
Publisher's PDF, also known as Version of record

License (if available): CC BY

Link to published version (if available): 10.1002/nau.24337

Link to publication record in Explore Bristol Research

PDF-document

This is the final published version of the article (version of record). It first appeared online via Wiley at https://onlinelibrary.wiley.com/doi/full/10.1002/nau.24337. Please refer to any applicable terms of use of the publisher.
Quality control of uroflowmetry and urodynamic data from two large multicenter studies of male lower urinary tract symptoms

Martino Aiello1 | Joseph Jelski1 | Amanda Lewis2,3 | Jo Worthington2,3 | Charlotte McDonald2,3 | Paul Abrams1 | Andrew Gammie1 | Chris Harding4 | Suzanne Biers5 | Hashim Hashim1 | J. Athene Lane2,3 | Marcus J. Drake4,6

1Bristol Urological Institute, Southmead Hospital, Bristol, UK
2Bristol Randomised Trials Collaboration (BRTC), Bristol Trials Centre, University of Bristol, Bristol, UK
3Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK
4Urology Department, Freeman Hospital, Newcastle upon Tyne, UK
5Urology Department, Addenbrooke’s Hospital, Cambridge, UK
6Translational Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

Correspondence
Marcus J. Drake, University of Bristol and Bristol Urological Institute, Level 3, Learning and Research Building, North Bristol NHS Trust, Southmead Hospital, Bristol BS10 5NB, UK.
Email: marcus.drake@bristol.ac.uk

Funding information
Health Technology Assessment Programme, Grant/Award Numbers: 12/140/01, 12/35/15

Abstract
Aim: The International Continence Society (ICS) has standardized quality control and interpretation of uroflowmetry and urodynamics. We evaluated traces from two large studies of male lower urinary tract symptoms (UPSTREAM and UNBLOCS) against ICS standards of urodynamic equipment and practice.

Methods: Ten percent of uroflowmetry and urodynamics traces were selected at random from hospital sites. A data capture template was designed from the ICS Fundamentals of Urodynamic Practice checklist. Two pre-trained blinded assessors extracted the data, with a third assessor to arbitrate. Departmental records of calibration checks and equipment maintenance were scrutinized.

Results: Seven out of twenty-five (28%) departments reported no calibration checks. Four sites (16%) could not provide annual service records. In 32 out of 296 (10.8%) uroflowmetry traces, findings were affected by artifact. One hundred ten urodynamic study traces were reviewed; in 11 records (10%), key pressure traces were incompletely displayed. In 30 (27.2%), reference zero was not set to atmospheric pressure. Resting pressures were outside the expected range for 36 (32.7%). Pressure drift was seen in 18 traces (16.4%). At pressure-flow study commencement, permission to void was omitted in 15 (13.6%). Cough testing after voiding was done in 71.2%, but the resulting cough spikes were significantly different in 16.5%. Erroneous diagnosis of bladder outlet obstruction (BOO) was identified in six cases (5.5%).

Conclusions: Erroneous diagnosis of BOO is a serious error of interpretation, as it could lead to unnecessary surgery. Other errors of standardization, testing, and interpretation were identified with lower risk of adverse implications.
Inconsistent documentation of service records mean equipment accuracy is uncertain.

KEYWORDS
LUTS, overactive bladder, standards, urodynamics, uroflowmetry

1 INTRODUCTION

Lower urinary tract symptoms (LUTS) in men are a common reason for referral to urologists, leading to consideration of surgery for benign prostatic obstruction (BPO). The diagnostic assessments for evaluating male LUTS are clearly set out in several good quality guidelines,1 based on a history and examination, symptom score, urinalysis, bladder diary, and uroflowmetry.2 Uroflowmetry is intended to measure maximum flow rate ($Q_{\text{max}}$), voided volume (VV) and post-void residual (PVR), with an assessment of flow pattern to give some initial impression of potential causative processes explaining LUTS.3 Additional testing in the form of urodynamic studies (UDS) may detect detrusor overactivity (DO) during bladder filling, and whether bladder outlet obstruction (BOO) or detrusor underactivity (DU) is responsible for the men’s voiding symptoms. Both the American Urological Association4 and the European Association of Urology2 Guidelines indicate that UDS are an optional part of assessment, but state that only pressure-flow studies (PFS) are able to differentiate DU from BPO as the cause of the men’s symptoms. Both uroflowmetry and UDS facilitate therapy decisions in men with LUTS, and this is especially significant when considering surgical intervention. Diagnostic assessment is mainly designed to focus therapy on causative mechanisms and therefore try to minimize the risk of “unnecessary surgery,” or to predict which men may have suboptimal outcomes if surgery is performed.

Any staff undertaking diagnostic tests need to have adequate training and experience in ensuring that testing is done to appropriate quality standards and understanding the implications of results for decision-making. The International Continence Society (ICS) working groups have proposed the current approaches to standardization, quality control, and interpretation of uroflowmetry and UDS.5,5,8 There are several key aspects with implications for test results:

1. Adherence to the ICS good urodynamic practices5,6 ensures standardization so that tests reliably derive the main observations and results can be compared.
2. Monitoring the technical aspects in real time during the test helps ensure prompt intervention to reduce the incidence of artifacts.9,10
3. Post-test processing is needed to ensure that key metrics to derive urodynamic observations are not taken at the time of an artifact and to calculate indices.
4. Sometimes a test is identified as unrepresentative of day-to-day life, notably as a result of patient anxiety or the artificial nature of the testing situation.
5. A department doing clinical testing regularly needs to ensure that the equipment measures values reliably, undertaking departmental calibration checks every 10 tests.11
6. Annual maintenance and checks of system performance should also be undertaken.7

These are complex issues requiring experience to achieve UDS testing in line with the required standards. Nonetheless, a wide range of staff come into contact with patients during the male LUTS assessment pathway, including doctors and nurses at various career stages. Thus, each department needs to ensure adequate training and scrutiny of performance.

UPSTREAM12 and UNBLOCS13 are two large UK-based studies set in 30 different urology departments, studying the therapeutic pathway for male LUTS. UPSTREAM is a randomized controlled trial assessing urodynamic testing for the diagnosis and management of BOO in men. UNBLOCS is a randomized controlled trial to determine the clinical and cost effectiveness of thulium laser vaporisation of the prostate vs standard transurethral resection. The aim of this paper is to assess the quality of a sample of the uroflowmetry and UDS data from the two studies.

2 METHODS

Uroflowmetry was undertaken in both UNBLOCS and UPSTREAM trials, with UDS testing also undertaken in the UPSTREAM trial. Two sites overlapped between the two trials. Ten percent of uroflowmetry (30 sites) and urodynamics traces (26 sites) were selected at random from sites. A data capture template was designed from the ICS Fundamentals of Urodynamic Practice checklist7 to examine parameters that could be identified from source traces and data (Table 1). Two pretrained blinded assessors (MA, JJ) independently
extracted the data for each trace. Where there was disagreement, a third assessor (MJD) arbitrated the conclusion. Trace scrutiny evaluated the presence of recognized urodynamic features and artifacts and adherence to the ICS good urodynamic practices; nonadherence was categorized as an error of standardization. Apparent failure to ensure real-time monitoring and correction of artifacts was categorized as an error of technique. Post-test processing was checked to ensure that urodynamic observations were not taken at the time of an artifact, and to validate the derived indices (BOO Index [BOOI] and Bladder Contractility Index [BCI]). Inaccuracy at this stage was categorized as an error of interpretation. Where applicable, patterns of flow rate were categorized as normal, intermittent, compressive (classically BPO), or constrictive (typically urethral stricture) as specified in the ICS good urodynamic practices documents. In addition, an “indeterminate” category was included, where a flow pattern could not be attributed to an established category.

In addition, each participating research site in the UPSTREAM trial was asked to provide routine departmental records of regular calibration checks to evaluate against ICS standards on urodynamic equipment. These were examined to look for evidence of the calibration of the flow meters (volume and flow rate measurement) and the pressure transducers used to measure vesical (Pves) and abdominal pressure (Pabd). Copies of annual equipment

### TABLE 1 Quality and interpretation recommendations (derived from Gammie and Drake and Schafer et al)

<table>
<thead>
<tr>
<th>Question/subject</th>
<th>Solution/guideline</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Uroflowmetry</strong></td>
<td></td>
</tr>
<tr>
<td>Is the trace clear of artifacts from the movement of</td>
<td>Adjust trace markers if possible and instruct patient for improved next flow</td>
</tr>
<tr>
<td>body, flow meter or urine stream?</td>
<td></td>
</tr>
<tr>
<td>Is Qmax marked at a point away from artifacts?</td>
<td>Move Qmax marker to smoothed maximum position</td>
</tr>
<tr>
<td>Are the markers for start and end of void away from</td>
<td>Move markers away from artifacts</td>
</tr>
<tr>
<td>artifacts or drops of urine?</td>
<td></td>
</tr>
<tr>
<td>Does the scale of printing make the flow trace clearly</td>
<td>Adjust scale of display/print</td>
</tr>
<tr>
<td>visible?</td>
<td></td>
</tr>
<tr>
<td>Has the residual urine volume been measured immediately</td>
<td>Measure volume, including comment on any time delay</td>
</tr>
<tr>
<td>after voiding?</td>
<td></td>
</tr>
<tr>
<td><strong>Cystometry and pressure-flow study</strong></td>
<td></td>
</tr>
<tr>
<td>Quality</td>
<td></td>
</tr>
<tr>
<td>Zero to atmosphere</td>
<td>Ensure taps are closed to patient and open to air when zero is pressed</td>
</tr>
<tr>
<td>Check resting pressures are normal</td>
<td>Supine: Pabd and Pves, 5 to 20 cmH2O</td>
</tr>
<tr>
<td></td>
<td>Seated: Pabd and Pves, 15 to 40 cmH2O</td>
</tr>
<tr>
<td></td>
<td>Standing: Pabd and Pves, 30 to 50 cmH2O</td>
</tr>
<tr>
<td></td>
<td>For all positions: Pdet, −5 to +5 cmH2O</td>
</tr>
<tr>
<td>Continuous monitoring</td>
<td>Check regularly that pressure transmission is equal on both lines, for example,</td>
</tr>
<tr>
<td></td>
<td>coughs, blowing</td>
</tr>
<tr>
<td></td>
<td>Check that live patient signal is present throughout</td>
</tr>
<tr>
<td></td>
<td>Check that baseline pressures do not drift</td>
</tr>
<tr>
<td></td>
<td>Troubleshoot above during the test, temporarily stopping recording/filling, if</td>
</tr>
<tr>
<td></td>
<td>necessary</td>
</tr>
<tr>
<td></td>
<td>Stop or reduce fill rate if urgency is excessive or compliance is poor</td>
</tr>
<tr>
<td></td>
<td>Change patent position as required (eg, discomfort, stress testing)</td>
</tr>
<tr>
<td></td>
<td>Consider repeating test if urodynamic question not answered</td>
</tr>
<tr>
<td>Interpretation</td>
<td>Place markers on the trace frequently, for example, sensation, patient position,</td>
</tr>
<tr>
<td></td>
<td>stress tests, permission to void</td>
</tr>
<tr>
<td></td>
<td>Adjust positions of markers after completion of test, if needed</td>
</tr>
<tr>
<td></td>
<td>Take care when interpreting, for example, rectal contractions, knocking of flow</td>
</tr>
<tr>
<td></td>
<td>meter</td>
</tr>
<tr>
<td><strong>Equipment records</strong></td>
<td></td>
</tr>
<tr>
<td>Documented calibration checks</td>
<td>Ensure calibration is checked and documented regularly</td>
</tr>
<tr>
<td>Equipment maintenance records</td>
<td>Ensure equipment maintenance is carried out and recorded</td>
</tr>
</tbody>
</table>
maintenance records were also requested. The anonymized database of UDS trace reviews and calibration/maintenance records is given in Supporting Information Material S1.

3 | RESULTS

3.1 | Equipment calibration and maintenance

Twenty-five of twenty-six UPSTREAM sites (96%) responded to a review of departmental calibration and annual maintenance records. The calibration of flow meter volumes was undertaken by 12 out of 25 (48%). Calibration of flow meters and pressure transducers was catalogued in 4 out of 25 (16%) and 8 out of 25 (32%), respectively. The method of transducer calibration was with a water column by 4 out of 8 (50%), and reliance on internal equipment “self-calibration” in the rest. Where undertaken, 66% of sites undertook them monthly or more often. Seven out of twenty-five (28%) urodynamic departments reported no calibration checks. Where undertaken, 66% of sites undertook them monthly or more often. Seven out of twenty-five (28%) urodynamic departments reported no calibration checks. In terms of urodynamic equipment maintenance, 19 out of 25 (76%) departments held an on-site annual service record. A further two sites (8%) reported an annual service contract with records held externally by the manufacturer of the urodynamic equipment. Four sites (16%) were unable to provide confirmation of annual servicing, of which two also did no regular department equipment calibration checks.

3.2 | Uroflowmetry testing

Three hundred thirteen uroflowmetry traces selected at random were reviewed from 30 sites. Of these, 14 traces from 10 centers (4.5%) could not be analyzed, as the patient was unable to do a flow test or did not comply with test requirements (10 of these had a bladder diary showing the largest void volume was above 150 mL). Accordingly, 299 traces were used for analysis. These came from 236 patients, as 48 patients from 15 sites did two flows, and another 15 from four sites did more than two flows.

The uroflowmetry traces did not display the volume or maximum flow rate ($Q_{\text{max}}$) values in 47 out of 299 (15.7%) from 17 sites. Voided volume was missing in 11 out of 299 (3.7%), $Q_{\text{max}}$ in 12 (4.0%), and PVR in 43 (14.4%). Flow pattern was categorized as normal in 55 out of 296 (18.6%), indeterminate in 49 (16.6%), intermittent in 3 (1.0%), compressive in 163 (55.1%), or constrictive in 5 (1.7%), and 21 were considered uninterpretable (7.1%).

In 32 out of 296 traces (10.8%), the $Q_{\text{max}}$ information was affected by an artifact. This took the form of an interruption of flow followed by a spiked increase of flow rate in 13, a spike without preceding interruption in 5, and suspected Valsalva or strain in 14. In only 2 out of 32 was a correction applied after the study to derive $Q_{\text{max}}$ data from a part of the trace unaffected by an artifact.

Consistency of information between the source trace and the test report was checked in 289 traces. Errors in values on the test report were identified in 90 (31%) traces from 14 sites, with each 3 of the 14 sites responsible for 10 or more consistency errors. The errors affected documentation of VV in 26 out of 90 (28.9%), $Q_{\text{max}}$ in 23 (25.6%), and PVR in 41 (45.6%).

3.3 | Urodynamics review

One hundred twenty-three traces were provided from 26 sites, of which 13 were excluded (1 due to poor quality
reproduction, 1 because the test was not done, 7 because only a typed report was provided, and 4 for unspecified/unknown reasons), giving an analysis set of 110 traces. In these, some features of the filling cystometry traces were not clearly reproduced for 11 traces. Features of the PFS could not be evaluated in three traces, including two patients who were unable to void during the test.

Three particular errors of standardization were identified:

1. Wrong display sequence of traces. In 11 records (10%), the pressure traces of key data ($P_{\text{ves}}$ or $P_{\text{det}}$) were incomplete as high values went off-scale, as a result of being placed too high in the display sequence.
2. Failure to set atmospheric pressure as the zero reference (Figure 1). In 30 out of 110 (27.2%) traces from 14 sites, the zero-reference pressure was not set to atmospheric pressure; instead, the “zero all” software command was given when transducers were recording from the patient.
3. Annotations to indicate aspects of the test were often not provided. For example, the position in which the test was run (seated or standing) was not marked in 90 (81.8%), and markers indicating provocation tests were not present in 98 (89.1%). Normal and abnormal subjective sensation markers (at least one of the first sensation of filling, normal desire to void, strong desire to void, or urgency) were generally included; they were missing in six (from four sites, with three from one site). At the start of the PFS, permission to void (indicated by markers such as “start void”/“voiding”/“PFS”) was not marked in 15 out of 110 (13.6%) from six sites.

Several errors of technique were identified in a high proportion of cases:

1. Unreliable pressure recording (Figure 1). Resting pressures, at the start of filling, were outside the expected range for 36 out of 110 traces (32.7%) from 16 sites (expected ranges used were 30 to 50 cmH$_2$O in the standing position and 15 to 40 cmH$_2$O if seated; where position was not stated when the range used was 15 to 50 cmH$_2$O). The resting detrusor pressure ($P_{\text{det}}$) at the start of filling cystometry was outside the range $-5$ to $+5$ cmH$_2$O in 14 (12.3%) from 7 sites. Regular transmission checks of pressure (individual coughs) were lacking in 14 (12.7%) from 7 sites. A drift in one of the pressure lines was seen in 18 traces (16.4%) from 11 sites. Only one trace did not appear to have a “live” signal.
2. Failure to interrupt filling in the event of urgency, DO or reduced compliance as per ICS recommendations. Due to limited annotation or lack of filling rate information, this feature could not be evaluated in 75 out of 110 (68.1%). In the 35 traces where it could be evaluated, it was done in 3 out of 35 (8.6%), not done in 15 (42.9%), and did not apply (no urgency, DO or impaired compliance present) in 17 (48.8%).
3. A cough test before voiding (ie, between stopping filling and start of voiding) was not done in 48 (43.6%); a prevoid cough showing bad subtraction (Figure 1), which was not dealt with before voiding, was seen in 1 case (bad pressure transmission of $P_{\text{abd}}$); and the cough test was apparent and showed good subtraction in 61 (55.5%). Cough testing after voiding was done in 79 (71.2%), but the resulting cough spikes were significantly different ($\geq30\%$ height discrepancy between the respective spikes in the abdominal and vesical pressure traces) in 13 out of 79 (16.5%), resulting in large subtraction artifacts on $P_{\text{det}}$. Both the pre- and post-voiding cough tests were missing in 24; for another 16 studies, it was missing in 1 and inaccurate in the other (without action taken to correct the poor quality).

### 3.4 Interpretation of urodynamics traces

PFS traces with derived voiding parameters (BOOI and BCI) were available for 107 patients (97.3%): central review diagnosis was assumed to be the correct calculation from the raw data. For diagnosis of BOO, there was an agreement between sites and central review for 71% of traces where BOO was present (BOOI $>40$), 79% where it was equivocal (BOOI, 20–40), and 42% of men who were unobstructed (BOOI < 20) (Table 2). Particularly notable was the group of 19 patients for whom central review was unable to derive a BOOI value, although sites gave a diagonsis for 15 of them, among whom 10 were given a diagnosis of BOO. The level of agreement was the lowest for the diagnosis of DU; being only $52\%$ where DU was present (BCI < 100) and $31\%$ where it was absent (BCI > 100). For 21 patients, central review was unable to derive a BCI value, for whom sites gave a diagnosis for 4 (2 of these being given a diagnosis of BOO). Data about the urodynamic observation of DO was available for 99 traces (90.0%). Fewer traces were evaluable for DO, as some sites did not provide traces that included filling cystometry. There was agreement between central review and site categorization in 57% of available traces where DO was present, and 51% where it was absent (Table 2).

The reviewers also identified additional factors needing caution when deriving BOOI and BCI, namely, straining at the time of $Q_{\text{max}}$ in 15 out of 99 (15.2%), voiding data derived from a DO incontinence episode in 4
out of 99 (4.0%), and a major drop in $P_{\text{abd}}$ sufficient to alter BOO diagnosis in one case. Annotation of key volumes (trace or attached data) was lacking for PVR in 37 out of 110 (33.6%), cystometric capacity in 35 (31.8%), and VV in 15 (13.6%). Twenty-four of those traces missing key volume data (21.8% of the 110-sample population) were missing two or more of the volumes.

### 4 DISCUSSION

Questions exist regarding the influence of technique on the utility and perceived value of urodynamics. The findings from the analysis of UDS and uroflowmetry data indicate that a detailed quality control process is essential for research involving urodynamics, and supports the importance of developing accreditation/reaccreditation schemes for all health care professionals running these types of test. The UK Continence Society recently reviewed its minimum standards for urodynamics, mindful of the fact that there is concern over the quality of urodynamic testing in practice, and the report has been endorsed by the ICS. Experience of central reading of urodynamic traces from multicenter studies on therapies have reinforced this concern and made it clear that this is an international issue (Abrams and Gammie, personal communication, December, 2019).

The need for caution when interpreting studies also arose because sites sometimes failed to identify a potential recording problem at the crucial moment of $Q_{\text{max}}$. For uroflowmetry across two major trials in the United Kingdom,
the $Q_{\text{max}}$ information was affected by an artifact in 11%, and in the majority no correction was applied after the study to derive flow rate data from a part of the trace unaffected by the artifact (“corrected $Q_{\text{max}}$”). For PFS, cough testing after voiding was omitted in 29% and was significantly inaccurate ($\geq 30\%$ height discrepancy between the pressure spikes) in 17%, making the reliability of the $P_{\text{det}}Q_{\text{max}}$ value uncertain.

Sites sometimes did not run tests in accordance with the ICS standardizations. For UDS tests, this included not setting the reference zero pressure to atmospheric pressure (27%) and limited annotation of traces to help interpretation (eg, permission to void not marked in 14%). This generally does not have direct implication for an individual patient test, though sometimes it may obscure a pressure recording error not picked up by discrepancy in the cough pressure spikes. However, it makes comparisons between centers very difficult, which is unhelpful for standard audit and exploratory research. In the United Kingdom, the Improving Quality in Physiological Services (IQIPS; https://www.ukas.com/services/accreditation-services/physiological-services-accreditation-iqips/) scheme allows sites to become accredited against good practice standards that encompass patient experience, facilities, safety, and technical quality. The results we present here show that there is a need for such accreditation to enable departments to demonstrate and maintain good practice for patient benefit.

The strengths of the current study were the use of quality control protocols derived from the ICS Fundamentals of Urodynamic Practice document,\(^7\) random trace-selection, and using specifically trained central readers blinded to the sites under review with expert adjudication for resolution of discrepancy. The main limitation was the necessity to sample 10% of traces, as this may give an unrepresentative picture of individual site performance, though this does not affect the overall evaluation.

5 | CONCLUSIONS

This quality control study has shown that there are significant issues with respect to equipment maintenance and calibration, testing procedures, and the interpretation of results for both uroflowmetry and urodynamics in a large number of urodynamic units. This has led to a small proportion of men being wrongly characterized as having BOO, and hence at risk of being counseled inappropriately to surgery.

ACKNOWLEDGMENTS

The UPSTREAM and UNBLOCs trials are funded by the National Institute for Health Research (NIHR) Health Technology Assessment (HTA) programme (project references 12/140/01 and 12/35/15, respectively). The views expressed are those of the author(s) and not necessarily those of the NIHR or the UK Department of Health and Social Care. The UPSTREAM and UNBLOCs trials were sponsored by North Bristol NHS Trust (NBT). The UPSTREAM and UNBLOCs trials were designed and delivered in collaboration with the Bristol Randomised Trials Collaboration (BRTC), a UKCRC registered clinical trials unit (CTU) which, as part of the Bristol Trials Centre (BTC), is in receipt of National Institute for Health Research CTU support funding. Study data were collected and managed using REDCap\(^7\) hosted at the University of Bristol.

REFERENCES


**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section.