Comparison of 2D and 3D ultrasound methods to measure serial bladder volumes during filling: Steps toward development of non-invasive ultrasound urodynamics

Anna S. Nagle1, Rachel J. Bernardo2, Jary Varghese3, Laura R. Carucci3, Adam P. Klausner4, and John E. Speich1,*

1Department of Mechanical and Nuclear Engineering, Virginia Commonwealth University, Richmond, VA, USA
2Department of Biomedical Engineering, Virginia Commonwealth University, Richmond, VA, USA
3Department of Radiology, Virginia Commonwealth University Medical Center, Richmond, VA, USA
4Department of Surgery/Division of Urology, Virginia Commonwealth University Medical Center, Richmond, VA, USA

Abstract

OBJECTIVES—Non-invasive methods to objectively characterize overactive bladder (OAB) and other forms of voiding dysfunction using real-time ultrasound are currently under development but require accurate and precise serial measurements of bladder volumes during filling. This study’s objective was to determine the most accurate and precise ultrasound-based method of quantifying serial bladder volumes during urodynamics (UD).

METHODS—Twelve female participants with OAB completed an extended UD procedure with the addition of serial bladder ultrasound images captured once per minute. Bladder volume was measured using three ultrasound methods: (1) V_{spheroid}: two-dimensional (2D) method calculated assuming spheroid geometry; (2) V_{bih}: 2D correction method obtained by multiplying V_{spheroid} by a previously derived correction factor of 1.375; and (3) V_{3D}: three-dimensional (3D) method obtained by manually tracing the bladder outline in six planes automatically reconstructed into a solid rendered volume. These volumes were compared to a control (V_{control}) obtained by adding UD infused volume and the volume of estimated urine production.

RESULTS—Based on linear regression analysis, both V_{bih} and V_{3D} were fairly accurate estimators of V_{control}, but V_{3D} was more precise. V_{spheroid} significantly underestimated V_{control}.

CONCLUSIONS—Although the V_{bih} and V_{3D} methods were more accurate than the more-commonly used V_{spheroid} method for measuring bladder volumes during UD, the V_{3D} method was the most precise and could best account for non-uniform bladder geometries. Therefore, the V_{3D}...
method may represent the best tool required for the continued development of non-invasive methods to diagnose OAB and other forms of voiding dysfunction.

**Keywords**

overactive bladder; transabdominal ultrasound imaging; urodynamics; volume calculations; volumetric ultrasound

---

**INTRODUCTION**

The ability to accurately and precisely measure the volume of the urinary bladder is an important tool in evaluating bladder function. For example, post-void residual (PVR) volume is utilized in the diagnosis of urinary retention, detrusor underactivity, and other bladder impairments [1–4]. In addition, measurements of both voided volume and PVR are needed for bladder training and other bladder assessments [5,6]. During urodynamics (UD), bladder volume during filling is assumed to be equal to the infused volume and does not typically include an estimation of urinary diuresis. The additional volume of urinary diuresis may be relatively small at a super-physiological fill rates which can be up to 100 ml/min [7]. However, considering the average diuresis rate of 10 ml/min in our accelerated hydration studies [8], it cannot always be considered negligible and may lead to inaccurate UD results.

Another important drawback to UD is its invasiveness which often causes anxiety and discomfort [9], has a risk of urinary tract infection [10], and can cause changes in the perception of bladder sensation [11]. The development of bladder sensation can be studied non-invasively in oral hydration studies that demonstrated significant differences between individuals with OAB and asymptomatic volunteers [8,12,13]. These studies mainly tracked volume by measuring voided volume at the end of a study and assuming a constant fill rate to linearly interpolate volume at specific time periods. However, the actual fill rate is unlikely to be constant [13], so a method to accurately and precisely measure serial, real-time bladder volumes in a non-invasive manner is essential.

Currently, techniques are under development where the addition of ultrasound during UD enables measurement of several new biomechanical properties of the bladder including wall tension, wall strain, wall stress, and dynamic elasticity [14]. Shearwave elastography and ultrasound lamb wave vibrometry have recently been used to estimate intravesical pressure [15,16], adding to the potential uses of non-invasive, ultrasound-based urodynamic methods. These novel metrics have the potential to improve the diagnosis and objective characterization of filling phase disorders, including OAB. More importantly, the use of serial ultrasound measurements during bladder filling may ultimately lead to the development of completely non-invasive “ultrasound urodynamics”. However, development of any non-invasive ultrasound-based UD methods requires utilization of the most accurate and precise measurement of serial bladder volumes during filling.

Therefore, a critical research objective is the development of non-invasive methods to more accurately measure serial bladder volumes during filling. There is a long history of using two-dimensional (2D) ultrasound imaging to estimate bladder volume [17,18] and more recently, three-dimensional (3D) methods have become available [19–21]. Thus, the
objective of this study was to compare three different ultrasound-based methods of calculating bladder volume to identify the most accurate and precise methodology and compare results to standard volume measurements obtained during UD.

METHODS

Experimental protocol

Women with bothersome symptoms of OAB were invited to participate in an Institutional Review Board approved prospective study using comparative-fill urodynamics [22,23] with ultrasound. Women clinically indicated for UD with high chronic urgency scoring 3–4 (“most of the time” or “all of the time”) on the International Consultation on Incontinence questionnaire (ICIQ)-OAB question 5 “do you have to rush to the toilet to urinate?” were included in this study. After giving informed consent, urodynamic filling was administered with a 7 Fr catheter on a Laborie Aquarius TT system (Toronto, Canada). During an initial UD fill, bladder cystometric capacity was defined as the volume at which the participant reported 100% bladder sensation using a sensation meter [8]. During a subsequent fill at an infusion rate of 10% capacity per minute, transabdominal images of the bladder were obtained every 60 s using a GE Voluson E8 system (Madison, WI) with a 3D convex 4–8.5 MHz transducer. All images were obtained by a trained ultrasound technologist with supervision from an attending radiologist fellowship trained in abdominal imaging. Infusion was paused for 5–10 min at 40% and 70% of bladder capacity and infusion was stopped when the participant reported that they had reached 100% sensation. The pauses were part of a separate study to quantify any low amplitude rhythmic contractions during the bladder filling phase [24].

At the end of the fill, participants voided and any PVR was extracted by syringe aspiration through the filling catheter. Ultrasound was used to confirm that the bladder was empty. The total bladder volume at the end of the fill was calculated as the sum of the voided volume and the PVR. Any positive difference between the total volume (void + PVR) and the infused volume ($V_{H2O}$) was considered to be due to urine production from the kidneys during the procedure. This urinary diuresis was assumed to have a constant rate. The control volume ($V_{control}$) at the moment at which each image was obtained was calculated as the instantaneous $V_{H2O}$ plus the proportion of urinary diuresis up to that time point in the fill, as previously calculated by Byun et al. [25].

Image calculations

Ultrasound images were exported and analyzed offline using GE’s 4D View software (Version 14, GE Healthcare) by a trained individual who was blinded to the UD results. As previously described in Nagle et al. [8], 3D images were manually traced in six planes 30° apart using the virtual organ computer-aided analysis (VOCAL) software which automatically combined these cross-sections into a continuous rendered volume (Fig. 1). The 30° step size was chosen for increased speed of analysis and was found to yield similar volume measurements as compared to smaller step sizes in our initial analyses and by others [20,26,27]. The resulting computed volume ($V_{3D}$) was recorded from the software’s display.
Two different 2D-image based calculations were used. Both utilized the height ($H$) in the transverse direction, the width ($W$) in the transverse direction, and the length ($L$) in the sagittal direction (Fig. 2). The first method assumed the bladder had a spheroid geometry and was calculated as

$$V_{\text{spheroid}} = \frac{\pi}{6} (W \ast H \ast D) \quad (\text{Eqn. 1})$$

The second 2D method was based on research by Bih et al. [28] that calculated a correction factor for $V_{\text{spheroid}}$ using linear regression on bladder images of healthy and spinal cord injured men and women as

$$V_{\text{bih}} = 0.72 \ast W \ast H \ast D = 1.375 \ast V_{\text{spheroid}} \quad (\text{Eqn. 2})$$

Volume comparisons

At each time point in which an image was obtained, $V_{\text{H2O}}$, $V_{\text{3D}}$, $V_{\text{spheroid}}$, and $V_{\text{bih}}$ were plotted as a function of $V_{\text{control}}$ and a linear regression line was fit to the data from each volume calculation method (Fig. 3). The percent root-mean-squared (%RMS) error for each patient was calculated so that the average error among all of the patients could be calculated without being skewed by the individuals with more images (Fig. 4A). Finally, to measure error as a function of bladder size, all control volumes were grouped into increasing bins of 100 ml until data from less than half of the participants were still available (1–100, 101–200, 201–300, 301–400, 401–500, and 501–600) and %RMS error was calculated for all bladder volume data points within each interval (Fig. 4B). The errors from the ultrasound methods were compared to the baseline error of $V_{\text{H2O}}$ by multiple comparison ANOVA. All calculations were performed in MATLAB (2017a, MathWorks, Natick, MA).

RESULTS

A total of 14 women participated in the study. Data from two individuals could not be used due to incomplete data. Mean, standard error of the mean (SEM), and range of demographic and experimental information on the remaining 12 women is presented in Table 1. In the table, the ICIq total score is the total of the four symptom questions and can range from 0 to 16. The initial capacity is the voided volume plus extracted PVR from the initial cystometric fill. The capacity is the voided volume plus extracted PVR from the fill during which ultrasound imaging was performed.

In Figures 3–5, data is shown with $V_{\text{H2O}}$ in cyan stars, $V_{\text{3D}}$ in green pentagons, $V_{\text{bih}}$ in red squares, and $V_{\text{spheroid}}$ in blue diamonds. Figure 3 shows the data points and linear regression lines with the line of ideal case (45° line with slope of 1 and $R^2$ of 1) in black. The slopes ($m$) and R-squared ($R^2$) values of the regression lines are shown in the legend. The 95% confidence bounds of the regression line’s slope were 0.981 to 1.005 for $V_{\text{H2O}}$, 0.905 to 0.983 for $V_{\text{3D}}$, 0.965 to 1.084 for $V_{\text{bih}}$, and 0.702 to 0.789 for $V_{\text{spheroid}}$. $V_{\text{spheroid}}$ was considered to be significantly different than $V_{\text{H2O}}$ because their 95% confidence intervals...
had no overlap while \( V_{3D} \) and \( V_{bih} \) were considered to not be significantly less accurate than \( V_{H2O} \) because there was overlap. Of the image-based methods, \( V_{3D} \) was considered the most precise because its \( R^2 \) value was the closest to unity. Figure 4A shows the %RMS error averaged for each method with data from all participants weighted equally. Again, \( V_{spheroid} \) was considered to have significantly higher error based on multiple comparison ANOVA comparing image based volumes to \( V_{H2O} \) as denoted by the star in the figure. Figure 4B shows %RMS error as a function of bladder volume in 100 ml increments. At the smallest volume increment, \( V_{3D} \) and \( V_{bih} \) had significantly more error than \( V_{H2O} \), and at all other volume increments, \( V_{bih} \) and \( V_{spheroid} \) had significantly more error than \( V_{H2O} \). Volume increments are only included up to 600 ml because less than half of the participants had bladder capacities higher than 600 ml.

Two example patterns of bladder filling are shown in Figure 5. In the example in Figure 5A and 5C, \( V_{3D} \) and \( V_{bih} \) were generally similar, which was the most common pattern. In the example in Figure 5B and 5D, \( V_{3D} \) was more accurate, and \( V_{bih} \) overestimated bladder volume. Note that the shape of the bladder in Figure 5D was less regular and extended beyond the image boundaries making it difficult to determine the actual location of the bladder walls at large volumes.

**DISCUSSION**

Of the three ultrasound methods studied, \( V_{3D} \) had the highest degree of precision based on the \( R^2 \) value and had the highest accuracy based on %RMS error. However, \( V_{bih} \) had the highest overall accuracy based on the slope of its linear regression. Considering that both methods had slopes and \( R^2 \) values close to unity and were not significantly different than the slope of \( V_{H2O} \), both methods could be considered acceptable ways measure bladder volume for many applications. However, when bladder volume is needed for fine calculations, such as to determine rate of physiologic bladder filling in hydration studies, the increased precision of \( V_{3D} \) will be beneficial. In contrast, \( V_{spheroid} \) consistently underestimated bladder volume, demonstrating that the bladder should not be assumed to be a sphere or spheroid when calculating volume.

Individual differences in bladder size and shape contribute to \( V_{bih} \) having a lower precision than \( V_{3D} \). An advantage of the 3D method is the ability to make no assumptions about the bladder shape as is done in the 2D methods. Additionally, by examining the bladder borders in several planes simultaneously, the ultrasound technologist was able to better estimate where the bladder border must be located even when it was not easily visible or extended beyond the image borders. The main disadvantage of the 3D method is that it is more time consuming, requiring 3–6 min per image to complete a volume tracing as opposed to only 10–20 s to define the three diameters needed to calculate the 2D volumes. The time to obtain the 3D volume could be lowered by using the available automatic and semi-automatic options in 4D View such as SonoAVC as used by Sætherhaug et al. [21] to measure fetal bladder volume or by developing custom software to automate the process. The only advantage the spheroid method had was that it was built into the ultrasound system’s volume measurements making it the fastest volume estimation to obtain during clinical studies, but
this number could easily be multiplied by a correction factor of 1.375 to obtain $V_{\text{bih}}$ yielding a more accurate measure of volume.

The highest errors in $V_{3D}$ and $V_{\text{bih}}$ were seen at the smallest bladder volumes (0–100 ml). There are two likely reasons for this. First, the bladder walls were more difficult to discern at very small volumes. Second, the difference between actual urinary diuresis rate and the constant rate assumption of filling is likely to be exaggerated in these small volumes. Interestingly, $V_{\text{spheroid}}$ had little variation in error with bladder size.

$V_{\text{control}}$ was calculated assuming that the urinary diuresis was constant. However in hydration studies, the diuresis rate was seen to vary depending on the quantity of fluid consumed and the duration of time after fluid was consumed [8,13]. Thus $V_{\text{control}}$ is not a perfect reflection of the instantaneous bladder volume. In seven (50%) participants, the total volume was more than 15 ml larger than the volume infused, indicating the need to compute the control volume rather than assuming $V_{\text{H2O}}$ accurately reflected instantaneous bladder volume. Surprisingly, in two (16.7%) participants, the final voided volume plus PVR was more than 15 ml less than the volume infused. This may have been due to unusual bladder shapes preventing complete emptying of the bladder or the UD system may have needed recalibration. In these individuals, the control volume was considered to be the instantaneous infused volume as this was the best estimate possible.

The protocol used a fill rate of 10% initial bladder capacity per minute and took one ultrasound image each minute, which would ideally result in ten images from each individual. However, in this varied population of overactive bladder patients, some sensed that they reached their bladder capacity at a substantially different volumes during the fill analyzed in this study. Some felt they had reached capacity as early as 30% of their initial capacity while others reached 130% of their initial capacity. As a result, some participants were imaged at more bladder volumes than others.

While there is a wealth of literature on the use of ultrasound to measure bladder volume [17–21,28–30], this study is unique in that it specifically focused on women with overactive bladder symptoms. Additionally, this study compared volumes to a control volume that included both urodynamic infusion and urinary diuretic filling while most other studies have relied only on the infusion volume or voided volume. Finally, most other studies have only looked at bladder volumes at void and/or the PVR volume while this study examined the whole range of bladder filling and calculated accuracy over a large range of volumes.

An alternative to using conventional ultrasound would be to use an automatic portable device such as the BladderScan® [4,31–33]. These devices have the advantages of being very simple and quick to use. They work by creating a 3D model of the bladder based on detection of fluid tissue borders, and often have comparable accuracy to conventional ultrasound [34]. However, they may be less reliable and perform poorly in individuals that have other fluid filled structures near the bladder or bladders of irregular shape [4,31–33]. Also, in a setting such as an emergency room where conventional ultrasound is already used to image multiple organs, using a second device is impractical. It is expected that in most cases, bladder volume can be measured using either system, but research specifically
comparing conventional 3D ultrasound to the results of portable devices would be necessary to determine this.

This study is limited by the small sample size and by the use only of women with relatively severe OAB symptoms. Further research will be required with larger numbers and on different patient populations both during UD and during non-invasive hydration studies to see if this methodology is more generalizable.

In conclusion, this study demonstrates that serial measurements of bladder volume during filling in women with OAB can be accurately computed during UD using both the 3D ultrasound and corrected 2D ultrasound (V\textsubscript{bih}) methods. However, the 3D method was more precise and better accounted for bladders with irregular shapes. Thus, the 3D ultrasound method of real-time serial bladder volume measurements may represent the best available tool for the continued development of non-invasive “ultrasound urodynamics”.

**Acknowledgments**

Research funding for this study was provided by the Virginia Commonwealth University Presidential Research Quest Fund, the Dean’s Undergraduate Research Initiative, and NIH grant R01DK101719. The authors would like to thank Stefan Harris, Rachel Wilbur, Ronie-Rafael Altejar, Zachary Cullingsworth, Ashley Carroll, Kimberly Bradley, Sandy Smith, and Pam Harrell.

**Abbreviations used**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICIq</td>
<td>international consultation on incontinence questionnaire</td>
</tr>
<tr>
<td>OAB</td>
<td>overactive bladder</td>
</tr>
<tr>
<td>PVR</td>
<td>post-void residual</td>
</tr>
<tr>
<td>SEM</td>
<td>standard error of the mean</td>
</tr>
<tr>
<td>2D</td>
<td>two-dimensions</td>
</tr>
<tr>
<td>3D</td>
<td>three-dimensions</td>
</tr>
<tr>
<td>UD</td>
<td>urodynamics</td>
</tr>
<tr>
<td>VOCAL</td>
<td>virtual organ computer-aided analysis</td>
</tr>
</tbody>
</table>

**References**


Figure 1. Three orthogonal planes of a 3D bladder image showing how the bladder is traced in VOCAL to construct the 3D volume measurement
The top left is the transverse plane with orange perimeter indicating manual tracing; the top right is the sagittal plane with yellow perimeter made automatically; the bottom left is the coronal plane with yellow perimeter made automatically with orange lines overlaid indicating the transverse cross-sections traced; the bottom right is the reconstructed volume based on tracings from the six cross sections.
Figure 2. Transverse (left) and sagittal (right) views of the same bladder image as Figure 1 showing diameters used to calculate 2D-image based volume measurements.
Figure 3. Plot of the four different volume estimators versus control volume (black 45° line)
A. The full data set. B. Zoomed plot of (A) with bladder volumes up to 600 ml. Shown are $V_{H2O}$ data points in cyan stars with linear regression line in solid cyan, $V_{3D}$ data points in green pentagons and regression line in dashed green, $V_{bih}$ data points in red squares and regression line in dotted red, and $V_{spheroid}$ data points in blue diamonds and regression line in dashed dotted blue. Slope ($m$) and $R$-squared ($R^2$) values of the linear regressions are shown in the legend in (A).
Figure 4. Mean and standard error of the normalized root-mean-squared (%RMS) error of each volume estimator compared to V\textsubscript{control}.

A. Mean %RMS for each participant. Significant difference based on ANOVA with multiple comparisons to V\textsubscript{H2O} is denoted by a star.

B. Normalized %RMS error as a function of bladder size. The ordinal location of each point represents the mean of all control bladder volumes of its 100 ml interval. Significant difference based on ANOVA with multiple comparisons to V\textsubscript{H2O} is denoted by a double dagger (V\textsubscript{3D}), dagger (V\textsubscript{bih}), or star (V\textsubscript{spheroid}).

*P < 0.05.
Figure 5. Two examples of bladder volume measurements throughout filling
The bladder with volumes shown in (A) and final ultrasound image in (C) followed the most typical pattern. The bladder with volumes shown in (B) and final ultrasound image in (D) was more difficult to measure using only 2D methods because of its borders extended beyond the image width.
Table 1
Demographic and experimental data on the 12 women who completed the study.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean ± SEM</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (year)</td>
<td>53.7 ± 11.1</td>
<td>28 to 67</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>33.5 ± 9.7</td>
<td>19 to 54</td>
</tr>
<tr>
<td>ICIq question 5a score</td>
<td>3.25 ± 0.13</td>
<td>3 to 4</td>
</tr>
<tr>
<td>ICIq total score</td>
<td>9.67 ± 2.93</td>
<td>5 to 15</td>
</tr>
<tr>
<td>Initial capacity (ml)</td>
<td>553.8 ± 89.6</td>
<td>132 to 1172</td>
</tr>
<tr>
<td>Capacity (ml)</td>
<td>523.5 ± 91.9</td>
<td>99 to 1225</td>
</tr>
<tr>
<td>Urine production (ml)</td>
<td>27.4 ± 9.52</td>
<td>0 to 109.1</td>
</tr>
<tr>
<td>Number of images</td>
<td>9.2 ± 0.9</td>
<td>3 to 13</td>
</tr>
</tbody>
</table>