Comparative Effectiveness of Dynamic Abdominal Sonography for Hernia vs Computed Tomography in the Diagnosis of Incisional Hernia

William C Beck, MD, Michael D Holzman, MD, MPH, FACS, Kenneth W Sharp, MD, FACS, William H Nealon, MD, FACS, William D Dupont, PhD, Benjamin K Poulose, MD, MPH, FACS

BACKGROUND: Diagnosis and characterization of incisional hernia are often established by CT, which incurs radiation exposure and substantial cost in clinical practice. The purpose of this study was to determine the comparative effectiveness of surgeon-performed Dynamic Abdominal Sonography for Hernia (DASH) vs CT for incisional hernia evaluation.

STUDY DESIGN: Patients with previous abdominal operations and recent CT imaging were enrolled prospectively; patients with stomas, fistula, or soft tissue infection were excluded. In the clinic setting, DASH was performed with prerequisite training of the American College of Surgeons Ultrasound for Surgeons Basic Course. Clinical evidence of hernia, results of DASH examination, and radiologist documentation of incisional hernia were compared with the gold standard of surgeon-interpreted CT. Testing characteristics of sensitivity and specificity were compared and predictive values were calculated. Inter-rater reliability was performed by comparing DASH results in a subgroup of patients with 3 different evaluators.

RESULTS: There were 181 patients enrolled, with a mean age of 54 years, and 68% were women. In patients in whom hernias were identified, the mean hernia size was 44.6 cm² (range 0.2 to 468.3 cm²). The DASH examination showed high sensitivity (98%) and specificity (88%). Hernia prevalence was 55% in this population, resulting in positive and negative predictive values of 91% and 97%, respectively. Four patients had clinically detectable hernias that were not seen on CT but were discovered with DASH. Inter-rater reliability for DASH was high, with an observed intraclass correlation coefficient of 0.79.

CONCLUSIONS: The DASH examination is an accurate alternative to CT scan for diagnosing abdominal wall hernias, with additional benefits of no radiation exposure and instant bedside interpretation. The use of DASH to detect hernia recurrence can greatly facilitate long-term follow-up of hernia patients. (J Am Coll Surg 2013;216:447–453. © 2013 by the American College of Surgeons)
for small hernias in obese patients. In this setting, follow-up CT scanning incurs prohibitively high cost, with some radiation risk to the patient.

Advances in ultrasound (US)-based imaging may offer a solution to these issues in the detection and characterization of incisional hernia. Den Hartog and colleagues described the use of US for detecting incisional hernia, with sensitivity of 71% and specificity of 100%. However, this study did not describe a standard technique for performance of the US examination, had a limited number of patients (40), and was restricted to the evaluation of vascular surgery patients. Although this study did provide critical initial data supporting the use of US for incisional hernia, these weaknesses limit the generalizability and validity of the results. In this study, we prospectively evaluated the comparative effectiveness of surgeon-performed Dynamic Abdominal Sonography for Hernia (DASH) with CT in the diagnosis and characterization of incisional hernia.

**METHODS**

**Design overview**

Patients were prospectively invited to participate in this study from a busy tertiary care referral center general surgery practice between 2010 and 2012. Baseline demographics, comorbid conditions, and surgical histories were recorded. For the comparative effectiveness portion of the study, DASH was first performed on patients, followed by review of abdominal and pelvic CT scans to evaluate the presence or absence of incisional hernia. Testing characteristics of sensitivity and specificity were calculated, using CT as the gold standard. A subgroup of patients were evaluated for the inter-rater reliability portion of the study, with 2 additional surgeons (total of 3 surgeons) performing DASH and evaluating the corresponding CT images. This study was approved by the Vanderbilt Human Research Protection Program.

**Inclusion and exclusion criteria**

Adult patients referred to the Vanderbilt University General Surgery Clinic were invited to participate in the study if they had any history of previous abdominal or pelvic operation performed via incision of the anterior abdominal wall (open or laparoscopic). Additionally, patients must have had a viewable CT scan of the abdomen and pelvis within 6 months of enrollment. Patients were not candidates if they had a current stoma, open wound, or active abdominal wall infection. In general, these patients were seen in clinic for issues related to foregut diseases, intra-abdominal malignancies, gallstones, or hernia.
performed as needed to help visualize subfascial structures passing through the hernia defect and hernia dimensions were recorded in 2 axes (cranio-caudal and transverse). Particular care was taken to visualize the umbilicus from different angles to ascertain the presence of hernia. The DASH examination was performed in both supine and upright positions. For the comparative effectiveness portion of the study, a single surgeon performed all DASH examinations used for analysis before review of CT scans. The CT scan images were evaluated by the same surgeon who performed the DASH examination. Transverse 3-mm cuts were visualized and incisional hernia presence was determined using the same criteria as for the DASH examination.

**Analysis: comparative effectiveness and sensitivity analysis**

Baseline descriptive statistics were determined for the study population. The main outcomes measure was identification of an incisional hernia, which was then recorded as a dichotomous outcome for each patient. The first comparison evaluated the effect of patient positioning on DASH results. We suspected that the yield of DASH would be increased with upright examination compared with supine examination. To evaluate testing characteristics of DASH, sensitivity and specificity were calculated with 95% CIs based on true positive and true negative examinations, as determined by the gold standard, surgeon-interpreted CT scan. Positive and negative predictive values were calculated based on the prevalence of incisional hernia within the study population. We did compare radiologist interpretation of CT scan for incisional hernia compared with surgeon-interpreted CT to confirm our assumption that surgeon-interpreted CT was the more reliable test.

Based on our clinical experience using the DASH technique, we suspected that some incisional hernias not detected by CT scan would be detectable by DASH. The typical scenario for this would be small hernias that are readily apparent on physical examination, but not easily discernible on CT scan. So, we would expect a different number of true positive CT readings had CT interpretation been able to detect these hernias. To better evaluate this, we performed sensitivity analysis evaluating the effect of an increasing true positive count on DASH specificity.

**Analysis: inter-rater reliability**

In addition to estimating the testing characteristics of DASH, we also wished to evaluate the consistency of results performed by multiple surgeons. This was especially important given the potential variability using a US-based modality. Two additional surgeons were used as examiners to determine the presence of incisional hernia based on DASH evaluation. These surgeons also viewed CT results after DASH was performed and recorded the presence of incisional hernia based on that modality. Inter-rater reliability was calculated by comparing the responses of all 3 surgeons and calculating an intraclass correlation coefficient. An intraclass correlation coefficient above 0.75 (scale of 0 to 1.0) indicates excellent inter-rater agreement. We estimated 20 patients as sufficient to determine inter-rater reliability. Data were processed and analyzed using IBM SPSS version 20 (IBM Corporation).

**Power calculation**

Using data from den Hartog and colleagues, the percent correlation outcome between CT and US was expected to be high, given 71% sensitivity and 100% specificity. To estimate the number of patients needed in this prospective study, we calculated 95% confidence intervals using Wilson’s procedure to estimate the true positive (sensitivity) and true negative (specificity) error rates assuming that CT was the gold standard. Based on these 95% CIs, power curves were plotted for a given number of true positive or true negative patients and the resultant 95% CIs widths were noted (Fig. 3). For this study, minimums of 100 true positives and 100 true negatives were set as target accrual parameters.
RESULTS

One hundred eighty-one patients were prospectively enrolled into the study, including 123 women (68%) and 58 men (32%). Mean age was 54 ± 1.0 years (mean ± standard error of the mean) with mean body mass index (BMI) of 31.3 ± 0.5 kg/m². Fifty-nine patients (33%) had a previous incisional hernia repair performed. The mean number of earlier hernia repairs performed was 1.7 ± 0.13. Comorbid conditions observed are summarized in Figure 4. The mean size of hernias evaluated in this study as calculated by CT was 44.6 ± 7.7 cm² (range 0.2 to 468.3 cm²).

The DASH evaluation was successfully performed in all 181 patients, with count results presented in Table 1. One hundred seven incisional hernias were detected by DASH. Upright performance of DASH did not change the number of hernias detected compared with supine DASH. After CT review, the testing characteristics of DASH were calculated (Table 2). The sensitivity of DASH for detection of incisional hernia was 97.8% (95% CI 92.2%, 99.6%), with specificity of 87.8% (95% CI 78.3%, 93.7%).

To evaluate the accuracy of surgeon-interpreted CT scan as the gold standard for detection of incisional hernia, we performed additional comparisons. Radiologist interpretation of CT scan failed to identify 16 incisional hernias detected by surgeon-interpreted CT (sensitivity of 78.8% [95% CI 69.2%, 86.1%]). Four clinically apparent hernias were not detected by surgeon-interpreted CT but were detected by DASH. The mean BMI for these patients was 30.7 ± 7.9 kg/m², with defect sizes ranging from 0.6 cm² to 55.0 cm². Because this group represented patients who were incorrectly classified by the surgeon-interpreted CT, we performed a sensitivity analysis to evaluate the effect of increasing true positive results on specificity (Fig. 5).

Seventeen patients were enrolled in the inter-rater reliability portion of the study. Identical results were obtained for 2 surgeons, with discordant results found in 3 patients with the third surgeon. The calculated level of agreement (as indicated by the intraclass correlation coefficient) was high, at 0.79.

Table 1. Dynamic Abdominal Sonography for Hernia (DASH) Compared with Computed Tomography for Detection of Incisional Hernia

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value (95% CI), %</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT Positive incisional hernia, n</td>
<td>97</td>
</tr>
<tr>
<td>CT Negative incisional hernia, n</td>
<td>2</td>
</tr>
<tr>
<td>Totals, n</td>
<td>99</td>
</tr>
</tbody>
</table>

Table 2. Testing Characteristics and Performance of Dynamic Abdominal Sonography for Hernia (DASH) Compared with Computed Tomography for Detection of Incisional Hernia

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value (95% CI), %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>97.8 (92.2, 99.6)</td>
</tr>
<tr>
<td>Specificity</td>
<td>87.8 (78.3, 93.7)</td>
</tr>
<tr>
<td>Positive predictive value</td>
<td>90.7 (83.1, 95.2)</td>
</tr>
<tr>
<td>Negative predictive value</td>
<td>97.3 (89.7, 99.5)</td>
</tr>
<tr>
<td>Prevalence</td>
<td>54.7 (47.1, 62.0)</td>
</tr>
</tbody>
</table>
DISCUSSION

In this study, DASH performed comparably to CT scan in the detection of incisional hernias. Along with the reduction in exposure to ionizing radiation, the potential advantage of DASH over CT scan includes the ability to detect hernias not readily apparent on CT. Additionally, DASH provides real-time results available to both the surgeon and patient for clinical decision making, and can be performed by different surgeons with consistent results.

The use of US is particularly suitable for evaluation of the anterior abdominal wall given its accessibility and ability to monitor structures in real-time as physical maneuvers (eg, Valsalva) are performed to assist in hernia detection. Bloe men and associates\(^9\) compared US with physical examination and determined that the addition of US to physical examination increased the yield of ventral hernia detection by more than 20%. In a case series, Cingi and coworkers\(^10\) noted an increased ability to detect hernia formation after ostomy takedown using US compared with physical examination alone. One of the best evaluations of the use of US for ventral hernia was performed by den Hartog and coauthors.\(^5\) This study evaluated 40 patients with a history of abdominal aortic aneurysm repair for incisional hernia. Compared with CT scan, US was observed to have a sensitivity of 71% and specificity of 100%. Our study observed a higher sensitivity for US (97.8%), likely for several reasons. First, DASH was performed by surgeons with extensive experience in the management of ventral hernia, affording effective integration of the patient’s physical examination into the determination of a possible hernia. The DASH technique also provided a systematic means of abdominal wall evaluation, potentially minimizing missed areas during the examination. For focused applications, surgeon-performed US appears comparable to that performed and interpreted by radiologists.\(^11\) Our study found a sizable discrepancy between radiologist-interpreted CT scans and surgeon-interpreted CT scans: surgeons detected more hernias with equal false positive rates. This was likely due to hernia surgeons performing a more focused evaluation of the anterior abdominal wall and having experience visualizing several different anterior abdominal wall abnormalities.

The gold standard used for this study was surgeon-interpreted CT scan, which has historically been found to be fairly reliable in detecting incisional hernia.\(^12\) To highlight the problem with the radiologic detection of ventral hernia, 4 patients in this study had clinically obvious hernias but normal surgeon-interpreted CT scans. This was likely due to the supine, static nature of CT. These particular patients had fairly thick hernia sacs that were difficult to distinguish radiographically from healthy adjacent fascia. The use of real-time DASH with Valsalva was able to identify these as hernias, emphasizing the utility of DASH over CT scanning. The observed specificity for DASH was somewhat compromised by these misclassified patients. The “corrected” specificity is shown in Figure 5. As the number of true positives increased, the specificity for DASH also increased, with narrowing 95% CI. It is important to emphasize that DASH serves to supplement the information gained from the history and physical examination, allowing a more “informed” radiographic evaluation of the abdominal wall. This is reflected by the “real world” performance of DASH with high positive and negative predictive values (Table 2). In our practice, we have found that detection of incisional hernias by history and physical examination is most accurate in thin patients and least accurate in obese patients with smaller hernias. This is supported by data from this study; we detected 28 incisional hernias by DASH in patients who were deemed not to have hernias based on history and physical examination alone.

Based on the results of this study, we propose that DASH should be the new gold standard for the radiographic identification and characterization of incisional hernia. The results are generalizable to ventral hernia and need not be limited to incisional hernia per se. This has important implications both clinically and from a research perspective. The DASH examination was relatively easy to perform in a real-world

\[\text{Figure 5. Sensitivity analysis to evaluate the effect of increasing true positive results on Dynamic Abdominal Sonography for Hernia (DASH) specificity (with 95\% CI). The red dot indicates data observed in this study: 97 true positives as determined by surgeon-interpreted CT, resulted in specificity of 87.8\% for DASH. Surgeon-interpreted CT failed to detect 4 clinically apparent incisional hernias, which were detected by DASH. Had these patients been classified correctly, the number of true positives would have increased, boosting specificity for DASH (101 true positives with resultant specificity of 92.3\% [95\% CI 83.4\%, 96.8\%]).}\]
clinical setting. The surgeons did find it critical to undergo formal US training via the American College of Surgeons Ultrasound for Surgeons Basic Course. This provided the fundamental information needed to understand US basics. Certainly this is not comparable to formalized training undertaken by radiologists, but it sufficed for this focused application of US. The US machine easily integrates into clinical practice and offers a means of reimbursement with appropriate documentation of the study. Based on our internal data, approximately 100 DASH examinations would need to be performed to recoup the initial cost of the machine and maintenance fees.

The results of this study have enormous potential from a research perspective. The “holy grail” of hernia research is identifying a reliable, cost-effective means of detecting hernias on a long-term basis. No clinical study in ventral hernia has been able to reliably calculate long-term recurrence rates. Clinical examination is too unreliable, and follow-up CT scanning is prohibitively expensive, with concomitant radiation risk. DASH can easily fulfill this role, even with the initial cost of a US machine. The cost of the US machine varies between $15,000 and $30,000, depending on the age of the machine and its features. Clinical practices can potentially offset this one-time cost through billing for the procedure itself. Practitioners would need to meet credentialing and billing requirements pertinent to their particular institution for performance and documentation of these procedures. We believe that DASH, combined with some measure of abdominal wall function and quality of life, should become the gold standard of long-term ventral hernia follow-up. Active investigation into this role for DASH is ongoing.

Several limitations to this study are apparent. Some bias was introduced by having examiners perform the DASH examination and review the CT images. Knowing the DASH results likely influenced the interpretation of the CT to some degree. This bias was deemed acceptable at the inception of the study; it favors increasing true positive and true negative detection for CT scan interpretation. Even with this bias, DASH was found to be at least comparable to CT—if not better—considering a fair number of false negatives by CT. Some bias may also have been introduced by the interaction between the patient and the surgeon performing the US by having the ultrasonographer aware of the patient’s diagnosis. Extreme efforts were taken to reduce this bias to keep the surgeon-ultrasonographer blinded to the patient’s diagnosis for referral. The study fell somewhat short of its intended accrual goal, with 82 true negative patients (goal of 100). Given the high probability of DASH providing a correct diagnosis, however, the 95% CI for the resultant testing characteristics was predicted to be fairly narrow, even with this 20% shortfall in accrual for true negative patients.

Characterization of abdominal wall hernias (beyond simple identification of the hernia) was more challenging with DASH compared with CT scan. This was due to the overwhelming familiarity that surgeons have with CT imaging compared with US for this particular use. We did find that mean hernia size was comparable with both modalities, but a 3-dimensional construct of the hernia was more difficult to visualize with DASH. Hopefully this will improve as experience with DASH is gained. The goal of this study was to evaluate the utility of DASH for identification of an incisional hernia rather than a more extensive characterization of the hernia itself. The most challenging hernias to detect with DASH were small umbilical hernias in obese patients. In difficult situations, supplemental information afforded by CT scan can help the diagnosis should this additional information be needed. We expected a fairly high learning curve with DASH, but results were surprisingly reproducible between surgeons, as reflected by the high inter-rater reliability (0.79).

CONCLUSIONS
The DASH examination provides a novel and accurate means of detecting and characterizing incisional and ventral hernias of the abdominal wall. Based on these results, DASH should be the new gold standard for detection of these types of hernias, reducing radiation risk without sacrificing diagnostic accuracy. The DASH evaluation may be more feasible than CT evaluation in the long-term follow-up of ventral hernia patients, which has significant implications for both clinical practice and research.

Author Contributions
Study conception and design: Beck, Holzman, Sharp, Nealon, Poulose
Acquisition of data: Beck, Holzman, Poulose
Analysis and interpretation of data: Holzman, Sharp, Nealon, Dupont, Poulose
Drafting of manuscript: Poulose
Critical revision: Beck, Holzman, Poulose

REFERENCES