Title
Critical Analysis of the Use of Uroflowmetry for Urethral Stricture Disease Surveillance.

Permalink
https://escholarship.org/uc/item/4p29r1st

Journal
Urology, 91

ISSN
0090-4295

Authors
Tam, Christopher A
Voelzke, Bryan B
Elliott, Sean P
etal.

Publication Date
2016-05-01

DOI
10.1016/j.urology.2015.12.070

Peer reviewed
Critical Analysis of the Use of Uroflowmetry for Urethral Stricture Disease Surveillance

Christopher A. Tam, Bryan B. Voelzke, Sean P. Elliott, Jeremy B. Myers, Christopher D. McClung, Alex J. Vanni, Benjamin N. Breyer, and Bradley A. Erickson for the Trauma and Urologic Reconstruction Network of Surgeons (TURNS)

OBJECTIVE
To critically evaluate the use of uroflowmetry (UF) in a large urethral stricture disease cohort as a means to monitor for stricture recurrence.

MATERIALS AND METHODS
This study included men that underwent anterior urethroplasty and completed a study-specific follow-up protocol. Pre- and postoperative UF studies of men found to have cystoscopic recurrence were compared to UF studies from successful repairs. UF components of interest included maximum flow rate ($Q_m$), average flow rate ($Q_a$), and voided volume, in addition to the novel post-UF calculated value of $Q_m$ minus $Q_a$ ($Q_m-Q_a$). Area under the receiver operating characteristic curves (AUC) of individual UF parameters was compared.

RESULTS
$Q_m-Q_a$ had the highest AUC (0.8295) followed by $Q_m$ (0.8241). UF performed significantly better in men ≤40 with an AUC of 0.9324 and 0.9224 for $Q_m$ and $Q_a$ respectively, as compared to 0.7484 and 0.7661 in men >40. Importantly, of men found to have anatomic recurrences, only 41% had a $Q_m$ of ≤15 mL/s at time of diagnostic cystoscopy, whereas over 83% were found to have a $Q_m-Q_a$ of ≤10 mL/s.

CONCLUSION
$Q_a$ rate alone may not be sensitive enough to replace cystoscopy when screening for stricture recurrence in all patients, especially in younger men where baseline flow rates are higher. $Q_m-Q_a$ is a novel calculated UF measure that appears to be more sensitive than $Q_m$ when using UF to screen for recurrence, as it may be a better numerical representation of the shape of the voiding curve. UROLOGY 91: 197–202, 2016. Published by Elsevier Inc.

Uroflowmetry (UF) is a simple, noninvasive method to evaluate voiding function in patients experiencing lower urinary tract symptoms. It is often combined with other metrics, including the International Prostate Symptom Score, in the initial diagnosis and follow-up of benign prostatic hyperplasia (BPH), and other causes of obstruction. In patients with urethral stricture disease (USD) who have undergone urethroplasty, UF is one of the most frequently used tests to monitor for stricture recurrence. However, UF’s use as a stand-alone tool to screen for recurrence following urethroplasty has never been rigorously validated.

It has been well established that the maximum flow rate ($Q_m$) in patients with USD is significantly diminished relative to age-matched normal controls. This knowledge has been extrapolated to the post-urethroplasty setting, where commonly used cutpoints of a postoperative $Q_m$ of less than 10 mL/s or a postoperative $Q_a$ of less than 15 mL/s are used as indicators of urethral stricture recurrence. Similarly, when UF data are available both pre- and postoperatively, a change in $Q_m$ following surgery of less than 10 mL/s has also been suggested as a predictor of recurrence. The goal for each of these UF parameters is to minimize the invasiveness of postoperative screening while maximizing the ability to find recurrences.

The purpose of this study is to rigorously evaluate the capability of individual UF parameters, such as $Q_m$ and average flow rate ($Q_a$), as well as a novel hybrid measure ($Q_m-Q_a$) to monitor for urethral stricture recurrence. Use of $Q_m-Q_a$ has not been described in prior literature and attempts to provide a simple method to quantify the shape of the voiding curve. The study tested two hypotheses: (1) when compared to the gold-standard cystoscopy, UF parameters will have high test (screening) sensitivity and specificity, and (2) the sensitivity and specificity of UF to screen for stricture recurrence will be diminished in older patients.

Financial Disclosure: The authors declare that they have no relevant financial interests.

Address correspondence to: Bradley A. Erickson, M.D., M.S., Department of Urology, University of Iowa, Iowa City, IA 52242. E-mail: brad-erickson@uiowa.edu

Submitted: September 25, 2015, accepted (with revisions): December 16, 2015

Published by Elsevier Inc.

http://dx.doi.org/10.1016/j.urology.2015.12.070 0090-4295
MATERIALS AND METHODS

Subjects

The Trauma and Urologic Reconstruction Network of Surgeons (TURNS) is a multi-institutional effort that aims to prospectively monitor urethroplasty outcomes. The shared, centrally located web-based TURNS database was retrospectively queried for all men who had undergone anterior urethroplasty between 2009 and 2014. Data for these men were prospectively collected under Institutional Review Board-approved protocols, with patient consent obtained prior to surgery. Study inclusion criteria included men who had a follow-up cystoscopy at 3, 6, or 12 months postoperatively and had a corresponding same-day UF study. In patients with multiple follow-up cystoscopies/UF studies, the most recent instance was used for analysis. Recurrence was defined as the inability to advance a 17 French cystoscope past the previously reconstructed portion of the urethral lumen with minimal force; neither symptoms nor requirement for secondary operations were considered in this definition.

UF

Interpretation of UF readouts was made by the surgeon of record as per study protocol. Basic parameters of UF included $Q_{\text{av}}$, voided volume (VV), postvoid residual (PVR), and shape of the voiding curve. A novel calculated value was $Q_{\text{av}} - Q_m$ ($Q_{\text{av}} - Q_m$). The changes ($\Delta$) between pre- and postoperative parameters were also calculated in a subset of men. UF studies with voided volumes of less than 150 mL were discarded from the analysis.

Statistics

Descriptive statistics were first used to characterize the patient demographics, location of urethral stricture, and nature of repair. Men were divided into either a cystoscopic recurrence or successful repair group, and $t$ tests were used to assess the differences in pre- and postoperative UF parameters between the two groups. Receiver operating characteristic (ROC) curves were constructed to determine the predictive value of each UF parameter in diagnosing urethral stricture recurrence relative to the cystoscopic gold standard. Sensitivity, specificity, positive predictive value, and negative predictive value of UF parameters to detect cystoscopic recurrence were calculated using predetermined, commonly cited cutoffs. The patients were further stratified into >40 years or ≤40 years of age, and similar analysis was repeated. Follow-up was determined as the time from surgery to the time of the last objective (UF or cystoscopy) data point. Statistical analysis was completed using SAS® 9.3 (Cary, NC), with statistical significance set at $P < .05$.

RESULTS

Demographics

Of the 1181 men in the TURNS database, 323 men met study criteria. The majority of men were excluded because of a lack of postoperative cystoscopy data ($n = 524$) or an absent or poor UF study ($n = 334$) from the same clinic visit. Urethroplasty was performed by 7 surgeons from different academic institutions. The mean age of included patients was $44.35 \pm 15.26$ with a mean follow-up time of $12.84 \pm 12.38$ months. The most common location of stricture repair was the bulbar urethra ($n = 272$), followed by the penile urethra ($n = 27$), and the mean intraoperative stricture length was $3.62 \pm 2.93$ cm. The most common repair was excision and primary anastomosis ($n = 139$), followed by substitution ventral onlay ($n = 55$) and substitution dorsal onlay ($n = 42$). Using cystoscopic criteria, 58 (18%) of the men in the study were noted to have recurrence.

Preoperative UF Data

Preoperative UF studies were available in 189 (59%) of the men. The mean preoperative $Q_m$ was $9.44 \pm 6.82$ mL/s, mean preoperative $Q_{\text{av}}$ was $5.87 \pm 4.40$ mL/s, mean VV was $258.12 \pm 176.50$ mL, and mean PVR was $162.26 \pm 198.64$ mL. Preoperative UF values were not predictive of operative success nor did they correlate with age, stricture length, or stricture location.

Postoperative UF Data

Comparison of postoperative UF data between men with and without evidence of cystoscopic recurrence is shown in Table 1. The mean postoperative $Q_{\text{av}}$, $Q_m$, and $Q_{\text{av}} - Q_m$ were significantly different between cohorts; there was no difference in postoperative VV ($398.91 \pm 204.33$ vs $365.33 \pm 205.62$ mL, $P = .2584$).

ROC analysis was performed comparing UF to cystoscopy (gold standard) (Fig. 1). Postoperative $Q_{\text{av}}$, $Q_m$, demonstrated the highest area under the receiver operating characteristic curves (AUC) of 0.8295 (95% confidence interval: 0.7426, 0.9164); postoperative $Q_{\text{av}}$ followed closely behind with an AUC of 0.8241 (0.7452, 0.9031). AUC values were not significantly different between $Q_{\text{av}}$ and $Q_m$. Postoperative PVR demonstrated an AUC of 0.6296.

Table 1. Comparison of UF parameters between successful repair and recurrence groups (ranked by ROC AUC)

<table>
<thead>
<tr>
<th>Successful Repair Group</th>
<th>Recurrence Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N$</td>
</tr>
<tr>
<td>Postoperative $Q_{\text{av}} - Q_m$ (mL/s)</td>
<td>253</td>
</tr>
<tr>
<td>Postoperative $Q_m$ (mL/s)</td>
<td>265</td>
</tr>
<tr>
<td>$\Delta Q_m$ (mL/s)</td>
<td>157</td>
</tr>
<tr>
<td>$\Delta (Q_{\text{av}} - Q_m)$ (mL/s)</td>
<td>146</td>
</tr>
<tr>
<td>Postoperative $Q_{\text{av}}$ (mL/s)</td>
<td>253</td>
</tr>
<tr>
<td>$\Delta Q_{\text{av}}$ (mL/s)</td>
<td>146</td>
</tr>
<tr>
<td>Postoperative PVR (mL)</td>
<td>244</td>
</tr>
<tr>
<td>Postoperative VV (mL)</td>
<td>265</td>
</tr>
</tbody>
</table>

AUC, area under the receiver operating characteristic curves; PVR, postvoid residual; $Q_{\text{av}}$, average flow rate; $Q_m$, maximum flow rate; ROC, receiver operating characteristic; UF, uroflowmetry; VV, voided volume.
Sensitivity tables were constructed with various cutpoints to further evaluate the predictive capabilities of each parameter (Table 2). A commonly used cutpoint of $Q_m < 10 \text{ mL/s}$ had a sensitivity for detecting cystoscopic recurrence of only 21%. A postoperative $Q_m-Q_a < 10 \text{ mL/s}$ was 83% sensitive and 58% specific.

Subgroup analysis stratified men into cohorts of $\leq 40$ and $>40$ years of age. The recurrence rates were similar between the groups (17% vs 18%, respectively, $P = .9016$). Men $\leq 40$ years had a higher postoperative mean $Q_m$ (31.45 ± 13.60 mL/s) compared to men $>40$ years (22.18 ± 10.16 mL/s, $P < .0001$). ROC analysis revealed significantly higher AUC in men $\leq 40$ years compared to men $>40$ for both $Q_m-Q_a$ (0.9324 vs 0.7484) and $Q_m$ (0.9224 vs 0.7661).

Men with preoperative UF studies available were used as a urethral stricture test cohort to validate the sensitivities of the cutpoints. A $Q_m-Q_a < 10 \text{ mL/s}$ was 94% sensitive (ie, 169 of 179 patients with preoperative UF had a $Q_m-Q_a$ of $<10 \text{ mL/s}$). Only 60% (115/189) of patients had a $Q_m$ of $<10 \text{ mL/s}$ and 84% (158/189) had a $Q_m$ of $<15 \text{ mL/s}$.

**DISCUSSION**

The purpose of this study was to critically evaluate and compare the ability of UF parameters to independently identify cystoscopic recurrence of urethral strictures following urethroplasty. Of specific interest were the commonly cited cutpoints of postoperative $Q_m < 10 \text{ mL/s}$, postoperative $Q_m < 15 \text{ mL/s}$, and $\Delta Q_m < 10 \text{ mL/s}$ as indicators of recurrence. In this study cohort, we did not find that these generic cutpoints were sensitive enough for use as reliable

---

**Figure 1.** (A-C)—ROC curves of UF parameters predicting cystoscopic urethral stricture recurrence. ROC, receiver operating characteristic; UF, uroflowmetry. (Color version available online.)

**Figure 2.** Uroflowmetry tracings from two postoperative patients with high (normal) maximum flow rates. Patient A had a normal cystoscopy, whereas Patient B was found to have recurrence. Note the differences in the $Q_m-Q_a$ between the two patients.
A postoperative $Q_m < 10 \text{ mL/s}$ was previously reported to have a sensitivity of only 41%. Although $Q_m$ alone is typically the parameter of interest when interpreting UF, its usage as a screening tool is hampered by the wide distribution of $Q_m$ in the recurrence group ($17.11 \pm 8.31 \text{ mL/s}$), likely the result of heterogeneous effects from bladder dysfunction and prostate size/obstruction. An improvement can be seen with $\Delta Q_m$, which allows for an individually normalized value. A prior study reported that an improvement of $\Delta Q_m < 10 \text{ mL/s}$ had a sensitivity of 94% with a specificity of 78%.

In this population, a similar improvement in sensitivity to 81% and specificity to 48% was seen with this threshold.

The novel $Q_m-Q_a$ parameter may be superior to $Q_m$ or $\Delta Q_m$ in monitoring for stricture recurrence (Fig. 2). ROC AUC for $Q_m-Q_a$ (0.8295) was similar to $Q_m$ (0.8241, $P = .0809$) but higher than $\Delta Q_m$ (0.7638, $P = .0492$). Using a cutpoint of $Q_m-Q_a < 10 \text{ mL/s}$, a sensitivity of 83% and a specificity of 58% were seen. Unlike $Q_m$ alone, the $Q_m-Q_a$ is able to capture the overall shape of the curve by factoring in $Q_a$. A patient with a cystoscopic recurrence on the higher end of the $Q_m$ spectrum may have a flow of 20 to 25 mL/s, yet still present with a flat voiding curve. Whereas the typical cutpoints of $Q_m$ will fail to capture this patient, the $Q_m-Q_a$ is more likely to identify the recurrence. Had a $Q_m-Q_a$ of $<10 \text{ mL/s}$ been used as a standalone method to screen for recurrence in this population, 154 fewer cystoscopies would have been performed, but 10 strictures would have been missed. If the entire cohort was
preoperatively considered as a group of strictures, the preoperative $Q_m - Q_a < 10 \text{ mL/s}$ would have identified 94% of strictures compared to only 60% for $Q_m < 10 \text{ mL/s}$.

The value of $Q_m$ correlates inversely with age, especially in the population over age 50, where there is a sharp drop off regardless of $VV$. Younger patients demonstrate more robust flow due to stronger bladder contractions and less prostatic obstruction. In this study, both $Q_m$ and $Q_m - Q_a$ demonstrated superior predictive capability in detecting stricture recurrence in patients ≤40 years of age (AUC of 0.9224 and 0.9324, respectively) compared to patients >40 years of age (AUC of 0.7661 and 0.7484, respectively). The stronger flow of healthy younger men allows for better discrimination between a patent and strictured urethra. In an older individual, this difference may be less pronounced. Overall, UF appears to have better predictive value in the younger patient population and thus, it may be a more useful stand-alone tool for stricture monitoring in the younger group. In older individuals, where UF parameters are more profoundly affected by the size of the prostate, monitoring of patient-specific subjective measures will likely always remain important.

Limitations to the study include the strict interpretation of a urethral stricture recurrence. This study focused on the anatomical recurrence, which was specifically defined as the inability to advance a standard 17 French cystoscope past the previously reconstructed portion of the urethral lumen with minimal force. Although this is an objective measure, it does not consider the functional outcome (ie, urinary symptoms, quality of life) for the patient. For example, some patients noted as recurrences in this study were relatively asymptomatic and did not undergo secondary repair. Currently, the clinical significance of asymptomatic stricture is unknown, and thus so is the clinical utility of diagnosing them. A second limitation is that the degree of stricture was not graded in this study; longer and tighter strictures likely have a stronger correlation with impaired flow. Finally, a large number of men were excluded from analysis, most of whom had inadequate UF studies. Whereas this exclusion does not diminish the studies’ ability to test UF as a stand-alone measure for diagnosing recurrence, it does highlight the fact that UF can oftentimes be difficult to administer in a busy clinic in which many men arrive with empty bladders. Thus, the clinical practicality of using UF alone must be studied further.

CONCLUSION

UF is a widely used test to monitor the integrity of the reconstructed urethra after urethroplasty, but the findings from this study suggest that when used alone, the sensitivity is unacceptably low to detect recurrences. Whereas UF appears to perform better in patients under 40 years old, utilization of a standard “cutpoint” (e.g., $Q_m < 15 \text{ mL/s}$) for all patients performed poorly in this group of individuals. A refined approach will likely need to include patient-specific UF parameters that monitor $Q_m$ over time, $Q_m - Q_a$ values (which may be a novel way to numerically describe the shape of the voiding curve), and the addition of patient-reported outcomes measures. If a standard, noninvasive approach to monitoring the urethra is adopted widely, as has been proposed by many, further refinement will be required.13,14

References


EDITORIAL COMMENT

The Trauma and Urologic Reconstruction Network of Surgeons (TURNS) group conducted this study with the aim to show the ability of a noninvasive test to monitor for recurrence of stricture after urethroplasty, introducing a novel new evaluation tool $Q_{\text{maximum}}/Q_{\text{average}}$ ($Q_m/Q_a$). The authors conducted a well-designed, multi-institutional study with 323 men to compare this new tool with what in the past was considered the gold standard for testing for recurrent urethral stricture disease—cystoscopy. They included only patients who underwent very well-conducted
and validated flow studies (those who had greater than 150 mL volume) and the surgeon interpreted every flow study personally, avoiding error introduced by computer interpretation.

When I was trained in reconstructive urology, we sometimes followed patients who underwent urethroplasty with invasive tests including the retrograde urethrogram and cystoscopy. These were in addition to noninvasive tests such as the American Urological Association symptom score, and the uro-flow and postvoid residual. However, often the noninvasive components of this evaluation did not tell us whether or not patients were actually developing recurrence often necessitating unnecessary invasive monitoring that may often have resulted in negative evaluations.

In today’s world of cost-consciousness and minimally invasive testing, it is important to continually look for new and novel ways to care for our patients. This study helps in multiple ways; first, it gives the urologist a validated and easily conducted noninvasive tool (Qm-Qa) to monitor the many patients seen yearly who have undergone definitive urethroplasty; second, it may help the clinician follow urethral stricture patients that have not had urethroplasty as well. Although the goals of the study concentrate on those who underwent open urethroplasty, I am convinced the data can be extrapolated and applied to those patients who have undergone various forms of endoscopic management for urethral stricture disease as well. The big question that persists in my mind is if this tool can be applied to this population who did not undergo definitive urethroplasty (presenting a nice opportunity for a follow-up study of the Qm-Qa tool).

The conclusions are derived from the well-conducted receiver operator curve analysis where they convincingly demonstrate that the tool Qm-Qa is better than anything described prior, having an 83% sensitivity and 58% specificity. Although not perfect, this is still one of the best things that we have in reconstructive urology. A caution is that this analysis did not include combination evaluation with patient-reported outcomes such as the American Urological Association symptom score nor the urethral stricture symptom score as described by Jackson and his group. However, my guess is that combining these other evaluation tools with Qm-Qa will only increase sensitivity and specificity. Overall, this is another great contribution to our field and is a tool that all of us should incorporate into our clinical practice to maximize the evaluation of patients while minimizing invasiveness and cost.

Andrew C. Peterson, M.D., Duke University, Durham, NC

References

http://dx.doi.org/10.1016/j.urology.2015.12.071
UROLOGY 91: 201–202, 2016. © 2016 Published by Elsevier Inc.