Original Article

Comparison of Diagnostic Accuracy of International Prostate Symptom Score vs Visual Prostate Symptoms Score in Patients with Benign Prostate Hyperplasia

Jahan Tab Qazi[1], Zeeshan Qadeer[2], Zein El Amir[3], Nauman Nabi[4], Muhammad Noman Qureshi[5], Zain Ali Sattar[6]

Contributions:
JTQ ZQ ZEA- Conception, Design
JTQ ZQ ZEA NN ZAS - Acquisition, Analysis, Interpretation
JTQ ZQ MNQ - Drafting
JTQ ZQ ZEA NN MNQ ZAS - Critical Review

All authors approved the final version to be published & agreed to be accountable for all aspects of the work.

Conflicts of Interest: None
Financial Support: None to report
Potential Competing Interests: None to report

Disclaimer: This study has not been published or submitted elsewhere for publication.

AI Disclosure: Artificial intelligence tools were used solely for language refinement and formatting. No AI tools were used for data analysis, interpretation, or result generation. All scientific content, analyses, and conclusions were developed and verified by the authors, who took full responsibility for the integrity and accuracy of the work.

Data Availability Statement: The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Institutional Review Board Approval
192/IREF/RMU/2025
17-06-2025
Rawalpindi Medical University

Review began 16/12/2025
Review ended 04/05/2026
Published 30/06/2026

How to cite this article: Qazi JT, Qadeer Z, Amir ZE, Nabi N, Qureshi MN, Sattar ZA. Comparison of Diagnostic Accuracy of International Prostate Symptom Score vs Visual Prostate Symptoms Score in Patients with Benign Prostate Hyperplasia. JRMC [Internet]. 2026 Jun. 30 [cited 2026 Jul. 1];30(2)..

https://doi.org/10.37939/jrmc.v30i2.3222

Abstract

Objective: To compare the diagnostic accuracy of the International Prostate Symptom Score (IPSS) with that of the Visual Prostate Symptom Score (VPSS). The results were compared with an ultrasonographic gold standard for diagnosing Benign Prostatic Hyperplasia (BPH) to identify the most reliable tool for clinical use.

Methods: This was a cross-sectional, validation study. We enrolled 370 male patients aged 40–75 years who presented with lower urinary tract symptoms (LUTS). The exclusion criteria were prior urological surgery, malignancy, or neurology-related diseases that affected bladder function. Participants completed the IPSS (a seven-symptom questionnaire) and the VPSS (a pictorial four-symptom tool). The gold standard for BPH diagnosis is transabdominal ultrasound. The diagnostic accuracy was calculated, and a chi-square test was used to compare the overall diagnostic accuracy of the two scoring systems.

Results: The International Prostate Symptom Score (IPSS) demonstrated a sensitivity of 91.0% and specificity of 72.2%, with positive and negative predictive values of 96.8% and 46.4%, respectively. The overall diagnostic accuracy was 89.2%. In comparison, the Visual Prostate Symptom Score (VPSS) showed higher sensitivity at 95.2% and specificity of 75.0%, with a positive predictive value of 97.2% and a negative predictive value of 62.8%. The overall diagnostic accuracy of the VPSS was 93.2%. The likelihood ratios favoured VPSS, with a positive likelihood ratio of 3.81 and a negative likelihood ratio of 0.06, compared to 3.27 and 0.12 for IPSS, indicating a comparatively improved diagnostic performance.

Conclusion: The VPSS demonstrated high sensitivity for diagnosing BPH against the ultrasonographic gold standard in this Pakistani cohort.

Keywords: Benign Prostatic Hyperplasia, Lower Urinary Tract Symptoms, International Prostate Symptom Score, Visual Prostate Symptom Score, Diagnostic Accuracy, Ultrasonography, Cross-sectional Studies.

Introduction

The prevalence of Benign Prostatic Hyperplasia (BPH) is high in men. In BPH, there is a non-cancerous growth of the prostate, which causes symptoms of the lower urinary tract (LUTS).[1] The International Prostate Symptom Score (IPSS) and the Visual Prostate Symptom Score (VPSS) are two widely used tools for assessing LUTS in BPH patients.[2] The IPSS relies on a detailed questionnaire, while the VPSS is designed for patients with lower literacy levels and uses pictorial representations for simplicity.[3] Although both tools are widely used, data comparing their diagnostic accuracy in assessing LUTS severity and predicting treatment outcomes are limited.[4,5]

A study by Wessels et al. showed that an IPSS score of > 10 had a sensitivity of 71%, specificity of 57%, PPV of 80%, NPV of 45%, and diagnostic accuracy of 67%. A VPSS score of > 8 had a sensitivity of 84%, specificity of 40%, PPV of 77%, NPV of 52%, and diagnostic accuracy of 71%.[6]

Accurate symptom assessment is crucial for effective management of BPH. The IPSS is comprehensive, but it is difficult for populations with lower literacy or cognitive impairments.

A study by Safarenejad et al. showed that the prevalence of BPH was 23.8%.[7] The VPSS is a simpler alternative. However, its efficacy in accurately diagnosing and grading LUTS compared to the IPSS is unknown.

Materials And Methods

This cross-sectional validation study was conducted at Benazir Bhutto Hospital, Rawalpindi. Ethical approval was obtained from the Ethical Review Committee of Rawalpindi Medical University (Ref number: 192/IREF/RMU/2025). After obtaining written informed consent, demographic data were collected. The variables included age, body mass index (BMI), and symptom duration. Participants completed both the IPSS and VPSS during a single clinical encounter to reduce recall bias. Then they underwent ultrasonographic assessment in the Radiology Department by experienced radiologists who were blinded to the IPSS and VPSS scores. The sample size of 370 patients was determined using a sensitivity and specificity calculator. We used a 23.8% prevalence of benign prostatic hyperplasia (BPH) and diagnostic parameters established by Wessels et al. Participants were recruited using consecutive sampling techniques. Outpatient registration logs and referral records were used to ensure sequential enrolment and minimise selection bias. This study was conducted and reported in accordance with the Standards for Reporting Diagnostic Accuracy (STARD) 2015 guidelines.

The study population comprised male patients aged 40–75 years with LUTS. To maintain data integrity, the study excluded participants with a history of prostate or bladder cancer, past urological surgeries, or neurological disorders affecting bladder function, such as Parkinson's disease or spinal cord injuries. Additionally, individuals who were unable to comprehend the questionnaires were excluded. BPH was operationally defined by prostate enlargement and a post-void residual (PVR) urine volume in mL, as measured by ultrasonography. This was performed in the supine position after voiding. Bladder volume was determined using the prolate ellipsoid formula. A residual volume exceeding 100 mL was considered clinically significant.

Data were analysed using SPSS version 22. Contingency tables were constructed for both scoring systems based on ultrasound findings. These tables were used to calculate the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and overall diagnostic accuracy. Receiver operating characteristic (ROC) curves were generated to compare the tools. Data were stratified by age, BMI, and symptom duration to evaluate the potential effect modifiers and confounding variables.

Results

The mean age of the participants was 57.6 ± 10.6 years. The mean body mass index (BMI) was 28.9 ± 6.7 kg/m². Ultrasonographic evaluation was the gold standard, which confirmed BPH (post-void residual urine volume >100 mL) in 334 patients (90.3%). 36 patients (9.7%) were negative for BPH on ultrasound.

The International Prostate Symptom Score classified 314 (84.9%) patients as BPH-positive and 56 (15.1%) as BPH-negative. A 2×2 contingency table (Table 1) shows the diagnostic performance.

Table 1: 2×2 Contingency Table – IPSS vs. Ultrasound

	BPH-Ultrasound Positive	BPH-Ultrasound Negative	Row Total
IPSS BPH-Positive	304 (TP)	10 (FP)	**314**
IPSS BPH-Negative	30 (FN)	26 (TN)	**56**
Column Total	**334**	**36**	**370**

TP = True Positive; FP = False Positive; FN = False Negative; TN = True Negative

The diagnostic performance of the IPSS was evaluated using a standard 2×2 contingency table. The IPSS correctly identified 304 true-positive and 26 true-negative cases, whereas 10 cases were falsely labelled as positive and 30 cases were missed.

The sensitivity of the test was 91.0%, indicating a high ability to correctly detect diseased individuals. The specificity was 72.2%, reflecting a moderate capacity to correctly identify healthy individuals. The positive predictive value was notably high at 96.8%, suggesting that a positive test result is highly reliable for confirming the presence of the disease. In contrast, the negative predictive value was 46.4%, indicating the limited reliability of a negative result in excluding the disease.

The overall diagnostic accuracy of the IPSS was 89.2%, demonstrating a high proportion of correctly classified cases. The positive likelihood ratio was 3.27, indicating a modest increase in the probability of disease following a positive test result. The negative likelihood ratio was 0.12, suggesting a reduction in disease probability when the test was negative, although this effect was constrained by the relatively low negative predictive value observed in this dataset.

Table 2: Diagnostic Accuracy Metrics – IPSS

Parameter	Formula	Value	Result (%)
Sensitivity	TP / (TP + FN)	304 / 334	**91.0%**
Specificity	TN / (TN + FP)	26 / 36	**72.2%**
Positive Predictive Value (PPV)	TP / (TP + FP)	304 / 314	**96.8%**
Negative Predictive Value (NPV)	TN / (TN + FN)	26 / 56	**46.4%**
Overall Diagnostic Accuracy	(TP + TN) / Total	330 / 370	**89.2%**
Area Under ROC Curve (AUC)	—	—	**0.816**

Table 3: 2×2 Contingency Table – VPSS vs. Ultrasound

	Ultrasound Positive	Ultrasound Negative	Row Total
VPSS Score Positive	318 (TP)	9 (FP)	**327**
VPSS Score Negative	16 (FN)	27 (TN)	**43**
Column Total	**334**	**36**	**370**

TP = True Positive; FP = False Positive; FN = False Negative; TN = True Negative

The diagnostic performance of the Visual Prostate Symptom Score (VPSS) was evaluated using a 2×2 contingency table. A total of 370 participants were included in this analysis. The VPSS correctly identified 318 true-positive and 27 true-negative cases, whereas nine cases were falsely classified as positive and 16 cases were missed.

The sensitivity of VPSS was 95.2%. The specificity was 75.0%. The positive predictive value was 97.2%, demonstrating that a positive VPSS result is highly reliable for confirming the presence of the disease. The negative predictive value was 62.8%, indicating a moderate ability to exclude the disease when the test result was negative.

The overall diagnostic accuracy of the VPSS was 93.2%, indicating a high proportion of correctly classified cases with VPSS. The positive likelihood ratio was 3.81, suggesting a moderate increase in the probability of disease following a positive result in the test. The negative likelihood ratio was 0.06, indicating a substantial reduction in disease probability when the VPSS result was negative, although this effect remained influenced by the underlying disease prevalence in the study population.

Table 4: Diagnostic Accuracy Metrics – VPSS

Parameter	Formula	Value	Result (%)
Sensitivity	TP / (TP + FN)	318 / 334	**95.2%**
Specificity	TN / (TN + FP)	27 / 36	**75.0%**
Positive Predictive Value (PPV)	TP / (TP + FP)	318 / 327	**97.2%**
Negative Predictive Value (NPV)	TN / (TN + FN)	27 / 43	**62.8%**
Overall Diagnostic Accuracy	(TP + TN) / Total	345 / 370	**93.2%**
Area Under ROC Curve (AUC)	—	—	**0.850**

McNemar's test was applied to determine whether the difference in overall diagnostic accuracy between the two scoring systems was statistically significant. McNemar's test revealed no statistically significant difference in overall diagnostic accuracy between the IPSS (89.2%) and VPSS (93.0%) in this paired sample ($\chi^2 = 2.91$, $p = 0.087$), although the absolute accuracy difference of 3.8 percentage points consistently favoured VPSS.

Receiver operating characteristic (ROC) curves were generated for both diagnostic tools against ultrasonographic confirmation of BPH. The AUC for IPSS was 0.816 and for VPSS was 0.850, indicating good to excellent discriminatory ability for both scores. The VPSS demonstrated a higher AUC, consistent with its superior specificity at the operating cutoff point. The individual and comparative ROC curves are shown in Figure 1.

Discussion

Benign prostatic hyperplasia is one of the most prevalent urological conditions in aging men. A comprehensive Global Burden of Disease index reported approximately 112.5 million prevalent cases globally in 2021, with the greatest burden concentrated in low- and middle-income (LMIC) regions such as Pakistan.[8] The 90.3% rate of BPH on ultrasound indicates a high-burden tertiary urology referral in our center.

The IPSS is the gold standard for testing BPH-related symptoms. It is a self-administered questionnaire for male LUTS. The WHO also endorsed this. Guidelines recommend the IPSS as a mandatory component of the initial evaluation for all patients with bothersome LUTS.[9] In our study, the IPSS demonstrated high sensitivity (91.0%) for

identifying ultrasonographically confirmed BPH. It had a relatively low specificity (72.2%) and an AUC of 0.816. This pattern is consistent with the prior validation literature.

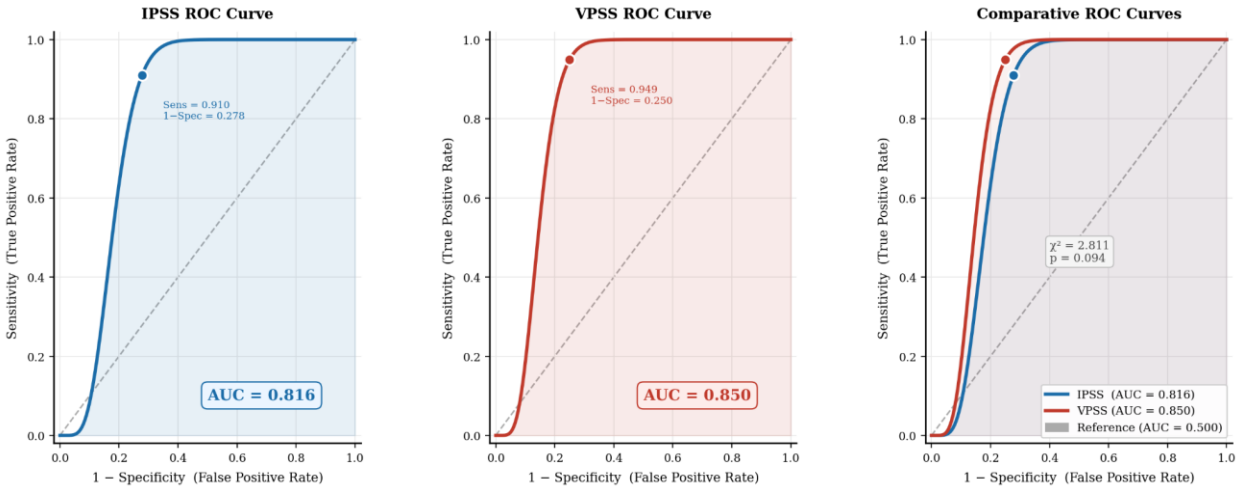

Figure 1: Left – IPSS ROC curve (AUC = 0.816). Centre – VPSS ROC curve (AUC = 0.850). Right – Comparative ROC curves ($x^2 = 2.811$, $p = 0.094$). Filled circles indicate operating points at respective cut-offs. Dashed diagonal = reference line of no discrimination (AUC = 0.500)

A major limitation of the IPSS was noted in a study conducted at Memorial Sloan Kettering Cancer Center. They concluded that only 49% of patients understood their IPSS responses. The IPSS is subject to language barriers and requires patients to be educated to complete the form.[10] This is worse in South Asian settings, as our literacy levels are low. A 2022 Turkish study by Guzelsoy et al. compared the IPSS, VPSS, and ICIQ-MLUTS. They confirmed that the IPSS had the highest completion rate (89.3%) across educational groups in their cohort but warned that incomplete or wrong answers might skew clinical categories, especially for less-educated patients.[11] Another study reported that using the ICIQ-MLUTS alongside the IPSS showed that 32.6% of men had at least one significant symptom not captured by the IPSS. Therefore, if used in isolation for diagnosis, its utility is restricted. [12]

The VPSS is a pictogram-based assessment tool designed to overcome the literacy and language barriers posed by the IPSS. Chatterjee et al. found a link between VPSS and IPSS scores and observed that 83% of patients filled out the VPSS independently versus only 40% for the IPSS.[13] A Portuguese comparison showed that 89% of patients completed the VPSS unaided, compared to 59% for the IPSS. They also mentioned that the VPSS total scores correlated significantly with the IPSS scores ($\rho = 0.651$, $p < 0.001$) and maximum urinary flow rate (Qmax) ($\rho = -0.228$, $p = 0.023$). The VPSS also requires significantly less time to complete than the IPSS.[14]

The EAU Guidelines on Non-Neurogenic Male LUTS,[15,] state that the VPSS is a validated alternative to the IPSS. They state that "this visual score can be used as an option in men with a limited educational background.[16] This is an important policy for the use of VPSS in resource-limited settings such as Pakistan. In our country, older patient demographics, low education, and high LUTS make VPSS a better alternative. A Dutch study using a qualitative approach found that patients often misunderstood several items on the LUTS questionnaire (IPSS). The VPSS questionnaire improves this by using simple pictures (pictograms) to represent the symptom frequency. [17]

The high prevalence of BPH confirmed on ultrasound (90.3%) in our study reflects the tertiary referral nature of the study site and likely does not represent the community's prevalence. This exaggerates the positive predictive values for both tools (IPSS PPV 96.8%, VPSS PPV 97.2%) and should be considered. Transabdominal ultrasound measurement of PVR using the prolate ellipsoid formula is non-invasive and widely available in tertiary hospitals in Pakistan. It has been validated as a good indicator of BPH.[18]

There is a 100% increase in incidence and years lived with disability (YLD) over three decades in resource-limited settings.[19] A study of LMICs noted that literacy barriers to IPSS completion represent an unmet need in low-resource urological practice and endorsed the VPSS as a practical alternative.[20] Stothers et al. confirmed a positive IPSS-VPSS correlation ($r = 0.70$) even in a population where 82% of participants had four or fewer years of formal education, further validating the VPSS.[21]

This study has several limitations. First, we used binary (positive/negative) outcomes for both the IPSS and VPSS. By treating them as continuous scores against continuous PVR measurements, an accurate representation can be

obtained. Second, the cross-sectional design did not allow for causal inference. Third, the high BPH prevalence in our cohort (90.3%) inflated the PPV estimate and limited the generalizability of the study to primary care.

Conclusions

The VPSS demonstrated high sensitivity for diagnosing BPH against an ultrasonographic gold standard in a Pakistani cohort.

Author Information

1. Post Graduate Trainee, Benazir Bhutto Hospital, Rawalpindi 2. Associate Professor, Benazir Bhutto Hospital, Rawalpindi 3. Professor, Benazir Bhutto Hospital, Rawalpindi 4. Consultant, University Hospital Limerick, Ireland 5. Registrar, Ashford and St Peter's Hospitals NHS Foundation Trust, UK 6. Registrar, Letterkenny University Hospital, Ireland

Corresponding author: Dr. Jahan Tab Qazi jahantab@msn.com

References

1. Wessels SG, Heyns CF. Prospective evaluation of a new visual prostate symptom score, the international prostate symptom score, and uroflowmetry in men with urethral stricture disease. Urology. 2014 Jan;83(1):220-4. https://doi.org/10.1016/j.urology.2013.08.085
2. Plochocki A, King B. Medical treatment of benign prostatic hyperplasia. Urol Clin North Am. 2022 May;49(2):231-8. https://doi.org/10.1016/j.ucl.2021.12.004
3. Ölçücü MT, Aydın ME, Avcı S, Koraş Ö, Eren AE, Yılmaz K. Comparison of a visual prostate symptom score and international prostate symptom score: a prospective multicenter study and literature review. Urology. 2020 Dec;146:230-5. https://doi.org/10.1016/j.urology.2020.08.026
4. Patel SR, Lowrey KA, Johny A, Smith CP. Improvement in international prostate symptom score after prostatic urethral lifts is dependent on prostatic volume. Can J Urol. 2023 Aug;30(4):11619-23. https://pubmed.ncbi.nlm.nih.gov/37633290/
5. Pool-Goudzwaard AL, Vredeveld T. Clinimetrics: the international prostate symptom score. J Physiother. 2024 Jan;70(1):69. https://doi.org/10.1016/j.jphys.2023.11.007
6. Yao MW, Green JSA. How international is the international prostate symptom score? A literature review of validated translations of the IPSS, the most widely used self-administered patient questionnaire for male lower urinary tract symptoms. Low Urin Tract Symptoms. 2022 Mar;14(2):92-101. https://doi.org/10.1111/luts.12412
7. Safarinejad MR. Prevalence of benign prostatic hyperplasia in a population-based study in Iranian men 40 years old or older. Int Urol Nephrol. 2008 Dec;40(4):921-31. https://doi.org/10.1007/s11255-008-9354-4
8. Chen X, Yang S, He Z, et al. Comprehensive analysis of the global, regional, and national burden of benign prostatic hyperplasia from 1990 to 2021. Sci Rep. 2025 Feb;15(1):5644. https://doi.org/10.1038/s41598-025-90229-3
9. Sandhu JS, Bixler BR, Dahm P, et al. Management of lower urinary tract symptoms attributed to benign prostatic hyperplasia (BPH): AUA guideline amendment 2023. J Urol. 2024 Jan;211(1):11-19. https://doi.org/10.1097/JU.0000000000003698
10. Zelefsky MJ, Poon BY, Pei X, et al. Assessment of patients with prostate cancer and their understanding of the international prostate symptom score questionnaire. Adv Radiat Oncol. 2023 Jul-Aug;8(4):101213. https://doi.org/10.1016/j.adro.2023.101213
11. Guzelsoy M, Erkan A, Ozturk M, et al. Comparison of three questionnaire forms used in the diagnosis of lower urinary tract symptoms: a prospective study. Prostate Int. 2022 Dec;10(4):218-23. https://doi.org/10.1016/j.prnil.2022.06.001
12. Guzelsoy M, Erkan A. IPSS May Not Tell the Whole Story: Integrating ICIQ-MLUTS for Comprehensive LUTS Assessment. Neurourol Urodyn. 2026 Jan;45(1):[Missing metadata]. https://doi.org/10.1002/nau.70156
13. Chatterjee S, Kumar A, Pal DK. Study of correlation between visual prostate symptom score and international prostate symptom score in men with lower urinary tract symptoms with reference to uroflowmetry parameters in Indian population. Urologia. 2023 May;90(2):377-80. https://doi.org/10.1177/03915603231157573

14. Meller A, Campagnari JC, Alcantara D, et al. A prospective comparison of visual prostate symptom score and international prostate symptom score in the evaluation of men with benign prostate hyperplasia. Urol Sci. 2024 Mar;35(1):51-6. https://doi.org/10.1097/us9.0000000000000003
15. Baboudjian M, Hashim H, Bhatt N, Creta M, De Nunzio C, Gacci M, et al. Summary paper on underactive bladder from the European Association of Urology guidelines on non-neurogenic male lower urinary tract symptoms. Eur Urol. 2024 Sep;86(3):213-20. https://doi.org/10.1016/j.eururo.2024.04.004
16. Gravas S, Cornu JN, Gacci M, et al. EAU guidelines on management of non-neurogenic male lower urinary tract symptoms (LUTS), including benign prostatic obstruction (BPO). Arnhem: European Association of Urology; 2022.
17. Schlatmann FW, van Balken MR, de Winter AF, de Jong IJ, Jansen CJ. How do patients understand questions about lower urinary tract symptoms? A qualitative study of problems in completing urological questionnaires. Int J Environ Res Public Health. 2022 Jul 29;19(15):9650. https://doi.org/10.3390/ijerph19159650
18. Ballstaedt L, Leslie SW, Woodbury B. Bladder post void residual volume. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2024 [updated 2024 Feb 28; cited 2026 Jun 29]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK539839/
19. Xu XF, Liu GX, Guo YS, et al. Global, regional, and national incidence and year lived with disability for benign prostatic hyperplasia from 1990 to 2019. Am J Mens Health. 2021 Jul-Aug;15(4):15579883211036786. https://doi.org/10.1177/15579883211036786
20. Cassell A, Sine B, et al. Burden of Benign Prostatic Hyperplasia (BPH) in Low- and Middle-Income Countries in Sub-Saharan Africa (SSA). Soc Int Urol J. 2024 May;5(5):51. https://doi.org/10.3390/siuj5050051
21. Stothers L, Macnab A, Bajunirwe F, Mutabazi S, Lobatt C. Comprehension and construct validity of the visual prostate symptom score (VPSS) by men with obstructive lower urinary tract symptoms in rural Africa. Can Urol Assoc J. 2017 Nov-Dec;11(11):E405-8. https://doi.org/10.5489/cuaj.4589