

ORIGINAL ARTICLE

Accuracy of diagnostic tests for clinically suspected upper extremity deep vein thrombosis: a systematic review

M. DI NISIO,* G. L. VAN SLUIS,† P. M. M. BOSSUYT,‡ H. R. BÜLLER,† E. PORRECA* and A. W. S. RUTJES*§¶

*Department of Medicine and Aging, Centre for Aging Sciences (Ce.S.I.), 'University G.D'Annunzio' Foundation, Chieti, Italy; †Department of Vascular Medicine and ‡Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands; §University of Berne, Institute of Social and Preventive Medicine (ISPM), Finkenhubelweg, Berne, Switzerland; and ¶Consorzio Mario Negri Sud, Laboratory of Clinical Epidemiology of Cardiovascular Disease, Department of Clinical Pharmacology and Epidemiology, Chieti, Italy

To cite this article: Di Nisio M, van Sluis GL, Bossuyt PMM, Büller HR, Porreca E, Rutjes AWS. Accuracy of diagnostic tests for clinically suspected upper extremity deep vein thrombosis: a systematic review. *J Thromb Haemost* 2010; **8**: 684–92.

Summary. *Background:* The best available test for the diagnosis of upper extremity deep venous thrombosis (UEDVT) is contrast venography. The aim of this systematic review was to assess whether the diagnostic accuracy of other tests for clinically suspected UEDVT is high enough to justify their use in clinical practise and to evaluate if any test can replace venography. *Methods:* MEDLINE and EMBASE databases were searched from inception to June 2009. Two reviewers independently evaluated study eligibility, extracted data, and assessed study quality. *Results:* We identified 17 papers, reporting on 793 patients. Overall, the methodological quality was poor, sample sizes were small, and large between-study differences were observed in spectrum and design. The summary estimates of sensitivity (95% confidence interval) were 97% (90–100%) for compression ultrasonography, 84% (72–97%) for Doppler ultrasonography, 91% (85–97%) for Doppler ultrasonography with compression, and 85% (72–99%) for phleboreography. The corresponding summary estimates of specificity were, respectively, 96% (87–100%), 94% (86–100%), 93% (80–100%), and 87% (71–100%). Clinical findings, a clinical score, D-dimer, magnetic resonance imaging, rheography and plethysmography were evaluated in one study each, involving a median number of 46 patients (range 21–214). Sensitivity and specificity ranged from 0% to 100% and from 14% to 100%. *Conclusions:* Methodological limitations, large between-study differences and small sample sizes limit the evidence of tests for clinically suspected UEDVT. Compression ultrasonography may be an acceptable alternative to

venography. The addition of (color) Doppler does not seem to improve the accuracy. Adequately designed studies are warranted to confirm these findings.

Keywords: deep vein thrombosis, sensitivity, specificity, upper extremity.

Introduction

Upper extremity deep vein thrombosis (UEDVT) may present with swelling, pain and functional impairment, although completely asymptomatic cases have been described [1–4]. The prompt recognition of UEDVT is important because patients may develop pulmonary embolism, while excluding the disease can avoid unnecessary anticoagulation. Therefore, diagnostic tests should have high sensitivity, to reduce the proportion of patients with UEDVT who remain untreated, and specificity as high as possible, to preserve the practical applicability.

The reference standard for the diagnosis of UEDVT is contrast venography. Unfortunately, venography is invasive, often difficult to perform, and it requires the use of ionizing radiation. Ultrasonography and magnetic resonance imaging have been considered as replacement tests for venography. Although commonly adopted as the sole diagnostic test for UEDVT, ultrasonography can nonetheless be difficult to perform because of the clavicle, which hinders and limits the possibility to image and compress the middle part of the subclavian vein, potentially resulting in false negative results. In general, the accuracy of ultrasonography for clinically suspected UEDVT is still unclear [1–4].

The diagnostic work-up of clinically suspected lower extremity deep vein thrombosis is well established and is based on a combined strategy, involving clinical probability, D-dimer and imaging tests. It remains unclear whether clinical probability assessment, the D-dimer test, ultrasonography or magnetic resonance imaging can be similarly used as triage tests in a diagnostic strategy to detect or exclude UEDVT.

Correspondence: Marcello Di Nisio, Reparto di Patologia Medica, Ospedale SS.ma Annunziata, via dei Vestini 31, 66013, Chieti Scalo, Italy.

Tel.: +39 871 357361; fax: +39 871 357361.

E-mail: mdinisio@unich.it

Received 9 September 2009, accepted 14 January 2010

The aim of this systematic review was to evaluate the accuracy of diagnostic tools for clinically suspected UEDVT and to evaluate whether any of these tests – alone or in combination – can be used as triage tests within a diagnostic strategy, or as a replacement test for venography. To safely exclude UEDVT without further testing, a triage test or test combination should have near-perfect sensitivity. Only if the proportion of false negative results is acceptably low, can further testing be safely omitted. Potential replacement tests should also have near-perfect specificity, minimizing any harm, inconvenience and strain on health resources induced by unnecessary further testing in patients with false positive test results.

Methods

Data sources and searches

A systematic search of the MEDLINE and EMBASE databases up to June 2009 was performed to identify studies reporting on the diagnostic accuracy of any diagnostic test in patients with suspected UEDVT. The following search terms (MeSH and textwords) were used: vein thrombosis, venous thrombosis, upper extremity, upper limb, and diagnosis. Reference lists of all included studies and of reviews were manually searched for other additional potentially eligible studies. No language restrictions were applied.

Study selection

Two investigators (MDN and AWSR) independently reviewed titles and abstracts from the initial search to determine whether the corresponding paper satisfied the inclusion criteria. Any article evaluating the diagnostic accuracy of a diagnostic test including clinical evaluation, D-dimer or an imaging test in patients with clinically suspected UEDVT was eligible, as long as all index test results were verified by an independent reference test. Articles were excluded if data could not be extracted to calculate a 2×2 table for symptomatic UEDVT. Case-reports were also excluded. Any disagreement was solved through discussion or by involvement of a third reviewer.

Data extraction and quality assessment

Two reviewers (M. Di Nisio, A. W. S. Rutjes or G. L. van Sluis) independently extracted study characteristics using standardized forms. Study characteristics had been identified based on their risk of bias and of variability, as listed in the QUADAS tool and in recent systematic reviews [5–8]. QUADAS is the only validated quality assessment tool for diagnostic accuracy studies. It was developed using formal consensus methods [7]. It consists of a check-list, accompanied by an instruction document, covering the domains of spectrum, verification of test results and handling of drop-outs and uninterpretable test results. The Cochrane handbook for diagnostic test accuracy studies recommends the use of 11 out of 14 QUADAS items; three items were dropped, as these concerned quality of

reporting only (<http://srdta.cochrane.org/en/authors.html>). Table 1 displays how QUADAS items were applied in this review. Where necessary, authors were contacted for additional information. Any disagreements were solved by consensus and, if necessary, by involving a third reviewer. No attempts to mask for authorship, journal name or institution were made.

Data synthesis and analysis

We used a bivariate random effects approach to obtain joint summary estimates of sensitivity and specificity of any test that was evaluated in at least three studies. We used a fixed effects approach if only two studies were identified for a given test. The bivariate model assumes that the logit transformed sensitivities and specificities of the included diagnostic studies follow a bivariate normal distribution around a common mean, allowing for a nonzero correlation between sensitivity and specificity [9]. Bivariate analyses with covariates were planned to explore the effects of differences in test type, and patient and design characteristics.

Meta-analysis was performed with the PROC NLMIXED module in SAS statistical software, version 9.2 (SAS Institute Inc, Cary, NC, USA). REVMAN 5 software (Nordic Cochrane Centre, Cochrane Collaboration, Copenhagen, Denmark) was used to calculate exact binomial 95% confidence intervals (95% CI) for sensitivities and specificities of primary studies and to make forest plots and methodological quality summary graphs. *P*-values below 0.05 were considered to indicate statistical significance. This study had no external funding source.

Results

A total of 1302 papers were identified with the initial search strategy, of which 87 were considered potentially eligible based on close reading of the title and abstract. Of these papers, 14 did not include patients with clinically suspected UEDVT, 16 were reviews, 27 were not diagnostic accuracy studies, and four included only cases. For another eight studies, it was impossible to construct a separate 2×2 table for UEDVT in symptomatic patients and we were unsuccessful in obtaining additional data from the investigators (Fig. 1). Despite involving librarians in several continents, and directly contacting investigators, one article could not be retrieved in full text [10]. Eventually, a total of 17 studies (793 patients) was available for analysis (Table 2) [11–27]. No study had evaluated a combination of tests within a diagnostic strategy. None of the studies explicitly reported the test under evaluation to be considered as a replacement for or a triage test before venography.

Methodological quality The spectrum was considered to be representative in five studies (35%, Fig. 2). Two additional articles reported inclusion to be consecutive but omitted the percentage of patients with cancer or central venous lines [18,23]. Venography was used as the reference standard in 12 studies (75%); in 10 (59%) it was applied as the sole reference standard, avoiding the risk of differential verification bias, and

Table 1 Assessment of methodological quality

1	Representative spectrum?	Yes, if selection of patients had been consecutive, and the percentages of patients with cancer and central venous lines were reported
2	Acceptable reference standard?	Yes, if venography was applied in some of the patients
3	Acceptable delay between tests?	Yes, if the delay between the test under evaluation and reference standard was 24 h or less
4	Partial verification avoided?	Yes, if all patients receiving the test under evaluation were verified by a reference standard
5	Differential verification avoided?	Yes if the same reference standard was applied in all patients, regardless of previous test results
6	Incorporation avoided?	Yes, if the reference standard was independent of the index test (i.e. the index test did not form part of the reference standard)
7	Reference standard results blinded?	Yes, if the reference standard results were interpreted without knowledge of the results of the index test
8	Index test results blinded?	Yes, if the index test results were interpreted without knowledge of the results of the reference standard
9	Relevant clinical information?	Not assessed*
10	Uninterpretable results reported?	Yes, if uninterpretable or intermediate test results were reported
11	Withdrawals explained?	Yes, if withdrawals from the study were explained
12	Prospective data-collection?†	Yes, if data collection, including gathering of test results, was conducted after formulation of the study objective

Items 1 to 11 are the QUADAS items recommended in the Cochrane handbook for diagnostic test accuracy studies (<http://srdta.cochrane.org/en/authors.html>). *In this field of research, clinical information is always used while interpreting the test under evaluation. †QUADAS and the Cochrane handbook advise adding items whenever relevant to a specific review topic. Item 12 was added in this review.

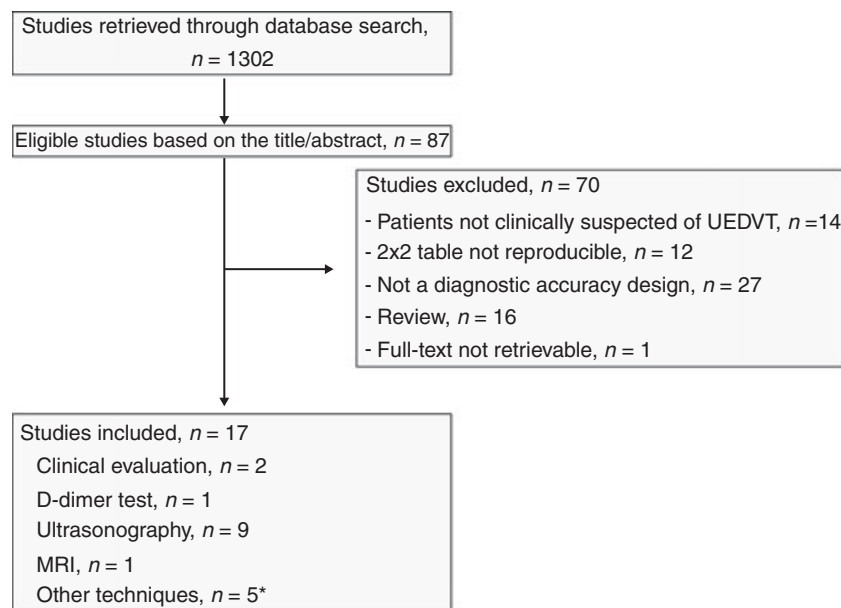


Fig. 1. Process of selection and assessment of primary studies. *One publication provided data separately for ultrasound and phleboreography. Abbreviations: MRI, magnetic resonance imaging; UEDVT, upper extremity deep vein thrombosis.

in 7 (41%) it was applied in all patients, avoiding the risk of partial verification bias. The existence of uninterpretable or intermediate results and withdrawals were explained in only four studies (23%). The proportion of differential and partial verification, withdrawals and uninterpretable results are given in Table 2. Only two studies fulfilled all quality criteria and were considered to be at low risk of bias [15,22].

Estimates of diagnostic accuracy

Clinical evaluation In one retrospective study, the accuracy of five clinical signs and symptoms was assessed in patients with ($n = 25$) and without ($n = 35$) a central venous catheter [11].

The sensitivity and specificity varied broadly, ranging, respectively, from 0% (95% CI, 0–26%) to 57% (95% CI, 18–90%) and from 69% (95% CI, 38–91%) to 92% (95% CI, 64–100%). The combination of pain and swelling carried the best trade-off for sensitivity and specificity: 50% (95% CI, 21–79%) and 85% (95% CI, 54–98%) for patients with and 57% (95% CI, 18–90%) and 57% (95% CI, 37–75%) for patients without a central venous catheter (Fig. 3).

One prospective study evaluated the accuracy of a clinical score based on the presence of a central venous catheter or a pacemaker, localized pain, unilateral pitting edema, and the plausibility of another diagnosis [12]. One point was given to each one of the first three items, one point subtracted for the

Table 2 Clinical features of included studies

Study	Index test	Incidence %	Patients, n	Age, years	In-outpatients, n / N	CVC, n / N	Cancer, n / N	Reference test: verified/total	Indeterminate index/reference test	Withdrawals, n / N
Clinical evaluation Stanley 1994 [11]	Clinical signs and symptoms	CVC + :50 CVC - :20	60	Nr	Nr	25/60	Nr	Doppler + CUS: 60/60	Nr	Nr
Constans 2008 [12]	Clinical score	30	214	59	In and outpatients	25/214	Nr	Doppler + CUS: 214/214	Nr	Nr
D-dimer Merminod 2006 [13]	Vidas D-dimer	77	52	61	In and outpatients	18/52	23/52	Doppler + CUS: 47/52; CT scan: 5/52	Nr	Nr
Ultrasonography Sullivan 1984 [14]	CUS	43	33	Nr	Nr	Nr	Nr	Venography: 7/7; Unverified*: 26/33	Nr	Nr
Prandoni 1997 [15]	CUS, Doppler US, Doppler + CUS	CUS:47 Doppler US: 45 Doppler + CUS: 56	58	UEDVT + 53, UEDVT-: 49	Outpatients	8/58	9/58	Venography: 58/58	CUS: Nr Doppler US:11/58 Doppler + CUS: Nr	4/62
Patel 1999 [16]	Doppler US†	29	21	42	Nr	16/21	15/21	Venography: 20/21; MRI: 1/21	Nr	Nr
Falk 1987 [17]	Doppler US	56	22	12-76	Nr	8/22	Nr	Venography: 14/18; CT scan: 4/18‡; Unverified*:3/21	0/22	1/22
Sottiturai 1982 [18]	Doppler + CUS	83	42	30	Nr	Nr	Nr	Venography: 23/23; Unverified*: 19/42	Nr	Nr
Haire 1991 [19]	Doppler + CUS	82	11	Nr	Inpatients	7/11	Nr	Venography: 11/11	Nr	Nr
Knudson 1990 [20]	Doppler + CUS	34	91 (130 arms)	49	Nr	44/130	39/130§	Venography: 22/121; CT scan: 1/121; MRI: 1/121; clinical follow-up: 99/121; Unverified*: 9/130	Nr	Nr
Baxter 1991 [21]	Doppler + CUS	27	19	52	Nr	14/19	Nr	Venography: 19/19	Nr	Nr
Baarslag 2002 [22]	Doppler + CUS	44	99	54	In and outpatients	24/99	28/99	Venography: 99/99	3/126	22/126
Magnetic resonance Baarslag 2004 [23]	TOF MRI Gd3D MRI	TOF MRI:44 Gd3D MRI:38	44	Nr	In- and outpatients	Nr	Nr	Venography: 44/44	TOF MRI:2 Gd 3D MRI:1	23/44
Other techniques Rheography Mukherjee 1991 [24]	Light reflection rheography	57	21	Nr	Inpatients	7/21	Nr	Venography: 21/21	Nr	Nr
Sottiturai 1982 [18]	Phleborheography	83	42	30	Nr	Nr	Nr	Venography: 23/23; Unverified*: 19/42	Nr	Nr

Table 2 Continued

Study	Index test	Incidence %	Patients, n	Age, years	In-outpatients, n	CVC, n/N	Cancer, n/N	Reference test: verified/total verified	Indeterminate index/reference test	Withdrawals, n/N
Sullivan 1983 [25]	Phleborheography	40	40	Nr	Nr	Nr	Nr	Venography: 16/16; Unverified*: 24/40	Nr	Nr
Plethysmography Patwardhan 1983 [26]	Impedance Plethysmography	68	46	44	Nr	4/46	5/46	Venography: 18/18; Unverified*: 28/46	Nr	Nr
Zufferey 1992 [27]	Strain gauge plethysmography	73	22	50	Outpatients	Cases: 12/16	Cases: 4/16	Venography or Doppler + CUS: 22/22	Nr	Nr

†Color and gray scale ultrasonography used to exclude the presence of the thrombus. If thrombus present, patient excluded. ‡Two patients were verified by chest X-rays and one was not verified. §Based on segment (upper extremity) level. *The investigators omitted patients with unverified test results from the 2-by-2 tables, leading to partial verification. ¶Results not provided separately for venography and ultrasound. CT, computed tomography; CUS, compression ultrasonography; CVC, central venous catheter; Gd, gadolinium enhanced; MRI, magnetic resonance imaging; Nr, not reported; TOF, time-of-flight; US, ultrasonography.

alternative diagnosis. A score of -1 or 0 indicated low probability, 1 intermediate probability, and a score of 2 or 3 high probability of UEDVT. In order to calculate 2 × 2 tables, the authors of the study presented a dichotomized score that classified patients with 0 or less as 'UEDVT unlikely' and with 1 or more points as 'likely'. This diagnostic score had a sensitivity of 78% (95% CI, 68–88%) and a specificity of 64% (95% CI, 57–72%) (Fig. 3).

D-dimer test

The accuracy of a rapid quantitative enzyme-linked immunosorbent assay was evaluated in 52 consecutive patients, of whom 23 had cancer [13]. The authors reported a sensitivity of 100% (95% CI, 78–100%) and a specificity of 14% (95% CI, 4–29%) (Fig. 3).

Ultrasonography

Nine studies evaluated in total 11 test comparisons, including compression ultrasonography (two studies) [14,15], Doppler ultrasonography (three studies) [15–17], and (color) Doppler ultrasonography with compression (six studies) [15,18–22]. The bivariate random effects approach resulted in summary estimates of sensitivity and specificity of 97% (95% CI, 90–100%) and 96% (95% CI, 87–100%) for compression ultrasonography, 84% (95% CI, 72–97%) and 94% (95% CI, 86–100%) for Doppler ultrasonography, and 91% (95% CI, 85–97%) and 93% (95% CI, 80–100%) for Doppler ultrasonography with compression. The differences between estimates of sensitivity and specificity of the three ultrasonographic techniques were not significantly different ($P = 0.3$ for sensitivity and $P > 0.7$ for specificity), and did not substantially explain the variation observed.

Magnetic resonance

One prospective study assessed the value of time-of-flight and Gadolinium-enhanced magnetic resonance imaging [23]. Of 44 consecutive patients initially included, about half were lost and not available for the final analysis. The calculated sensitivity was 71% (95% CI, 29–96%) for time-of-flight and 50% (95% CI, 12–88%) for gadolinium-enhanced magnetic resonance imaging. The corresponding specificities were 89% (95% CI, 52–100%) and 80% (95% CI, 44–97%).

Other imaging tests

Volumetric techniques including strain-gauge or impedance plethysmography and phleborheography were evaluated in several venography-controlled studies. In a retrospective study of 21 inpatients, light reflection rheography had a sensitivity of 92% (95% CI, 61–100%) and a specificity of 100% (95% CI, 66–100%) [24]. Phleborheography was evaluated in two studies (82 patients) reporting a relatively high incidence of UEDVT (Table 2). The pooled sensitivity and specificity were,

	Representative spectrum?	Acceptable reference standard?	Acceptable delay between tests?	Partial verification avoided?	Differential verification avoided?	Incorporation avoided?	Reference standard results blinded?	Index test results blinded?	Uninterpretable results reported?	Withdrawals explained?	Prospective data-collection?
Baarslag 2002	+	+	+	+	+	+	+	+	+	+	+
Baarslag 2004	?	+	+	+	+	+	+	+	+	+	+
Baxter 1991	?	+	+	+	+	+	?	?	-	-	+
Constans 2008	?	-	?	+	+	+	-	+	-	-	+
Falk 1987	?	+	+	-	-	+	+	+	+	+	+
Haire 1991	?	+	?	+	+	+	+	+	-	-	+
Knudson 1990	?	?	?	?	-	+	?	?	-	-	+
Merminod 2006	+	-	?	+	-	+	?	?	-	-	+
Mukherjee 1991	?	+	?	+	+	+	?	?	-	-	-
Patel 1999	+	+	+	+	-	+	+	+	-	-	+
Patwardhan 1983	?	+	?	-	+	+	?	+	-	-	-
Prandoni 1997	+	+	+	+	+	+	+	+	+	+	+
Sottiurai 1982	?	+	?	-	+	+	+	+	-	-	?
Stanley 1994	?	-	?	+	+	+	?	+	-	-	-
Sullivan 1983	?	+	?	-	+	+	?	?	-	-	-
Sullivan 1984	?	+	?	-	+	+	+	+	-	-	?
Zufferey 1992	+	-	?	+	-	+	?	?	-	-	+

Fig. 2. Methodological quality summary of included studies. Risk of bias summary: Review authors’ judgements about each risk of bias item for each included study. Green circles with plus indicate ‘Yes’, referring to low risk of bias, yellow circles with a question mark ‘Unclear’, and red circles with a minus indicate ‘No’, referring to higher risk of bias.

respectively, 85% (95% CI, 72–99%) and 87% (95% CI, 71–100%) (Fig. 3) [17,25].

Impedance plethysmography was assessed in a retrospective study of 46 patients, of whom 39% were verified by venography [26]. The authors reported a sensitivity of 100% (95% CI, 78–100%) and a specificity of 28% (95% CI, 4–71%). Strain gauge plethysmography was evaluated in 22 patients with suspected UEDVT and found to have a sensitivity of 100% (95% CI, 79–100%) and a specificity of 83% (95% CI, 36–100%) [27].

Discussion

This systematic review found a relative paucity of studies on the diagnosis of clinically suspected UEDVT. There were no studies on a combination of tests within a diagnostic strategy. None of the studies was specifically designed to evaluate a diagnostic method as a triage test or as a replacement test for venography. The accuracy indexes, both the ones derived from single studies and the summary estimates, need to be interpreted with caution due to the small sample sizes and

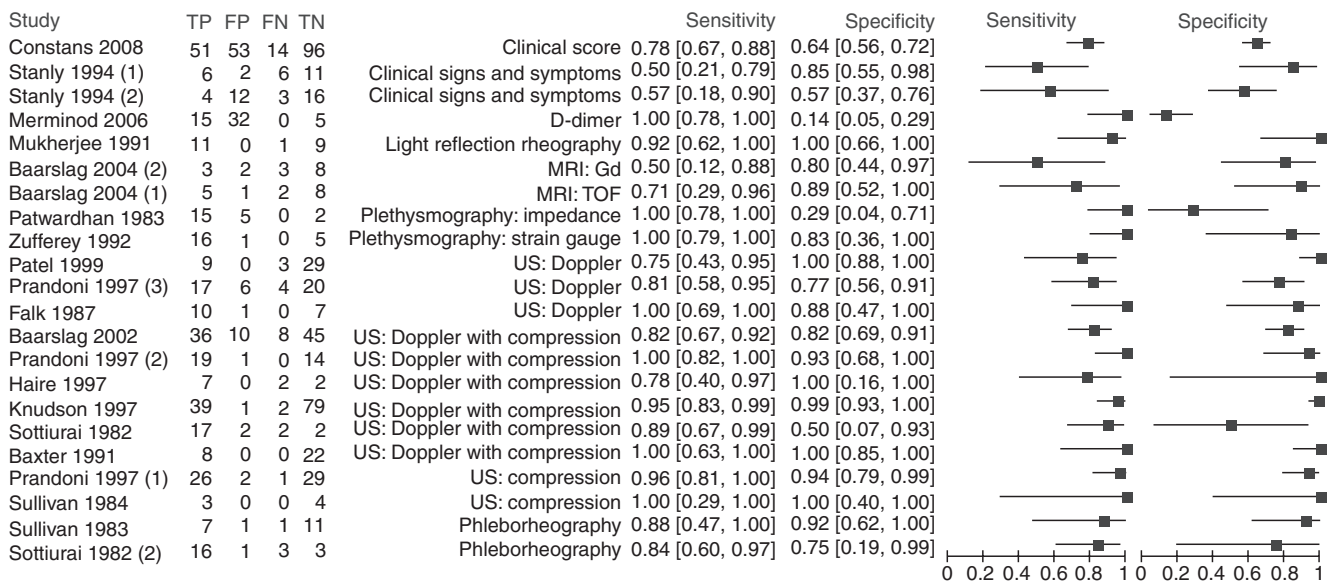


Fig. 3. Forest plots of sensitivity and specificity of included studies. The squares represent the sensitivity and specificity of one study, the black line its confidence interval. Studies are grouped by type of index test. If a study reported accuracy data for more than one type of test comparison, its results are included more than once. *Sensitivity and specificity based on segment (upper extremity) level. Abbreviations: FN, false negative; FP, false positive; Gd, gadolinium enhanced; MRI, magnetic resonance imaging; TN, true negative; TOF, time-of-flight; TP, true positive; US, ultrasonography. Stanley 1994 (1) included only patients with a central catheter, Stanley (2) included only patients without a central catheter.

the suboptimal methodological quality applied in the primary studies.

Most of the studies were found to have methodological limitations, which may have introduced substantial bias and additional variability. As shown in meta-epidemiological studies, suboptimal design choices influence the accuracy estimates, but it is difficult to predict the magnitude and direction [28,29]. We could not evaluate the influence of design characteristics in our bivariate model, due to the limited number of studies available for each specific test.

The initial diagnostic work-up of suspected lower extremity deep vein thrombosis includes the estimation of the pretest clinical probability [30]. The usefulness of signs and symptoms can be increased by combining them into a clinical score, as has been done for lower extremity deep vein thrombosis and for pulmonary embolism [30]. In this regard the score proposed by Constans and colleagues looks promising although the 13% incidence of thrombosis in the low probability group suggests that it may work less efficiently than the Wells' score for lower extremity DVT [12,30].

In clinically suspected lower extremity deep vein thrombosis, the D-dimer test is performed after the pretest clinically probability has been determined. When combined with a low clinical probability it safely excludes thrombosis, avoiding unnecessary further tests in up to 50% of patients [31]. Only one small study evaluated the D-dimer in the exclusion of UEDVT [13]. The negative predictive value was 100% but the specificity only 14%. The low specificity could be the result of the high percentage of cancer patients included, although the estimate is still 4-fold lower than previously reported in patients with malignancy and clinically suspected lower deep vein thrombosis [32].

Ultrasonography for the diagnosis of UEDVT embraces several methods, including compression and Doppler ultrasonography used alone or in combination [33]. B-mode imaging with or without the addition of color permits the visualization of the venous vessels, while their patency is verified by means of compression and/or pulsed-wave Doppler. Compression ultrasonography assesses the presence or absence of vein compressibility and the echogenicity within the vein lumen; Doppler ultrasonography evaluates the characteristics of venous flow, including phasicity, pulsatility, and variation with physiologic maneuvers. The diagnostic accuracy of ultrasonography for suspected UEDVT has been summarized in a number of narrative reviews [1–3] and recently evaluated by Mustafa and colleagues [4] who attempted a more quantitative evaluation. The lack of a systematic assessment and the availability of recent data have prompted this review to carefully re-evaluate the method [22]. Ultrasonography was evaluated in a few small studies, often with major methodological limitations. Two studies (157 patients) considered at low risk of bias evaluated the same ultrasonographic method and reached opposite conclusions [15,22]. The specificity of ultrasonography for symptomatic lower limb DVT is around 94%, whereas sensitivity is 89%, 97% and 73% for overall, proximal and distal lower DVT, respectively [34]. Although the figures observed in the current review appear, at least for some compression ultrasonography methods, comparable to the accuracy of ultrasonography for lower limb DVT, the evidence they rely upon is not as strong and convincing. Compression ultrasonography may be an acceptable alternative to venography whereas the addition of (color) Doppler does not seem to improve the accuracy. The potential adverse events and the non-feasibility of venography in approximately 20% of

patients due to renal dysfunction, poor vascular access or contrast allergies should be taken into consideration. In the absence of large accuracy or management studies that could clarify whether the benefits gained from the exclusive use of venography would justify the associated risks, it seems reasonable not to suggest venography for all clinically suspected UEDVT, but to limit its use to the clinical situations where there is a concern for false positive or false negative ultrasound results. Several urgent questions for clinical practise remain unanswered, such as the use of a single test versus serial ultrasonography, the combination of ultrasonography with pretest clinically probability and/or D-dimer test within diagnostic algorithms, and the role of Doppler in the poorly compressible area of the clavicle.

Preliminary data have suggested that magnetic resonance imaging can accurately diagnose lower extremity deep vein thrombosis [35]. Although magnetic resonance imaging enables a visualization of the entire deep venous system of the arm, the central thoracic veins as well as the surrounding tissues to help in differential diagnosis [36], only one study evaluated this technique for UEDVT [23]. Its small size and extreme partial verification do not allow us to make conclusive statements about its accuracy.

Strain-gauge plethysmography, impedance plethysmography and phleborheography are non-invasive volumetric techniques that attempt to indirectly estimate the patency or the occlusion of the veins through the measurements of pressure changes in the limb. All these techniques have undergone relatively limited evaluation in small sized studies with a substantial risk of bias. In clinical practise these methods have all largely been replaced by ultrasonography.

In summary, the evidence on the accuracy of diagnostic tests for suspected UEDVT is largely inconclusive. Clinicians should be aware of the relatively small size and poor quality of the studies that have evaluated diagnostic tests for UEDVT, including ultrasonography. Given the available evidence, we might suggest ultrasonography as the first choice and consider venography in the case of an indeterminate result on ultrasonography or in patients with a high clinical suspicion and an initial normal ultrasound. Additional studies are needed before venography is fully replaced by ultrasonography as the test of choice for UEDVT. Adequately conducted diagnostic accuracy or managements studies are required to clarify the sensitivity and specificity of all other tests and strategies, while demonstrating how these may vary over different spectra of the disease.

Disclosure of Conflict of interests

The authors state that they have no conflict of interest.

References

- Joffe HV, Goldhaber SZ. Upper-extremity deep vein thrombosis. *Circulation* 2002; **106**: 1874–80.
- Bernardi E, Pesavento R, Prandoni P. Upper extremity deep venous thrombosis. *Semin Thromb Haemost* 2006; **32**: 729–36.
- Fraser JD, Anderson DR. Deep venous thrombosis: recent advances and optimal investigation with US. *Radiology* 1999; **211**: 9–24.
- Mustafa BO, Rathbun SW, Whitsett TL, Raskob GE. Sensitivity and specificity of ultrasonography in the diagnosis of upper extremity deep vein thrombosis a systematic review. *Arch Intern Med* 2002; **162**: 401–4.
- Whiting P, Rutjes AWS, Reitsma JB, Glas AS, Bossuyt PM, Kleijnen J. Sources of variation and bias in studies of diagnostic accuracy: a systematic review. *Ann Intern Med* 2004; **140**: 189–202.
- 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.
- Whiting PF, Weswood ME, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. *BMC Med Res Methodol*. 2006; **6**: 6–9.
- Leeflang MMG, Deeks JJ, Gatsonis C, Bossuyt PMM, On behalf of the Cochrane Diagnostic Test Accuracy Working Group. Systematic reviews of diagnostic test accuracy. *Ann Intern Med* 2008; **149**: 889–97.
- Reitsma GB, Glas AS, Rutjes AWS, Scholten RJPM, Bossuyt PMM, Zwiderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005; **58**: 982–90.
- Foley WD. Rapid diagnosis of a venous thrombosis by angiodynography. *MedicaMundi* 1988; **33**: 32.
- Stanley JM, McGrath RS, Freeman MB, Stevens SL, Goldman MH. Predictability of symptoms of upper extremity deep venous thrombosis in patients with central venous catheters with colour duplex imaging. *J Vasc Tech* 1994; **18**: 71–3.
- Constans J, Salmi L-R, Sevestre-Pietri M-A, Perusat S, Nguon M, Degeilh M, Labarere J, Gattoliat O, Boulon C, Laroche J-P, Le Roux P, Pichot O, Quere I, Conri C, Bosson J-L. A clinical prediction score for upper extremity deep venous thrombosis. *Thromb Haemost* 2008; **99**: 202–7.
- Merminod T, Pellicciotta S, Bounameaux H. Limited usefulness of D-dimer in suspected deep vein thrombosis of the upper extremities. *Blood Coagul Fibrinolysis* 2006; **17**: 225–7.
- Sullivan ED, Peter DJ, Cranley JJ. Real-time B-mode venous ultrasound. *J Vasc Surg* 1984; **1**: 465–71.
- Prandoni P, Polistena P, Bernardi E, Cogo A, Casara D, Verlato F, Angelini F, Simioni P, Signorini GP, Benedetti L, Girolami A. Upper-extremity deep vein thrombosis: risk factors, diagnosis, and complications. *Arch Intern Med* 1997; **157**: 57–62.
- Patel MC, Berman LH, Moss HA, McPherson SJ. Subclavian and internal jugular veins at Doppler US: abnormal cardiac pulsatility and respiratory phasicity as a predictor of complete central occlusion. *Radiology* 1999; **211**: 579–83.
- Falk RL, Smith DF. Thrombosis of upper extremity thoracic inlet veins: diagnosis with duplex Doppler sonography. *AJR Am J Roentgenol* 1987; **149**: 677–82.
- Sotturrai VS, Towner K, McDonnell AE, Zarins CK. Diagnosis of upper extremity deep venous thrombosis using non-invasive technique. *Surgery* 1982; **91**: 582–5.
- Haire WD, Lynch TG, Lund GB, Lieberman RP, Edney JA. Limitations of resonance magnetic imaging and ultrasound-directed (duplex) scanning in the diagnosis of subclavian vein thrombosis. *J Vasc Surg* 1991; **13**: 391–7.
- Knudson GJ, Wiedmeyer DA, Erickson SJ, Foley WD, Lawson TL, Mewissen MW, Lipchik EO. Color Doppler sonographic imaging in the assessment of upper-extremity deep venous thrombosis. *AJR Am J Roentgenol* 1990; **154**: 399–403.
- Baxter GM, Kincaid W, Jeffrey RF, Millar GM, Porteous C, Morley P. Comparison of colour Doppler ultrasound with venography in the diagnosis of axillary and subclavian vein thrombosis. *Br J Radiol* 1991; **64**: 777–81.
- Baarslag H-J, van Beek EJR, Koopman MMW, Reekers JA. Prospective study of color duplex ultrasonography compared with con-

- trast venography in patients suspected of having deep venous thrombosis of the upper extremities. *Ann Intern Med* 2002; **136**: 865–72.
- 23 Baarslag HJ, van Beek EJR, Reekers JA. Magnetic resonance venography in consecutive patients with suspected deep vein thrombosis of the upper extremity: initial experience. *Acta Radiol* 2004; **45**: 38–43.
 - 24 Mukherjee D, Andersen CA, Sado AS. Use of light reflection rheography for diagnosis of axillary or subclavian venous thrombosis. *Am J Surg* 1991; **161**: 651–6.
 - 25 Sullivan ED, Reece CI, Cranley JJ. Phleborheography of the upper extremity. *Arch Surg* 1983; **118**: 1134–6.
 - 26 Patwardhan NA, Anderson FA, Cutler BS, Wheeler HB. Noninvasive detection of axillary and subclavian venous thrombosis by impedance plethysmography. *J Cardiovasc Surg* 1983; **24**: 250–5.
 - 27 Zufferey P, Pararas C, Monti M, Depairon M. Assessment of acute and old deep venous thrombosis in upper extremity by venous strain gauge plethysmography. *VASA* 1992; **21**: 263–7.
 - 28 Lijmer JG, Mol BW, Heisterkamp S, Bossel GJ, Prins MH, van der Meulen JH, Bossuyt PM. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999; **282**: 1061–6.
 - 29 Rutjes AW, Reitsma JB, Di Nisio M, Smidt N, van Rijn JC, Bossuyt PM. Evidence of bias and variation in diagnostic accuracy studies. *CMAJ* 2006; **174**: 469–76.
 - 30 Wells PS. Integrated strategies for the diagnosis of venous thromboembolism. *J Thromb Haemost* 2007; **5**(Suppl. 1): 41–50.
 - 31 Di Nisio M, Squizzato A, Rutjes AW, Buller HR, Zwinderman AH, Bossuyt PM. Diagnostic accuracy of d-dimer test for exclusion of venous thromboembolism: a systematic review. *J Thromb Haemost* 2007; **5**: 296–304.
 - 32 Di Nisio M, Rutjes AWS, Buller HR. Combined use of clinical pretest probability and D-dimer test in cancer patients with clinically suspected deep-venous thrombosis. *J Thromb Haemost* 2005; **4**: 52–7.
 - 33 Cronan JJ. Venous thromboembolic disease: the role of US. *Radiology* 1993; **186**: 619–30.
 - 34 Kearon C, Julian JA, Newman TE, Ginsberg JS. Noninvasive diagnosis of deep venous thrombosis. *Ann Intern Med* 1998; **128**: 663–77.
 - 35 Fraser DGW, Moody AR, Morgan PS, Martel AL, Davidson I. Diagnosis of lower-limb deep venous thrombosis: a prospective blinded study of magnetic resonance direct thrombus imaging. *Ann Intern Med* 2002; **136**: 89–98.
 - 36 Kelly J, Hunt BJ, Moody A. Magnetic resonance direct thrombus imaging: a novel technique for imaging venous thromboembolism. *Thromb Haemost* 2003; **89**: 773–82.