

Accuracy of Conventional Transthoracic Echocardiography for the Diagnosis of Intracardiac Right-to-Left Shunt: A Meta-Analysis of Prospective Studies

Mohammad Khalid Mojadidi, M.D.,* Jared S. Winoker, M.D.,* Scott C. Roberts, M.D.,* Pavlos Msaouel, M.D., Ph.D.,* Muhammad Omer Zaman, M.D.,† Rubine Gevorgyan, M.D.,‡ and Jonathan M. Tobis, M.D., F.A.C.C.‡

*Department of Medicine, Jacobi Medical Center, Albert Einstein College of Medicine, Bronx, New York; †Department of Family Medicine, William Beaumont Hospital, Grosse Pointe, Michigan; and ‡Program in Interventional Cardiology, Division of Cardiology, David Geffen School of Medicine, University of California at Los Angeles, Los Angeles, California

Background: Paradoxical embolization through a right-to-left shunt (RLS), often from a patent foramen ovale (PFO), has been associated with cryptogenic stroke. While transesophageal echo (TEE) bubble study is the current standard reference for diagnosing PFO, transthoracic echo (TTE) remains the most commonly used screening test for RLS due to its noninvasiveness and easy availability. The aim of this meta-analysis was to determine the accuracy of TTE compared to TEE as the reference. **Methods and Results:** A systematic review of Medline, Cochrane, and Embase was done to look for all the prospective studies assessing for intracardiac RLS using conventional TTE compared to TEE as the reference; both TTE and TEE were performed with a contrast agent and a maneuver to provoke RLS in all studies. A total of 13 studies with 1436 patients fulfilled the inclusion criteria. The weighted mean sensitivity and specificity for TTE were 46% and 99%, respectively. Likewise, the positive likelihood ratio and negative likelihood ratio were 20.85 and 0.57, respectively. Using different contrast agents, different microbubble cutoffs for a positive TTE/TEE, and different cardiac cycle cutoffs for a positive TTE/TEE did not affect the accuracy of TTE. In a population of patients with cryptogenic stroke, a TTE that tests positive for RLS has a 95% probability of being a true positive. **Conclusion:** Transthoracic echocardiogram has a low sensitivity and extremely high specificity, making it a poor rule out test but an excellent rule in test for the detection of intracardiac RLS. (Echocardiography 2014;31:1036–1048)

Key words: transthoracic echocardiogram, transesophageal echocardiogram, right-to-left shunt, patent foramen ovale

Paradoxical embolization through a right-to-left shunt (RLS), often from a patent foramen ovale (PFO), is a recognized mechanism for ischemic stroke,^{1,2} especially in patients without any identifiable cause of stroke.³ Although the CLOSURE 1, RESPECT, and PC trials failed to meet their primary endpoints by intention-to-treat analysis, recent meta-analyses of these trials and observational studies are encouraging.^{4,5} The totality of evidence suggests that PFO occluding devices may reduce the recurrence of stroke and transient ischemic attack compared to medical treatment in patients with cryptogenic stroke, particularly in those with greater shunting. These

data along with the evaluation of patients with severe migraines or other conditions associated with PFO^{6–9} make it essential to accurately diagnose RLS in patients being considered for transcatheter PFO closure.

While contrast-enhanced transesophageal echocardiogram (TEE) is considered the gold standard for diagnosing PFO,^{10,11} conventional transthoracic echocardiogram (TTE) with bubble study is the most commonly used initial screening test for the detection of RLS due to its noninvasiveness and low cost.^{12–24} The aims of this study were to:

- (a) Expand on prior reviews of TTE and provide the first meta-analysis on this topic that methodically assesses the diagnostic accuracy of conventional TTE in the evaluation of patients for an intracardiac RLS.

Address for correspondence and reprint requests: Mohammad Khalid Mojadidi, M.D., Department of Medicine, Albert Einstein College of Medicine, 1400 Pelham Parkway South Building 1, Rm 3N1, Bronx, NY 10461, USA. Fax : (310) 267-0384; E-mail: mkmojadidi@gmail.com

- (b) Determine the utility of TTE for the diagnosis of RLS in the general population and in a population of patients with cryptogenic stroke.
- (c) Perform a sensitivity analysis of different TTE protocols to determine the best methodology for diagnosing intracardiac RLS.

Methods:

Literature Review:

Relevant citations were searched for on Medline, Cochrane, and Embase. The search was completed in August 2013, yielding published literature since 1913. The terms that were used in the search were "PFO" OR "patent foramen ovale" OR "right to left shunt" OR "atrial septal defect" AND "TTE" OR "transthoracic echocardiography" OR "transthoracic echo" OR "echo" OR "echocardiography."

The references of all of the primary studies as well as those of other known prior reviews were analyzed to find cited articles that were not found by the initial searches. No restrictions were used regarding publication language. Abstracts lacking peer-reviewed manuscripts were omitted since they would not have enough data required for the meta-analysis.

Selection of Studies:

Articles that were identified were analyzed by three independent reviewers (M.K.M., J.S.W., and S.C.R.). Each article was screened for preset inclusion criteria:

- (1) Original prospective studies (reviews, abstracts, isolated cases, commentaries, editorials, and letters were excluded)
- (2) Subject age ≥ 18
- (3) Studies were selected for the review if they included at least 20 patients with suspected intracardiac RLS who were screened by TTE bubble study and were confirmed by TEE bubble study as a reference gold standard. If a study conducted both TTE and TEE but did not consider TEE as the gold standard technique, we calculated the appropriate parameters assuming TEE as the reference comparison.
- (4) Able to interpret diagnostic accuracies by adequate demonstration of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). Studies that did not provide TP, TN, FP, and FN data were automatically excluded as they would not have sufficient data required to calculate diagnostic accuracies.
- (5) If studies utilized different protocols for TTE or TEE (for example: utilizing different contrast agents, different cutoffs for a

positive TTE/TEE study, etc.), a subanalysis was then performed, where possible, to determine the effect of changing the study protocol on accuracy of TTE.

Data Extraction:

Data were extracted onto a specifically designed spreadsheet with information regarding study design, number of patients, age, sex, TTE/TEE indication, type of contrast used, method of provocation (Valsalva maneuver or cough), microbubble cutoff used for a positive TTE/TEE study, test accuracy results (TP, FP, FN, and TN), and any other comments that should be considered in the analysis such as disagreements between the reviewers, any difficulties with interpreting the data and reasons for exclusion.

Quality Assessment:

The assessment of quality of each study was done by evaluating 14 items considered relevant to the review topic, based on the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) instrument.²⁵ Three reviewers (S.C.R., J.S.W., and M.K.M.) independently assessed the quality items, and discrepancies were resolved by consensus.

Statistical Analysis:

Inter-rater agreement was assessed by calculating Fleiss' kappa using R.²⁶ Diagnostic odds ratio (DOR) was used as a measure of effect size to assess for publication bias. Publication bias analysis was then performed using Comprehensive Meta-analysis version 2²⁷ by visual inspection of the funnel plot and statistically using fail-safe N and the trim-and-fill method. The Meta-DiSc software (Version 1.4; Unit of Clinical Biostatistics, Ramón y Cajal Hospital, Madrid, Spain) was used for all other diagnostic accuracy analyses.²⁸ Potential variations due to threshold effect were assessed statistically by computing the spearman correlation coefficient between the logit of sensitivity and the logit of 1-specificity as well as graphically by visually inspecting the accuracy estimates pairs in forest plots and summary receiver operating characteristic (sROC) curves.^{28,29} Between-study heterogeneity (other than threshold effect) and between-study inconsistency were assessed by the Cochran Q statistic and the inconsistency index (I^2), respectively, and the level of significance for the corresponding P-value was set at $P = 0.10$. Due to anticipated inter-study heterogeneity, a random effects analysis model (DerSimonian Laird)³⁰ was used for all meta-analyses because it provides more conservative estimates of the pooled data. The stability of the diagnostic accuracy results was assessed by one-way sensitivity analysis which was

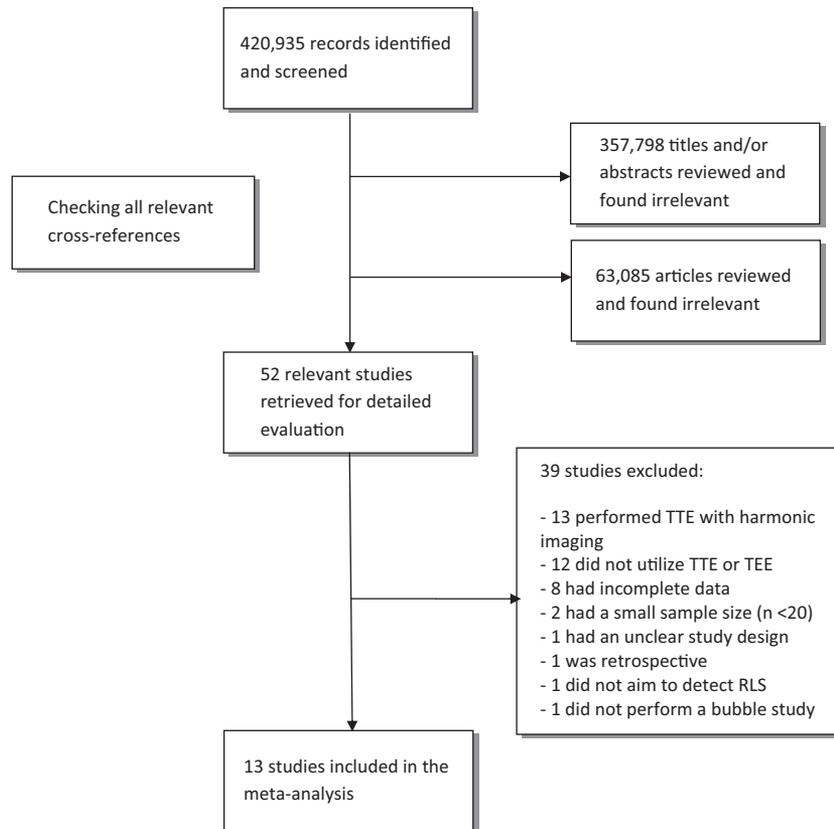


Figure 1. Selection of studies.

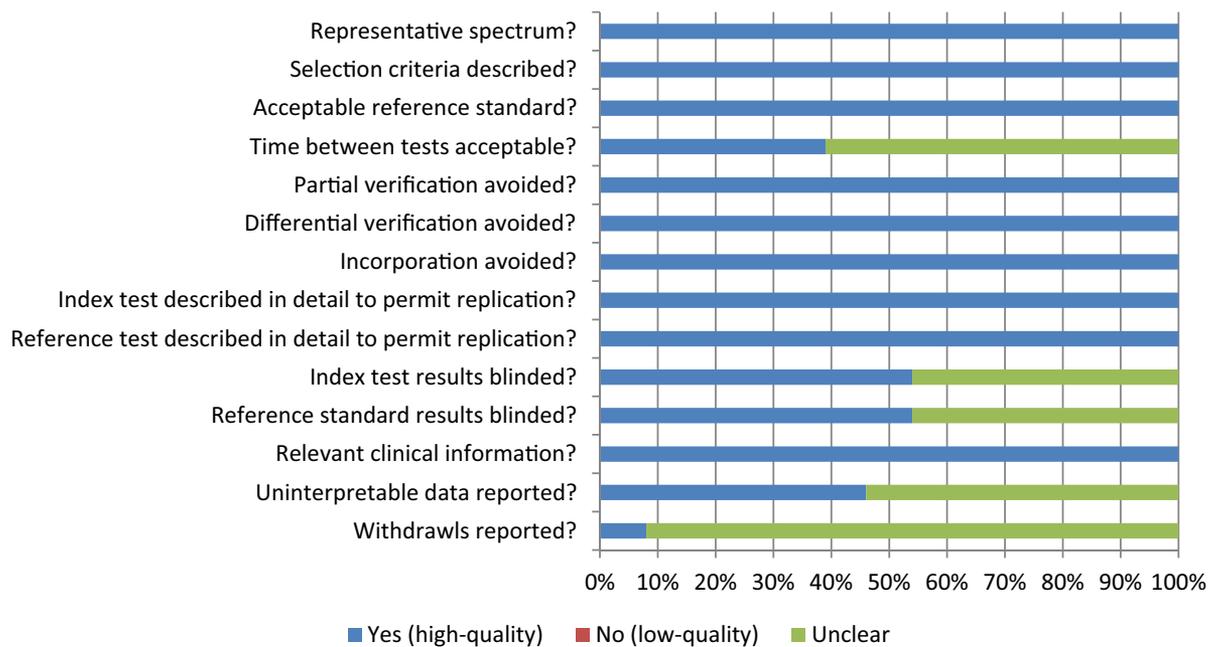


Figure 2. Methodological quality summary.

performed by omitting every study (one at a time) from the meta-analysis. To investigate the effects of potential sources of heterogeneity in

the pooled calculations, subgroup analyses were performed considering more homogenous set of studies that adopted similar design variables.

sROC curves were constructed using the DerSimonian Laird random effects model. The area under the curve (AUC) and index Q* were used to assess and summarize the discriminating ability of the sROC curve.³¹ Posttest probabilities were calculated and illustrated using Fagan nomograms. Subgroups were constructed only when ≥3 studies could be included. Differences between subgroups were assessed by tests of interaction.³² To correct for multiplicity of comparisons in subgroup analyses, P-values of paired comparisons between subgroups were adjusted with the Bonferroni–Holm procedure.³³ Values of 95% confidence intervals (CI) were used for all pooled data, all P-values are two tailed and an

adjusted P-value of <0.05 was considered statistically significant unless otherwise specified.

Results:

Characteristics of the Included Studies:

Of the 52 potential studies identified, 13 prospective studies comprising 1436 patients met inclusion criteria and formed the dataset.^{12–24} Overall reviewer agreement with regard to study inclusion or exclusion was substantial, with a kappa of 0.83. 39 studies did not meet the inclusion criteria and were excluded; the excluded studies along with the reasons for exclusion are provided in the supplementary results. Figure 1 describes the study selection method used for this analysis.

	Representative spectrum?	Selection criteria described?	Acceptable reference standard?	Time between tests acceptable?	Partial verification avoided?	Differential verification avoided?	Incorporation avoided?	Index test described in detail to permit replication?	Reference test described in detail to permit replication?	Index test results blinded?	Reference standard results blinded?	Relevant clinical information?	Uninterpretable data reported?	Withdrawals reported?
Belkin 1994 [12]	(+)	(+)	(+)	(?)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(?)	(?)
Chen 1992 [13]	(+)	(+)	(+)	(?)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(?)	(?)
de Bruijn 2006 [14]	(+)	(+)	(+)	(?)	(+)	(+)	(+)	(+)	(+)	(?)	(?)	(+)	(?)	(?)
Ha 2001 [15]	(+)	(+)	(+)	(?)	(+)	(+)	(+)	(+)	(+)	(?)	(?)	(+)	(?)	(?)
Hausmann 1992 [16]	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(?)	(?)	(+)	(?)	(?)
Kuhl 1999 [17]	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(?)
Madala 2004 [18]	(+)	(+)	(+)	(?)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(?)
Monte 2010 [19]	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)
Nemec 1991 [20]	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(?)	(?)	(+)	(+)	(?)
Rahmouni 2008 [21]	(+)	(+)	(+)	(?)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(?)
Siostrzonek 1988 [22]	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(?)	(?)	(+)	(+)	(?)
Stendel 2000 [23]	(+)	(+)	(+)	(?)	(+)	(+)	(+)	(+)	(+)	(?)	(?)	(+)	(?)	(?)
Zito 2009 [24]	(+)	(+)	(+)	(?)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(?)	(?)

Figure 3. Methodological quality table.

Quality Assessment:

Using the recommended 14-item checklist for evaluating imaging studies using QUADAS, items 4, 10, 11, 13, and 14 either scored poorly or were considered unclear: item 4: (“time between tests acceptable?”), item 10: (“index test results blinded?”), item 11: (“reference standard results blinded?”), item 13: (“uninterpretable data reported?”), and item 14: (“withdrawals reported?”). Item 4 refers to the time interval between the index and reference tests and may lead to disease progression bias. Items 10 and 11 refer to blinding and may affect diagnostic accuracy, otherwise known as review bias. Item 13 refers to any segments that may have been uninterpretable, and can, therefore, result in false elevations of test accuracy and item 14 may lead to test performance bias, as patients unfit are removed to improve accuracy. Otherwise, all studies demonstrated high-quality scoring on the remaining 9 items (Figs. 2 and 3).

Transthoracic Echo Diagnostic Value:

A total of 13 studies met all inclusion criteria and were used for further meta-analytic calculations. Figure 4 shows that the actual combined effect size does not significantly differ from the theoretical combined effect size (one-sided P-value = 0.20). The fail-safe N was 415. Table I

describes the characteristics of the included studies and Table II describes the diagnostic accuracies of the studies. The major clinical indication for performing a TTE in most of the studies was stroke. Of the 13 studies that performed TTE and TEE with contrast, 6 (46%) used agitated saline as the contrast agent, 2 (15%) used agitated saline with blood, 4 (31%) used a gelatin-based solution, and 1 (8%) used more than one contrast agent. The majority of the comparisons used detection of ≥ 1 microbubble in the left atrium as the embolic cut-off for a positive TTE and TEE (46%; 6/13). Forty-six percent (6/13) of the studies performed the contrast injection during the Valsalva maneuver. The most commonly used cutoff for a positive intracardiac RLS was visualization of bubbles in the left atrium within 3 cardiac cycles (38%; 5/13).

When all eligible studies were pooled into the diagnostic accuracy meta-analysis, the overall sensitivity of TTE for the diagnosis of intracardiac RLS was 46.4% (95% CI: 41.1–51.8%; $I^2 = 71.1\%$; Fig. 5A), the specificity was 99.2% (95% CI: 98.4–99.7%; $I^2 = 64.3\%$; Fig. 5B), the positive likelihood ratio (LR+) was 20.85 (95% CI: 10.37–41.92; $I^2 = 23.3\%$; Fig. 5C), and the negative likelihood ratio (LR-) was 0.57 (95% CI: 0.48–0.66; $I^2 = 63.6\%$; Fig. 5D). The included studies were significantly heterogeneous in their estimates of sensitivity, specificity, and LR- (Q

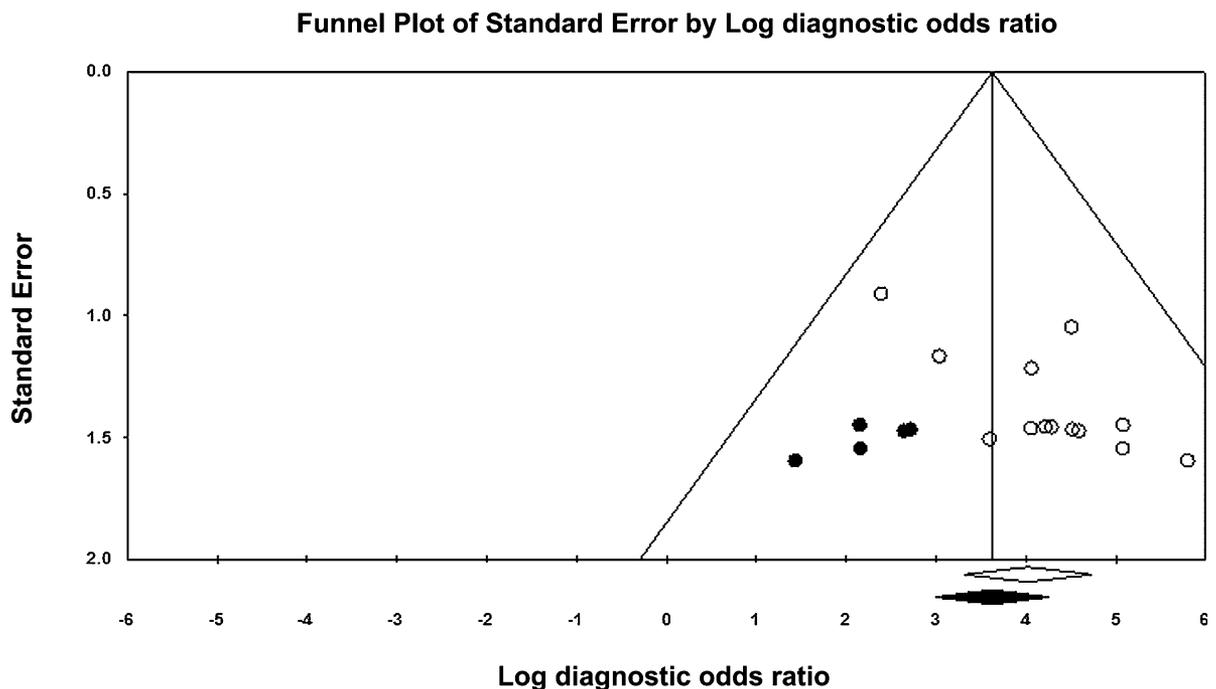


Figure 4. Funnel plot of the 13 eligible studies. White flakes represent the published articles used in the meta-analysis, whereas black flakes represent potentially missed, unpublished articles. White and black diamond’s represent actual and theoretical combined effect sizes, respectively (one-sided P-value = 0.20).

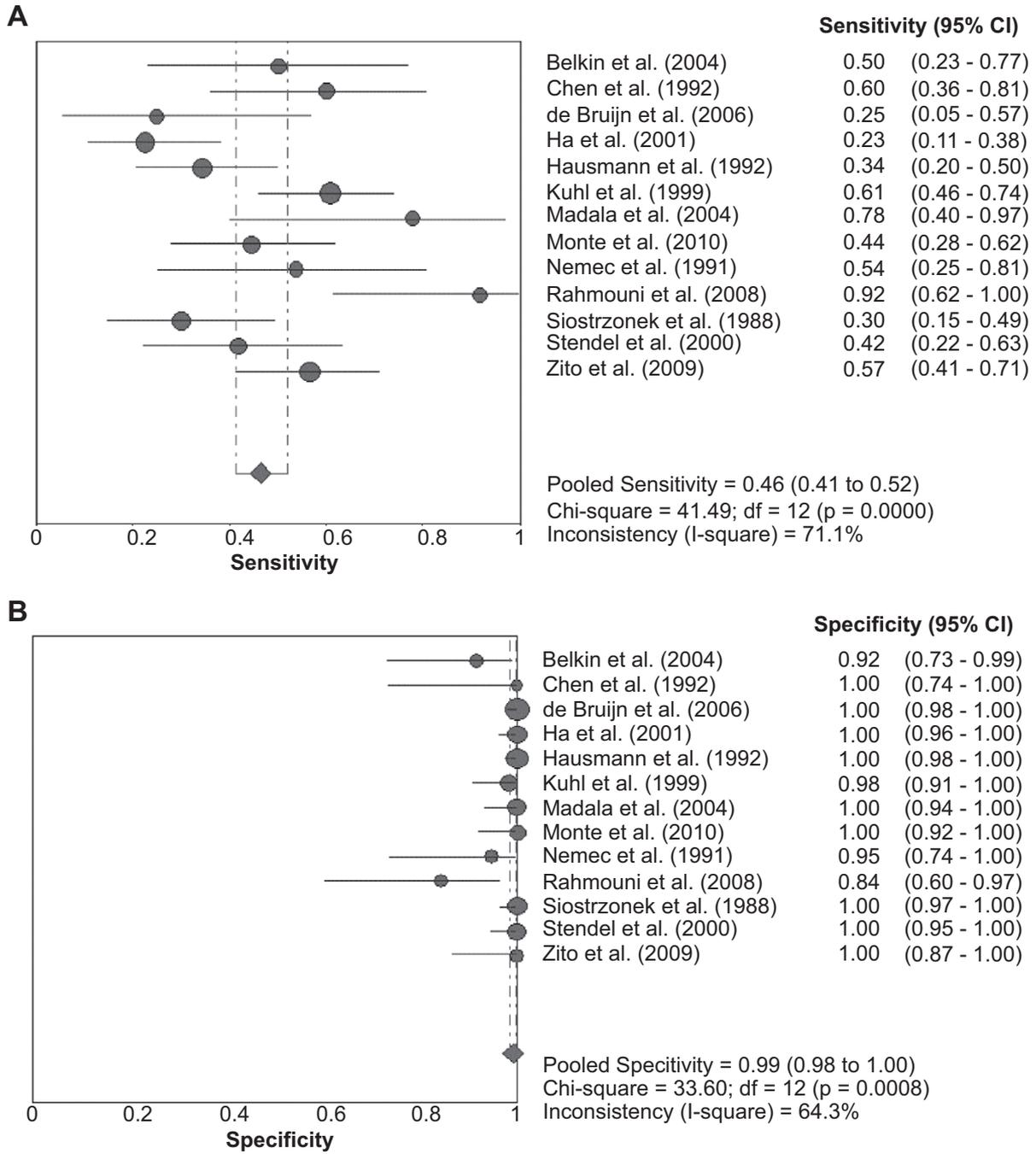
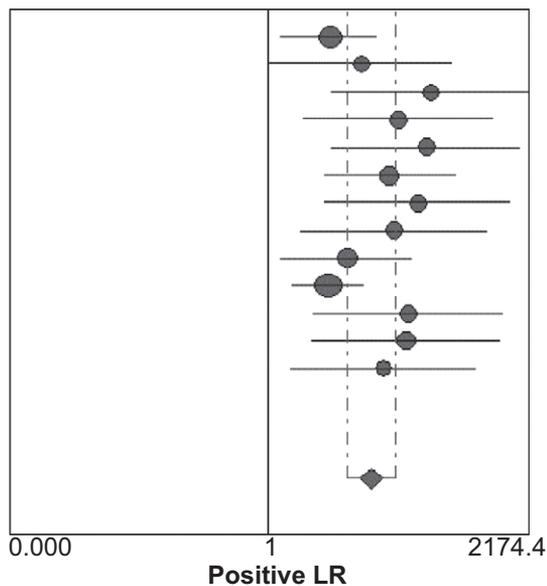


Figure 5. Diagnostic accuracy forest plots. Forest plots of the overall **A.** sensitivity, **B.** specificity, **C.** positive likelihood ratio, and **D.** negative likelihood ratio of intracardiac right-to-left shunts by transthoracic echocardiogram (TTE). The size of each square is proportional to sample size. The horizontal lines in each square show the corresponding 95% confidence intervals (CI). The center of the diamond indicates the overall sensitivity, specificity, positive likelihood, and negative likelihood ratios, respectively and the ends correspond to the 95% CI.

statistic P-values <0.1). Threshold effect was statistically significant (spearman $r = 0.742$; $P = 0.004$), indicating significant heterogeneity between the studies. The stability of our model

was confirmed by the leave-one-out sensitivity analysis which generated pooled estimates close to those obtained with all eligible studies (mean sensitivity 46.4%, range 44.0–49.5%; mean

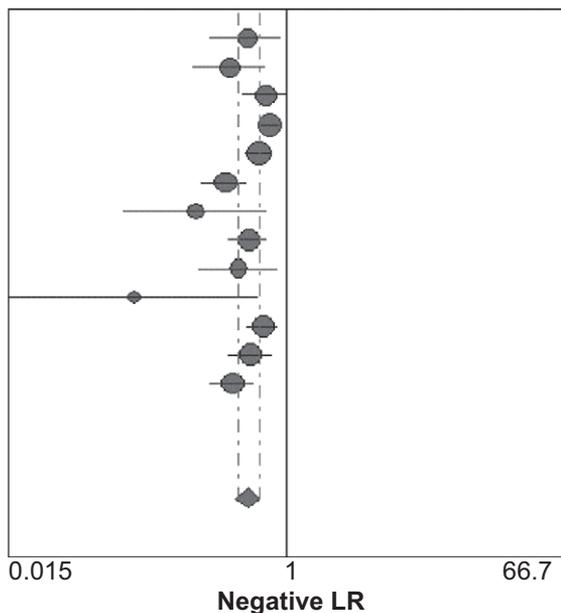
C



Study	Positive LR (95% CI)
Belkin et al. (2004)	6.00 (1.44 - 24.99)
Chen et al. (1992)	15.48 (1.00 - 239.87)
de Bruijn et al. (2006)	118.46 (6.45 - 2,174.38)
Ha et al. (2001)	44.95 (2.68 - 754.36)
Hausmann et al. (1992)	106.78 (6.52 - 1,749.86)
Kuhl et al. (1999)	36.47 (5.16 - 257.89)
Madala et al. (2004)	84.00 (5.20 - 1,357.41)
Monte et al. (2010)	41.03 (2.55 - 661.32)
Nemec et al. (1991)	10.23 (1.42 - 73.58)
Rahmouni et al. (2008)	5.81 (2.03 - 16.63)
Siostrzonek et al. (1988)	64.35 (3.85 - 1,074.78)
Stendel et al. (2000)	57.96 (3.52 - 953.04)
Zito et al. (2009)	30.45 (1.93 - 479.86)

Random Effects Model
 Pooled Positive LR = 20.85 (10.37 to 41.92)
 Cochran-Q = 15.65; df = 12 (p = 0.2079)
 Inconsistency (I-square) = 23.3%
 Tau-squared = 0.3618

D



Study	Negative LR (95% CI)
Belkin et al. (2004)	0.55 (0.32 - 0.93)
Chen et al. (1992)	0.42 (0.25 - 0.72)
de Bruijn et al. (2006)	0.73 (0.53 - 1.02)
Ha et al. (2001)	0.77 (0.65 - 0.91)
Hausmann et al. (1992)	0.66 (0.53 - 0.81)
Kuhl et al. (1999)	0.40 (0.28 - 0.56)
Madala et al. (2004)	0.25 (0.09 - 0.74)
Monte et al. (2010)	0.56 (0.42 - 0.75)
Nemec et al. (1991)	0.49 (0.27 - 0.88)
Rahmouni et al. (2008)	0.10 (0.02 - 0.65)
Siostrzonek et al. (1988)	0.70 (0.55 - 0.88)
Stendel et al. (2000)	0.58 (0.42 - 0.82)
Zito et al. (2009)	0.44 (0.32 - 0.62)

Random Effects Model
 Pooled Negative LR = 0.56 (0.48 to 0.66)
 Cochran-Q = 32.98; df = 12 (p = 0.0010)
 Inconsistency (I-square) = 63.6%
 Tau-squared = 0.0474

Figure 5. (Continued).

specificity 99.2%, range 99.0–99.5%; mean LR+ 21.08, range 18.32–25.24; mean LR– 0.56, range 0.55–0.59). The sROC curve is shown in Figure 5. The pooled AUC and index Q* were 0.94 (95% CI: 0.92–0.97) and 0.88 (95% CI: 0.85–0.92), respectively (Fig. 6).

Figure 7 demonstrates the pre- and posttest probabilities of detecting an intracardiac RLS with TTE in the general population, and in our

study cohort consisting mainly of patients with cryptogenic stroke. Since RLS through a PFO is present in 20% of the adult population³⁴ and in approximately 50% of patients with cryptogenic stroke,^{35,36} these respective prevalences were assumed to demonstrate the likelihood of detecting an intracardiac RLS by TTE in the 2 populations. Figure 6 illustrates the posttest probabilities of diagnosing intracardiac RLS

TABLE I
Characteristics of the Included Studies

Study/Year	Majority Disease	Subjects (n)	Mean Age	Males (%)	Contrast	Timing of Injection Relative to VM	MB Cutoff for Positive TTE/TEE Study	Cardiac Cycle Cutoff for Positive TTE/TEE Study
Belkin 1994 ¹²	Stroke	38	45	23 (61)	Saline	Inj during VM	NS	4 cardiac cycles
Chen 1992 ¹³	Valvular heart disease	32	39	17 (53)	Saline	Inj during VM	≥5	NS
de Bruijn 2006 ¹⁴	Stroke	231	NS	NS	Saline	Inj after VM	NS	NS
Ha 2001 ¹⁵	Stroke	136	59	NS	Saline	Inj during VM	NS	3 cardiac cycles
Hausmann 1992 ¹⁶	Stroke	198	NS	NS	Gelatin	Inj during VM	≥5	3 cardiac cycles
Kuhl 1999 ¹⁷	Stroke	111	56	64 (58)	Gelatin	NS	≥5	3 cardiac cycles
Madala 2004 ¹⁸	Stroke	71	21-88	40 (56)	Saline	Inj during VM	≥1	5 cardiac cycles
Monte 2010 ¹⁹	Migraine	84	40	18 (21)	Saline with blood	NS	≥1	3-5 cardiac cycles
Nemec 1991 ²⁰	Stroke	32	50	14 (44)	Saline with blood	NS	≥1	1-2 cardiac cycles
Rahmouni 2008 ²¹	Stroke	189	59	103 (54)	Saline	Inj during VM	≥1	3 cardiac cycles
Siostrzonek 1988 ²²	Stroke	150	52	88 (59)	Gelatin	NS	NS	3 cardiac cycles
Stendel 2000 ²³	Pre-neurosurgery	92	51	47 (51)	D-galactose or echovist	Inj before VM	≥1	3-10 cardiac cycles
Zito 2009 ²⁴	Stroke & migraine	72	49	33 (46)	Gelatin	Inj before VM	≥1	3-5 cardiac cycles

NS = not specified; MB = microbubbles; VM = Valsalva maneuver; Inj = injection; TTE = transthoracic echocardiogram; TEE = transesophageal echocardiogram.

TABLE II
Accuracies of the Included Studies

Study/Year	TP	FP	FN	TN	Sen (95% CI)	Spec (95% CI)	LR+ (95% CI)	LR- (95% CI)
Belkin 1994 ¹²	7	2	7	22	0.50 (0.23-0.77)	0.92 (0.73-0.99)	6.00 (1.44-24.99)	0.55 (0.32-0.93)
Chen 1992 ¹³	12	0	8	12	0.60 (0.36-0.81)	1.00 (0.74-1.00)	15.48 (1.00-239.87)	0.42 (0.25-0.72)
de Bruijn 2006 ¹⁴	3	0	9	219	0.25 (0.05-0.57)	1.00 (0.98-1.00)	118.46 (6.45-2174.38)	0.73 (0.53-1.02)
Ha 2001 ¹⁵	9	0	31	96	0.23 (0.11-0.38)	1.00 (0.96-1.00)	44.95 (2.68-754.36)	0.77 (0.65-0.91)
Hausmann 1992 ¹⁶	15	0	29	154	0.34 (0.20-0.50)	1.00 (0.98-1.00)	106.78 (6.52-1749.86)	0.66 (0.53-0.81)
Kuhl 1999 ¹⁷	31	1	20	59	0.61 (0.46-0.74)	0.98 (0.91-1.00)	36.47 (5.16-257.89)	0.40 (0.28-0.56)
Madala 2004 ¹⁸	7	0	2	55	0.78 (0.40-0.97)	1.00 (0.94-1.00)	84.00 (5.20-1357.41)	0.25 (0.09-0.74)
Monte 2010 ¹⁹	16	0	20	45	0.44 (0.28-0.62)	1.00 (0.92-1.00)	41.03 (2.55-661.32)	0.56 (0.42-0.75)
Nemec 1991 ²⁰	7	1	6	18	0.54 (0.25-0.81)	0.95 (0.74-1.00)	10.23 (1.42-73.58)	0.49 (0.27-0.88)
Rahmouni 2008 ²¹	11	3	1	16	0.92 (0.62-1.00)	0.84 (0.60-0.97)	5.81 (2.03-16.63)	0.10 (0.02-0.65)
Siostrzonek 1988 ²²	9	0	21	104	0.30 (0.15-0.49)	1.00 (0.97-1.00)	64.35 (3.85-1074.78)	0.70 (0.55-0.88)
Stendel 2000 ²³	10	0	14	68	0.42 (0.22-0.63)	1.00 (0.95-1.00)	57.96 (3.52-953.04)	0.58 (0.42-0.82)
Zito 2009 ²⁴	26	0	20	26	0.57 (0.41-0.71)	1.00 (0.87-1.00)	30.45 (1.93-479.86)	0.44 (0.32-0.62)

TP = true positive; FP = false positive; FN = false negative; TN = true negative; Sen = sensitivity; Spec = specificity; LR+ = positive likelihood ratio; LR- = negative likelihood ratio; CI = confidence interval.

using TTE. As Figure 7A shows, with an RLS prevalence of 20%, a positive TTE result will have an 84% probability (95% CI: 76-89%) of being a true positive. Furthermore, a negative TTE will have a 12% probability (95% CI: 11-14%) of being a false negative. If RLS prevalence is 50% (Fig. 7B) then a positive TTE result will have a 95% probability (95% CI:

92-97%) of being a true positive. In addition, a negative TTE result will have a 36% probability (95% CI: 35-38%) of being a false negative.

Table III summarizes the results of the subgroup analyses. There were no statistically significant differences in sensitivity, specificity, LR+ or LR- between the studies that utilized different contrast agents, different microbubble cutoffs for

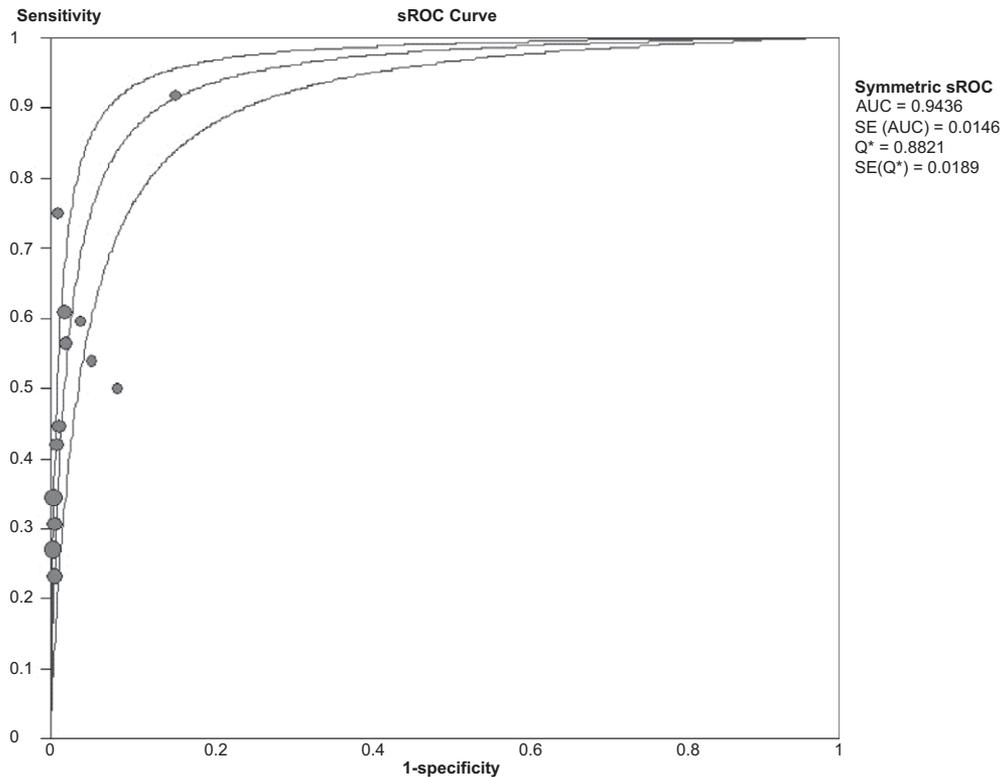


Figure 6. Summary receiver operating characteristic (sROC) curve. Individual study estimates of sensitivity and 1-specificity are represented by the circles. Circle sizes are proportional to study weights. The lateral lines represent 95% confidence intervals.

a positive TTE/TEE, and different cardiac cycle cutoffs for a positive TTE/TEE (all adjusted P-values >0.1).

Discussion:

Our study demonstrates that conventional TTE with contrast detects intracardiac RLS with a sensitivity of 46% and specificity of 99% when contrast TEE is used as the reference technique. In addition, TTE has a LR+ of 20.85 and LR– of 0.57 which specifically make it an excellent rule in test but a poor rule out test for intracardiac RLS in a population of patients with cryptogenic stroke (Fig. 7). A subanalysis of different study protocols revealed that the accuracy of TTE was not significantly affected by the use of different contrast agents, different microembolic cutoffs for a positive TTE/TEE or different cutoffs for visualization of bubbles in the left atrium within a certain number of cardiac cycles after contrast injection. This study is the first meta-analysis to determine the accuracy of conventional TTE in the detection of intracardiac RLS while also comparing different protocols to demonstrate the optimal methodological technique. This is also the first meta-analysis to demonstrate the posttest probabilities of TTE in the general population and in a population of patients with cryptogenic stroke.

While contrast TEE currently remains the gold standard for the diagnosis of RLS through a PFO,^{10,11} TEE may be an uncomfortable and time-consuming procedure. Although extremely uncommon, life-threatening complications such as esophageal bleeding or perforation may occur. Contraindications of TEE such as esophageal or gastric varices, Barrett's esophagus, Zenker's diverticulum, esophageal or pharyngeal carcinoma, strictures, Mallory–Weiss tears or patients with a serious risk of bleeding make a reliable alternative imaging modality a growing need in contemporary clinical practice.³⁷

Transthoracic echocardiogram with bubble study remains the most commonly used initial screening test for RLS due to its easy availability, low cost and noninvasiveness.^{12–24} However, our meta-analysis confirms that the low sensitivity and high LR– of TTE make it a poor rule out test for intracardiac RLS. As an initial screening test, conventional TTE may need to be replaced by more sensitive alternatives such as TTE with second harmonic imaging or transcranial Doppler, which have reported sensitivities that are comparable to TEE.^{18,20} Two recent meta-analyses of prospective studies comparing TTE harmonic imaging and transcranial Doppler to TEE for the diagnosis of intracardiac RLS demonstrated a sen-

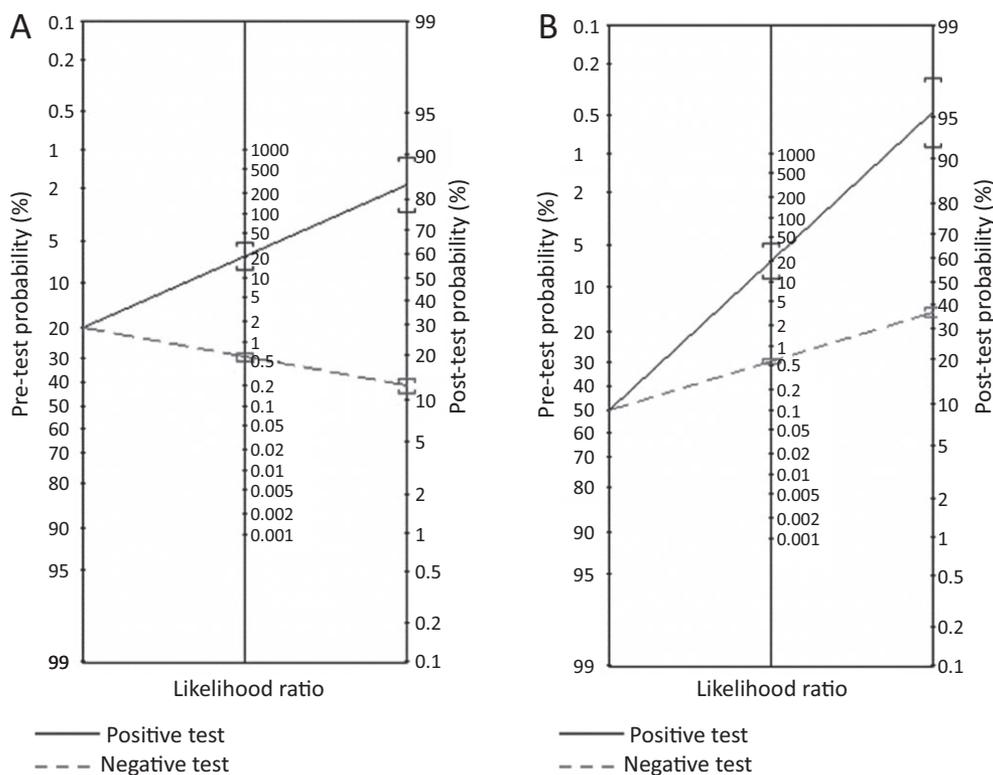


Figure 7. Fagan nomograms demonstrating the pre- and posttest probabilities of intracardiac right-to-left shunts (RLS) after trans-thoracic echocardiogram (TTE). Assuming an RLS prevalence of 20%, a positive TTE result will have a 84% probability (95% confidence intervals [CI] 76–89%) of being a true positive, whereas a negative TTE will have a 12% probability (95% CI: 11–14%) of being a false negative **A**. Assuming an RLS prevalence of 50%, a positive TTE result will have a 95% probability (95% CI: 92–97%) of being a true positive, whereas a negative TTE result will have a 36% probability (95% CI: 35–38%) of being a false negative **B**.

sitivity of 91% and specificity of 93% with TTE harmonic imaging (unpublished data), and a sensitivity of 97% and specificity of 93% with transcranial Doppler.³⁸ However, the excellent specificity and LR+ of conventional TTE make a positive result highly useful, especially in a population of patients with cryptogenic stroke (Fig. 7).

There is currently no standardized protocol for performing a TTE bubble study and the methodology often varies depending on the institution. Thus, there was considerable variability in the diagnostic protocols of the included studies. Considering the heterogeneity of the studies, we sought to examine the effect of different protocols on the accuracy and utility of TTE. According to our secondary analysis, using different contrast agents, different microembolic cutoffs for a positive TTE, and different cutoffs for the appearance of bubbles in the left atrium within a different number of cardiac cycles, did not affect the accuracy of TTE. Thus, while the exact protocol of TTE bubble study is variable among different institutions, our study indicates that changing these parameters may neither enhance nor reduce the accuracy of TTE.

Limitations:

In performing our analysis, we recognize the presence of several limitations. Differences in study design and diagnostic threshold for detecting a positive test are sources of heterogeneity that limit our calculation of diagnostic accuracy. We attempted to perform a subanalysis on different protocols, where possible, to assess the effect of changing the TTE protocol on accuracy of the test. However, we were unable to perform a subanalysis on some parameters such as utility of agitated saline with blood as the contrast agent or comparing the timing of contrast injection relative to Valsalva maneuver, due to the lack of the minimum number studies required to make these comparisons.

The methodological quality of the included studies was assessed using quality assessment of diagnostic accuracy studies (QUADAS) to determine the effect of methodology, and inherent shortcomings, on diagnostic accuracy. Most of the studies in this meta-analysis had high methodological quality with very minimal concerns regarding study biases (Figs. 2 and 3). Of the 13 included studies, 62% (8/13) did not mention the time interval between the index and refer-

TABLE III
Subgroup Analyses of Diagnostic Accuracy Variables

Parameter	Subgroups	Sensitivity*	Specificity*	Positive Likelihood Ratio*	Negative Likelihood Ratio*
Microembolic cutoff for positive TTE/TEE study	≥1 microbubble (n = 6)	55.0% (46.4–63.4%); I ² = 62.3%	98.3% (95.6–99.5%); I ² = 68.8%	17.31 (5.98–50.09); I ² = 31.9%	0.493 (0.391–0.622); I ² = 32.3%
	≥5 microbubble (n = 3)	50.4% (41.0–59.9%); I ² = 74.1%	99.6% (97.6–100%); I ² = 24.9%	38.13 (9.56–152.12); I ² = 0	0.497 (0.339–0.727); I ² = 72.7%
Cardiac cycle cutoff for positive TTE/TEE study	Within 3 cardiac cycles (n = 5)	42.4% (35.0–50.0%); I ² = 86.8%	99.1% (97.7–99.7%); I ² = 78.6%	27.26 (6.49–114.55); I ² = 56.2%	0.605 (0.458–0.798); I ² = 79.1%
	Within more than 3 cardiac cycles (n = 5)	51.2% (42.2–60.1%); I ² = 16.2%	99.1% (96.7–99.9%); I ² = 55.5%	21.57 (6.60–70.49); I ² = 21.1%	0.520 (0.439–0.617); I ² = 0
Type of Contrast	Saline (n = 6)	45.8% (36.1–55.7%); I ² = 82.6%	98.8% (97.3–99.6%); I ² = 79.2%	14.84 (5.18–42.52); I ² = 38.4%	0.534 (0.360–0.792); I ² = 74.9%
	Gelatin Based (n = 4)	47.4% (39.7–55.1%); I ² = 75.4%	99.7% (98.4–100%); I ² = 14.4%	48.56 (14.01–168.38); I ² = 0	0.549 (0.422–0.713); I ² = 73.0%
	Other (n = 3)	45.2% (33.5–57.3%); I ² = 0	99.2% (95.9–100%); I ² = 28.4%	22.32 (5.53–90.08); I ² = 0	0.560 (0.456–0.688); I ² = 0

TTE = transthoracic echocardiogram; TEE = transesophageal echocardiogram.

*Data in parentheses are 95% confidence intervals.

ence tests. Although this could potentially lead to disease progression bias, this limitation may not apply in our study as PFO is a congenital condition and the results of the tests are unlikely to be affected by a delayed time period between the 2 tests. Uninterpretable results and withdrawals, which represented the “flow and timing” section, were unclear in 54% (7/13) and 92% (12/13) of the studies, respectively. These 2 data are often not reported in diagnostic accuracy studies with the uninterpretable results and withdrawals simply removed from the analysis. This may lead to a biased assessment of the test characteristics. Whether or not bias will arise depends on the possible correlation between uninterpretable test results and the true disease status. Uninterpretable results frequently occur randomly and are not related to the true disease status of the individual. Therefore, in theory, these should not have any effect on test performance. Likewise, 6 of 13 studies (46%) did not clearly specify whether or not blinding occurred in the index or reference tests. This may potentially lead to review bias resulting in inflated measures of diagnostic accuracy.

Finally, TEE itself may have limitations in the detection of RLS in some patients, due to sedation and the difficulty of performing an adequate Valsalva maneuver with a probe in the esophagus. Studies that have compared TEE to autopsy or intraoperative detection of PFO have demonstrated that the diagnosis can sometimes be missed by TEE.^{39,40} Other studies that determined accuracy of TEE in the detection of PFO using catheterization and/or surgery as the reference demonstrated a sensitivity of 91–100%.^{13,41} Thus, our results may have overestimated the sensitivity of TTE. In addition, some false-positive TTE or TEE studies may occur due to a pulmonary arteriovenous fistula that can be mistaken for an intracardiac shunt. Since our study cohort mainly consisted of patients with cryptogenic stroke, the presence of observed FP due to pulmonary shunting may be higher as pulmonary shunting occurs more commonly in patients with cryptogenic stroke compared to age-matched controls.⁴²

Conclusion:

Although TTE bubble study remains the most commonly used initial screening test for RLS, it has a poor sensitivity which makes it an unreliable rule out test. With a specificity of 99% and LR+ of 20.85, a positive TTE result yields a 95% posttest probability of RLS in patients with cryptogenic stroke. The accuracy of TTE is not affected by different contrast agents, and different cutoffs used for a varying number of microbubbles appearing in the left atrium within a varying number of cardiac cycles after contrast injection.

References

- Jones HR, Caplan LR, Come PC, et al: Cerebral emboli of paradoxical origin. *Ann Neurol* 1983;13:14–19.
- Biller J, Johnson MR, Adams HP: Further observations on cerebral or retinal ischemia in patients with right-to-left intra-cardiac shunts. *Arch Neurol* 1987;44:740–743.
- Webster MW, Chancellor AM, Smith HJ, et al: Patent foramen ovale in young stroke patients. *Lancet* 1988;2:11–12.
- Agarwal S, Bajaj NS, Kumbhani DJ, et al: Meta-analysis of transcatheter closure versus medical therapy for patent foramen ovale in prevention of recurrent neurological events after presumed paradoxical embolism. *JACC Cardiovasc Interv* 2012;5:777–789.
- Rengifo-Moreno P, Palacios IF, Junpaparp P, et al: Patent foramen ovale transcatheter closure vs. medical therapy on recurrent vascular events: A systematic review and meta-analysis of randomized controlled trials. *Eur Heart J* 2013;34:3342–3352.
- Azarbal B, Tobis JM, Suh W, et al: Association of interatrial shunts and migraine headaches: Impact of transcatheter closure. *J Am Coll Cardiol* 2005;45:489–492.
- Shanoudy H, Soliman A, Raggi P, et al: Prevalence of patent foramen ovale and its contribution to hypoxemia in patients with obstructive sleep apnea. *Chest* 1998;113:91–96.
- Billinger M, Zbinden R, Mordasini R, et al: Patent foramen ovale closure in recreational divers: Effect on decompression illness and ischemic brain lesions during long-term follow-up. *Heart* 2011;97:1932–1937.
- Cheng TO: Mechanisms of platypnea-orthodeoxia: What causes water to flow uphill? *Circulation* 2002;105:e47.
- Pearson AC, Labovitz AJ, Tatineni S, et al: Superiority of transesophageal echocardiography in detecting cardiac sources of embolism in patients with cerebral ischemia of uncertain etiology. *J Am Coll Cardiol* 1991;17:66–72.
- De Belder MA, Tourikis L, Griffith M: Transesophageal contrast echocardiography and color flow mapping: Methods of choice for the detection of shunts at atrial level? *Am Heart J* 1992;124:1545–1550.
- Belkin RN, Pollack BD, Ruggiero ML, et al: Comparison of transesophageal and transthoracic echocardiography with contrast and color flow Doppler in the detection of patent foramen ovale. *Am Heart J* 1994;128:520–525.
- Chen WJ, Kuan P, Lien WP, et al: Detection of patent foramen ovale by contrast transesophageal echocardiography. *Chest* 1992;101:1515–1520.
- de Bruijn SF, Agema WR, Lammers GJ, et al: Transesophageal echocardiography is superior to transthoracic echocardiography in management of patients of any age with transient ischemic attack or stroke. *Stroke* 2006;37:2531–2534.
- Ha JW, Shin MS, Kang S, et al: Enhanced detection of right-to-left shunt through patent foramen ovale by transthoracic contrast echocardiography using harmonic imaging. *Am J Cardiol* 2001;87:669–671.
- Hausmann D, Mügge A, Becht I, et al: Diagnosis of patent foramen ovale by transesophageal echocardiography and association with cerebral and peripheral embolic events. *Am J Cardiol* 1992;70:668–672.
- Kühl HP, Hoffmann R, Merx MW, et al: Transthoracic echocardiography using second harmonic imaging: Diagnostic alternative to transesophageal echocardiography for the detection of atrial right to left shunt in patients with cerebral embolic events. *J Am Coll Cardiol* 1999;34:1823–1830.
- Madala D, Zaroff JG, Hourigan L, et al: Harmonic imaging improves sensitivity at the expense of specificity in the detection of patent foramen ovale. *Echocardiography* 2004;21:33–36.
- Monte I, Grasso S, Licciardi S, et al: Head-to-head comparison of real-time three-dimensional transthoracic echocardiography with transthoracic and transesophageal two-dimensional contrast echocardiography for the detection of patent foramen ovale. *Eur J Echocardiogr* 2010;11:245–249.
- Nemec JJ, Marwick TH, Lorig RJ, et al: Comparison of transcranial Doppler ultrasound and transesophageal contrast echocardiography in the detection of interatrial right-to-left shunts. *Am J Cardiol* 1991;68:1498–1502.
- Rahmouni HW, Keane MG, Silvestry FE, et al: Failure of digital echocardiography to accurately diagnose intra-cardiac shunts. *Am Heart J* 2008;155:161–165.
- Siostrzonek P, Zangeneh M, Gössinger H, et al: Comparison of transesophageal and transthoracic contrast echocardiography for detection of a patent foramen ovale. *Am J Cardiol* 1991;68:1247–1249.
- Stendel R, Gramm HJ, Schröder K, et al: Transcranial Doppler ultrasonography as a screening technique for detection of a patent foramen ovale before surgery in the sitting position. *Anesthesiology* 2000;93:971–975.
- Zito C, Dattilo G, Oretto G, et al: Patent foramen ovale: Comparison among diagnostic strategies in cryptogenic stroke and migraine. *Echocardiography* 2009;26:495–503.
- Whiting P, Rutjes AW, Reitsma JB, et al: The development of QUADAS: A tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;3:25–38.
- R Development Core Team. R: *A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing; 2012
- Comprehensive Meta-Analysis version 2. Available at <http://www.meta-analysis.com/index.html>
- Zamora J, Abraira V, Muriel A, et al: Meta-DiSc: A software for meta-analysis of test accuracy data. *BMC Med Res Methodol* 2006;6:31.
- Deeks JJ: Systematic reviews in health care: Systematic reviews of evaluations of diagnostic and screening tests. *BMJ* 2001;323:157–162.
- DerSimonian R, Laird N: Meta-analysis in clinical trials. *Control Clin Trials* 1986;7:177–188.
- Walter SD: Properties of the summary receiver operating characteristic (SROC) curve for diagnostic test data. *Stat Med* 2002;21:1237–1256.
- Altman DG, Bland JM: Interaction revisited: The difference between two estimates. *BMJ* 2003;326:219.
- Nakanishi Y, Toyozumi S, Nakajima A, et al: Subgroup analysis in meta-analysis: A comparison of different methods. *Jpn J Pharmacoepidemiol* 2007;12:13–24.
- Hagen PT, Scholz DG, Edwards WD: Incidence and size of patent foramen ovale during the first 10 decades of life: An autopsy study of 965 normal hearts. *Mayo Clin Proc* 1984;59:17–20.
- Job FP, Ringelstein EB, Grafen Y, et al: Comparison of transcranial contrast Doppler sonography and transesophageal contrast echocardiography for the detection of patent foramen ovale in young stroke patients. *Am J Cardiol* 1994;74:381–384.
- Tobis JM, Azarbal B: Does patent foramen ovale promote cryptogenic stroke and migraine headache? *Tex Heart Inst J* 2005;32:362–365.
- Hilberath JN, Oakes DA, Shernan SK, et al: Safety of transesophageal echocardiography. *J Am Soc Echocardiogr* 2010;23:1115–1127.
- Mojadidi MK, Roberts SC, Winoker JS, et al: Accuracy of transcranial Doppler for the diagnosis of intracardiac right-to-left shunt: A bivariate meta-analysis of prospective studies. *J Am Coll Cardiol Img* 2014;7:236–250.
- Schneider B, Zienkiewicz T, Jansen V, et al: Diagnosis of Patent Foramen Ovale by Transesophageal Echocardiog-

- raphy and Correlation with Autopsy Findings. *Am J Cardiol* 1996;77:1202–1209.
40. Konstadt SN, Louie EK, Black S, et al: Intraoperative detection of patent foramen ovale by transesophageal echocardiography. *Anesthesiology* 1991;74:212–216.
 41. Konstantinides S, Kasper W, Geibel A, et al: Detection of left-to-right shunt in atrial septal defect by negative contrast echocardiography: A comparison of transthoracic and transesophageal approach. *Am Heart J* 1993;126:909–917.
 42. Abushora MY, Bhatia N, Alnabki Z, et al: Intrapulmonary shunt is a potentially unrecognized cause of ischemic

stroke and transient ischemic attack. *J Am Soc Echocardiogr* 2013;26:683–690.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Data S1. Supplementary results.