Butterfly™

Butterfly iQ Auto Bladder Volume

Development and validation of a Deep Learning Algorithm for Calculating Bladder Volume

A White Paper by

Erin Girard Ph.D.

Swami Sankaranarayanan Ph.D.

of Butterfly Network, Inc.

In collaboration with

Kevin Cline, MD,
Regional Urology, Shreveport LA

Lev Elterman, M.D.,
Elterman Center for Women's and Men's Health, Skokie IL

Bernadette Greenwood, BSc., PG Cert., RT (R)(MR)(ARRT),
HALODx, Indian Wells, CA

Part number 980-20085-00 Rev A

Scan. Analyze. Verify.

"With smart AI tools, we help you make informed clinical decisions."

Key Take-Aways

1. The Auto Bladder Volume tool in the Butterfly iQ app can provide accurate and reliable bladder volume measurements in <6 seconds, regardless of your setting—hospital, urology practice, or nursing home.
2. The Butterfly iQ costs less than existing bladder scanners and can also perform additional imaging of the pelvis and guide IV access. In general, facilities have improved ROI using an ultrasound bladder scanner when compared to the cost of catheterization.
3. This paper is a summary of the novel technology supporting the Auto Bladder Volume tool and details the rigorous performance tests used for validation.

Introduction

Nurses and physicians need accurate, reliable, and fast bladder volume measurements for bladder care programs, urologic evaluations, and post-operative care. In the in-patient care setting, bladder volume is used to avoid complications with urinary retention and to avoid unnecessary urinary catheterizations that may lead to urinary tract infections (UTI's). In the urology practice, bladder volume measurement is integral in the assessment of lower urinary tract symptoms, bladder training, and ensuring consistent filling before prostate cancer treatment. Here we introduce the Butterfly iQ with the Auto Bladder Volume tool, which reliably calculates the bladder volume from an automated 3D ultrasound sweep (Figure 1).

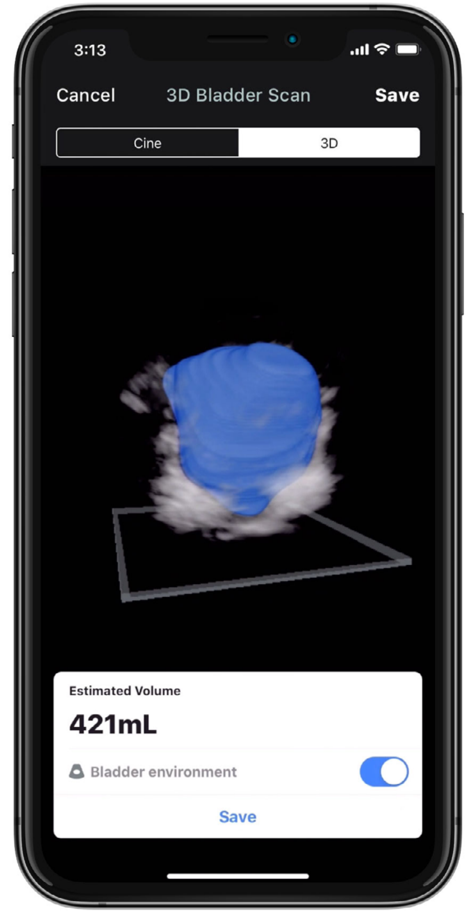


Fig. 1

The reliability and simplicity of the Auto Bladder Volume tool provides healthcare professionals a measurement that they trust.

We approach the discussion in four phases. We start by describing the technology behind the Auto Bladder Volume tool. We then follow with a definition of our performance and reliability standards and methods for validating this new technology to give transparency to the rigor of our process.

We then describe the user experience, including the full workflow if your iQ is connected to an electronic medical record system. Finally, we share some of the clinical applications that can be summarized from prior literature on ultrasound for bladder volume measurements and extend the value proposition to illustrate how the iQ is more than just a Bladder Scanner. **We value our commitment to delivering portable and affordable ultrasound to all users.**

"Butterfly iQ's Bladder Volume tool provides an accurate, rapid, and inexpensive new method for non-invasive bladder volume measurements in a community urology setting."

— Richard David, MD

The Auto Bladder Volume tool

Overview

The Auto Bladder Volume tool provides a non-invasive calculation of bladder volume using transabdominal ultrasound and can be accessed through the **Bladder** preset in the Butterfly iQ app. The tool has the unique capability to acquire a 3D sweep of the bladder region in seconds, while you hold the probe steady. After the sweep, a bladder volume is automatically calculated and shown on screen. This tool can be used with the probe positioned either in transverse (axial) or longitudinal (sagittal) directions.

Guidance feature

The Auto Bladder Volume tool includes a guidance feature that ensures the user is centered over the bladder to optimize accuracy. As soon as the probe is placed in the midline over the symphysis, the bladder appears on the screen highlighted in blue, as shown in Figure 2. The bladder is highlighted in the image using the deep learning algorithm described below. In the center of the bladder a square appears as a guide. An in-plane vertical guideline also appears and the user should adjust the position of the probe until the guideline intersects the square to ensure optimal positioning. Once the bladder is centered, select **Calculate**.

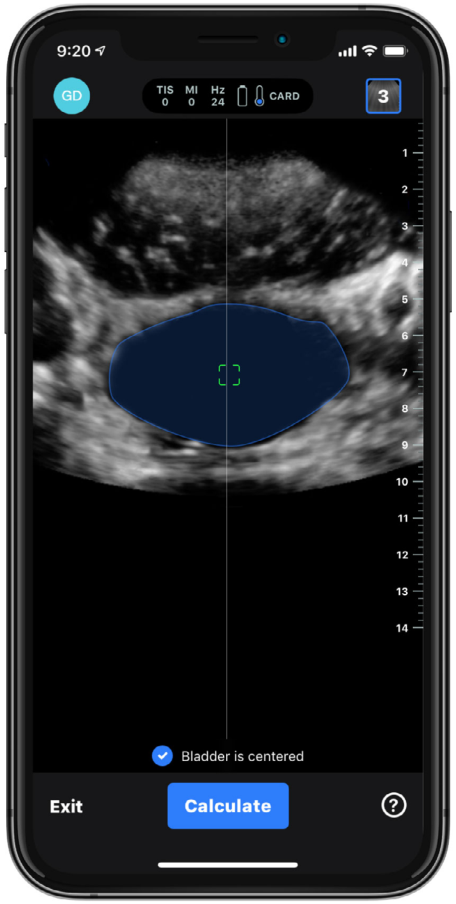


Fig. 2

3D sweep and calculation

After pressing **Calculate**, an electronic 3D sweep is automatically acquired by steering the 100° B-mode bladder image through a 120° sector to obtain 25, 2D images (Figure 3). The Butterfly iQ uses a 9000-element, 2D capacitive micromachined ultrasound transducer (CMUT) array operating at ~3MHz. A border is then automatically traced around the fluid inside the bladder on each 2D image, using the algorithm described below. Finally, we combine the segmentation results on each image to determine the volume of the bladder and construct a 3D model of the bladder. With the 3D model we can also help you verify the volume and flag potential issues of the scan. For instance, when the captured bladder touches the edge of the image area, we know the scan is very likely to be 'clipped' and the true volume is greater than measured. In such a case, a lower bound of bladder volume is reported (≥ XmL) to bring caution to the user.

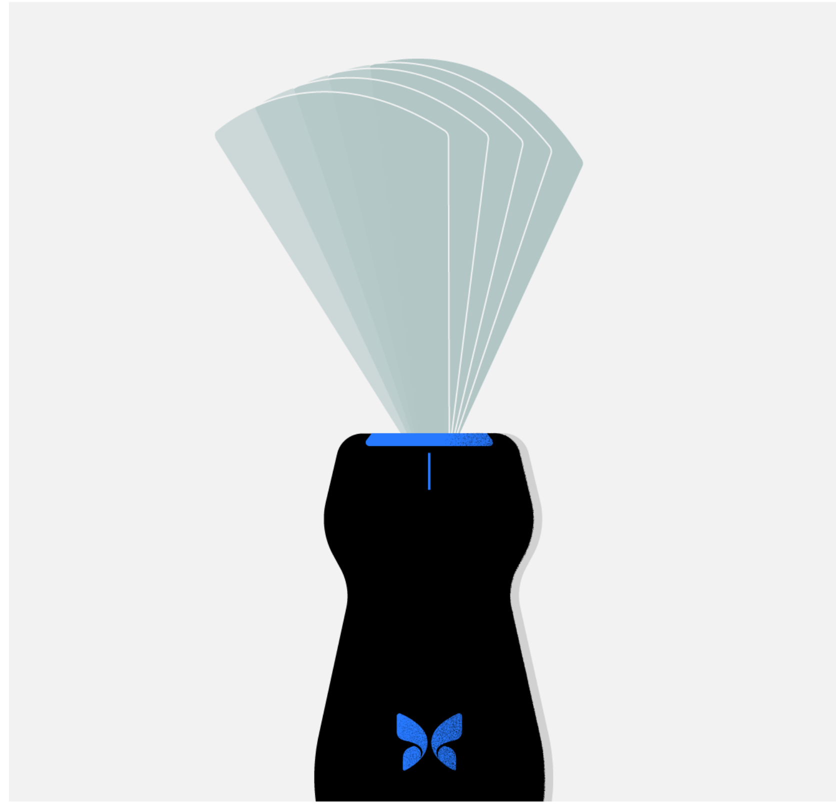


Fig. 3

Deep learning segmentation

The Auto Bladder Volume tool is driven by a deep neural network based artificial intelligence (AI) algorithm that is trained to trace the contour of the fluid in the bladder appearing in every frame of the ultrasound scan. A deep neural network (DNN) is an artificial neural network with multiple hierarchical layers between the input and output. In our case, the input is a 2D ultrasound image and the output is the contour of the bladder contained in the input image. Each layer in the hierarchy becomes an expert in identifying different parts of the bladder image. For the architecture of our algorithm, we use a variant of a well known algorithm called U-Net [1], which has been extensively applied to medical image segmentation tasks in the research literature. The parameters of the algorithm are optimized using a learning procedure that involves “showing” the network several thousand ultrasound images, whose contours are traced manually by expert sonographers. The algorithm is then trained to be able to distinguish bladder from other structures and organs in the abdominal region and subsequently to imitate the contours of the expert sonographers.

Training data

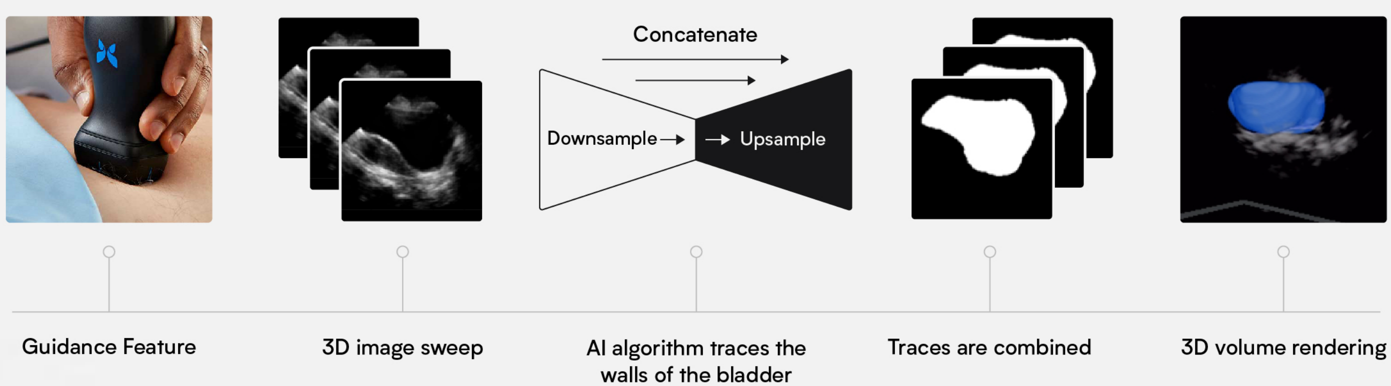
Following standard machine learning practices, the manually marked contour data used in developing the bladder tool is broadly divided into two sets: training and validation. The training set consists of 70,000 images from 700 patients and the validation set consists of 35,000 images from 100 patients. The patients that appear in the training set are not included in the validation set. The performance of each model is measured by a metric called the intersection over union (IoU), which denotes how close the contours predicted by the model are compared to expert traces. Each trained model is scored against the validation set and the highest scoring model, with respect to the IoU metric, is chosen for evaluation. The higher the IoU, the more accurate is the predicted contour.

IoU = # true-positives / (# true-positives + # false positives + # false negatives)

Summary

In summary, the iQ with the Auto Bladder Volume tool works by automatically performing a 3D image sweep through the bladder (Figure 4). The images are then sent to an artificial intelligence algorithm that traces the walls of the bladder. These traces are combined to give a resulting volume measurement and a 3D image of the bladder for visualization. The 3D visualization provides assurance that the whole bladder was captured.

Fig. 4



Ensuring performance and reliability

The performance of any new measurement method is a combination of its accuracy and precision. The accuracy of the measurement refers to the closeness of measured values to a true value. Precision refers to the variability that arises from repeated measurements made in identical circumstances - and although there should be identical results, the operator, equipment, and patient all contribute to variations measurements.

The following sections describe studies performed in the clinical environment and at the bench, using bladder phantoms, to test and ensure a high level of performance of the Auto Bladder Volume tool. In summary, the following assessments were performed:

1. **Accuracy on bladder phantoms**
2. **Precision on bladder phantoms**
3. **Accuracy on human subjects**
4. **Precision on human subjects**

Accuracy and Precision: Bladder Phantoms

Methods

Accuracy and precision were measured in the laboratory using custom bladder phantoms, having a known bladder volume covering the range of expected volumes. To determine the accuracy of the tool, the bladder phantoms were measured by five experienced sonographers. The mean error and tolerance are reported.

Precision is demonstrated through repeatability and reproducibility. Auto Bladder Volume estimation must be repeatable on the same subject and reproducible on the same subject using different devices. We collected a dataset using the phantoms where each phantom was scanned five times using five different probes. To calculate the repeatability and reproducibility, we follow the procedure outlined in ISO-5725-2, which defines the repeatability standard deviation and the reproducibility standard deviation.

For the accuracy and precision measurements, we stratify the phantoms into two bins, containing volumes less than 100 mL and more than 100 mL, respectively. For the volumes >100 mL, we consider the fractional (%) repeatability/reproducibility standard deviations (relative to the mean of the scans), while for the volumes <100 mL we report the absolute standard deviations (in mL).

Results

The accuracy is reported in the table below as an error range for each phantom.

Phantom	Error range
23.6 mL	1.58 to 3.38 mL
129.7 mL	-2.9 to 5.9%
294.5 mL	-4.8 to 1.6%
741.1 mL	-5.9 to -3.8%

The following table reports the results of our precision study, specifically the repeatability and reproducibility.

Phantom	Number of iQ devices tested	Repeatability standard deviation	Reproducibility standard deviation
23.6 mL	10	0.45ml	0.45ml
129.7 mL	10	2.2%	2.2%
294.5 mL	10	1.6%	1.6%
741.1 mL	10	0.5%	0.5%

Accuracy and Precision: Human Subjects

Methods

The clinical validation was a prospective, multi-center study under institutional ethical approval conducted between August - October 2019 at 4 locations to represent geographic diversity. The sites included 3 urology clinics (Regional Urology, Shreveport LA, Elterman Center for Women's and Men's Health, Skokie IL, and Desert Medical Imaging, Indian Wells CA) and the Butterfly office in Guilford, CT. 165 subjects were recruited and all subjects gave written informed consent. All subjects were older than 18 and there were 91 females, 59 males, and 15 unreported genders. Exclusion criteria included pregnancy, age younger than 18 years, vulnerable, and agitated subjects.

Subjects were both patients at the urology clinics and healthy volunteers.

Bladder volume was assessed by the Butterfly iQ Auto Bladder Volume tool and a state-of-the art predicate device and both were compared to the gold standard measurement of urine void. For volume measurement by ultrasound, trained sonographers measured pre-void and post-void volumes. For the iQ and predicate device, the probe was placed superior to the symphysis pubis and directed towards the bladder with the subject in the supine position. The probe was centered over the bladder for the acquisition. Scanning was done in both transverse and sagittal planes, and if the ultrasound device indicated clipping as indicated by the “>” symbol on either device, up to three volume measurements were made. There was no delay between measurements and the subjects did not change their position.

Once the pre-void scanning was complete, the subjects were asked to void in a beaker and the volume was recorded. All subjects underwent repeat ultrasound scanning to measure the post-void residual.

For each subject, if multiple scans were performed due to clipping, then the maximum of the scan volumes was considered as the corresponding pre/post volume for that patient.

The difference in pre- and post-void estimates were then compared to the volume of urine, where the error was is defined as:

$$\text{error} = \left| (V_{pre} - V_{post}) - V_{measure} \right|$$

for volumes between <100 mL

and

$$\text{error} = \left| 100 \; * \; \frac{(V_{pre} - V_{post}) - V_{measure}}{V_{measure}} \right|$$

for volumes between 100-1000 mL,

V_{pre} and V_{post} are the pre-void and post-void volumes as measured on ultrasound, respectively, and $V_{measure}$ is the volume from urine output.

Auto Bladder Volume estimation must be repeatable on the same subject. We collected a dataset of 100 subjects where each subject was scanned more than once and report the repeatability standard deviation.

As in the bladder phantom analysis, all metrics report error on volumes <100 mL in absolute values (in milliliters) and for volumes between 100-1000 mL, the percentage (%) error is used.

Results

The mean error and 95% confidence interval of bladder volume measured on the iQ compared to urine volume is illustrated in the table below.

	Mean error	95% confidence interval
In-vivo (0-100ml)	15.4 mL	10.8 to 19.9 mL
In-vivo (100-1000ml)	16.8%	14.1 to 19.4%

Figure 5 shows the plot between the true voided volumes on the x-axis and the estimated volumes on the y-axis. The color indicates the device whose proprietary algorithm was used to estimate the volume, with blue indicating the Butterfly iQ and yellow indicating a state of the art predicate device that is exclusively used as a bladder scanner. The plot shows that the volume estimates from the Butterfly iQ are highly correlated with both the predicate device and the true volume, with the pearson correlations values 0.967 and 0.955 at $p < 0.001$, indicating high level of significance.

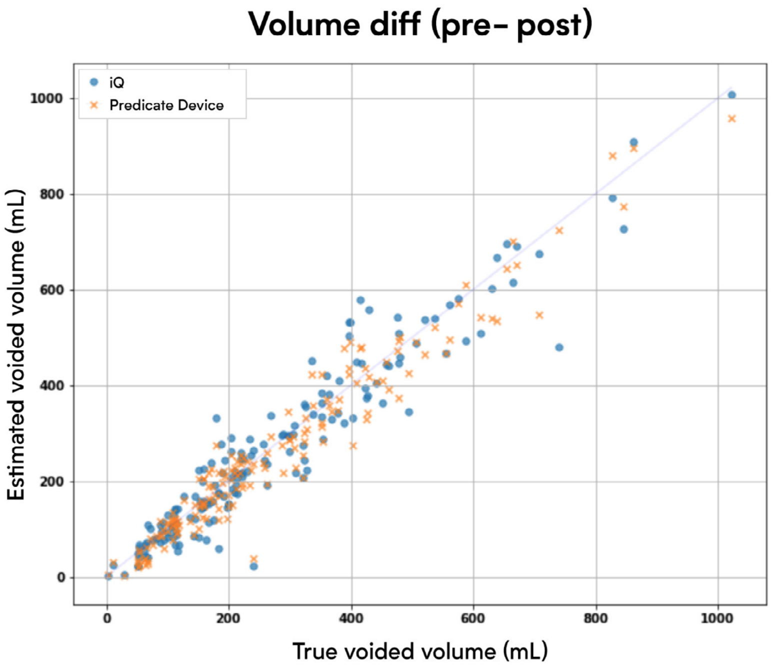


Fig. 5

The Bland-Altman plot (Figure 6) illustrates the agreement between the volume estimates from the Butterfly iQ and the predicate device. The tight clustering indicates good agreement between both the devices and it shows the effectiveness of the iQ as a stand along bladder scanner.

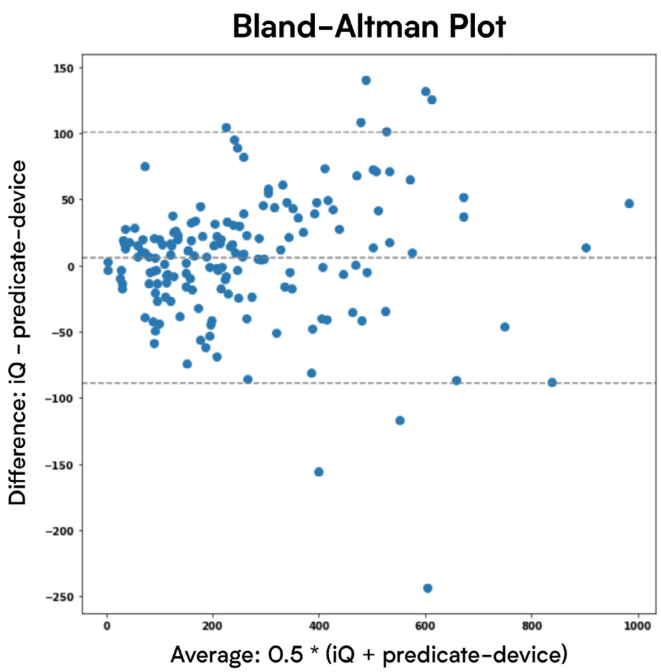


Fig. 6

Summary

Our results across both the bladder phantoms and clinical subjects show excellent accuracy and precision. Our comparison with the true patient urine volumes and a start of the art predicate shows the effectiveness of our device to function on the field as a stand along bladder scanning product.

User Experience/Workflow

The Auto Bladder Volume tool was designed with the simplicity that our users expect. Within 6 seconds, you can have an accurate and reliable bladder volume measurement. To access the tool, go to the Bladder preset and tap or swipe the tools icon located at the bottom right corner. Then select **Volume**.

Training is provided within the Butterfly app through a first-time user experience guide that pops up and walks the user through the steps, as shown in Figure 7. Once positioned over the bladder, the volume calculation is automatically calculated when the user presses **Calculate.**

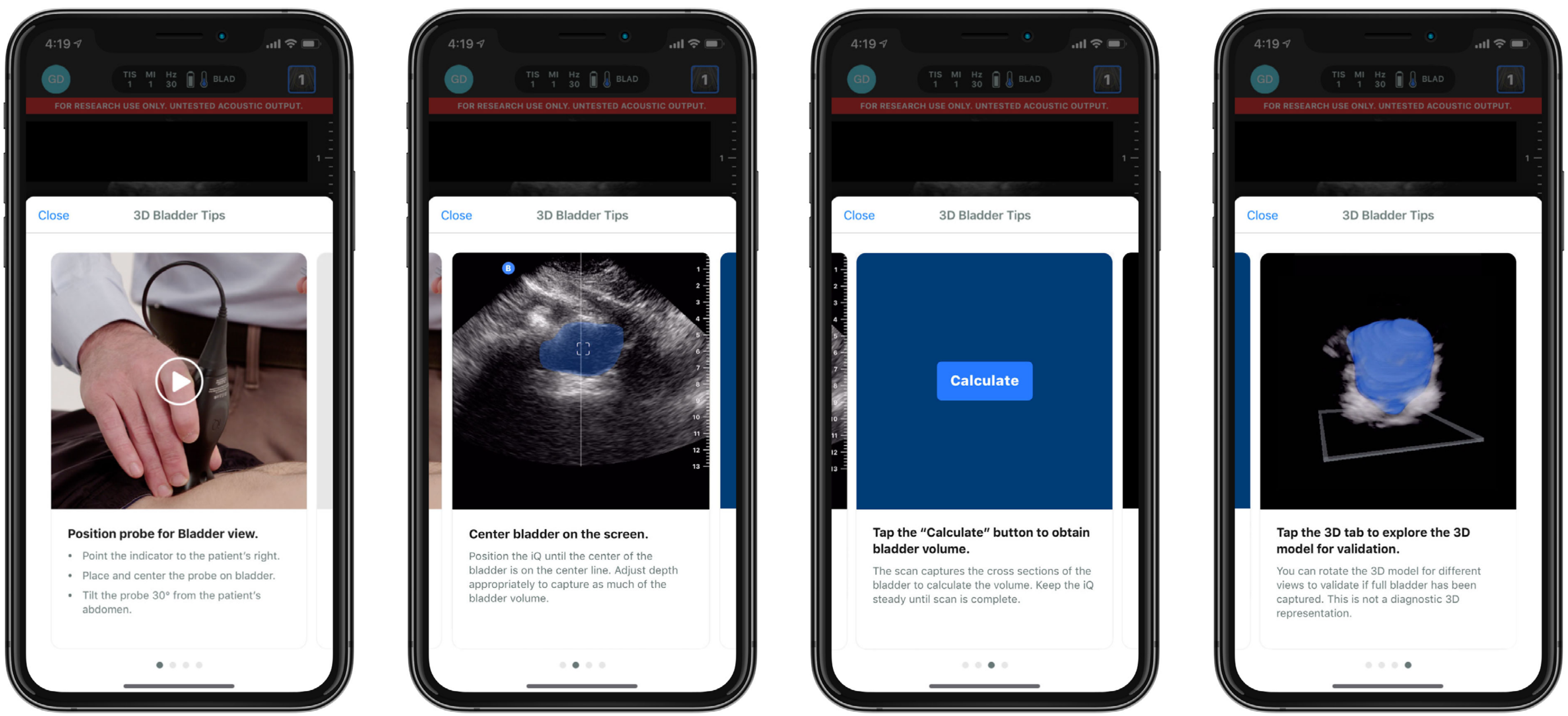


Fig. 7

With the Butterfly iQ placed over the center of the bladder, press **Calculate** and hold the probe still while a 3D sweep of the bladder area is automatically acquired and the volume is calculated (~5 seconds). After successfully calculating the volume, the following three items are displayed on the screen:

1. **The calculated Volume**
2. **A 3D render of the bladder**
3. **The cine of the images acquired during the sweep**

Once the measurement has been made, the image and volume results can be saved to the Butterfly Cloud. Once saved, your study can be viewed on the Cloud website as well as on your iPhone.

The Butterfly Cloud is HIPAA compliant and integrates directly into your EMR or to PACS.

Lastly, Butterfly iQ with the Auto Bladder Volume tool has been designed to replace more expensive bladder scanners, in order to democratize access to this technology.

Clinical applications

The Auto Bladder Volume tool may be used whenever there is a need to non-invasively measure bladder volume. The following are a few examples where healthcare professionals may find use in knowing their patient's bladder volume.

1. CAUTI Prevention

A catheter-associated urinary tract infection (CAUTI) is one of the most common infections a person can contract in the hospital. Indwelling catheters are associated with high infection rates and intermittent catheterization has become preferable. The CDC Guideline for Prevention of Catheter-Associated Urinary Tract Infections (CAUTI) from 2009 [2] suggests to

"Consider using a portable ultrasound device to assess urine volume in patients undergoing intermittent catheterization to assess urine volume and reduce unnecessary catheter insertions."

They also recommend implementing a quality improvement program that includes

"Protocols for management of postoperative urinary retention, such as nurse-directed use of intermittent catheterization and use of bladder ultrasound scanners"

2. Evaluating lower urinary tract symptoms (LUTS)

The relationship between LUTS and urodynamic findings is complex. Uroflowmetry and measurement of postvoid residual (PVR) urine volume are standard noninvasive tests to evaluate voiding dysfunctions.

Determining the cause of LUTS requires a differential diagnosis between urological and nonurological conditions that span the breadth of Parkinson's, diabetes, congestive heart failure, prostate cancer, benign prostatic hyperplasia and more [3].

PVR volume is a primary component of the clinical evaluation and monitoring and can be obtained quickly and non-invasively with the Auto Bladder Volume tool.

"...bladder ultrasound can be performed with a portable device, is noninvasive and time-efficient, minimizes medical waste and supplies, and determines when catheterization is medically appropriate..." [4]

3. Post-Operative Urinary Retention

Postoperative urinary retention (POUR) is associated with risk of overdistention and permanent detrusor damage. POUR can be prevented by monitoring bladder volume with an ultrasound scanner or by using an indwelling catheter, although the latter may increase the risk of urinary infection. It has been suggested to routinely evaluate the bladder content in the PACU on entry and before discharge [5]. Additionally, with training, nurses can operate an ultrasound bladder scanner [6] for this evaluation, under physician supervision.

As with other bladder scanning devices, our Auto Bladder Volume tool was designed with the ease and simplicity that nurses expect from a bladder scanner.

4. Urinary incontinence

The elderly suffer from bladder dysfunctions, including urinary incontinence, urinary retention, and urinary tract infection. For example, urinary incontinence occurs in 50% of nursing home residents and is the second leading cause of institutionalization among the elderly [7]. Nursing staff are requested to perform a comprehensive assessment of bladder dysfunction, and this can be supported by ultrasound bladder volume measurements.

Bladder volume measurement by ultrasound has demonstrated

"...more efficient resident care because of the reduced frequency of bladder distention and catheterization and increased resident satisfaction with bladder management routines. [8]"

5. Ensuring a full bladder for radiotherapy

Changes in bladder volume from the time of radiotherapy planning to the treatment period affects tissue toxicity and the radiotherapy dose. Patients are asked to maintain a full bladder to displace the dome of the bladder and small bowel from the target volume and to minimize influence on prostate movement [9]. Even with instruction, there are large variations in bladder volume throughout treatment period. Bladder volume assessment with the Auto Bladder Volume tool before radiotherapy treatment could be a cost effective screening tool to optimize irradiation time based on bladder volume [10].

Conclusion

We have demonstrated an AI powered Auto Bladder Volume tool that provides accurate and precise results in <6 seconds. The onscreen guidance walks the user through the acquisition, simplifying the learning curve.

We have described here our development, reliability, and reproducibility testing of the Auto Bladder Volume tool that were performed to ensure product integrity. We are committed to developing a robust yet simple bladder volume feature that can perform in all clinical conditions, from the operating room to the nursing home.

We have shown that the Butterfly iQ Auto Bladder Volume tool can meet the rigorous clinical accuracy and operational ease of use that our customers demand. It is paramount to understand that we provide this tool as a no-cost addition to the broad, general use of the Butterfly iQ, which is a fraction of the cost of our competitors. In an effort to equip all of our users with ultrasound for clinical decision making, we believe that having smart AI tools helps you make informed clinical decisions.

We continue to challenge and improve our device in the development process and look forward to providing more updates soon. Future work consists of a focus on small volumes to address the pediatric market and drive improved accuracy for measuring post void residual.

References

1. Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. “U-net: Convolutional networks for biomedical image segmentation.” International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015.
2. Gould, Carolyn V., et al. “Guideline for prevention of catheter-associated urinary tract infections 2009.” Infection Control & Hospital Epidemiology 31.4 (2010): 319-326.
3. Lepor, Herbert. “Pathophysiology of lower urinary tract symptoms in the aging male population.” Reviews in urology 7.Suppl 7 (2005): S3.
4. Kelly, Christopher E. “Evaluation of voiding dysfunction and measurement of bladder volume.” Reviews in Urology 6.Suppl 1 (2004): S32.
5. Keita, Hawa, et al. “Predictive factors of early postoperative urinary retention in the postanesthesia care unit.” Anesthesia & Analgesia 101.2 (2005): 592-596.
6. Rosseland, L. A., A. Stubhaug, and H. Breivik. “Detecting postoperative urinary retention with an ultrasound scanner.” Acta Anaesthesiologica Scandinavica 46.3 (2002): 279-282.
7. Fantl, J. Andrew. Urinary incontinence in adults: acute and chronic management. No. 96. Department of Health and Human Services, 1996.
8. Wooldridge, Leslie. “Ultrasound technology and bladder dysfunction.” The American Journal of Nursing 100.6 (2000): 3-14.
9. O’Doherty, Úna M., et al. “Variability of bladder filling in patients receiving radical radiotherapy to the prostate.” Radiotherapy and Oncology 79.3 (2006): 335-340.
10. Stam, Marcel R., et al. “Bladder filling variation during radiation treatment of prostate cancer: can the use of a bladder ultrasound scanner and biofeedback optimize bladder filling?.” International Journal of Radiation Oncology Biology Physics 65.2 (2006): 371-377.

For questions on the information described here or for additional resources, contact us at support@butterflynetwork.com