

# Systematic Review of Emergency Physician–performed Ultrasonography for Lower-Extremity Deep Vein Thrombosis

Patrick R. Burnside, MD, Michael D. Brown, MD, MSc, Jeffrey A. Kline, MD

---

## Abstract

**Objectives:** The authors performed a systematic review to evaluate published literature on diagnostic performance of emergency physician–performed ultrasonography (EPPU) for the diagnosis and exclusion of deep venous thrombosis (DVT).

**Methods:** Structured search criteria were used to query MEDLINE and EMBASE, followed by a hand search of published bibliographies. Relevance and inclusion criteria required prospective investigation of emergency department (ED) outpatients with suspected DVT; diagnostic evaluations had to consist of EPPU followed by criterion standard (radiology-performed) imaging. Two authors independently extracted data from included studies; study quality was assessed utilizing a validated tool for quality assessment of diagnostic accuracy studies (QUADAS). Pooled data were analyzed using an unweighted summary receiver-operating-characteristic (SROC) curve; sensitivity and specificity were estimated using a random effects model.

**Results:** The initial search yielded 1,162 publications. Relevance screening and selection yielded six articles including 936 patients. Four of the six studies reported adequate blinding but a number of other methodologic flaws were identified. A random effects model yielded an overall sensitivity of 0.95 (95% confidence interval [CI] = 0.87 to 0.99) and specificity of 0.96 (95% CI = 0.87 to 0.99).

**Conclusions:** Systematic review of six studies suggests that EPPU may be accurate for the diagnosis of DVT compared with radiology-performed ultrasound (US). However, given the methodologic limitations identified among the primary studies, the estimates of diagnostic test performance may be overly optimistic. Further research into EPPU for suspected DVT is needed before it can be adopted into routine clinical practice.

ACADEMIC EMERGENCY MEDICINE 2008; 15:493–498 © 2008 by the Society for Academic Emergency Medicine

**Keywords:** deep vein thrombosis, ultrasound, systematic review, emergency department

---

**D**eep venous thrombosis (DVT) causes death and disability in ambulatory patient populations.<sup>1,2</sup> Approximately 200,000 outpatients each year are diagnosed with DVT, and experts estimate that many more outpatients have DVT that is not diagnosed.<sup>2</sup> At least one-third of patients with untreated DVT also experience clinically significant pulmonary embolism (PE), and the short-term mortality rate from untreated PE

probably exceeds 20%.<sup>3</sup> Emergency medicine (EM) practitioners have become accustomed to using the D-dimer assay as a method to exclude DVT; however, the D-dimer only allows exclusion of DVT in fewer than one-half of patients with suspected DVT, and the D-dimer cannot confirm the diagnosis of DVT.<sup>4,5</sup> Venous ultrasonography remains the main comprehensive modality to evaluate for DVT, but at present, venous ultrasonography requires significant time, patient transport, and availability of a medical sonographer and radiologist to perform and interpret.

The majority of EM residencies in the United States train their residents to use ultrasound (US) in the emergency department (ED).<sup>6</sup> Several original research reports have focused on emergency physician–performed US (EPPU) to evaluate for lower-extremity DVT.<sup>7</sup> Hence, we pursued the following research question: What is the accuracy of EPPU for suspected lower-extremity DVT?

---

From the Department of Emergency Medicine, Carolinas Medical Center (PRB, JAK), Charlotte, NC; and the Department of Emergency Medicine, Grand Rapids Medical Education & Research Center/Michigan State University (MDB), Grand Rapids, MI.

Received December 3, 2007; revision received January 18, 2008; accepted January 25, 2008.

Address for correspondence and reprints: Jeffrey A. Kline, MD; e-mail: jkline@carolinas.org.

A related commentary appears in the June issue.

## METHODS

### Search Techniques

Two physicians (PRB, JAK) independently performed structured searches of MEDLINE, and a librarian searched EMBASE, including in-process and other nonindexed citations (January 1988 to December 2007). The Medical Subject Headings (MeSH) string utilized was (ultrasonography) AND (venous thrombosis) with both terms exploded followed by free text searches including the terms "ultrasound" and "venous thrombosis." The search was limited to English language and human subjects.

We also pursued data from research studies that were not published as full-length articles. We conducted online bibliographic searches of abstract submissions to *Academic Emergency Medicine* and *Annals of Emergency Medicine* from January 1994 to July 2007. The query string used to search both publications was "ultrasound" AND "venous thrombosis." Additionally, we read all bibliographies of studies that passed relevance screening and attempted personal communication with authors to further pursue unpublished studies.

### Study Selection

Two independent reviewers read all abstracts for relevance. Criteria for relevance required all of the following: 1) original research reports of ED patients, 2) patient population with signs and symptoms suggestive of DVT, 3) venous US performed by nonradiology personnel, and 4) second venous US performed by a radiology department or vascular laboratory.

### Inclusion Criteria

We obtained and reviewed full-length reprints of all studies that met relevance criteria for inclusion. For inclusion, studies had to report on a prospective sample of predominantly outpatients (>50%). Patients could self-refer or could have been referred from another medical facility. We specifically sought to exclude studies performed solely for DVT surveillance in high-risk populations; hence, studies had to specify that patients manifested clinical findings suggestive of DVT as the basis for enrollment. The "diagnostic test" had to be an US of one or both legs, performed by an emergency physician (EP). The criterion standard required a second US to be performed by a US technician with images interpreted by a radiologist or vascular physician sonographer.

### Final Inclusion

Following the relevance search, two reviewers (JAK, PRB) compared exclusion logs for discordance, reaching consensus by conference. Systematic data extraction was completed via a predesigned data collection sheet (available as an online Data Supplement at <http://www.blackwell-synergy.com/doi/suppl/10.1111/j.1553-2712.2008.00101.x>) for studies meeting relevance screen and inclusion criteria. Authors were individually contacted as needed for data and inclusion criteria clarification. If it was determined after full article review and clarification from the authors that a study was retro-

spective, the study was excluded. Two reviewers (JAK, MDB) independently confirmed numeric calculations.

### Quality Assessment

We elected to include higher quality studies by using a few key quality measures as inclusion criteria; as described under "Inclusion Criteria," studies had to include an appropriate patient spectrum and use an acceptable reference standard.<sup>8</sup> We then graded each study based on adequate blinding. Grade A defined prospective studies in which the EP US performer was blinded to the criterion standard. Grade B defined studies wherein blinding measures were not explicitly stated or not performed. A recently validated tool for quality assessment of diagnostic accuracy studies (QUADAS)<sup>8,9</sup> was subsequently applied by two independent observers. Final agreement was reached by consensus regarding potential study limitations.

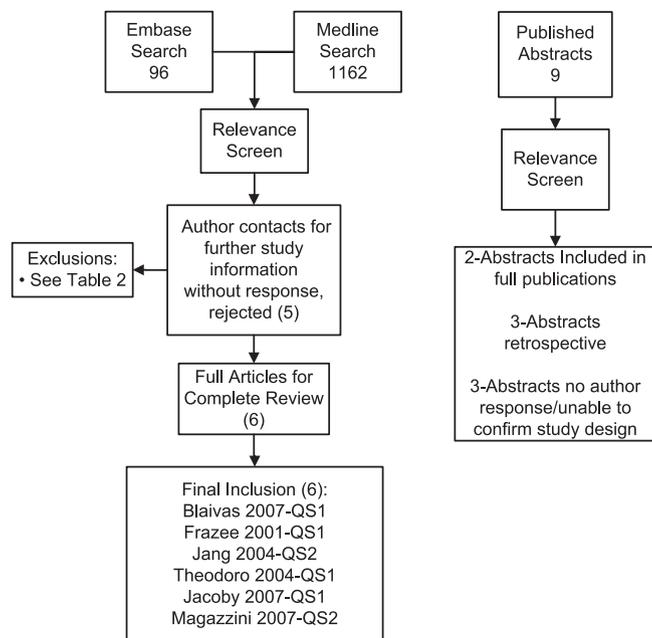
### Data Analysis

Agreement between reviewers was assessed with Cohen's kappa ( $\kappa$ ). Study US performance was assessed using summary receiver-operating-characteristic (SROC) curve analysis, pooled diagnostic odds ratios (DORs), and pooled sensitivity and specificity values.<sup>10</sup> The SROC curve analysis was based on an unweighted least-squares regression model, which has been fully described previously.<sup>10-14</sup> A random effects model was used to pool estimates; a correction factor of 0.05 was added to each cell. The SROC curve analysis was performed using Meta-Test (Version 0.9, Tufts-New England Medical Center, Boston, MA) and dr-ROC (Version 2.0, Diagnostic Research Design & Reporting, Glenside, PA) software. Given the questionable validity of using funnel plots or statistical models to detect publication bias for diagnostic test meta-analysis, we did not formally test for the presence of publication bias.<sup>15,16</sup>

## RESULTS

Figure 1 shows the results of the search and article selection. The Medline search yielded 1,162 titles. After screening the abstracts for relevance criteria, 1,156 were excluded from further review. Most of these were publications of traditional US performed by radiology. Five studies had insufficient abstract information, namely, the specification of the training of the study sonographers. The listed corresponding authors of these five studies were contacted via e-mail and/or telephone, but none responded. Thirteen publications reported that the US was performed by EPs; 7 of these 13 were rejected for reasons stated in Table 1. Six articles remained eligible for SROC curve analysis (Figure 1). Search of EMBASE yielded the 6 identified articles, and no further publications were eligible by relevance criteria. Comparison of agreement between two independent reviewers for the relevance screen results yielded a Cohen's  $\kappa = 0.65$  (95% confidence interval [CI] = 0.50 to 0.78).

Search of published non-full-length articles (i.e., abstracts) yielded eight potential studies. Two abstracts were part of the above full publication articles (eligible



**Figure 1.** Flow diagram for study selection (updated Week 4, December 2007).

**Table 1**  
Primary Reasons for Study Exclusion (n = 1,156)

Unable to discern US study operator or no author response	5
Comments/letters to the editor	2
US use consisted of Doppler stethoscope	1
Clinical policy statements/practice guideline	2
Case reports	2
Study sonographer was "radiology-based" sonographer	1,144
<b>Total</b>	<b>1,156</b>

US = ultrasound.

for full review), three abstracts met the relevance criteria, and three abstracts had insufficient information to ascertain study methods. These three authors did not respond to repeated e-mail inquiries regarding their methods. Personal communications with experts in the emergency US realm did not yield studies beyond those identified in the aforementioned strategies. After correspondence with the authors of the three abstracts that passed the relevance screen, quality assessment revealed retrospective designs; these articles were thus excluded.

Table 2 shows quality scoring for each of the six included studies<sup>17-22</sup> and summaries of extracted data, with four of the studies reporting adequate blinding (Grade A). Also within Table 2 are potential limitations identified using the QUADAS tool. Given the presence of statistical heterogeneity and the imprecision of all these estimates, caution must be taken with interpretation.

**Study Descriptions**

The prevalence of DVT within the six studies ranged from 7% to 32%. Gender and age information was provided in only two trials. Authors defined a positive result in all studies as the inability to compress the common femoral vein or popliteal vein as demonstrated by gray scale B-mode ultrasonography. Of the six studies, two included routine color flow and one had discretionary use of color flow and augmentation techniques. Studies that included indeterminate and equivocal findings classified them as a positive test. Within each study, the mean number of EPs who performed test US was 5.3 (range 2-8).

**Statistical Analysis**

The sensitivity and specificity of each included study were calculated with the 95% CI displayed (Figure 2). The pooled summary estimate using a random effects model produced a sensitivity of 0.95 (95% CI = 0.87 to 0.99) and a specificity of 0.96 (95% CI = 0.87 to 0.99). The  $\beta$  (slope) of 0.33 (95% CI = -2.5 to 3.1) indicated no

**Table 2**  
Study Characteristics and Diagnostic Test Performance

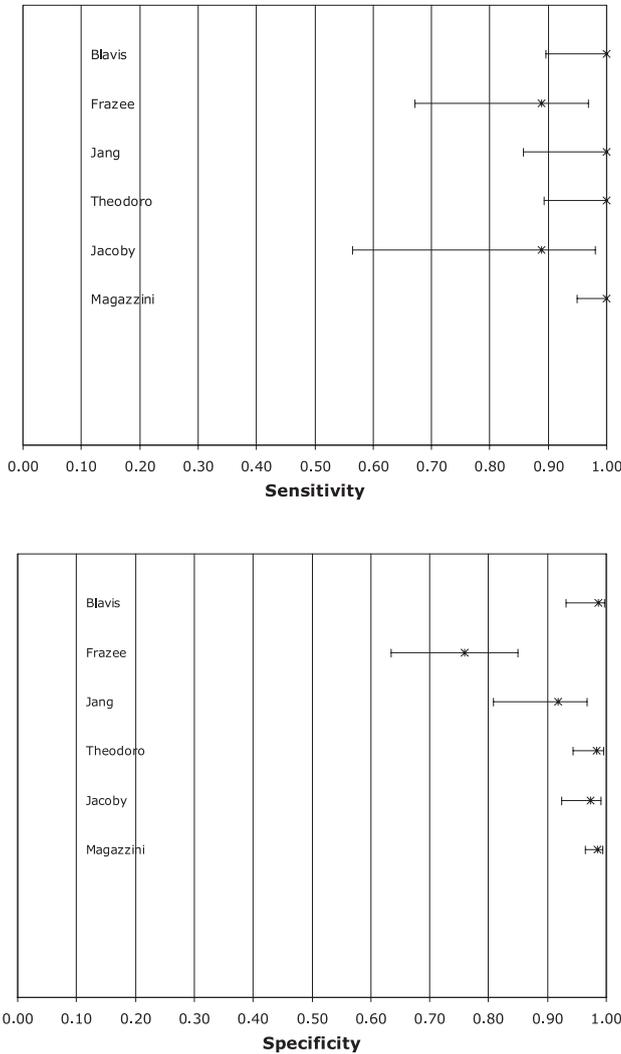
Study (year)	DVT+/sample size (%)	Sensitivity (95% CI)	Specificity (95% CI)	Quality Grade	Potential Limitations* (QUADAS)
Blaivas (2000) <sup>17</sup>	33/112 (30)	1.0 (0.89, 1.0)	0.99 (0.92, 1.0)	A	a, f, g
Frazee (2001) <sup>18</sup>	18/76 (24)	0.89 (0.64, 0.98)	0.76 (0.63, 0.86)	A	b
Jang (2004) <sup>19</sup>	23/72 (32)	1.0 (0.85, 1.0)	0.92 (0.80, 0.97)	B	c, d, e, f, h
Theodoro (2004) <sup>20</sup>	32/156 (21)	1.0 (0.89, 1.0)	0.98 (0.94, 1.0)	A	f
Jacoby (2007) <sup>21</sup>	9/121 (7)	0.89 (0.51, 0.99)	0.97 (0.92, 0.99)	A	a, b, i, j
Magazzini (2007) <sup>22</sup>	72/399 (18)	1.0 (0.95, 1.0)	0.98 (0.96, 0.99)	B	b, h
<b>Pooled results†</b>	<b>132/936 (14)</b>	<b>0.95 (0.87, 0.99)</b>	<b>0.96 (0.87, 0.99)</b>		

CI = confidence interval; DVT = deep venous thrombosis; QUADAS = quality assessment of diagnostic accuracy studies; US = ultrasound.

a = sample possibly not representative of typical emergency department US population; b = selection criteria not clearly described; c = disease progression bias possible; d = inconsistent application of reference standard; e = inconsistent reference standard; f = reference standard insufficiently described; g = reference test result not specifically stated as blinded to performer of index test; h = index test result not specifically stated as blinded to performer of reference standard; i = clinical data available during performance of index test not consistent with typical information available during test application in routine practice; j = indeterminate results not reported.

\*Limitations determined by the QUADAS tool.<sup>8</sup>

†Using random effects model.

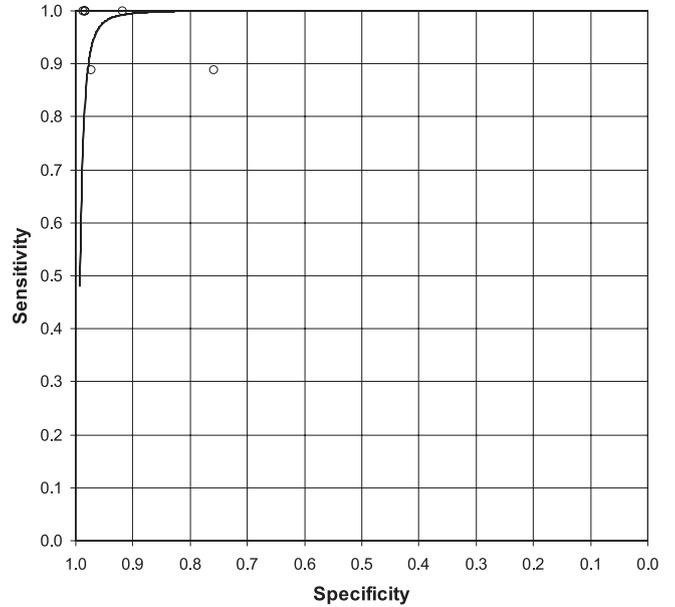


**Figure 2.** Sensitivity and specificity plots for emergency physician-performed ultrasound (EPPUS) to detect deep venous thrombosis (DVT) using radiologist-performed US as the reference standard. \*Point estimates; 95% confidence intervals (CIs) are displayed as horizontal bars.

significant threshold effect. Statistical testing for heterogeneity indicated substantial variability among the results ( $I^2 = 77\%$ ) but this variability is difficult to appreciate upon visual inspection of the SROC curve (Figure 3). Although the pooled DOR was 591 (95% CI = 70 to 4,940), caution must be taken with interpretation given the presence of statistical heterogeneity and lack of precision.

**DISCUSSION**

We believe that this report is the first systematic review of EPPU for DVT. In view of the fact that most residencies in EM incorporate US as a required curriculum, the results have relevance to the academic EM community. Although our SROC curve analysis suggests the high potential value for EPPU, the estimates are imprecise due to the small sample size. More



**Figure 3.** Summary receiver-operating-characteristic (SROC) curve analysis ( $n = 6$ ) for emergency physician-performed ultrasound (EPPU) to detect deep venous thrombosis (DVT) using radiologist performed US as the reference standard. Individual studies are depicted as circles. The unweighted SROC curve is limited to the range where data are available.

importantly, we found a number of methodologic issues that raise caution regarding the validity and generalizability of these results. First, the low number of EP sonographers and their high level of expertise raise concern. The ultrasonography skills of EPs at academic medical centers where EM US research is conducted likely exceed the capabilities of most EPs in community practice.<sup>23</sup> Second, the lack of details regarding patient enrollment methods and patient clinical characteristics makes it difficult to compare patients enrolled in these studies with other populations. Additionally, our six included studies provided little information about the anatomic location of the DVTs identified and do not address the issue of missed calf vein thrombosis. Current literature demonstrates a growing concern over the clinical significance of calf vein DVT that may not be detected by our criterion standard of radiology-performed US.<sup>24-26</sup>

In this report, we essentially quantify the discordance between EPPU and formal ultrasonography. We can speculate on the influences that may cause discordance between the results of EPPU for DVT versus radiology technician-performed, radiologist-interpreted venous ultrasonography. These influences might include US equipment, technique, patient location, experience, and the impact of preexisting knowledge of the overall clinical picture held by the EP, which may not be available to the sonographer and radiologist.

Our data show promise for EPPU as a method to evaluate for DVT. However, the limitations in the literature comprising this systematic review point toward the need for further research before widespread adoption of EPPU for DVT.

## LIMITATIONS

We performed a comprehensive search but did not include foreign language studies. Although we attempted to contact authors to further clarify and describe data and methods, the response rate was low, which limited our ability to adequately perform data extraction and quality assessment.

The criterion standard we used was a second US. Thus, to some degree, the SROC results represent agreement data, as opposed to a measurement of diagnostic test performance compared against clinical follow-up or an alternative imaging modality. Only two studies included patient follow up, one at 30 days<sup>22</sup> and the other at 1 year.<sup>19</sup> This lack of clinical follow-up among the majority of included studies may have resulted in significant measurement bias. The number of false-negative studies may have been underestimated, since it is unknown how many patients with DVT may have been missed by both EPPU and the radiology department-based US. Estimates for test sensitivity may have been further biased given that only 25% of the included studies reported that radiology department-based sonographers either were registered diagnostic medical sonographer (RDMS) credentialed or had specifically defined expertise. However, few EP sonographers are RDMS-certified, and the minority have sufficiently documented US examinations to meet the suggested training guidelines.<sup>23</sup>

The quality assessment of individual studies was limited since none of the six studies provided detailed descriptions of the study populations, and three studies provided no clinical information whatsoever. The limitations of using quality scores have been well described; therefore, the quality grade was only used to provide a summary of the blinding methodology reported in each included publication, rather than attempting to use scores in a weighted regression analysis or sensitivity analysis.<sup>1,3,27</sup> Finally, given the low number of EP examiners (range 2–8) among the six studies, the external validity of this systematic review is limited.

We did not use funnel plots or statistical models to detect publication bias since there is a lack of empirical evidence validating the use of these methods for diagnostic test meta-analysis.<sup>15,16</sup> Even if we were to use the recently recommended effective sample size funnel plot, the power to detect publication bias is low when there is significant heterogeneity among the DORs.<sup>28</sup> If publication bias was present, it would be anticipated that our estimates for test sensitivity and specificity may be inflated.

## CONCLUSIONS

Based on the results of six studies comprising 936 patients, the overall sensitivity and specificity of EPPU for DVT appear to be excellent. However, we identified a number of potential study biases that suggest the need for a properly designed study that includes a larger numbers of EP ultrasonographers with fully described methods of patient selection and clinical follow-up to assess clinically important outcomes.

## References

1. Heit JA. Venous thromboembolism: disease burden, outcomes and risk factors. *J Thromb Haemost.* 2005; 3:1611–7.
2. Courtney DM, Kline JA. Identification of prearrest clinical factors associated with outpatient fatal pulmonary embolism. *Acad Emerg Med.* 2001; 8:1136–42.
3. Calder KK, Herbert M, Henderson SO. The mortality of untreated pulmonary embolism in emergency department patients. *Ann Emerg Med.* 2005; 45:302–10.
4. Snow V, Qaseem A, Barry P, et al. Management of venous thromboembolism: a clinical practice guideline from the American College of Physicians and the American Academy of Family Physicians. *Ann Intern Med.* 2007; 146:204–10.
5. Stein PD, Hull RD, Patel KC, et al. D-dimer for the exclusion of acute venous thrombosis and pulmonary embolism: a systematic review. *Ann Intern Med.* 2004; 140:589–602.
6. Moore CL, Molina AA, Lin H. Ultrasonography in community emergency departments in the United States: access to ultrasonography performed by consultants and status of emergency physician-performed ultrasonography. *Ann Emerg Med.* 2006; 47:147–53.
7. Kendall JL, Hoffenberg SR, Smith RS. History of emergency and critical care ultrasound: the evolution of a new imaging paradigm. *Crit Care Med.* 2007; 35(5 Suppl):S126–S130.
8. 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.
9. 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.
10. Dinnes J, Deeks J, Kirby J, Roderick P. A methodological review of how heterogeneity has been examined in systematic reviews of diagnostic test accuracy. *Health Technol Assess.* 2005; 9:1–113.
11. Irwig L, Macaskill P, Glasziou P, Fahey M. Meta-analytic methods for diagnostic test accuracy. *J Clin Epidemiol.* 1995; 48:119–130.
12. Jones AE, Fiechtel JF, Brown MD, Ballew JJ, Kline JA. Procalcitonin test in the diagnosis of bacteremia: a meta-analysis. *Ann Emerg Med.* 2007; 50:34–41.
13. Mitchell AM, Brown MD, Menown IB, Kline JA. Novel protein markers of acute coronary syndrome complications in low-risk outpatients: a systematic review of potential use in the emergency department. *Clin Chem.* 2005; 51:2005–12.
14. Moses LE, Shapiro D, Littenberg B. Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations. *Stat Med.* 1993; 12:1293–316.

15. Tatsioni A, Zarin DA, Aronson N, et al. Challenges in systematic reviews of diagnostic technologies. *Ann Intern Med.* 2005; 142:1048–55.
16. Terrin N, Schmid CH, Lau J. In an empirical evaluation of the funnel plot, researchers could not visually identify publication bias. *J Clin Epidemiol.* 2005; 58:894–901.
17. Blaivas M, Lambert MJ, Harwood RA, Wood JP, Konicki J. Lower-extremity Doppler for deep venous thrombosis—can emergency physicians be accurate and fast? *Acad Emerg Med.* 2000; 7:120–6.
18. Frazee BW, Snoey ER, Levitt A. Emergency Department compression ultrasound to diagnose proximal deep vein thrombosis. *J Emerg Med.* 2001; 20: 107–12.
19. Jang T, Docherty M, Aubin C, Polites G. Resident-performed compression ultrasonography for the detection of proximal deep vein thrombosis: fast and accurate. *Acad Emerg Med.* 2004; 11:319–22.
20. Theodoro D, Blaivas M, Duggal S, Snyder G, Lucas M. Real-time B-mode ultrasound in the ED saves time in the diagnosis of deep vein thrombosis (DVT). *Am J Emerg Med.* 2004; 22:197–200.
21. Jacoby J, Cesta M, Axelband J, Melanson S, Heller M, Reed J. Can emergency medicine residents detect acute deep venous thrombosis with a limited, two-site ultrasound examination? *J Emerg Med.* 2007; 32:197–200.
22. Magazzini S, Vanni S, Toccafondi S, et al. Duplex ultrasound in the emergency department for the diagnostic management of clinically suspected deep vein thrombosis. *Acad Emerg Med.* 2007; 14:216–20.
23. Moore CL, Gregg S, Lambert M. Performance, training, quality assurance, and reimbursement of emergency physician-performed ultrasonography at academic medical centers. *J Ultrasound Med.* 2004; 23:459–66.
24. Lagerstedt CI, Olsson CG, Fagher BO, Oqvist BW, Albrechtsson U. Need for long-term anticoagulant treatment in symptomatic calf-vein thrombosis. *Lancet.* 1985; 2:515–8.
25. Pinede L, Cucherat M, Duhaut P, Ninet J, Boissel JP. Optimal duration of anticoagulant therapy after an episode of venous thromboembolism. *Blood Coagul Fibrinolysis.* 2000; 11:701–7.
26. Schwarz T, Schmidt B, Beyer J, Schellong SM. Therapy of isolated calf muscle vein thrombosis with low-molecular-weight heparin. *Blood Coagul Fibrinolysis.* 2001; 12:597–9.
27. Whiting P, Harbord R, Kleijnen J. No role for quality scores in systematic reviews of diagnostic accuracy studies. *BMC Med Res Methodol.* 2005; 5:19.
28. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol.* 2005; 58: 882–93.

### Supplementary Material

The following supplementary material is available for this article:

**Data Supplement S1.** Data collection form.

This material is available as part of the online article from: <http://www.blackwell-synergy.com/doi/abs/10.1111/j.1553-2712.2008.00101.x>

(This link will take you to the article abstract).

Please note: Blackwell Publishing are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

### Where to Find AEM Instructions for Authors

For complete instructions for authors, see the January or July issue of *Academic Emergency Medicine*; visit <http://www.blackwellpublishing.com/aemj> and click on “Author Guidelines”; or contact SAEM via e-mail at [aem@saem.org](mailto:aem@saem.org), via phone at 517-485-5484, or via fax at 517-485-0801.