We thank the reviewers for their helpful feedback and comments. Please find our responses below.

Reviewer 1:

1. “It is important to clarify that the study refers only to children with a non-neurogenic voiding disorder. I would like to suggest adding this information to title and objective.”

Response: We agree with the reviewer that this information is important. However, because it is clearly stated in the study aims and the Methods section, we prefer to leave it out of the title, which is already considerably complex.

1. “Although results were stratified by gender, they were not stratified by age, and it varied from 4 to 18 years of age.

Response: As suggested by reviewers 1 and 3, we analyzed the data stratified by age (using a cutoff point of 10 years). The analysis is presented in the Results section and in the tables of the revised manuscript. We found no differences between the two age groups in children who managed to urinate in either of the two tests. Also, there were no differences in tests’ sensitivity between the two age groups.

The results shown that 41 (39%) children had vesicoureteral reflux, 9 (9%) underwent excision of a posterior urethral valve and 11 (10%) children underwent ureteral reimplantation surgery. What about the other 43 children?

Response: These numbers were mentioned in the context of the participants' medical history. However, not all children had been diagnosed before the study; some children presented with symptoms and were tested here for diagnostic purposes. We agree that the presentation of medical history and symptoms in the first part of the Results section was unclear, and we have revised it.

1. “However, the voiding curves of the UDS and non-invasive EMG uroflow tests were classified and compared by “three pediatric urologists highly experienced with interpreting such studies”. It is not clear if all three revised all exams and how did you decided which one to use, if a different interpretation raised among them.”

Response: Not all three pediatric urologists were present during the tests. However, all three attended routine weekly departmental meetings to discuss test results and reach a consensus on diagnosis, classification, and continued follow-up and treatment. We have added this information to the relevant part of the Method section.

1. “Also, it was interesting to learn that, in children with normal micturition curves, the results were the same in both tests. If the catheter was the only villain, one would expect a different conclusion. Is it possible that other factors, beside the catheter, are influencing the results?”

Response: The fact that children with normal, bell-shaped micturition curves appeared to be less affected by the presence of the catheter is discussed in the Discussion section. We believe the catheter has a critical effect on the quality of the test, although it appears that having a normal micturition curve can compensate for the catheter’s detrimental effect.

In children with an abnormal baseline micturition curve, the catheter exacerbates the existing problem and interferes with pelvic muscle relaxation, creating noticeable, unfavorable changes in the micturition curves. The only difference between the PF and the non-invasive tests is the presence of the catheter in the urethra. Therefore, even if other factors contribute to the difference in the results between the two tests, they are unlikely to be independent. In addition, we tested the hypothesis that age and gender affect the examination in the presence of a catheter, and therefore, we stratified the data according to these variables.

1. “Although the conclusion gives us good information, it does not elucidate if the authors felt confident, based on the results, to say that EMG uroflowmetry utilizing a urethral catheter at the pressure flow phase IS, or IS NOT, a reliable test in children? Do you recommend EMG uroflowmetry WITHOUT a urethral catheter instead, as a more reliable test in children with non-neurogenic bladder?”

Response: Our opening question is indeed provocative; therefore, we exercise caution and do not categorically recommend a complete revision of operating procedures and protocols. However, we conclude that the PF test with a catheter is less accurate, and any pathological result must be further confirmed by a non-invasive micturition test.

Reviewer 2:

1. “I recommend the authors to examine the study by Fugaru et al. and discuss it with their own findings.”

Response: We agree with the reviewer that the study by Fugaru et al., although smaller than the current study, presents important findings that can complement our own findings. The paragraph below has been added to the Discussion section.

As mentioned above, studies that validate the PF test in children and compare it to uroflowmetry without a catheter are largely missing from the clinical literature. One exception is a recent study by Fugaru et al. (2023), which examined the effect of urethral catheterization in PF compared to uroflowmetry without a catheter in a group of 46 children. The study did not exclude children with spinal issues, which could indicate neurogenic voiding disorders, but excluded children who could not void spontaneously with the catheter in place. Fugaru et al. found reduced Qmax and flow index during PF compared to uroflowmetry without a catheter. These findings are congruent with studies in adults and with our current study.

1. “I would like to see a comparison between the two studies, not only of voiding pattern and EMG activity, but also of flow rates, voiding times, voiding volumes, PVR.”

Response: We have performed these comparisons and added them to the Results section.

Comparison of additional parameters between the PF test and the non-invasive EMG uroflowmetry test found significant differences in voided volume (209 vs. 165 ml, respectively, p-value 0.02), Qmax (11.9 vs. 16.6 ml/second, respectively, p-value 0.01) and flow time (32.4 vs. 17.9 seconds, respectively, p-value 0.03). In contrast, residual volume was similar in the two tests (33 vs. 31 ml, respectively, p-value 0.6).

1. “It was stated that 9 of the children included in the study had PUV. This is a very different pathology compared to others and makes the study group heterogeneous. I recommend removing the PUV patients and reanalyzing.”

Response: The heterogeneity of the study group is mentioned in the Discussion section. We consider heterogeneity a strength of the study because the inclusion of a diverse population presenting with different urological disorders (apart from the excluded neurological urinary disorders) provides a more reliable validation of the PF test for different pathologies. It is important to note that analysis was performed on the differences between the two tests for each subject. Therefore, even if PUV represents a unique pathology, the results from children with PUV are unlikely to affect the group analysis. Furthermore, the inclusion of this group provides useful information about the differences between the tests in these patients.

1. “The authors used the chi-square test to compare categorical values between two tests. However, it is understood that these two groups are dependent groups. Chi-square test is used to compare categorical values between independent groups. Mc-Nemar test is recommended for comparing categorical values between two dependent groups. I recommend to the authors reconsider the statistical analysis.”

Response: We thank the reviewer for pointing out this error. The statistical test used to compare the categorical values between the dependent groups was indeed McNemar’s test. The ‘chi-square test’ mistakenly mentioned for this analysis in the Methods section has now been corrected.

Response: We would like to note that the data analysis was supported by professional statistical consulting. A follow-up consultation to clarify the issue raised by the reviewer and further analysis confirmed that the test used for the categorical dependent variables was indeed McNemar’s test. The mistake originated from the chi-square distribution of the data. The error has been corrected.

Reviewer 3:

1. “Although I found a wide range of samples in the study in regards to underlying conditions, I could not cover if there were neurogenic lower urinary tract dysfunction cases. It would be very interesting to compare neurogenic to non-neurogenic cases, of course the ones who can still urinate.”

Response: Only children with non-neurogenic voiding disorders were included in this study. We agree that it would be interesting to compare neurogenic to non-neurogenic cases, but this was beyond the scope of the current study and should be investigated in the future.

1. “The other interesting point is the ages of the patients since 8 Fr double lumen seems disturbing to children younger than 10 years old during urination. I would be very surprised if the age would not be found a confounding factor for a disturbed voiding.”

Response: As suggested by reviewers 1 and 3, we analyzed the data stratified by age (using a cutoff point of 10 years). The analysis is presented in the Results section and in the tables of the revised manuscript. We found no differences between the two age groups in children who managed to urinate in either of the two tests. Also, there were no differences in tests’ sensitivity between the two age groups.

1. “Lastly, the pressure measurements during the start of the voiding and at the maximum flow during disturbed voiding compared to free flow patterns would be very interesting to study.”

Response: We did not analyze bladder pressures because they can only be measured in the PF test and not in the uroflow test and, therefore, cannot be compared between the two tests. Bladder pressure is a unique and important value provided by the PF test but not by the uroflow test.

1. “I noticed that almost no patients had high amounts of residual urine. Would it cause a bias observing only clinically normal n-aggressive cases?”

Response: Children with high amounts of residual urine were excluded from the study because, according to ICCS, the PF and uroflow tests are reliable only when the voided volume exceeds 50% of the expected bladder capacity. On balance, the increased reliability of our analysis was more important than the loss of information caused by the exclusion of these children from the study.

1. “Besides I would like to see if any abnormal voider would turn into a normal voider during pressure flows? This would open the dilemma of the gold standard validity of free uroflowmetrics.”

Response: According to our data, no children had a normal bell-shaped micturition curve in the PF test if they had an abnormal one in the non-invasive uroflow test. Thus, the choice of the non-invasive test as the “gold standard” is justified.