Prostatic Diseases and Male Voiding Dysfunction

Baseline Symptom Score and Flow Rate Can Predict Failure of Medical Treatment of Lower Urinary Tract Symptoms: Prospective 12-year Follow-up Study

Said Fadel Mishriki, Omar Aboumarzouk, John T. Graham, Thomas B. Lam, and Bhaskar K. Somani

OBJECTIVE
To assess predictors of failure of medical treatment of lower urinary tract symptoms (LUTS) and evaluate long-term outcome.

METHODS
Between January 1993 and September 1994, 178 men referred with LUTS were prospectively recruited. Assessments included maximum urine flow (Qmax), postvoiding residuals (PVR), transrectal ultrasound (TRUS) prostate volumes, American Urological Association symptom score, and validated quality of life (QOL) and bother scores. Treatment failure was defined as need for transurethral resection of the prostate (TURP). Data were collected at baseline, with final follow-up at 12 years. Univariate and multivariate analyses used Kaplan-Meier and the Cox proportional hazards regression model, respectively, to assess covariates on risk of failure and independent variable prognostic values.

RESULTS
Median follow-up was 7.9 years (range, 0-12 years). The mean QOL baseline score of 7.1 improved to 3.6 at 6 years and to 3.3 at 12 years (P < .05 for all). Fifty patients (28%) underwent TURP, with 36 undergoing TURP within 3 years. By univariate analysis, Qmax, AUA and bother scores, and PVR were significantly associated with treatment failure. Independent predictors of failure by multivariate Cox regression were Qmax (>15 vs <15 mL/s; hazard ratio, 3.37; 95% confidence interval, 1.74-6.52; P < .0001) and bother score (<13 vs >13; hazard ratio, 2.37; 95% confidence interval, 1.29-4.35; P = .005). At 12 years, AUA, QOL, and bother scores statistically improved compared with baseline (13 vs 8, 10 vs 6, and 5 vs 2, respectively). Limitations included attrition bias from nonresponders.

CONCLUSION
The beneficial effect of medical treatment persisted for up to 12 years. Treatment is more likely to fail within the first 3 years in patients with low baseline Qmax and high bother scores.

Lower urinary tract symptoms (LUTS) are a major cause of impairment in health-related quality of life (QOL) and a driving factor for patients to inquire about treatment options. Improved QOL and enhanced patient satisfaction by alleviation of LUTS should have long-lasting outcomes and minimal side effects. Management options include watchful waiting, medical treatment, and surgical interventions, and the delicate balance between efficacy, long-lasting therapeutic outcomes, and minimal side effects dictates option choices. Trends in surgical interventions for benign prostatic hypertrophy (BPH) have been declining during the past 2 decades. With robust evidence demonstrating safety, efficacy, and tolerability of α-blockers and 5α-reductase (5AR) inhibitors, BPH has increasingly become a medically managed condition.

Randomized controlled trials (RCTs), such as Medical Therapy of Prostatic Symptoms (MTOPS) and Combination of Avodart and Tamsulosin (ComAT), provide invaluable information on the efficacy of medical treatments for LUTS/BPH and evaluate predictors of progressive disease. However, by their very design, they have
specific participant entry criteria and are restricted in length of follow-up, with the longest follow-up period of recent large RCTs limited to 6 years. There is paucity of data to provide evidence to practicing clinicians of what happens after this period of time.

This is a real-life practice study with 12 years of follow-up of patients and a final notes review at 17 years. The 2 objectives were:

1. To prospectively assess the long-term outcomes of medical treatment of LUTS/BPH in clinical effectiveness, patient satisfaction rate, and QOL
2. To identify potential predictors for failure of medical treatment using baseline variables

METHODS

This study was approved by the local Research Ethics Committee. Between January 1993 and September 1994, 458 men referred to a university hospital for evaluation and treatment of LUTS due to benign prostatic enlargement or hyper trophy were recruited into the study after counseling and informed consent. Management options were discussed with each patient. Men diagnosed with prostate carcinoma or presenting acutely with urinary retention were excluded.

Data were prospectively collected in this single-institution study. The routine clinic assessment included a digital rectal examination, maximum urine flow (Qmax), postvoiding residuals (PVR), and transrectal ultrasound (TRUS) assessment for prostate size. Prostate-specific antigen was not a routine test in Scotland in 1993. Patients completed the American Urological Association (AUA) symptom score, QOL, and bother questionnaires (Appendix 1 and 2, available online).

After the initial evaluation, surgical vs nonsurgical treatment options were discussed with the patients. At the outset, 280 men (61%) were treated by transurethral resection of the prostate (TURP) based on symptom score, age (<66 or >66 years), Qmax (>15 or <15 mL/s), PVR (<70 or >70 mL), prostate volume (<36 or >36 mL), AUA total score (<16 or >16), QOL total score (<7 or >7), and bother total score (<13 or >13). Treatment failure was defined as worsening LUTS despite being on dual therapy with \( \alpha \)-blockers and 5αAR inhibitors.

TURP with patient consent was offered to those with worsening LUTS despite receiving dual therapy. Covariates that were significantly associated with risk of treatment failure on univariate analysis were entered into the Cox model in a backward, stepwise fashion in which all relevant variables were entered into the model, and subsequently, nonsignificant variables were sequentially removed. Cox proportional hazards regression model assumptions (eg, effects of predictor variables and hazard ratios were constant over time) were assessed using survival curves and complementary log-minus-log curves. A value of \( P = .05 \) was considered statistically significant.

RESULTS

Of 458 men recruited and assessed, 280 underwent TURP according to their initial symptoms and flow rates. Data on the TURP group were published previously. Of the remaining patients, 178 were deemed suitable for nonsurgical management. Their baseline characteristics are reported in Table 1. Patients had a mixture of irritative and obstructive symptoms. During the subsequent 12 years, medical treatment failed in 50 men, and TURP was required (Fig. 1, Table 2). All men had been tried on an \( \alpha \)-blocker or 5αAR inhibitor, or a combination. Medical treatment was continued in 128 men. At the 12-year follow-up, 85 men of this elderly population were alive, of which 61 were receiving medical treatment and 24 had TURP. Of 61 still on medical treatment, 35 (57%) attended the 12 years of follow-up. Of the 24 patients who underwent TURP because medical treatment failed, 17 (70.8%) attended the 12 years of assessment (Fig. 1).

Statistically significant improvements in AUA, bother, and QOL scores were recorded, suggesting long-term efficacy in those patients (Table 2). This also shows a comparison between the baseline characteristics of the 2 groups that attended at 12 years (medical vs failed medical). Statistically worse baseline AUA, bother, and QOL scores were recorded in the group that required surgical intervention.

Table 3 reports predictor variables determined at recruitment that affected risk of treatment failure by

---

**Table 1. Baseline characteristics of patients**

<table>
<thead>
<tr>
<th>Baseline Characteristics</th>
<th>Medical Treatment (n = 178)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>65 ± 8.6 (47-84)</td>
</tr>
<tr>
<td>AUA/IPSS</td>
<td>14.65 ± 7.3 (4-30)</td>
</tr>
<tr>
<td>QOL score</td>
<td>6.6 ± 4.05 (4-17)</td>
</tr>
<tr>
<td>Bother score</td>
<td>12.04 ± 7.46 (2.27)</td>
</tr>
<tr>
<td>Maximum flow rate, mL/s</td>
<td>12.54 ± 4.97 (2.27-6)</td>
</tr>
<tr>
<td>Prostate size, cm³</td>
<td>31.8 ± 68.0 (19-123)</td>
</tr>
<tr>
<td>Residual volume, mL</td>
<td>30.77 ± 61.67 (0-500)</td>
</tr>
</tbody>
</table>

AUA, American Urological Association; IPSS, International Prostate Symptom Score; QOL, quality of life; SD, standard deviation.
univariate analysis. Four variables among the 7 tested demonstrated predictive value for treatment failure: $Q_{\max}$ (>15 or <15 mL/s; $P < .0001$), AUA total score (<16 or >16; $P = .002$), bother total score (<13 or >13; $P = .002$), and PVR (<70 or >70 mL; $P = .028$). The remaining 3 variables of age, prostate volume, and QOL score did not show statistical significance between baseline and final readings.
Multivariate analysis showed 2 independent predictors of treatment failure, which were Qmax (>15 vs <15 mL/s; hazard ratio, 3.37, 95% confidence interval, 1.74-6.52; \( P < .0001 \)) and the bother score (<13 vs >13; hazard ratio, 2.37, 95% confidence interval, 1.29-4.35; \( P = .005 \)). The AUA and bother scores were highly correlated (\( P = .01 \)), with the AUA score being dependant on the bother score. At 17 years, no patient continued to require the input of a secondary care physician.

**COMMENT**

**Principle Findings and Strengths of the Study**

The uniqueness of this study is the longevity of the follow-up, which is 12 years. The first objective of this study was to assess the long-term effectiveness, patient satisfaction, and QOL of patients who continued on medical treatment. At 12 years, when the patients were last objectively assessed, those who persisted with medical treatment had statistically significant improved outcome in AUA, QOL, and bother scores compared with their baseline values. During the 12-year period of this study, 50 men (28%) required surgical intervention because medical treatment failed; of these, 36 with medical treatment failures needed TURP within the first 3 years.

The second objective of this study was to identify potential predictors for medical treatment failure in the long-term, using baseline variables. When predictors for failure of medical treatment were investigated in this cohort, similar predictors of progression were found, although some appeared to be more important than others. Flow of <15 mL/s, an International Prostate Symptom Score or AUA of >16, and a bother score >13 were all highly significant factors (\( P < .001 \)), whereas a PVR of >70 mL just reached statistical significance (\( P = .028 \)). Although age, prostate volume, and QOL score were not statistically significant factors, the trends shown appeared to suggest that they approached significance.

**Strengths and Weaknesses in Relation to Other Studies**

In the 2 most recent large RCTs investigating combination therapy for BPH, most patients who failed treatment did so because of progressive BPH symptoms rather than because of urinary retention or need for surgery.\(^8,10\) Of those who “progressed” in the placebo arm of the MTOPS trial, 79% did so because of symptom progression.\(^9\) That study reported that doxazosin, finasteride, and combination therapy all resulted in significant improvements in symptom scores vs placebo; however, what happens after the 4-year trial period is not clear. Few if any other studies have supported the efficacy of medical treatment of LUTS/BPH during this period of time. In 2003, the safety and efficacy of tamsulosin over 6 years was reported.\(^13\) The Proscar Long-Term Efficacy and Safety Study noted a sustained decrease of acute urinary retention incidence and need for surgery with finasteride during a 6-year duration.\(^14\) Combination studies reveal the benefit of dual treatment with \( \alpha \)-blockers and 5AR inhibitors, but again, the evidence does not exceed 6 years.\(^15\) Cumulative incidence of the need for surgery in MTOPS was nearly 5%.\(^8\) The cumulative incidence of the need for surgery in CombAT was 7.8%, 3.5%, and 2.8% in the \( \alpha \)-blocker alone, 5AR inhibitor alone, and combination arms, respectively.\(^10\)

During the first 3 years of this study, 36 men (20.2%) required TURP. Although this figure is higher than figures reported by many other published series, certain factors may explain the apparent discrepancy. Out study was performed over an unparalleled period of follow-up. It also commenced in the early 1990s, when TURP was the most common treatment of BPH. Therapy with finasteride was launched in Scotland in 1993, which was the year this study began; as such, awareness of this medication in the early years of the study might have been limited. This study was also commenced before any evidence of the benefit of combination therapy was available.\(^8,10\) These factors may account for the high rate of surgical intervention in this study. A profound transition of medical and surgical treatment modalities has occurred during these 12 years. Less invasiveness of laser prostatectomy may facilitate surgery, whereas a 5AR inhibitor would abrogate immediate surgery for an enlarged prostate.

The 72% failure within the first 3 years would suggest that the 4- to 6-year follow-up duration undertaken in large RCTs and observational studies could capture almost all of the patients in whom medical treatment is destined to fail. After this initial period, when most failures occur because of nonresponsive symptom improvement, the risk of failure appears to be relatively low. Only 14 men required TURP in the latter 9 years of this study. Evaluating predictors of progressive disease and stratifying patients into risk groups influences treatment options and aids appropriate follow-up regimens. Current European Association of Urology guidelines suggest annual review of every patient who has been referred with LUTS/BPH, regardless of risk factors. This is not practical within all health care systems. There is no guidance about how long these patients should be monitored.\(^4\)

Observational studies and the placebo arms of RCTs provide evidence regarding predictors of progressive disease in BPH. Age, prostate volume, and prostate-specific antigen are often reported as the strongest predictors of progressive disease.\(^16\) However, most RCTs have a relatively short duration of follow-up. A reasonable conclusion from the results of our study is that those men commenced on medical treatment with significant risk factors should be followed up in secondary care for potentially 3 years. Thereafter, with subsequent low incidence of treatment failure, it would appear to be appropriate for the primary care physician to continue annual review in the community. Larger comparative studies with an equal length of follow-up may be required.
**Study Limitations**

Limitations include those patients switching various medical treatments. It is also possible that the results were biased by the selection of patients who completed follow-up, who might be more motivated than others to persist with medical therapy. Nonattendance at 12 years is likely to present attrition bias; however, these patients were calculated in the final analyses based on an intention-to-treat bias and the assumption that they did not attend because their symptoms were controlled.

Data from observational studies suggest that BPH is a progressive condition; as such, intuitively one could reasonably assume that the incidence of the need for surgical intervention would have been higher had patients not been medically treated in this study. It is theoretically conceivable that some men would have had symptomatic relief, even without medication, after 12 years.

Despite the limitations, this study represents the longest follow-up to date of patients medically treated for LUTS/BPH. Furthermore, it underscores 4 variables that can be used to predict therapy failure or success. A RCT with long follow-up is required to further evaluate the role of medical therapy for these patients.

**CONCLUSIONS**

In those patients in whom medical treatment for LUTS/BPH was successful, the objective symptomatic beneficial effect remained for up to 12 years by comparison with their baseline assessment, thereby supporting the long-term efficacy of medical treatment of LUTS/BPH. Patients with worse baseline flow rates and AUA and bother scores were more likely to fail medical treatment for LUTS/BPH and require TURP, the majority within the first 3 years of follow-up.

**References**


**SUPPLEMENTARY DATA**

Supplementary data related to this article can be found at 10.1016/j.urology.2012.08.104.

**EDITORIAL COMMENT**

The authors should be congratulated for completing a long-term prospective follow-up study (median; 7.9 years, range, 0-12 years) of 178 men who were first treated medically for lower urinary tract symptoms due to benign prostate hyperplasia. Their major findings are (1) approximately one-fourth underwent surgery (thus classified as failure), (2) two-thirds of the failures occurred within 3 years, and (3) predictors for failure included lower flow rate and higher bother scores at baseline. Although these predictors have been repeatedly reported in large-scale studies lasting for 3 to 6 years, the study is of value in confirming them in a much longer run.

Limitations, which are recognized by the authors, would be inherent to such a study, including attrition bias by nonattendance and transition of assessment and therapeutic modalities during the study period. Examples of the transition from the study commencement in 1993 to the current days are routine testing of the prostate-specific antigen level and much commoner use of anticholinergics for overactive bladder symptoms or 5α-reductase inhibitors for an enlarged prostate gland. Development of less invasive surgeries, such as laser prostatectomy or improvement in surgical skills among urologists, should have lessened the threshold to indicate surgery to men with persistent...
symptoms. Thus, their conclusion, “medical treatment beneficial effect persisted for up to 12 years,” should be conditioned.

The study is not a randomized study of fixed medical treatments sponsored by big pharmaceutical companies but an observational study of multiple choice of treatment in clinical practice. This is advantageous for us to inform patients of expected outcomes in real-life daily practice.

A concern is understandability of data by the patients. Therapeutic efficacy is usually presented in the mean or the median of the scores or score changes. The figures are compared between treatment groups or between the baseline and the last visit. A statistically significant difference is automatically translated into clinically significant efficacy. However, does a remark “the score will decrease by 4.5 points on average” hold any meaning to the patients? They should like to hear an explanation such as, “the symptoms will slightly improve in most cases and will much improve in one of two cases.” This explanation would be feasible by defining reduction of symptom score by 25% as slight improvement and a reduction by 50% as moderate improvement, for example. Data presentation primarily should be scientifically valid but secondarily readily understandable from the point view of patient education, which is a perspective in the future.

Yukio Homma, M.D., Ph.D., Department of Urology, The University of Tokyo, Tokyo, Japan

References


http://dx.doi.org/10.1016/j.urology.2012.08.108

REPLY

We are grateful for the useful and thought-provoking comments and largely concur with all the comments made, but would like to clarify a few points. Firstly, regarding conditioning the conclusion that the beneficial effect of medical treatment persisted for up to 12 years is entirely valid. However, there is merit in emphasizing the point that in the early 1990s, when medical treatment for benign prostatic obstruction and lower urinary tract symptoms can be considered suboptimal by today’s standards, the fact that so many men did not progress to require surgery is extremely striking. Consequently, with the improvement in medical treatment and patient assessment, whereby those destined to fail can be identified early on, one would intuitively expect an even lower rate of failure of medical therapy in the long run.

Secondly, we agree with the assertion that for any observed differences to be of clinical benefit, even if they are statistically significant, there is a need to correlate them to direct patient benefit and to present the benefits in a way that patients find meaningful. However, this is not always possible in manuscript submissions to journals, due to word limit restrictions.

Said Fadel Mishriki, M.D., F.R.C.S. Urol, F.E.B.U., F.R.C.S.,
Omar Aboumarzouk, F.R.C.S., John T. Graham, F.R.C.S.,
Thomas B. Lam, F.R.C.S. Urol, Department of Urology,
Aberdeen Royal Infirmary, Aberdeen, Scotland

Bhaskar K. Somani, F.R.C.S. Urol, Department of Urology,
Southampton General Hospital, Southampton, United Kingdom

http://dx.doi.org/10.1016/j.urology.2012.08.109