation

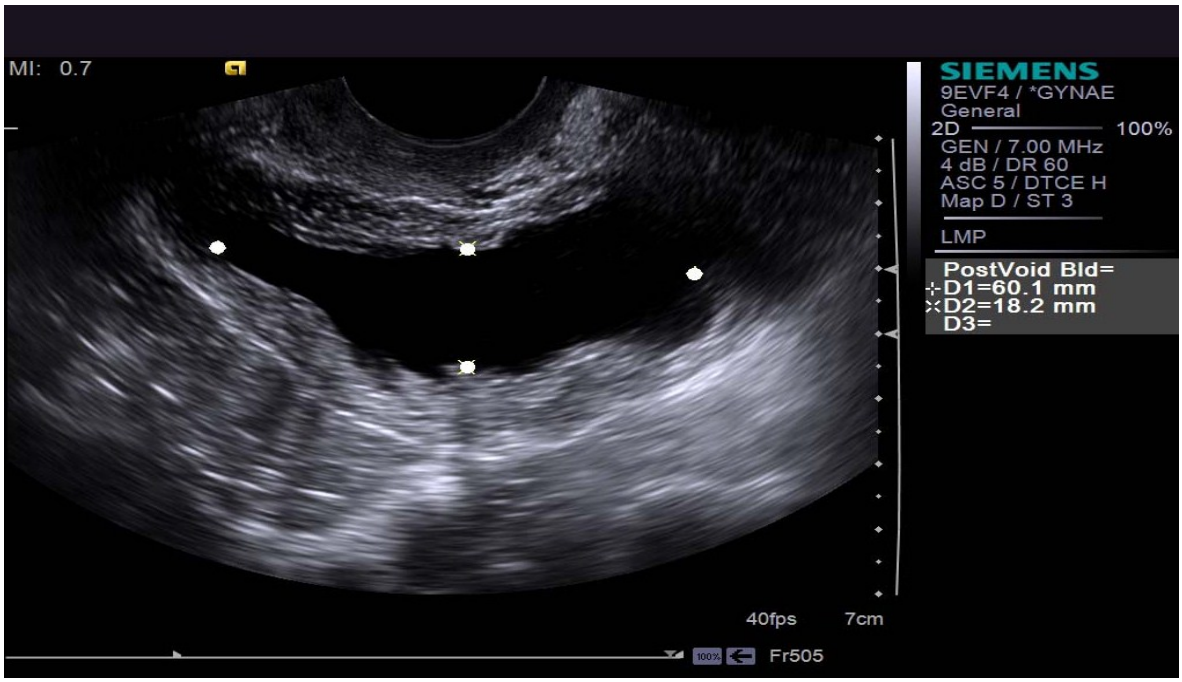
Position probe such that the VUJ is close to top of screen

Ensure image sizing is appropriate for screen, between 5-7cm depth

Focal zone positioned at region of interest; multiple focal zones if possible to give good definition at various levels.



PVR and BWT should be measured before any other assessment:



PVR:

Identify entire bladder in sagittal plane, measure longest AP dimension and then CC dimension perpendicular to this

Rotate probe through 90° and measure axial dimension

Most machines are now able to generate volumes automatically (need to select this before starting with callipers)

If not available, use the following to standardize calculations: cranio-caudal (H) x anterior-posterior (D) x transverse diameter (W) x 0.5233 (prolate ellipsoid) = PVR volu*me*

ELIGIBILITY: Proceeding to BWT measurement:

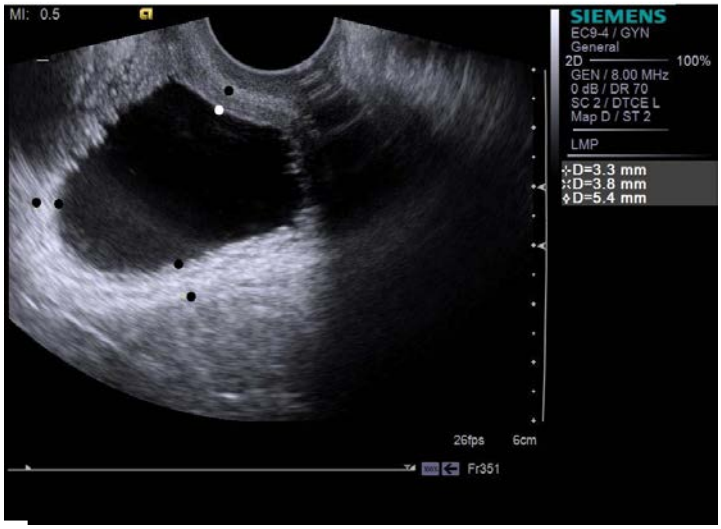
If PVR <=30mls proceed with BWT measurement.

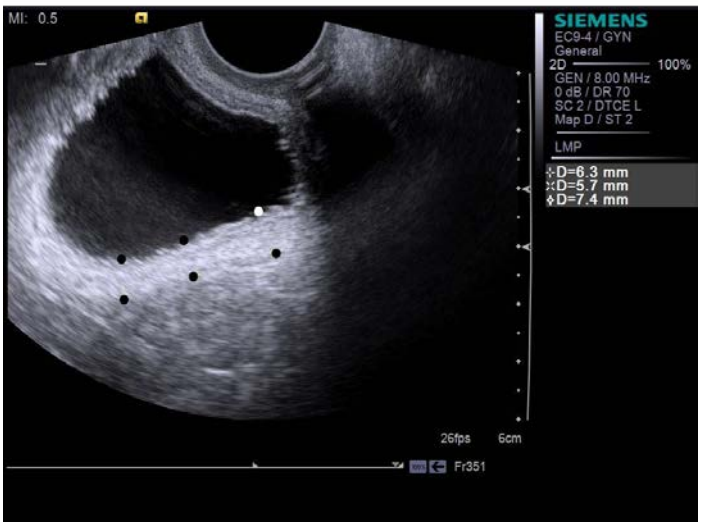
If PVR >=30ml, ask patient to re-void

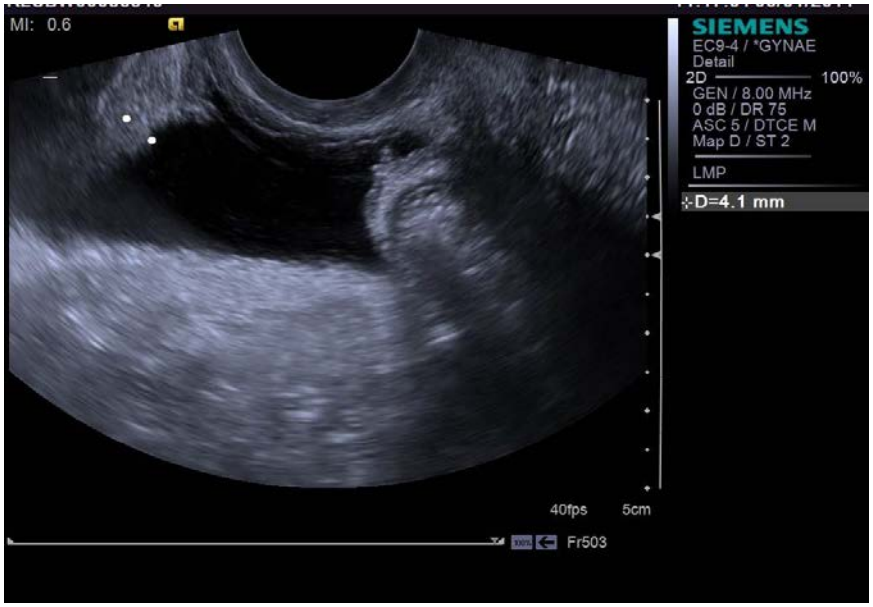
If after re-void PVR is >30 ml but <100 ml measure the BWT.

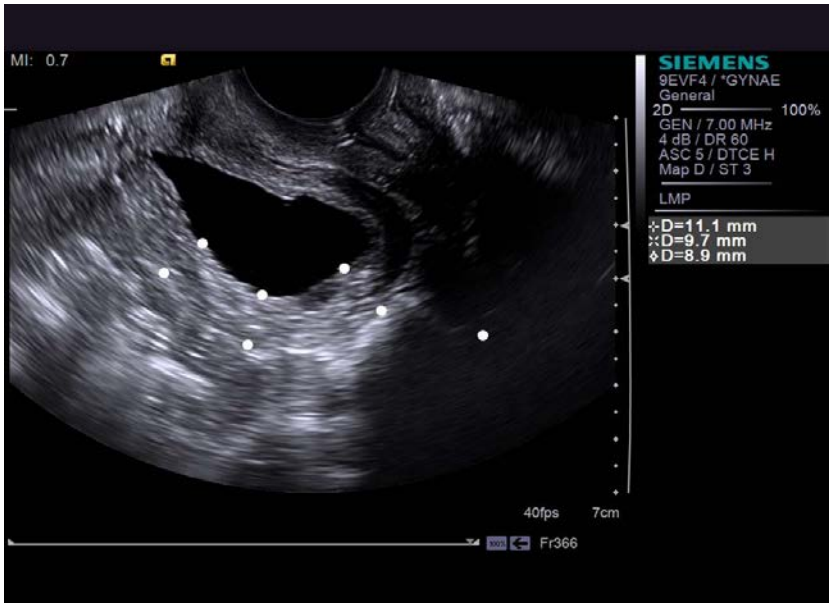
If the PVR is >100 ml exclude patient from BUS

BWT assessment:







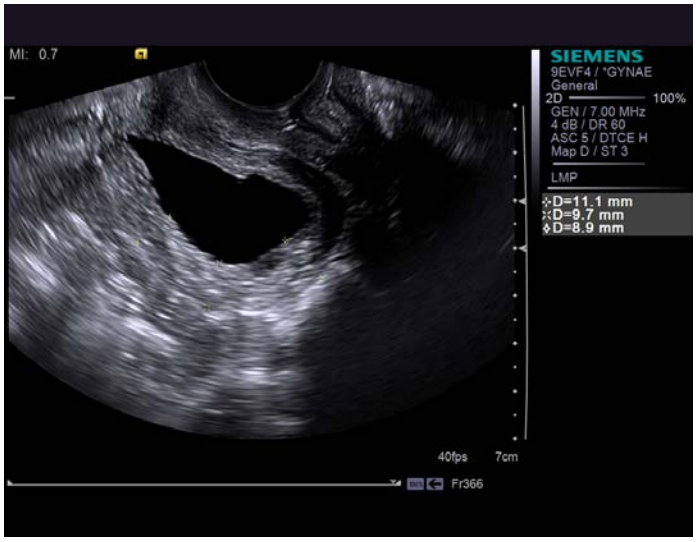


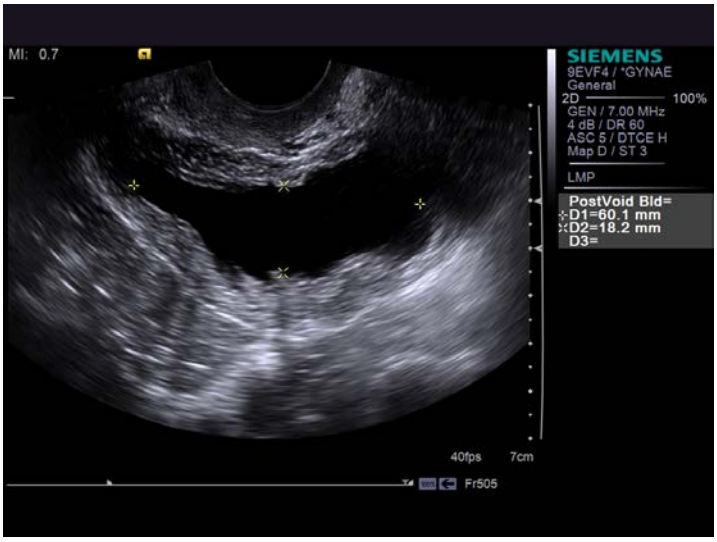
Measurement should be obtained in sagittal plane with VUJ on margin of screen (probe may need to be introduced slightly further)

3 x measurements: at the dome, anterior wall and trigone need to be obtained

The same image should be stored twice; with and without callipers

Any focal areas of thickening need separate assessment and evaluation





It is not always possible to see the anterior wall or trigone well, as many women with bladder problems often have irregular bladder outline. If the whole of the bladder is not visible on one image (as is preferable), there may be the need to angle anteriorly or posteriorly and measure the trigone or anterior wall thickness. Patients with significant cystocoeles may find complete emptying difficult and trigone may be difficult to visualise adequately

It is important to document any additional findings (diverticuli, cystocoeles, focal masses)

Providing Electronic Images:

Anonymised electronic images for each patient should be provided, labelled clearly with Study ID as the only identifier. These images should be as follows:

PVR image 1 WITH callipers

PVR image 1 WITHOUT callipers

PVR image 2 WITH callipers

PVR image 2 WITHOUT callipers

Bladder Wall Thickness image 1 WITH callipers

Bladder Wall Thickness image 1 WITHOUT callipers

Bladder Wall Thickness image 2 (optional) WITH callipers

Bladder Wall Thickness image 2 (optional) WITHOUT callipers

For optimal quality these should be sent to Birmingham Clinical Trials Unit on disk or memory stick in DICOM format ideally.

Failing this JPEG images will be accepted.

If downloadable images are not possible please supply a print out of each image listed above, ensuring that the print quality is optimal (what is seen on screen).

SUMMARY

Prepare patient (ensure voiding takes place just prior to scan and measure in supine position)

Prepare machine (use vaginal probe and frequency 7-9 MHz for optimal image)

USS assessment

Important to obtain good quality images

Callipers placed on margin of bladder wall for thickness; within bladder lumen (on wall) for volumes

Store images with and without callipers for cross-referencing and evaluation

Measure if PVR is <=30ml, revoid if PVR is >=30ml and measure if PVR is >30 but <100ml, exclude if PVR is>100ml.

6b) Standard Operating Procedure for Urodynamics

The urodynamics should be performed with aseptic precautions, counselling and verbal consent and according to the Good Urodynamics Practice Guidelines (ref- Schafer).

The equipment needed for the running of the Urodynamic Clinic include:

Catheter Pack

Filling Catheter

Abdominal and Bladder Pressure Catheters

Instillagel/Sterile Lubricant Gel

4x 3 way taps (depending on the type of transducers being used)

2x Fluid Filled Domes

1x 500ml bag of Normal Saline used for irrigation

1x Pump Infusion Set

1x Set Guard

2x Giving Sets

2x 100ml bag Normal Saline to flush the domes

3x pieces of tape (micropore etc, to attach once catheters inserted; ensuring they stay in place during filling)

1x Pair of Sterile Gloves

Non-Sterile Gloves

2x Incontinence Pads (1 used for the floor and 1 for patient to sit on)

Paper Roll to cover the couch

Sharps Box

Plastic Apron

Towel or Cover for the patient

Clean Trolley with Antiseptic Wipes

Please note: The above items may vary depending on the type and make of equipment used in each clinic, supplies used at the trust and also in accordance to infection control and hospital policies.

Ensure all equipment is set up and the person performing the test has not also undertaken the ultrasound scan on the patient.

Ensure the urodynamics test form is to hand and the patient registration number is entered.

If, for any reason the test had to be abandoned, note this on the test form.

Uroflowmetry (Initial Voiding Test)

The patient is asked to attend clinic with a comfortably full bladder. The patient should be encouraged to sit in order to void into the voiding flow/volume transducer funnel mounted under the commode.

- The patient should be instructed to dispose of any tissues/wipes into the bin/bag provided and not into the flow meter.
- The utmost privacy must be maintained during the test and the patient should be made to feel comfortable and relaxed, enabling a usual voided pattern to be established.
- The maximum void flow rate and volume should then be recorded.
- The post void residual volume should then be recorded using a drainage catheter and measuring container.

Filing Cystometry (Catheterisation)

It is essential that the machine is calibrated, set at zero at atmospheric pressure and a reference level for pressures should be established.

- Ideally the patient should be in the sitting position for the test. If this is not possible it should be recorded on the test form. A sheet should be provided for covering, maintaining dignity.

- Under aseptic technique, introduce catheters up through the urethra into the bladder and one into the rectum.

- Prior to filling, ask the patient to cough so that the traces can be observed. The spikes on the intravesical and intra-abdominal lines should be identical. Any necessary adjustments should be made and the cough repeated.

- Fill rate should be recorded on the test form, but is recommended as 100ml per minute.

- Ask the patient to cough every minute to ensure continued subtraction. If the lines slip, then stop the filling and rectify the problem.

- Complete test form with the ml at which the patient reports first, normal and strong desire to urinate, pain and volume leaked (if applicable).

- Total volume in the bladder at end of filling should be recorded.

- Detail any rise in detrusor pressure with or without urgency.

- At the end of filling, the large catheter used for filling the bladder is removed. The small catheter remains in the bladder to record voiding pressures (if using two separate catheters in the bladder).

Provocation Test (whilst bladder is still filled)

Whilst the intravesical and intra-abdominal lines are in situ, the patient should stand up on the incontinence sheet provided and the provocation tests like running taps, coughing, etc should be performed.

- Complete methods used and observations on the test form.

Flow Cystometry (voiding)

- Allow patient to void into commode, recording peak flow rate, max void pressure and residual volume with the pressure lines still in. During this voiding phase, the patient's dignity and privacy must be maintained and staff should leave the room if necessary.
- Ask the patient to cough pre and post void to ensure adequate subtraction.

Diagnosis

On completion of the investigation the results may be explained to the patient and fluid advice should be given.

- Record diagnosis in red section of test form.
- Any 'optional'/additional tests undertaken should be noted at the end of the form.
- If video urodynamics is being done, then it is recorded at the end of the form.
- If patient is scheduled for ambulatory urodynamics, then please give this information on the form.

Advice for patients

All women who have undertaken the test should be advised to expect some dysuria for up to 72 hours, their fluid intake should be increased during this time.

The occurrence of systemic symptoms, pyrexia and malaise should be advised as an indication to seek medical advice i.e. from their GP.

A contact number should be provided if problems occur.