SAEM Academic Emergency Medicine
Official Journal of the Society for Academic Emergency Medicine

EVIDENCE-BASED DIAGNOSTICS

Systematic Review: Emergency Department Bedside Ultrasonography for Diagnosing Suspected Abdominal Aortic Aneurysm

Elizabeth Rubano, MD, Ninfa Mehta, MD, William Caputo, MD, Lorenzo Paladino, MD, and Richard Sinert, DO

Abstract

Background: The use of ultrasound (US) to diagnose an abdominal aortic aneurysm (AAA) has been well studied in the radiology literature, but has yet to be rigorously reviewed in the emergency medicine arena.

Objectives: This was a systematic review of the literature for the operating characteristics of emergency department (ED) ultrasonography for AAA.

Methods: The authors searched PubMed and EMBASE databases for trials from 1965 through November 2011 using a search strategy derived from the following PICO formulation: *Patients*—patients (18+ years) suspected of AAA. *Intervention*—bedside ED US to detect AAA. *Comparator*—reference standard for diagnosing an AAA was a computed tomography (CT), magnetic resonance imaging (MRI), aortography, official US performed by radiology, ED US reviewed by radiology, exploratory laparotomy, or autopsy results. AAA was defined as $\geq$3 cm dilation of the aorta. *Outcome*—operating characteristics (sensitivity, specificity, and likelihood ratios [LR]) of ED abdominal US. The papers were analyzed using Quality Assessment of Diagnostic Accuracy Studies (QUADAS) guidelines.

Results: The initial search strategy identified 1,238 articles; application of inclusion/exclusion criteria resulted in seven studies with 655 patients. The weighted average prevalence of AAA in symptomatic patients over the age of 50 years is 23%. On history, 50% of AAA patients will lack the classic triad of hypotension, back pain, and pulsatile abdominal mass. The sensitivity of abdominal palpation for AAA increases as the diameter of the AAA increases. The pooled operating characteristics of ED US for the detection of AAA were sensitivity 99% (95% confidence interval [CI] = 96% to 100%) and specificity 98% (95% CI = 97% to 99%).

Conclusions: Seven high-quality studies of the operating characteristics of ED bedside US in diagnosing AAA were identified. All showed excellent diagnostic performance for emergency bedside US to detect the presence of AAA in symptomatic patients.

ACADEMIC EMERGENCY MEDICINE 2013; 20:128–138

CLINICAL SCENARIO

A 60-year-old man presents to your emergency department (ED) with left flank and back pain that started 30 minutes ago and radiates to the left groin. The pain is making him nauseated and he is slightly diaphoretic. He has a history of hypertension, for which he takes a thiazide diuretic and a beta-blocker, and has smoked a half-pack of cigarettes per day for 25 years. He states his wife has had a kidney stone that presented similar to this and wonders if he is suffering from the same ailment. The heart rate is 90 beats/min, and the blood pressure is 109/70 mm hg. You order pain medication, a urinalysis, and a noncontrast computed tomography (CT) scan, when your senior resident reminds you that abdominal aortic aneurysms (AAA) are sometimes misdiagnosed as renal colic. A lively discussion among the house staff ensues and you recall that a hypertensive smoking male in his 60s is at risk for AAA and that your patient's blood pressure is in fact a little lower than you would expect, even if his hypertension is well controlled. Re-reviewing his vital signs you also recall that the patient's beta-blocker will blunt his tachycardic response to hemorrhage.

From the Department of Emergency Medicine, SUNY Downstate Medical Center, Brooklyn NY.
Received January 18, 2012; revision received March 29, 2012; accepted August 29, 2012.
The authors have no relevant financial information or potential conflicts of interest to disclose.
Supervising Editor: Christopher Carpenter, MD, MSc.
Address for correspondence and reprints: Ninfa Mehta, MD; Ninfa.mehta@gmail.com.

ISSN 1069-6563
PII ISSN 1069-6563583
doi: 10.1111/acem.12080

You are working at a busy urban ED; there are a number of medical and trauma patients awaiting CT scans. You perform a bedside ultrasound (US) to assess for an AAA; the abdominal aorta is well visualized from the proximal segments to the bifurcation. The aortic diameter measures less than 3 cm. You wonder if a bedside ED US would be sufficient to rule out an AAA while he awaits his CT scan.

INTRODUCTION

The prevalence of AAA ranges from 1.3% for men aged 45 to 54 years to up to 12.5% for men 75 to 84 years of age.[1–11] For women, the prevalence ranges from 0% in the youngest, to 5.2% in the oldest age groups.[1] Ruptured AAA causes approximately 9,000 deaths a year in the United States.[12] Although the exact number of missed AAAs is difficult to estimate, it is most frequently misdiagnosed as nephrolithiasis.[13] The use of US for the detection of an AAA has been well studied and validated in radiologic and surgical literatures. Lindholt et al.[14] have demonstrated 98.9% sensitivity and 99.9% specificity for detecting AAA (defined as $\geq$ 3 cm in diameter) in asymptomatic patients. The rate of misdiagnosis of AAA is reported to be 30% to 60% secondary to delays due to nonspecific clinical presentations that mimic AAA: renal colic, diverticulitis, and gastrointestinal hemorrhage.[15,16] This is especially important when considering the differential diagnosis because more than 80% of AAAs have not been previously diagnosed at the time of rupture.[16] The mortality rate of ruptured AAA approaches 90%.[17] Hoffman et al.[18] noted a large decrease in mortality in vascular surgery patients when AAAs were diagnosed and treated earlier. A retrospective review of AAA by Plummer et al.[19] demonstrated a more expedient diagnosis and better outcomes when ED US was used. These findings are compelling for developing an accelerated protocol in the ED using US for suspected AAA.

Bedside US has been incorporated into the training of emergency physicians (EPs).[20] Research has increasingly focused on the diagnostic accuracy of EPs using bedside US for the diagnosis of AAA. We systematically reviewed the literature for the diagnostic accuracy of ED US to rule out AAA in suspected patients.

METHODS

Search Strategy

The recommendations from the Meta-analysis Of Observational Studies in Epidemiology (MOOSE) statement were followed in this review.[21] A search strategy was derived from the following PICO formulation of our clinical question: Is ED-performed US sufficiently accurate to rule out an AAA in a suspected patient? *Patients*—patients (18+ years) suspected of having an AAA. *Intervention*—bedside ED US performed by EPs to detect AAA. *Comparator*—reference standard for diagnosing AAA was a CT, magnetic resonance imaging (MRI), aortography, ED US reviewed by radiology, or official US performed by radiology, exploratory laparotomy, or autopsy results. AAA was defined as $\geq$ 3 cm dilation of the aorta. *Outcome*—operating characteristics (sensitivity, specificity, and likelihood ratios [LR]) of ED abdominal US.

To be included in this review, prospective studies were required to have 1) bedside US performed by EPs, 2) enrollment of adult patients with symptoms/signs suggestive of AAAs, and 3) comparison/confirmation of results. We searched MEDLINE and EMBASE with the PubMed interface for articles from 1965 through November 2011 (see Appendix A for complete MEDLINE and EMBASE search strategies). We also searched the Cochrane Central Register of Controlled Trials and the Cochrane Review addressing the topic of emergency bedside US in the diagnosis of AAA. The searches were conducted with the assistance of a medical librarian. Review of the titles and abstracts of the search results were conducted independently by two authors (ER and NM) and disagreements were adjudicated by a third author (RS). Bibliographies of the included articles were also reviewed.

Individual Evidence Quality Appraisal

Two authors (ER and NM) independently extracted data from included studies; study quality was assessed using a validated tool for Quality Assessment of Diagnostic Accuracy Studies (QUADAS).[22,23] This tool was designed to assess studies for potential for bias, applicability, and quality of reporting. It consists of a 14-item checklist, each rated as yes, no, or unclear. Agreement between reviewers was assessed using kappa (SPSS version 18, SPSS, Inc., Chicago, IL).

Data Analysis

We defined "disease" as the presence of an AAA, ruptured or unruptured, and "no disease" as the absence of an AAA. A "true positive" was a diagnostic test that correctly identified an aortic diameter of $\geq$3 cm, whereas a "false positive" indicated an abnormal test result suggesting AAA or rupture when the criterion standard did not demonstrate one. Similarly, a "true-negative" test indicated the absence of an AAA when the criterion standard confirmed none, while a "false-negative" test suggested no AAA or rupture when in fact an AAA was identified by the criterion standard. In many AAA studies, presence of clot, rupture and free fluid, and other alterations of aortic anatomy are frequently cited as reasons for indeterminate scans. Indeterminate scans were coded as false positives because they trigger a need for further testing without deescalating fears of the presence of a life-threatening condition. This mimics real-life experience because the patient with an indeterminate scan would need a more definitive modality while mobilizing and expediting many (but not all) of the same resources as a positive scan would have.

Meta-analysis was conducted to obtain more precise estimates for diagnostic measures (sensitivity, specificity, LR, and 95% confidence intervals [CIs]) of ED abdominal US. A DerSimonian-Laird[24] random-effects model was used to combine studies, while accounting for variation among studies. The presence of statistical heterogeneity among the studies was assessed using the chi-square test, and the magnitude of heterogeneity was assessed using the I^2 statistic[25] using Meta-DiSc (Hospital Universitario Ramón y Cajal, Madrid, Spain).[26]

Test–Treatment Threshold

A recent Cochrane Review by Cosford and Leng[27] found a significant reduction in mortality (odds ratio [OR] = 0.60; 95% CI = 0.47 to 0.78) by US screening for AAA in asymptomatic men ages 65 to 79 years. This finding supports the recommendations of the U.S. Preventive Services Task Force that male patients greater than 65 years of age with a history of smoking should receive a one-time US screening for AAA even if asymptomatic. Unfortunately, the ED is not the appropriate venue to do this AAA US screening of all asymptomatic patients who meet the above criteria. The prevalence of asymptomatic AAA in men is between 5% and 10%.[28] The patients included in this meta-analysis were selected because they were symptomatic and therefore underwent sonography and CT scan. From the prevalence data of AAA, we feel that all patients with abdominal or flank pain age 50 years or older should be screened for AAA in the ED. Using a weighted average of prevalence from six studies, we calculated a pretest probability of an AAA to be 23% in symptomatic patients over the age of 50 years. We have chosen not to compute the test–treatment threshold since per the evidence-based diagnostics series recommendations, the intent of an emergency bedside US for AAA is to advance higher risk patients in how promptly they obtain definitive CT imaging.

Emergency bedside US for the evaluation of a suspected AAA in a hemodynamically stable patient is to prioritize for further imaging; if an AAA is detected, the patient would be expedited to CT or MRI. If an AAA is not detected by bedside US, further imaging will still need to be obtained. Since the mortality of ruptured AAA approaches 90%,[17] there is no combination of history and physical examination findings that are sensitive enough to rule out the diagnosis of AAA. In a hemodynamically unstable patient with a suspected AAA, the US finding of free intraperitoneal fluid should prompt emergent consultation for vascular surgery for definitive surgical treatment, even in the absence of a detectable aneurysm.

RESULTS

Our MEDLINE search returned 1,195 studies. EMBASE provided an additional 43 for a total of 1,238 studies. Search of the Cochrane trial registry and reviewed bibliographies did not return any additional relevant studies. Through a review of the titles and abstracts 1,193 studies were rejected for relevance. Forty-five articles were selected for in-depth full review. Twenty-seven articles were excluded because enrolled patients were non-ED, four papers were excluded because patients were asymptomatic,[29–32] four articles were excluded because they were not prospective studies,[15, 33–35] and two articles were excluded because there were no confirmatory diagnostic tests/procedures.[36, 37] One study by Lanoix et al.[38] was excluded because the data were republished in a later study, which was included in our review. Seven studies met our selection criteria[39–45] (Figure 1). The seven included studies were read in their entirety, and two authors (ER and NM) independently abstracted the data and each author's results were checked for accuracy by a third author (RS). All seven studies met our definition of high-quality, prospective studies, and comparison to a reference standard.

The papers by Kuhn et al.[40] and Rowland et al.[41] may have overlapping data, possibly leading to counting some patients twice when the two studies are pooled. The studies share many similar characteristics (overlapping enrollment dates, total patients entered, etc.). We were unable to contact the corresponding author for further clarification. We did, however, include these two studies in our analysis as separate studies.

Description of Included Studies

A full description of reviewed studies is shown in Table 1. Five of the seven studies reported the age of included patients: in Kuhn et al.,[40] Rowland et al.,[41] and Knaut et al.[44] study age was greater than 50 years; in Costantino et al.[45] study age was greater than 55 years; and in Tayal et al.[43] study mean age was reported as 66 years. Jones et al.[42] and Lanoix et al.[39] did not report

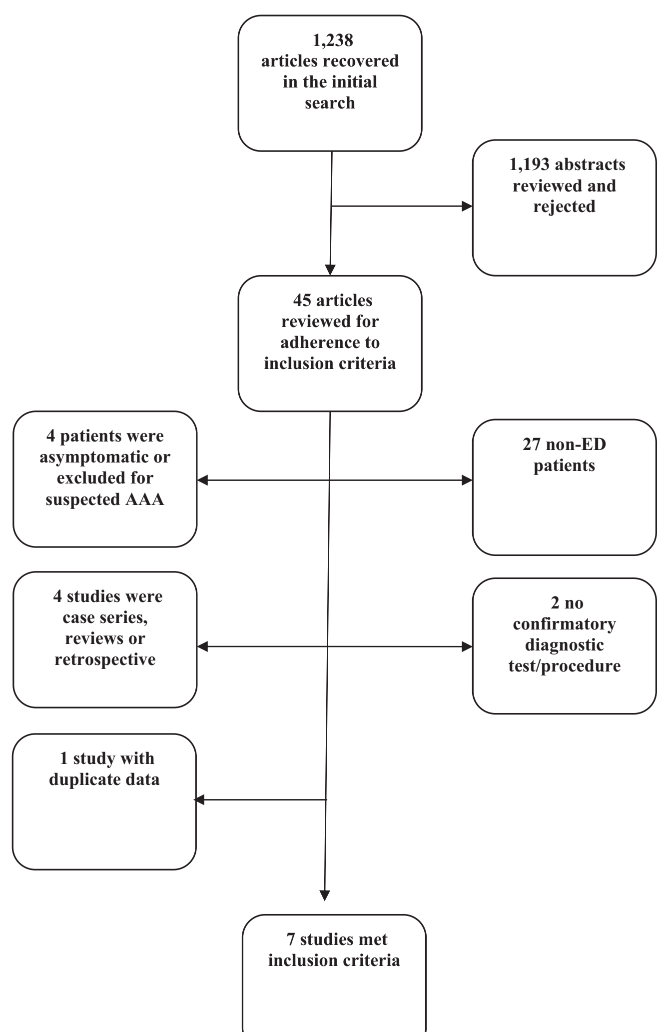

Figure 1. Flow of reviewed studies.

Table 1
Summarizing the Characteristics of the Seven Trials Under Review

Studies	Characteristics	Intervention	Operator/Interpreter	Reference Standard	Outcomes
Lanoix et al., 2000[39]	Inclusion criteria: • Pulsatile abdominal mass • Unexplained abdominal pain or hypotension Exclusion criteria: • None stated Sample size: N = 21	Bedside US machine: Diasonics DRF400 1986 Transducer: 3.5 MHz Cardiac (curved sector) probe Views: Not stated	• Thirty-nine EPs • No prior US experience but attended a 7-hour didactic	• Radiology review of ED images	• Positive for AAA defined as aortic diameter > 3 cm
Kuhn et al., 2000[40]	Inclusion criteria: • Patients suspected of AAA • Patients > 50 years • Abdominal/back pain of unclear origin or presumed renal colic Exclusion criteria: • Presence of an AAA had already been established by prior radiologic investigation Sample size: *N* = 68	Bedside US machine: Akola Flexus SSD 1100 Transducer: 3.5–5.0 MHz convex transducer Views: Aorta measured at region of maximal dilation in longitudinal and transverse plane	• EP with 3 or more years of postgraduate experience • No prior US experience but attended three day US training course prior to start of the study	• Radiologist US • Abdominal CT • Angiography • Laparotomy • Radiology review of ED images	• Positive for AAA defined as aortic diameter > 3 cm • Improvement in patient care
Rowland et al., 2001[41]	Inclusion criteria: • Patients > 50 years presenting with abdominal pain/back pain with unclear etiology or renal colic • Adult patient with pulsatile abdominal mass Exclusion criteria: • Pregnant women Sample size: *N* = 33	Bedside US machine: Akola SSD 1100 Transducer: 3.5–5.0 MHz convex transducer Views: Aorta measured at region of maximal dilation in longitudinal and transverse plane	• EP with 3 or more years of postgraduate experience • No prior US experience but attended three day US training course prior to start of the study	• Radiologist US • Abdominal CT • Angiography • Laparotomy • Radiology review of ED images	• Positive for AAA defined as aortic diameter > 3 cm • Evidence of leaking blood • Accuracy of ED US interpretation
Jones et al., 2003[42]	Inclusion criteria: • Not defined but examination was performed only if clinically indicated Exclusion criteria: • Not defined Sample size: *N* = 66	Bedside US Machine: B-K Medical Cheetah 2003 US system Transducer: Unknown Views: Three aortic views, longitudinal, proximal and distal, maximal transverse	• Three EPs or a critical care fellow • 4-hour US workshop	• Radiologist US • Abdominal CT • Laparotomy	• Positive for AAA • Accuracy of ED US interpretation
Tayal et al., 2003[43]	Inclusion criteria: • Adult patients suspected of AAA • Patients presenting with abdominal, flank and/or back pain, or syncope Exclusion criteria: • Patients with known diagnosis of AAA Sample size: *N* = 125	Bedside US machine: Shimadzu 400 and Shimadzu 450 gray-scale Transducer: low-frequency probe Views: Four aortic views proximal and distal in longitudinal and transverse plane	• Senior ED residents and board certified EPs • Minimum introductory emergency US education who performed at least 50 emergency US	• Radiologist US • Abdominal CT • Abdominal MRI • Laparotomy	• Positive for AAA defined as aortic diameter > 3 cm

Table 1
Summarizing the Characteristics of the Seven Trials Under Review

Studies	Characteristics	Intervention	Operator/Interpreter	Reference Standard	Outcomes
Knaut et al., 2005[44]	Inclusion criteria: • Patients > 50 years • Patients presenting with abdominal pain who were scheduled for abdominal/pelvic CT Exclusion criteria: • Patients with ruptured AAA who went to OR without CT scan being performed Sample size: *N* = 104	Bedside US machine: Toshiba 140-A Transducer: Unknown Views: Aortic diameter at SMA, bifurcation and longitudinal view	• ED second-, third-, and fourth-year residents and EPs • Annual formal didactic 5-hour training	• Aortic diameter from abdominal CT measured by two separate radiologist at SMA, bifurcation and longitudinal	• Difference in measured diameter at SMA, bifurcation and longitudinal view
Constantino et al., 2005[45]	Inclusion criteria: • Patients >55 years • Patients presenting with one of the following: abdominal, back, flank or chest pain or hypotension as well as clinical suggestion of AAA Exclusion criteria: • Patients with known diagnosis of AAA Sample size: *N* = 238	Bedside US machine: Sonosite 180plus or a Siemens Adara Transducer: 2.0–5.0 MHz abdominal probe Views: Entire abdominal aorta measured at maximum diameter	• Third-year EM residents • 23-day US rotation completing at least 150 emergency US	• Radiologist US • Abdominal CT • Abdominal MRI • Laparotomy	• Positive for AAA defined as aortic diameter > 3 cm • Evidence of other abdominal aortic abnormalities

AAA = abdominal aortic aneurysm; SMA = superior mesenteric artery; US = ultrasound.

Table 2
QUADAS Analysis

Item		Lanoix et al., 2000[39]	Kuhn et al., 2000[40]	Rowland et al., 2001[41]	Jones et al., 2003[42]	Tayal et al., 2003[43]	Knaut et al., 2005[44]	Costantino et al., 2005[45]
1	Was the spectrum of patients described in the paper and was it chosen adequately?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2	Were selection criteria described clearly?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
3	Is the reference standard likely to classify the target condition?	No	Yes	Yes	Yes	Yes	Yes	Yes
4	Was there an abnormally long time period between the performance of the test under evaluation and the confirmation of the diagnosis with the reference standard?	No	No	No	No	No	No	No
5	Did the whole sample, or a random selection of the sample, receive verification using a reference standard of diagnosis?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
6	Did all patients receive the same reference standard regardless of the index test result?	Yes	No	No	No	No	Yes	No
7	Were the results of the index test incorporated in the results of the reference standard?	No	No	No	No	No	No	No
8	Was the execution of the index test described in sufficient detail to permit replication of the test?	No	Yes	Yes	Yes	Yes	Yes	Yes
9	Was the execution of the reference standard described in sufficient detail to permit replication of the test?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
10	Were the index test results interpreted blind to the results of the reference standard?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
11	Were the reference standard results interpreted blind to the results of the index test?	Yes	Yes	Yes	Unclear	Unclear	Yes	Unclear
12	Was clinical data available when test results were interpreted?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
13	Were uninterpretable/indeterminate/intermediate results reported and included in the results?	No	Yes	Yes	Yes	Yes	Yes	Yes
14	Were reasons for drop-out from the study reported?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Kappa	1	1	1	1	1	1	1

QUADAS = Quality Assessment of Diagnostic Accuracy Studies.

Table 3
Operating Characteristics of US to Detect AAA

Studies	Sample size	AAA, *n* (%)	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)	LR– (95% CI)
Lanoix et al., 2000[39]	21	4 (19.0)	100	94.1 (71–100)	10.8 (2.3–51.4)	0
Kuhn et al., 2000[40]	68	26 (38.2)	100	95.2 (89–100)	20.9 (5.4–81.2)	0
Rowland et al., 2001[41]	33	12 (36)	100	100	∞	0
Jones et al., 2003[42]	66	40 (60.6)	97.5 (92.6–100)	100	∞	0.025 (0.004–0.173)
Tayal et al., 2003[43]	125	27 (21.6)	100	98 (95–100)	48.99 (12.43–193.16)	0
Knaut et al., 2005[44]	104	5 (4.8)	100	97 (94–100)	33.00 (10.8–100.5)	0
Costantino et al., 2005[45]	238	36 (15.1)	100	100	∞	0

AAA = abdominal aortic aneurysm; LR+ = positive likelihood ratio; LR– = negative likelihood ratio; US = ultrasound.

the patient ages. Sex distribution was only reported by Tayal et al.[43] and Knaut et al.[44] to be 51 and 54% men, respectively.

The level of training of the ED ultrasonographer varied from second-year residents in Knaut et al.,[44] to 3 years' postgraduate experience by Kuhn et al.[40] and

Rowland et al.[41] The level of US training also varied; five studies reported that EPs underwent US didactic courses and workshops ranging from 4 to 7 hours (Lanoix et al.,[39] Jones et al.,[42] and Knaut et al.)[44] to 3 days (Kuhn et al.[40] and Rowland et al.).[41] Tayal et al.[43] and Costantino et al.[45] reported that EPs had completed 50 to 150 emergency US scans prior to the start of the study. There was variation in the US machines used for the studies, as shown in Table 1.

Five of the seven studies used multiple modalities as their reference standards for comparison, such as CT scans, MRIs, radiology-performed US, operative reports, and autopsies. Knaut et al.[44] defined one modality, CT scan, as their reference standard. Lanoix et al.[39] used only the radiologist interpretation of the ED US as their reference standard.

The primary outcome for six of the seven studies was the detection of an AAA by ED US. The primary outcome of Knaut et al.[44] was the degree to which US measurements of aortic diameter by EPs agree with corresponding measurements obtained by CT scan; the detection of a AAA was a secondary outcome measurement. Sample size ranged between 21 patients[39] and 238 patients.[45]

METHODOLOGIC QUALITY EVALUATION

The methodologic quality of the included trials was tested using the QUADAS tool. The QUADAS results between the two authors (ER and NM) were checked for inter-rater reliability by a third author (RS). There was a 100% correlation between the two authors on the methodologic quality assessment of the included studies using the kappa statistic. Table 2 represents the adjudicated QUADAS results.

Blinding of the treating physician and criterion standard reviewer to the results of the diagnostic test prevents bias. Knowledge of the diagnostic test result by the treating physician can influence whether the criterion standard test is ordered, creating verification bias. Verification bias most commonly leads to an overestimation of sensitivity.[46] Knowledge of the diagnostic test result by the reviewer of the definitive test may influence his or her interpretation, resulting in test review bias.[47, 48] Test review bias could lead to a deceptive increase in the diagnostic test's performance.[46, 49, 50]

Kuhn et al.[40] specifically mention that the treating physician was only informed of the results of the bedside US if an unexpected AAA was found, a certain risk of verification bias. Verification bias was also found to be possible in the studies by Lanoix et al.,[39] Jones et al.,[42] Tayal et al.,[43] and Costantino et al.,[45] who did not give clear statements of blinding the US test results to the treating physicians. This does not seem to be the case in studies by Rowland et al.[41] and Knaut et al.,[44] which state explicitly that the referring physicians were blinded to the results of the US exams for AAA.

Lanoix et al.,[39] Rowland et al.,[41] Knaut et al.,[44] and Kuhn et al.[40] avoided test review bias by blinding of the US results to the radiologists reviewing the criterion standard exams. Jones et al.,[42] Tayal et al.,[43] and Costantino et al.[45] did not give descriptions of blinding of the interpreters of the criterion standard tests to the results of the US, leaving open the possibility of test review bias.

In six of the seven studies, the initial US images were re-reviewed by attending radiologists or experienced ED ultrasonographers. Only Knaut et al.[44] did not explicitly state any re-reads of initial US images. None of the studies that had their US images reinterpreted reported inter- or intra-rater reliability. None of the studies reported inter- or intra-rater reliability of the outcome CTs, MRIs, aortographies, or radiology-performed US studies.

The wide range of US experience among the operators/interpreters, coupled with the absence of inter- and intra-rater reliability in many of our reviewed studies, exposes these studies to unknown risks of observer variability. This is especially important in US studies where the skills of the operators in obtaining high-quality images may not be synonymous with the ability to accurately interpret the images.[51] High observer variability in either the index or reference tests has been shown to affect measures of diagnostic accuracy.[52, 53] The unstudied effect of US experience on the operating characteristics of US in AAA detection limits the generalizability, and thus the external validity, of many of our reviewed studies.

The identification of inadequate US scans was mentioned by Lanoix et al.,[39] Kuhn et al.,[40] and Jones et al.[42] Lanoix et al.[39] removed one indeterminate scan from their analysis; we recoded this scan as a false positive. Kuhn et al.[40] removed two indeterminate scans from their analysis; we recoded these scans as false positives in our analysis. Jones et al.[42] identified six inadequate scans; one scan was coded as a false negative and the remaining five were coded as true negatives. The removal of indeterminate scans from analysis disguises the difficulty in obtaining the images, and the misclassification of these images affects the veracity of the diagnostic performance of the diagnostic test.[54] For example, if in the study by Jones et al.[42] we used the most conservative approach to the data and recoded the five inadequate scans that were CT negative for AAA as false positive, the sensitivity would be unchanged but the specificity would decrease from 100% to 81%, with a decrease LR+ from infinity to 5.1 and LR– unchanged at 0.03.

Consecutive sampling was only stated by Tayal et al.[43] Convenience sampling was used by Lanoix et al.,[39] Kuhn et al.,[40] Rowland et al.,[41] and Knaut et al.[44] In two studies, Jones et al.[42] and Costantino et al.,[45] the authors failed to mention their sampling methodology. Since none of our reviewed studies used a rigorous sampling methodology, the potential for context bias exists[55] as evidenced by the wide range of observed AAA prevalence (4.8% to 61%) among our reviewed studies. Context bias occurs in studies with high prevalences of positive results where the examiners expect and usually find positive studies.[53] The sonographers would tend to lower their decision thresholds for calling AAAs, and this would tend to increase sensitivity and decrease specificity relative to what would be seen in an unselected group of patients.

DIAGNOSTIC PERFORMANCE OF BEDSIDE US

Table 3 represents the diagnostic performance of US for detection of AAA in our reviewed trials. The operating characteristics for the detection of AAA were as follows: sensitivity 97.5% to 100%, specificity 94.1% to 100%, LR+ 10.8 to ∞, and LR– 0.00 to 0.025. The prevalence of our studies' primary outcome for the detection of AAA had significant heterogeneity, ranging from 4.8%[44] to 61%.[42]

The studies reviewed have considerable heterogeneity in terms of operator training and experience, as well as the number of participating EPs, ranging from four to 40 EPs. However, statistical heterogeneity was assessed for the reviewed studies and was moderate (chi-square > 0.05 and $I^2 < 50\%$). The moderate heterogeneity allowed for pooled analysis of the operating characteristics. Figure 2 depicts the Forest plot and the pooled data: sensitivity 99.0% (95% CI = 96.0% to 100%) and specificity 99.0% (95% CI = 97.0% to 99.0%).

DISCUSSION

Returning to the patient from our clinical scenario, after reviewing the medical literature we feel confident, with a LR+ (10.8 to ∞) and LR– (0.00 to 0.025), in using US to rule out AAA. All seven studies showed excellent operating characteristics for both ruling in and ruling out AAA in the ED. This is consistent with previous studies in the diagnosis of AAA by a radiologist.[14] A study by Singh et al.[56] assessed the variability of ultrasonographic measurements at different levels of the abdominal aorta and concluded inexperienced sonographers might achieve acceptable performance given appropriate training and surveillance. All of the studies in this review reported limited experience with minimal training of the ED providers performing the bedside US, with favorable outcomes as well. Present-day emergency medicine residency programs incorporate US in the resident curriculum; however, many practicing physicians graduated before this integration. The American College of Emergency Physicians (ACEP) policy statement on US supports a practice-based pathway that "allows those emergency physicians not previously exposed to training in emergency ultrasound during residency to become proficient in utilizing this technology requiring didactic lessons, hands-on skill sessions, and a quality assurance program set up to review examinations at least until the physician has the ability to integrate this skill safely into clinical practice."[51] Still, there remains a question of how much training is required to be proficient at diagnosing AAA. A recent study by Hoffman et al.[32] suggests that credentialed ED sonographers with less than 3 years of experience were significantly less likely to identify AAA in asymptomatic patients than their more experienced colleagues. One must keep in mind that most of the literature evaluates the identification of AAA, whether the AAA is symptomatic or not. The identification of a ruptured AAA is a different endeavor in that retroperitoneal ruptures, which are not amenable to transabdominal sonographic visualization, are common, and there is also often a loss

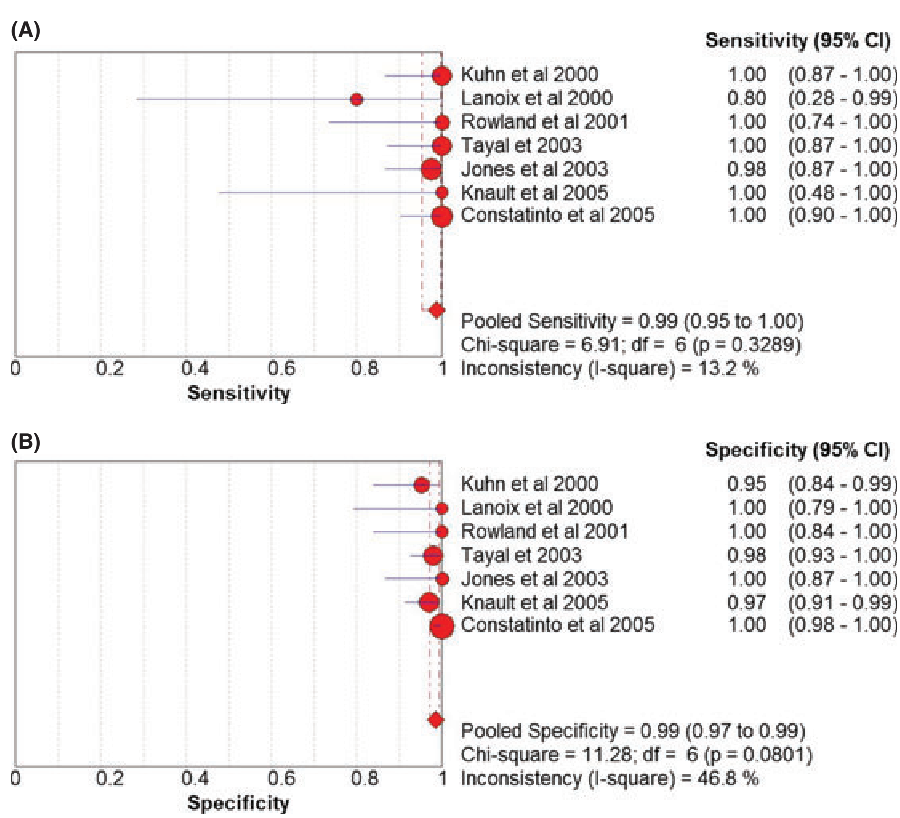

Figure 2. Diagnostic performance of bedside US for detection of AAA. (**A**) Sensitivity (95% CI); (**B**) specificity (95% CI). AAA = abdominal aortic aneurysm; US = ultrasound.

of integrity of the tubular structure the sonographer is trying to identify.[57] At this point, the only finding on sonography may be free fluid if rupture is peritoneal.

IMPLICATIONS FOR FUTURE RESEARCH

There is a need for an increase in outcome-based research such as randomized controlled trials to assess if the decreased time to diagnosis has an effect on AAA mortality and cost. Other areas of future investigation include optimal training time for proficiency in diagnosing AAA and the use of contrast-enhanced US for diagnosing leakage.

Future research should focus on properly designed trials using the Standards for the Reporting of Diagnostic accuracy studies (STARD) criteria.[58] This diagnostic research should examine history and physical examination findings in conjunction with bedside US findings to better understand the accuracy of these tests in isolation and in combination. This systematic review helps future researchers to understand the complementary roles of history, physical examination, and bedside US for the diagnosis of AAA.

Lack of inter-rater reliability and the effect of operator and interpreter experience on US evaluation is a deficiency common in US for diagnostic studies reviewed in the systematic review. Future US studies should be designed to include measurements of inter-rater reliability among EPs performing bedside US.

Indeterminate scans in this review were coded as false positives for reasons stated. The most conservative approach would be to code indeterminate scans as false negatives or "missed" disease. Future researchers can perform a sensitivity analysis that includes the indeterminate scans coded as false negatives and compare the recalculated sensitivity to the reported sensitivity. Any significant difference would be a limitation of that study and would need to be addressed.

LIMITATIONS

The literature search was limited to studies published in English and published manuscripts. There was a lack of assessment for inter- and intra-observer variability in all of the studies reviewed. Different US machines were used in each study, and resolution or image quality may have affected interpretation. The use of multiple reference standards, such as radiology US, intravenous pyelogram, or autopsy, sometimes within the same study, is also a limitation. Operator experience and training varied, further contributing to the heterogeneity of the studies. The unique limitations of the ED environment, such as multitasking, nonfasted patients, and sonographers with heterogeneous training and skill maintenance, are all factors that may affect the validity and reliability of estimates for the diagnostic accuracy of US to assess for AAA. There was also a wide range of prevalence of AAA across studies. It has been shown that increased body mass index decreases the sensitivity of the physical examination for the detection of AAA[59]; Elkouri et al.[60] also reported difficulty in obtaining adequate images in patients with increased body mass index. Only Rowland et al.[41] mentioned patients' body mass indices, but did not specifically analyze accuracy within the AAA subgroup.

CONCLUSIONS

The pretest probability of abdominal aortic aneurysm in symptomatic ED patients is 23%. History and physical examination are inaccurate assessment tools to diagnose abdominal aortic aneurysm. Emergency physician bedside ultrasound can be used to rule in (positive likelihood ratio 10.8 to infinity) or rule out (negative likelihood ratio 0 to 0.025) the need for emergent CT and/or vascular surgery consultation. Emergency bedside ultrasound can be used with great accuracy to detect the presence of abdominal aortic aneurysm in symptomatic patients.

The authors thank medical librarian Christopher Stewart for his help with the literature searches.

References

1. Hirsch AT, Haskal ZJ, Hertzer NR, et al. ACC/AHA 2005 Practice Guidelines for the management of patients with peripheral arterial disease (lower extremity, renal, mesenteric, and abdominal aortic): a collaborative report from the American Association for Vascular Surgery/Society for Vascular Surgery, Society for Cardiovascular Angiography and Interventions, Society for Vascular Medicine and Biology, Society of Interventional Radiology, and the ACC/AHA Task Force on Practice Guidelines (Writing Committee to Develop Guidelines for the Management of Patients With Peripheral Arterial Disease): endorsed by the American Association of Cardiovascular and Pulmonary Rehabilitation; National Heart, Lung, and Blood Institute; Society for Vascular Nursing; TransAtlantic Inter-Society Consensus; and Vascular Disease Foundation. Circulation. 2006; 113:e463.
2. Bengtsson H, Sonesson B, Bergqvist D. Incidence and prevalence of abdominal aortic aneurysms, estimated by necropsy studies and population screening by ultrasound. Ann N Y Acad Sci. 1996; 800:1–24.
3. Jamrozik K, Norman PE, Spencer CA, et al. Screening for abdominal aortic aneurysm: lessons from a population-based study. Med J Aust. 2000; 173:345–50.
4. Lederle FA, Johnson GR, Wilson SE, et al. The aneurysm detection and management study screening program: validation cohort and final results. Aneurysm Detection and Management Veterans Affairs Cooperative Study Investigators. Arch Intern Med. 2000; 160:1425–30.
5. Singh K, Bonaa KH, Jacobsen BK, et al. Prevalence of and risk factors for abdominal aortic aneurysms in a population-based study: the Tromso Study. Am J Epidemiol. 2001; 154:236–44.
6. Pleumeekers HJ, Hoes AW, van der Does E, et al. Aneurysms of the abdominal aorta in older adults. The Rotterdam Study. Am J Epidemiol. 1995; 142:1291–9.
7. Vazquez C, Sakalihasan N, D'Harcour JB, et al. Routine ultrasound screening for abdominal aortic

aneurysm among 65- and 75-year-old men in a city of 200,000 inhabitants. Ann Vasc Surg. 1998; 12: 544–9.
8. Boll AP, Verbeek AL, van de Lisdonk EH, et al. High prevalence of abdominal aortic aneurysm in a primary care screening programme. Br J Surg. 1998; 85:1090–4.
9. Wilmink AB, Quick CR. Epidemiology and potential for prevention of abdominal aortic aneurysm. Br J Surg. 1998; 85:155–62.
10. Takei H, Ishikawa S, Otaki A, et al. Screening for abdominal aortic aneurysm and occlusive peripheral vascular disease in Japanese residents. Surg Today. 1995; 25:608–11.
11. Adachi K, Iwasawa T, Ono T. Screening for abdominal aortic aneurysms during a basic medical checkup in residents of a Japanese rural community. Surg Today. 2000; 30:594–9.
12. Fleming C, Whitlock EP, Bell TL, Lederle FA. Screening for abdominal aortic aneurysm: a best-evidence systematic review for the U.S. Preventive Services Task Force. AHRQ Pub. No. 05-0569-B. 2005.
13. Borrero E, Queral LA. Symptomatic abdominal aortic aneurysm misdiagnosed as nephroureterolithiasis. Ann Vasc Surg. 1988; 2:145–9.
14. Lindholt JS, Vammen S, Juul S, Henneberg EW, Fasting H. The validity of ultrasonographic scanning as screening method for abdominal aortic aneurysm. Eur J Vasc Endovasc Surg. 1999; 17:472–5.
15. Akkersdijk GJ, van Bockel JH. Ruptured abdominal aortic aneurysm: initial misdiagnosis and the effect on treatment. Eur J Surg. 1998; 164:29–34.
16. Marston WA, Ahlquist R, Johnson G Jr, Meyer AA. Misdiagnosis of ruptured abdominal aortic aneurysms. J Vasc Surg. 1992; 16:17–22.
17. Ernst CB. Abdominal aortic aneurysm. N Engl J Med. 1993; 328:1167–72.
18. Hoffman M, Avellone JC, Plecha FR, et al. Operation for ruptured abdominal aortic aneurysms: a community-wide experience. Surgery. 1982; 91:597–602.
19. Plummer D, Clinton J, Matthew B. Emergency department ultrasound improves time to diagnosis and survival in ruptured abdominal aortic aneurysm [abstract]. Acad Emerg Med. 1998; 5:417.
20. Counselman FL, Sanders A, Slovis CM, Danzl D, Binder LS, Perina DG. The status of bedside ultrasonography training in emergency medicine residency programs. Acad Emerg Med. 2003; 10:37–42.
21. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA. 2000; 283:2008–12.
22. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol. 2003; 3:25.
23. Whiting PF, Weswood ME, Rutjes AW, Reitsma JB, Bossuyt PN, Kleijnen J. Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. BMC Med Res Methodol. 2006; 6:9.
24. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986; 7:177–88.
25. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003; 327:557–60.
26. Zamora J, Abraira V, Muriel A, Khan K, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. BMC Med Res Methodol. 2006; 6:31.
27. Cosford PA, Leng GC. Screening for abdominal aortic aneurysm. Cochrane Database Syst Rev. 2007; 2:CD002945.
28. Vardulaki KA, Prevost TC, Walker NM, et al. Incidence among men of asymptomatic abdominal aortic aneurysms: estimates from 500 screen detected cases. J Med Screen. 1999; 6:50–4.
29. Salen P, Melanson S, Buro D. ED screening to identify abdominal aortic aneurysms in asymptomatic geriatric patients. Am J Emerg Med. 2003; 21:133–5.
30. Moore CL, Holliday RS, Hwang JQ, Osborne MR. Screening for abdominal aortic aneurysm in asymptomatic at-risk patients using emergency ultrasound. Am J Emerg Med. 2008; 26:883–7.
31. Hoffmann B, Um P, Bessman ES, Ding R, Kelen GD, McCarthy ML. Routine screening for asymptomatic abdominal aortic aneurysm in high-risk patients is not recommended in emergency departments that are frequently crowded. Acad Emerg Med. 2009; 16:1242–50.
32. Hoffmann B, Bessman ES, Um P, Ding R, McCarthy ML. Successful sonographic visualisation of the abdominal aorta differs significantly among a diverse group of credentialed emergency department providers. Emerg Med J. 2011; 28:472–6.
33. Chu NH. Ruptured abdominal aortic aneurysm: a 7-year retrospective observational study in an emergency department of Hong Kong. Hong Kong J Emerg Med. 2010; 17:118–25.
34. Blaivas M, Theodoro D. Frequency of incomplete abdominal aorta visualization by emergency department bedside ultrasound. Acad Emerg Med. 2004; 11:103–5.
35. Walker A, Brenchley J, Sloan JP, Lalanda M, Venables H. Ultrasound by emergency physicians to detect abdominal aortic aneurysms: a UK case series. Emerg Med J. 2004; 21:257–9.
36. Schlager D, Lazzareschi G, Whitten D, Sanders AB. A prospective study of ultrasonography in the ED by emergency physicians. Am J Emerg Med. 1994; 12:185–9.
37. Dent B, Kendall RJ, Boyle AA, Atkinson PR. Emergency ultrasound of the abdominal aorta by UK emergency physicians: a prospective cohort study. Emerg Med J. 2007; 24:547–9.
38. Lanoix R, Baker WE, Mele JM, Dharmarajan L. Evaluation of an instructional model for emergency ultrasonography. Acad Emerg Med. 1998; 5:58–63.
39. Lanoix R, Leak LV, Gaeta T, Gernsheimer JR. A preliminary evaluation of emergency ultrasound in the setting of an emergency medicine training program. Am J Emerg Med. 2000; 18:41–5.
40. Kuhn M, Bonnin RL, Davey MJ, Rowland JL, Langlois SL. Emergency department ultrasound scanning for

abdominal aortic aneurysm: accessible, accurate, and advantageous. Ann Emerg Med. 2000; 36:219–23.
41. Rowland JL, Kuhn M, Bonnin RL, Davey MJ, Langlois SL. Accuracy of emergency department bedside ultrasonography. Emerg Med (Fremantle). 2001; 13:305–13.
42. Jones PG, Peak S, McClelland A, Holden A, Higginson I, Gamble G. Emergency ultrasound credentialing for focused assessment sonography in trauma and abdominal aortic aneurysm: a practical approach for Australasia. Emerg Med (Fremantle). 2003; 15:54–62.
43. Tayal VS, Graf CD, Gibbs MA. Prospective study of accuracy and outcome of emergency ultrasound for abdominal aortic aneurysm over two years. Acad Emerg Med. 2003; 10:867–71.
44. Knaut AL, Kendall JL, Patten R, Ray C. Ultrasonographic measurement of aortic diameter by emergency physicians approximates results obtained by computed tomography. J Emerg Med. 2005; 28:119–26.
45. Costantino TG, Bruno EC, Handly N, Dean AJ. Accuracy of emergency medicine ultrasound in the evaluation of abdominal aortic aneurysm. J Emerg Med. 2005; 29:455–60.
46. Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. N Engl J Med. 1978; 299:926–30.
47. Loy CT, Irwig L. Accuracy of diagnostic tests read with and without clinical information: a systematic review. JAMA. 2004; 292:1602–9.
48. Worster A, Carpenter C. Incorporation bias in studies of diagnostic tests: how to avoid being biased about bias. Can J Emerg Med. 2008; 10:174–5.
49. Begg CB, McNeil BJ. Assessment of radiologic tests: control of bias and other design considerations. Radiology. 1988; 167:565–9.
50. Kelly S, Berry E, Roderick P, et al. The identification of bias in studies of the diagnostic performance of imaging modalities. Br J Radiol. 1997; 70:1028–35.
51. American College of Emergency Physicians. Policy Statement Emergency Ultrasound Guidelines. 2008. Available at: www.acep.org/workarea/downloadasset.aspx?id=32878. Accessed Nov 9, 2012.
52. Quinn MF. Relation of observer agreement to accuracy according to a two-receiver signal detection model of diagnosis. Med Decis Making. 1989; 9:196–206.
53. Reid MC, Lachs MS, Feinstein AR. Use of methodological standards in diagnostic test research. Getting better but still not good. JAMA. 1995; 274:645–51.
54. Begg CB, Greenes RA, Iglewicz B. The influence of uninterpretability on the assessment of diagnostic tests. J Chronic Dis. 1986; 39:575–84.
55. Egglin TK, Feinstein AR. Context bias. A problem in diagnostic radiology. JAMA. 1996; 276:1752–5.
56. Singh K, Bonaa KH, Solberg S, Sorlie DG, Bjork L. Intra- and interobserver variability in ultrasound measurements of abdominal aortic diameter. The Tromso Study. Eur J Vasc Endovasc Surg. 1998; 15:497–504.
57. Lederle FA. In the clinic. Abdominal aortic aneurysm. Ann Intern Med. 2009; 150:ITC5–1.
58. Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. Ann Intern Med. 2003; 138:40–44.
59. Fink HA, Lederle FA, Roth CS, Bowles CA, Nelson DB, Haas MA. The accuracy of physical examination to detect abdominal aortic aneurysm. Arch Intern Med. 2000; 160:833–6.
60. Elkouri S, Panneton JM, Andrews JC, et al. Computed tomography and ultrasound in follow-up of patients after endovascular repair of abdominal aortic aneurysm. Ann Vasc Surg. 2004; 18:271–9.

APPENDIX A

SEARCH TERMS

“aortic aneurysm, abdominal”[MeSH Terms] OR (“aortic”[All Fields] AND “aneurysm”[All Fields] AND “abdominal”[All Fields]) OR “abdominal aortic aneurysm”[All Fields] OR (“abdominal”[All Fields] AND “aortic”[All Fields] AND “aneurysm”[All Fields]) and AND “ultrasonography”[Subheading] OR “ultrasonography”[All Fields] OR “ultrasound”[All Fields] OR “ultrasonography”[MeSH Terms] OR “ultrasound”[All Fields] OR “ultrasonics”[MeSH Terms] OR “ultrasonics”[All Fields].