Necrotizing Soft Tissue Infection: Diagnostic Accuracy of Physical Examination, Imaging, and LRINEC Score

A Systematic Review and Meta-Analysis

Shannon M. Fernando, MD, MSc,† Alexandre Tran, MD, MSc,‡§ Wei Cheng, PhD,¶ Bram Rochwerg, MD, MSc,|| Kwadwo Kyeremanteng, MD, MHA,†¶ Andrew J. E. Seely, MD, PhD,†‡§¶ Kenji Inaba, MD,** and Jeffrey J. Perry, MD, MSc*‡¶*

Objective: We sought to summarize accuracy of physical examination, imaging, and Laboratory Risk Indicator for Necrotizing Fasciitis (LRINEC) score in diagnosis of necrotizing soft tissue infection (NSTI) in adults with a soft tissue infection clinically concerning for NSTI.

Summary of Background Data: NSTI is a life-threatening diagnosis. Delay to diagnosis and surgical management is associated with increased mortality.

Methods: We searched 6 databases from inception through November 2017. We included English-language studies reporting diagnostic accuracy of testing or LRINEC Score. Outcome was NSTI confirmed by surgery or histopathology. Two reviewers screened all citations and extracted data independently. Summary measures were obtained from the Hierarchical Summary Receiver Operating Characteristic model.

Results: From 2,290 citations, we included 23 studies (n = 5982). Of physical examination signs, pooled sensitivity and specificity for fever was 46.0% and 77.0% respectively, for hemorrhagic bullae 25.2% and 95.8%, and for hypotension 21.0% and 97.7%. Computed tomography (CT) had sensitivity of 88.5% and specificity of 93.3%, while plain radiography had sensitivity of 48.9% and specificity of 94.0%. Finally, LRINEC $\geq$ 6 had sensitivity of 68.2% and specificity of 84.8%, while LRINEC $\geq$ 8 had sensitivity of 40.8% and specificity of 94.9%.

Conclusions: Absence of any 1 physical examination feature (eg, fever or hypotension) is not sufficient to rule-out NSTI. CT is superior to plain radiography. LRINEC had poor sensitivity, and should not be used to rule-out NSTI. Given the poor sensitivity of these tests, a high clinical suspicion warrants early surgical consultation for definitive diagnosis and management.

Keywords: necrotizing soft tissue infection, necrotizing fasciitis, Laboratory Risk Indicator for Necrotizing Fasciitis, septic shock, critical care, computed tomography

(*Ann Surg* 2018;xx:xxx–xxx)

Necrotizing soft tissue infection (NSTI, commonly referred to a "necrotizing fasciitis") is a life-threatening skin and soft tissue diagnosis that is characterized by widespread tissue necrosis.[1,2] NSTI is often severe, rapidly progressive, and associated with sepsis and multisystem organ failure. Despite advances in care, mortality from NSTI remains high, estimated between 20% and 30%.[3,4] Rapid identification of NSTI and urgent surgical debridement of necrotic tissue are critical,[1,2] and delays to surgical intervention are associated with increased mortality.[5] NSTI is a rare disease, with an incidence of 0.3 to 5 per 100,000,[4,6] and therefore differentiation of NSTI from other more common clinical entities such as cellulitis can be difficult. Commonly described risk factors such as intravenous drug use, immunosuppression, and diabetes mellitus may also be seen in other severe skin infections.[7] To assist in making the diagnosis of NSTI, clinicians often rely upon physical examination, diagnostic imaging, and clinical decision instruments; however, little evidence validates the diagnostic utility of these assessments.

Some classic physical examination signs have been described to differentiate NSTI from other skin and soft tissue infections. These include erythema, soft-tissue edema, severe pain (often described as "pain out of proportion"), fever, and hemorrhagic bullae.[8] The presence of hypotension and shock are also thought to indicate those likely to have NSTI.[5] Various imaging modalities have also been utilized to help make the diagnosis. Plain radiography may demonstrate gas in the soft tissues.[1,2] Computed tomography (CT) performed with contrast may demonstrate fascial air or gas, soft tissue edema, or enhancement of the fascia.[9] Although thought to be more accurate, CT is time-consuming and can delay definitive surgical management.

Finally, laboratory values are often utilized to aid in the diagnosis of NSTI. The Laboratory Risk Indicator for Necrotizing Fasciitis (LRINEC) score is a diagnostic clinical decision instrument validated for differentiating NSTI from other soft tissue infections.[10] LRINEC utilizes 6 laboratory serum parameters including white blood cell (WBC) count, hemoglobin, sodium, glucose, creatinine, and C-reactive protein. A score $\geq$ 6 (traditional threshold for diagnosis of NSTI) indicates a "moderate" risk of NSTI

From the *Department of Emergency Medicine, University of Ottawa, Ottawa, Ontario, Canada; †Division of Critical Care, Department of Medicine, University of Ottawa, Ottawa, Ontario, Canada; ‡School of Epidemiology and Public Health, University of Ottawa, Ottawa, Ontario, Canada; §Department of Surgery, University of Ottawa, Ottawa, Ontario, Canada; ¶Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada; ||Department of Medicine, Division of Critical Care, and Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada; and **Division of Acute Care Surgery, Department of Surgery, University of Southern California, Los Angeles, CA.

SMF, AT, and JJP conceived the study idea. SMF, AT, WC, and JJP coordinated the systematic review. SMF and AT designed the search strategy. SMF and AT screened abstracts and full texts. SMF and AT acquired the data and judged risk of bias in the studies. WC performed the data analysis. BR created the GRADE evidence profiles. SMF, AT, WC, BR, KK, AJES, KI, and JJP interpreted the data analysis and critically revised the manuscript. All authors have had the opportunity to review the final manuscript, and provided their permission to publish the manuscript. All authors agree to take responsibility for the work.

AJES holds patents related to multiorgan variability analysis, and has shares in Therapeutic Monitoring Systems Inc. None of the other authors report any conflict of interest.

Registration: PROSPERO CRD42017081976.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.annalsofsurgery.com).

Reprints: Shannon M. Fernando, MD, MSc, Department of Emergency Medicine, The Ottawa Hospital, Civic Campus, 1053 Carling Ave., Ottawa, ON, Canada K1Y 4E9. E-mail: sfernando@qmed.ca.

ISSN: 0003-4932/16/XXXX-0001

DOI: 10.1097/SLA.0000000000002774

(50–75% probability), whereas a score ≥8 indicates a ''high'' risk (greater than 75% probability).

We conducted a systematic review and meta-analysis with the primary objective of obtaining summary estimates of diagnostic performance (including sensitivity and specificity) across studies of physical examination, imaging, and LRINEC score for the diagnosis of NSTI in patients where the diagnosis was being considered.

METHODS

We structured this systematic review according to PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines,[11,12] the Cochrane Handbook for Diagnostic Test Accuracy,[13] and existing guidelines for reviews of diagnostic accuracy.[14] The study protocol was registered with the PROSPERO registry (CRD42017081976).

Search Strategy

We searched MEDLINE, PubMed, EMBASE, Scopus, Web of Science, and the Cochrane Database of Systematic Reviews from their respective inception to November 13, 2017. An experienced health sciences librarian assisted in the development of the search strategy. The search was conducted using the terms ''necrotizing fasciitis,'' ''necrotizing skin and soft tissue infection,'' ''necrotizing soft tissue infection,'' ''gas gangrene,'' and ''fournier's gangrene'' (Search strategy is depicted in eFigure 1; http://links.lww.com/SLA/B411). We utilized Science Citation Index to retrieve reports citing the relevant articles identified from our search, and then entered them into PubMed. We then conducted further surveillance searches using the previously described 'Related Articles' feature[15] to identify further reports.

Study Selection

We included all English-language abstracts and full-text articles describing retrospective and prospective observational studies, as well as randomized controlled trials and quasi-randomized controlled trials. We included studies meeting the following criteria: (1) enrolled adult patients (≥16 years) with suspected NSTI; (2) conducted in the emergency department (ED), the hospital wards or intensive care unit (ICU); and (3) evaluated the test characteristics of: physical examination, imaging modalities, or LRINEC score for diagnosis of NSTI. Diagnosis of NSTI had to be defined by any of the following: Operative findings (presence of grayish necrotic fascia, demonstration of a lack of resistance to normally adherent muscular fascia to blunt dissection, lack of bleeding of the fascia during dissection, or the presence of foul smelling ''dishwater'' pus), histopathologic tissue examination, or death from suspected NSTI. We excluded case reports, case series, animal studies, pediatric studies and observational studies evaluating prognosis in cohorts of patients with confirmed NSTI only (ie, without controls). Each study was required to have a 2 × 2 table of true positive, false negative, true negative, and false positive counts, either extracted from the original article or calculated from other reported information such as declared sensitivity and specificity. In instances where these values could not be obtained from the reported data, we contacted authors. If the corresponding author did not respond after 3 attempts, the study was excluded.

We screened studies using Covidence software (Melbourne, Australia). Titles were imported into Covidence directly from the search databases, and duplicates were removed. In phase 1, 2 reviewers (SMF and AT) independently screened the titles and abstracts of all identified studies. Disagreements regarding citation inclusion were resolved by consensus, and no third-party adjudication was necessary. In phase 2, the same 2 reviewers independently assessed full texts of the selected articles from phase 1. Disagreements regarding citation inclusion were resolved by consensus.

Data Extraction

One investigator (SMF) collected the following variables from the included articles: author information, year of publication, study design, eligibility criteria, details regarding CT imaging technique, number of patients included, mean or median age, and number of deaths. We used a pre-designed data extraction sheet (eTable 1; http://links.lww.com/SLA/B411) to minimize the risk for transcriptional errors. Subsequently, 2 investigators (SMF and AT) independently collected the true positive, false positive, false negative, and true negative counts, total number of diagnosed NSTI cases, and stated sensitivity and specificity of diagnostic tests from all included trials.

Quality Assessment

Two reviewers (SMF and AT) independently assessed the risk of bias of the included studies, using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool.[16] Disagreements regarding risk of bias among citations were resolved through consensus. The QUADAS-2 assesses 4 potential areas for bias and applicability of the research question: (1) patient selection: risk of bias is considered high if there is non-consecutive enrollment, the use of case-control study design, or inappropriate exclusions; (2) index test: risk of bias is considered high if the index test results were interpreted without explicit blinding to the reference standard (ie, definitive diagnosis of NSTI); (3) reference standard (operative or histopathologic diagnosis of NSTI): risk of bias is considered high if the reference standard could misclassify the target condition; and (4) Flow and timing: risk of bias is considered high if not all patients had the diagnostic test applied using the same criteria, if the diagnostic test was calculated at an inappropriate time interval prior to definitive operative or histopathologic diagnosis, or if patients were excluded from analyses.

Evidence Synthesis

We presented individual study results graphically by plotting sensitivity and specificity estimates on one-dimensional forest plots (ordered by sensitivity) as well as on the Receiver Operating Characteristic (ROC) space, to visually assess for heterogeneity. To pool the results, we applied the Hierarchical Summary Receiver Operating Characteristic (HSROC) model[17] and obtained summary point estimates of the pairs of sensitivity and specificity, as well as Diagnostic Odds Ratios (OR) and likelihood ratios, with their 95% confidence intervals (CI). The HSROC model appropriately incorporates both within-study and between-study variability. Summary estimates of test accuracy were plotted in the ROC space together with the summary ROC curve. The analyses were conducted using MetaDAS (Version 1.3),[18] as recommended by the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy.[13] We conducted sensitivity analyses for parameters that had 3 or more studies remaining after excluding studies with high risk-of-bias. Univariate tests for heterogeneity in sensitivity and specificity are not recommended by the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy, as they do not account for heterogeneity explained by phenomena such as positive threshold effects.[13] Instead, it is preferable to demonstrate heterogeneity graphically through the scatterplot surrounding the summary ROC curve, and the confidence/prediction regions of the summary point in addition to the forest plots, as we have done in previous systematic reviews of diagnostic test accuracy.[19]

We assessed the overall confidence in pooled diagnostic effect estimates using the Grading of Recommendations, Assessments, Development and Evaluation (GRADE) approach.[20] Assessments were based on the following criteria: risk-of-bias of the included

studies, precision, consistency, directness of the evidence, and risk of publication bias. The overall confidence in effect estimates were categorized into 1 of 4 levels, which included high, moderate, low, or very low. A GRADE evidence profile was created using the guideline development tool (gradepro.org).

RESULTS

Search Results

A total of 2290 citations were identified through the relevant searches (Fig. 1). Following removal of duplicates, 1661 studies were screened, and 30 studies underwent full-text review. Following this, we included 24 cohorts from 23 studies in the meta-analysis.[10,21–42] One study examining the LRINEC score[10] contained both an internal derivation cohort and an external validation cohort. Only the validation cohort was included in our analysis of LRINEC diagnostic accuracy. Only 3 physical examination findings had at least 3 relevant studies allowing for meta-analyses: fever, hemorrhagic bullae, and hypotension. Four studies evaluated the diagnostic accuracy of fever (defined as body temperature $\geq$38.0°C),[10,21,22,24] 5 studies evaluated the presence of hemorrhagic bullae,[21,24,32,38,39] and 6 studies evaluated the presence of hypotension (defined as a systolic blood pressure $\leq$90 mmHg).[10,21,24,25,38,39] Four studies investigated the diagnostic accuracy of plain radiography,[29,37–39] while 7 studies evaluated the presence of fascial gas on CT.[23,29,31,32,37,40,41] Six of the 7 studies investigating fascial gas also evaluated accuracy of the presence of any additional subtle findings on CT, namely fascial enhancement or fascial edema.[23,29,31,32,40,41] Finally, LRINEC was analyzed at 2 thresholds. 14 studies evaluated the diagnostic accuracy of a LRINEC score $\geq$6,[10,22–24,26–28,30,33–37,42] while 9 studies also evaluated a LRINEC score $\geq$8.[10,22,23,27,28,33,34,36,37] One study required contact with the corresponding author in order to obtain 2×2 table counts.[29]

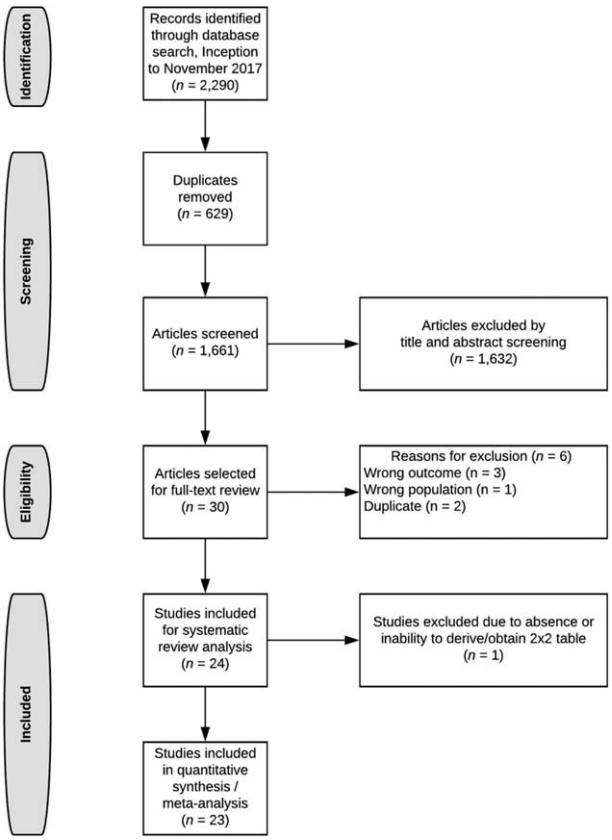

FIGURE 1. Flow chart summarizing evidence search and study selection.

TABLE 1. Characteristics of the 23 Included Studies

Description	Frequency (%)
Continent of Study	
North America	10 (43.5)
Asia	7 (30.4)
Europe	4 (17.4)
Australia/Oceania	2 (8.7)
Year of Publication	
2000–2004	4 (17.4)
2005–2009	1 (4.3)
2010–2014	7 (30.4)
2015–2017	11 (47.8)
Publication	
Full-Text Article	22 (95.7)
Published Conference Abstract	1 (4.3)
Study Design	
Retrospective Cohort	16 (69.6)
Prospective Cohort	2 (8.7)
Retrospective Case-Control	5 (21.7)
Definition of 'Suspicion of NSTI'	
All Skin and Soft Tissue Infections	6 (26.1)
Patients undergoing Imaging to rule-out NSTI	5 (21.7)
Patients taken to Operating Room to rule-out NSTI	2 (8.7)
Physician diagnosis of suspected NSTI	3 (13.0)
Case-Control design	5 (21.7)
Other	2 (8.7)

NSTI indicates necrotizing soft tissue infection.

Study Characteristics

Table 1 describes the 23 included studies, and eTable 2; http://links.lww.com/SLA/B411 provides more details on individual study characteristics. Of the studies included, 43.5% were conducted in North America, while 30.4% were conducted in Asia, and 17.4% were conducted in Europe. 16 studies (69.6%) were retrospective cohort studies, while 2 (8.7%) were prospective cohort studies, and 5 (21.7%) were retrospective case-control studies. There were no randomized controlled trials included. The included studies used variable definitions for 'suspected NSTI'. Six studies (26.1%) recruited all consecutive patients presenting with a skin and soft tissue infection.[30,35,37,39,40,42] A further 5 studies (21.7%) only included patients who underwent imaging for suspected NSTI.[23,27,31,32,41] Two studies (8.7%) included only patients taken to the operating room for suspected NSTI,[28,29] and 3 other studies (13.0%) included patients with a physician diagnosis of suspected NSTI.[26,34,36] Five studies (21.7%) utilized a case-control design,[10,21,22,33,38] including all consecutive cases of NSTI, and comparing them to a random selection of control cases with a non-necrotizing skin and soft tissue infection. Associative comparisons of patient demographic and risk factors between NSTI and non-NSTI control patients for each study are depicted in Table 2. Diabetes was found to be a significant NSTI risk factor in 4 of 8 studies, immunocompromised status in 4 of 6 studies and intravenous drug use in 2 of 3 studies. For physical exam findings, the classical "pain out of proportion" was a significant risk factor in 1 of 3 studies. Of the LRINEC score components, white blood cell count was most commonly found to be a significant predictor, noted in 6 of 8 studies.

TABLE 2. Comparisons of Demographic and Risk Factors in Included Studies

	Demographics		Comorbidities						Physical Exam Findings					Vital Signs		
	Age	Sex	Diabetes	Renal Failure	Liver Disease	Immune Compro-mised	Alcohol Abuse	IV Drug Use	Pain out of Proportion	Erythema	Edema	Necrosis	Crepitus	HR	T	SBP
Alayed (2015)		NS	NS	NS	NS	NS	NS	NS								
Borschitz (2015)	NS	NS	✓	✓			✓		✓	NS	NS	NS	✓	✓	✓	
Chao (2012)	NS	NS	NS	✓	✓	✓								✓	✓	
Cranendonck (2017)	NS	NS	NS	NS		✓										
Liao (2012)	NS	✓	✓	✓	✓	✓										
Martinez (2017)	NS	NS														
McGillicuddy (2011)	NS	NS	✓			✓	NS	✓		NS		✓	NS	NS		NS
Narasimhan (2017)	✓	NS	NS	NS	NS											
Neeki (2017)	✓		NS													
Wall (2000)a			NS						NS	NS	✓	✓	NS	NS	NS	✓
Wall (2000)b	NS	NS	NS		✓	NS		✓	NS	NS	✓	NS	NS	NS	NS	✓
Wang (2004)		NS	NS	NS	NS											
Wong (2004)		NS	✓												NS	✓

	LRINEC Components					
	White Blood Cell Count	Hemoglobin	Sodium	Creatinine	C-Reactive Protein	Glucose
Borschitz (2015)	NS	NS	NS	✓	✓	NS
Chao (2012)	✓	✓	✓	✓	✓	✓
Cranendonck (2017)				NS	NS	
Kim (2013)	NS	NS	NS	✓	✓	NS
Liao (2012)	✓	NS	✓	✓	✓	✓
Martinez (2017)	✓	✓	✓	NS	NS	✓
McGillicuddy (2011)	✓	✓	NS	✓		✓
Neeki (2017)	✓	NS	✓	✓	✓	✓
Wall (2000)a	✓		✓	NS		NS

✓ indicates significant difference; HR, heart rate; LRINEC, laboratory risk indicator for necrotizing fasciitis; NS, no significant difference between NSTI and control patients; SBP, systolic blood pressure; T, temperature.

TABLE 3. Summary Estimates of the Performance of Physical Examination Features, Imaging, and LRINEC Score in Diagnosing Necrotizing Soft Tissue Infection

	No. of Cohorts (No. of Patients)	Sensitivity (%)	Specificity (%)	Diagnostic Odds Ratio	Positive Likelihood Ratio	Negative Likelihood Ratio
Physical examination						
Fever	4 (647)	46.0 (38.9 to 53.2)	77.0 (59.7 to 88.1)	2.81 (1.34 to 5.88)	1.98 (1.12 to 3.51)	0.70 (0.59 to 0.84)
Hemorrhagic bullae	5 (951)	25.2 (12.8 to 43.7)	95.8 (87.3 to 98.7)	7.64 (3.81 to 15.32)	5.97 (2.89 to 12.32)	0.78 (0.66 to 0.93)
Hypotension	6 (1014)	21.0 (9.4 to 40.4)	97.7 (91.4 to 99.4)	11.38 (5.00 to 25.90)	9.20 (3.87 to 21.86)	0.81 (0.68 to 0.96)
Imaging						
Plain Radiography	4 (478)	48.9 (24.9 to 73.4)	94.0 (63.8 to 99.3)	15.03 (3.69 to 61.22)	8.17 (1.61 to 41.47)	0.54 (0.36 to 0.82)
Computed tomography (fascial gas only)	7 (787)	88.5 (55.5 to 97.9)	93.3 (80.8 to 97.9)	107.64 (12.32 to 940.18)	13.27 (4.24 to 41.50)	0.12 (0.03 to 0.62)
Computed tomography (fascial edema OR fascial enhancement OR fascial gas)	6 (700)	94.3 (81.2 to 98.5)	76.6 (21.3 to 97.5)	54.29 (5.51 to 534.73)	4.04 (0.62 to 26.47)	0.07 (0.02 to 0.24)
LRINEC Score						
≥6	14 (4339)	68.2 (51.4 to 81.3)	84.8 (75.8 to 90.9)	11.95 (5.32 to 26.83)	4.49 (2.74 to 7.35)	0.38 (0.24 to 0.60)
≥8	9 (1905)	40.8 (28.6 to 54.2)	94.9 (89.4 to 97.6)	12.71 (4.71 to 34.28)	7.94 (3.44 to 18.32)	0.62 (0.50 to 0.78)

LRINEC indicates laboratory risk indicator for necrotizing fasciitis.

Quality Assessment

Quality assessments using QUADAS-2 criteria are summarized in eFigure 2; http://links.lww.com/SLA/B411. 11 articles (47.9%) had unclear risk-of-bias in the utilization of the Index Test (either physical examination, imaging or LRINEC score), as it was not explicitly stated whether the Index Tests were interpreted without knowledge of the results of the reference standard (operative or histopathological diagnosis of NSTI),[22,24,26–30,34,35,38,39] 10 (43.5%) studies were noted for potential high risk-of-bias, and were therefore excluded in a sensitivity analysis. Five of these studies utilized a case-control design.[10,21,22,33,38] Another 5 were considered high risk-of-bias for applicability in patient selection, as 1 study only included patients admitted to the ICU,[25] 1 only included patients with confirmed *Vibrio vulnificus* infection,[24] and 3 only included patients with cervical NSTI.[35,37,42]

Results of Synthesis

Summary estimates of all diagnostic accuracy measures from the HSROC model are tabulated in Table 3. All summary estimates described are pooled values. GRADE evidence profiles are included in the supplemental data (eTables 3–9; http://links.lww.com/SLA/B411).

Physical Examination

The forest plots describing the reported sensitivity and specificity for fever, hemorrhagic bullae, and hypotension from the included studies are presented in Figure 2. Presence of fever had a sensitivity of 46.0% (95% CI 38.9%–53.2%) and a specificity of 77.0% (95% CI 59.7%–88.1%) for diagnosis of NSTI. Presence of hemorrhagic bullae was associated with a sensitivity of 25.2% (95% CI 12.8%–43.7%) and specificity of 95.8% (95% CI 87.3%–98.7%) for diagnosis of NSTI. Finally, hypotension had a sensitivity of

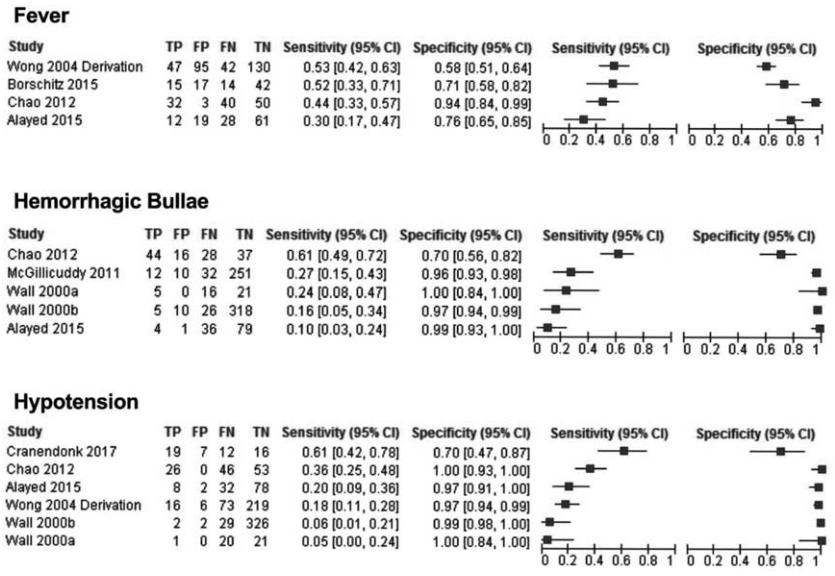

FIGURE 2. Forest plots of sensitivity and specificity for fever, hemorrhagic bullae, and hypotension for diagnosis of necrotizing soft tissue infection. CI indicates confidence interval; FN, false negative; FP, false positive; TN, true negative; TP, true positive.

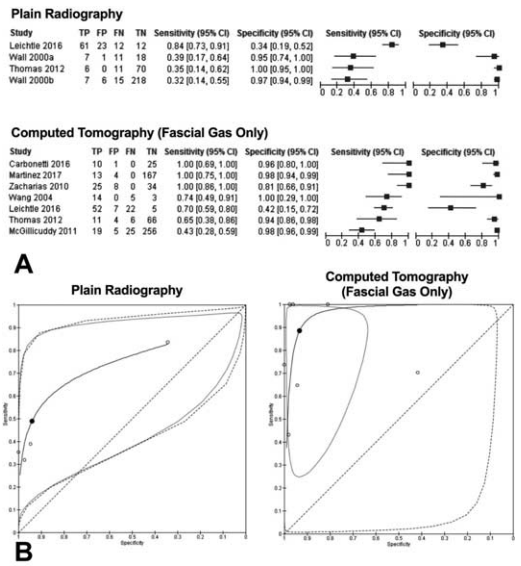

FIGURE 3. (A) Forest plots of sensitivity and specificity and (B) hierarchical summary receiver operating characteristic curves and bivariate summary points of (specificity, sensitivity), their 95% confidence regions (dotted lines), and 95% prediction regions (dashed lines) for diagnosis of necrotizing soft tissue infection. CI indicates confidence interval; FN, false negative; FP, false positive; TN, true negative; TP, true positive.

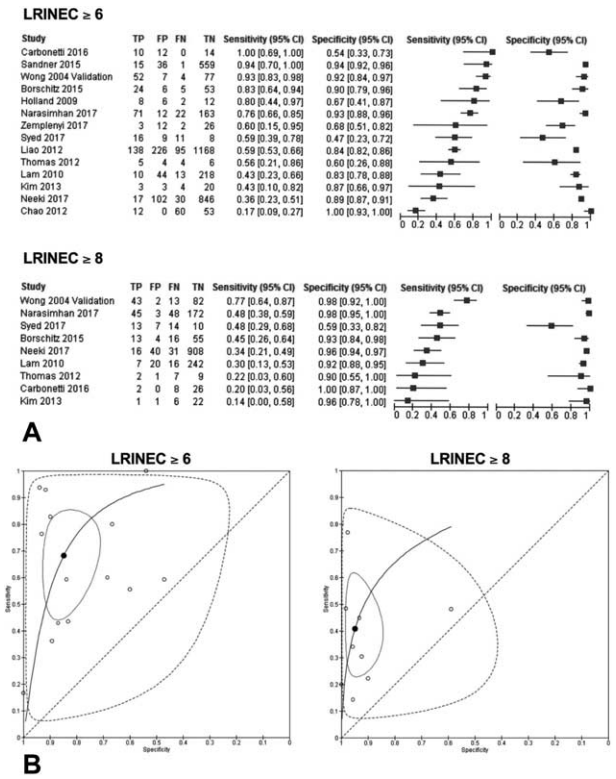

FIGURE 4. (A) Forest plots of sensitivity and specificity and (B) hierarchical summary receiver operating characteristic curves and bivariate summary points of (specificity, sensitivity), their 95% confidence regions (dotted lines), and 95% prediction regions (dashed lines) for LRINEC $\geq$ 6 and LRINEC $\geq$ 8 for diagnosis of necrotizing soft tissue infection. CI indicates confidence interval; FN, false negative; FP, false positive; LRINEC, laboratory risk indicator for necrotizing fasciitis; TN, true negative; TP, true positive.

21.0% (95% CI 9.4%–40.4%) and specificity of 97.7% (95% CI 91.4%–99.4%) for diagnosis of NSTI.

Imaging

The diagnostic accuracy of plain radiography and CT for diagnosis of NSTI were compared, and forest plots and HSROC curves describing reported sensitivity and specificity for each are presented in Figure 3. Visualization of soft tissue gas on plain radiography was associated with a sensitivity of 48.9% (95% CI 24.9%–73.4%) and specificity of 94.0% (95% CI 63.8%–99.3%) for diagnosis of NSTI. In comparison, visualization of fascial gas on CT was associated with a sensitivity of 88.5% (95% CI 55.5%–97.9%) and specificity of 93.3% (95% CI 80.8%–97.9%) for diagnosis of NSTI. Forest plot and HSROC curve for the composite findings of fascial enhancement, fascial edema, or fascial gas on CT are depicted in eFigure 4; http://links.lww.com/SLA/B411. The presence of any of these findings on CT was associated with a sensitivity of 94.3% (95% CI 81.2%–98.5%) and specificity of 76.6% (95% CI 21.3%–97.5%) for diagnosis of NSTI.

LRINEC

The LRINEC score was evaluated at 2 different thresholds, and forest plots and HSROC curves describing reported sensitivity and specificity for each are depicted in Figure 4. A LRINEC $\geq$6 was associated with a sensitivity of 68.2% (95% CI 51.4%–81.3%) and specificity of 84.8% (95% CI 75.8%–90.9%) for diagnosis of NSTI. In comparison, a LRINEC $\geq$8 had a sensitivity of 40.8% (95% CI 28.6%–54.2%) and specificity of 94.9% (95% CI 89.4%–97.6%).

Sensitivity Analyses: Excluding High Risk-of-Bias Studies

The results of the sensitivity analyses excluding high risk-of-bias studies (by QUADAS-2 criteria) for accuracy of physical examination, imaging and LRINEC for diagnosis of NSTI are depicted in eTable 10 and eFigures 5–7; http://links.lww.com/SLA/B411. Only CT with fascial gas, LRINEC $\geq$6, and LRINEC $\geq$8 had at least 3 included studies for meta-analysis. In the sensitivity analysis, CT had a pooled sensitivity and specificity of 93.3% (95% CI 48.7%–99.5%) and 93.1% (95% CI 80.9%–98.2%), respectively. LRINEC $\geq$6 had a sensitivity of 62.6% (95% CI 43.7%–78.3%) and specificity of 78.7% (95% CI 67.0%–87.1%), while LRINEC $\geq$8 had a sensitivity of 32.4% (95% CI 22.0%–45.1%) and specificity of 93.9% (95% CI 80.9%–98.2%).

DISCUSSION

We performed a systematic review and meta-analysis to evaluate the accuracy of physical examination findings, imaging, and LRINEC score in diagnosis of NSTI among adult patients with suspected NSTI. Given the clinical implications of delayed or missed NSTI diagnosis, and that waiting for imaging or laboratory results may delay time to definitive surgical management,[1,2] it is important to gain an understanding of the diagnostic accuracy of these tests in order to appropriately weigh the risks and benefits of using them. Taken together, this study comprehensively summarizes the available

literature and synthesizes the best available assessment of these various tools for diagnosis of NSTI.

NSTI is classically described as a clinical diagnosis based on patient risk factors and so-called pathognomonic physical examination features.[2] However, as this review illustrates, this dogma is not founded in high-quality evidence. We were unable to identify any historical risk factors suitable for meta-analysis, noting mixed findings for diabetes, immunocompromised status and intravenous drug use based on descriptive tests of association. Classic physical manifestations of NSTI may include a variety of findings,[1] however the available literature only allowed for meta-analysis of fever, hemorrhagic bullae, and hypotension. We found that all 3 physical examination findings had poor sensitivity for diagnosis of NSTI. In many cases, the development of physical signs such as hemorrhagic bullae or hypotension and shock appear to be evidence of more advanced disease.[43] Therefore, their absence should not be individually used to rule out the disease. This point is highlighted by examining the effect of these physical examination features in deriving post-test probability of NSTI from physician-determined pre-test probability (eTable 11; http://links.lww.com/SLA/B411). For example, a patient with a pre-test probability of NSTI of 50% but with absence of fever, hemorrhagic bullae, or hypotension, still retains a post-test probability of 41.3%, 43.9%, and 44.7%, respectively. Such patients should still undergo further testing or immediate surgical consultation.[44] It is important to note that in the clinical context, it is often not a single physical examination finding that is used to make the diagnosis, but rather a combination of findings. Unfortunately, such combinations were not evaluated in the available literature.

Imaging modalities are commonly used for diagnosis of NSTI. Plain radiography is readily available at most centers, and can often be obtained at the bedside. We found that plain radiography had poor sensitivity for diagnosis of NSTI, and therefore should not be used to rule out the diagnosis. In comparison, presence of fascial gas on contrast CT had far superior sensitivity and specificity than plain radiography. While the presence of fascial gas on CT was associated with a sensitivity of 88.5%, this finding had a specificity of 93.3% for NSTI. We performed a sensitivity analysis and pragmatically broadened acceptable CT criteria to include more subtle signs of NSTI (including fascial edema and enhancement), which increased the sensitivity to 94.3%, but decreased the specificity to 76.6%. The effects of these CT findings on physician derived pre-test probability of NSTI are depicted in eTable 10; http://links.lww.com/SLA/B411. In a patient with an equivocal pre-test probability of NSTI of 50%, the presence of fascial gas increases the probability to 93%, while the absence of fascial enhancement, edema or gas decreases the probability to less than 7%. The pragmatic approach of pooling studies with varying CT criteria for NSTI resulted in more heterogeneity of test specificity, as demonstrated by the wider confidence intervals for the pooled specificity estimate. This illustrates the importance of not simply accepting "positive" or "negative" CTs for NSTI at face value, but rather understanding that much like the physical examination, CT findings encompass a variety of specific components with a range of potential diagnostic utility. In fact, these findings may highlight the need for universal reporting checklists for CT requests querying the possibility of NSTI. Importantly, not all hospitals have access to CT imaging, and even if available, CT imaging may delay definitive surgical management. Therefore, despite the relatively strong accuracy of CT in diagnosis of NSTI, surgical consultation and intervention should never be delayed, particularly in cases of severe systemic illness. Diagnostic accuracy of Magnetic Resonance Imaging (MRI) was not evaluated in enough studies for meta-analysis, but existing work suggests that MRI can recognize subtle signs of NSTI, potentially allowing for earlier diagnosis.[27] However, the lack of availability of MRI in many centers may limit its practical utility, and since it may result in significant delay to surgical intervention, its use in the diagnosis of NSTI cannot be recommended at present. Point-of-care ultrasound, a newer bedside diagnostic modality, is available in many centers, and may play a role in NSTI diagnosis in the future; however, no data currently exists.[45] Future research investigating the accuracy of point-of-care ultrasound in diagnosis of NSTI is warranted, given the ubiquitous availability of this modality, and its ability to be used without significant delay to surgical consultation.

Finally, we evaluated the diagnostic accuracy of the LRINEC score,[10] which has become the most widely used clinical decision instrument for the diagnosis of NSTI.[1,2] We found that a LRINEC score $\geq$6 ("moderate" risk of NSTI) was poorly sensitive for diagnosis of NSTI, and only moderately specific. These performance characteristics are markedly worse than reported in the external validation population of the original study.[10] A LRINEC score $\geq$ 8 ("high" risk of NSTI) increased the specificity, but at the cost of substantially decreased sensitivity. Recognizing the limitations in sensitivity of the LRINEC score is extremely important, as computation of the score requires laboratory values, and therefore can delay definitive surgical management and result in worse outcomes.[5] A low LRINEC score (<6) does not significantly reduce post-test probability of NSTI in a patient with moderate risk of the diagnosis (eTable 13; http://links.lww.com/SLA/B411), as a patient with a pre-test probability of 50% but a LRINEC <6 still retains a 27.3% risk of NSTI. While the LRINEC score itself was associated with poor diagnostic accuracy, it is possible that individual elements of the score (such as WBC or sodium) may have better individual accuracy on their own. This represents an important avenue for future research.

This review was performed using a comprehensive search with clear inclusion and exclusion criteria, and it synthesizes the best available evaluation of the available assessments for diagnosis of NSTI. Limitations of this review relate primarily to the quality and heterogeneity of included studies. First, many included studies did not mention whether the diagnostic tests (namely the LRINEC score) were interpreted by individuals who were blinded to knowledge of the final diagnosis. Five of the included studies were case-control design, which represents a potential high risk-of-bias.

With regard to clinical heterogeneity, 3 articles specifically looked at cervical NSTI, which may represent a distinct subtype whose findings may not be extended to all types of NSTI. Unfortunately, none of the other studies evaluated diagnostic accuracy of affected body site, or total body surface area, which are known indicators of prognosis in patients with NSTI.[8] We did perform a sensitivity analysis removing high risk-of-bias studies, which did not substantially alter the conclusions. Finally, we sought to include studies that differentiated NSTI from control patients with clinical suspicion of NSTI, and as a result, there was variability in inclusion criteria between studies, with many of them including relatively "high-risk" patients (i.e. undergoing imaging or surgery to rule out NSTI), and only a minority including all consecutive patients presenting with a skin and soft tissue infection. Therefore, these studies may be biased towards prioritizing tests that are more strongly associated with severe or late disease, as opposed to tests that would be more useful for screening. Tests derived and evaluated in high-risk settings do not necessarily project their performance in low-risk settings.

This review demonstrates that the vast majority of diagnostic accuracy literature for NSTI is based on CT or LRINEC findings of high-risk populations. However, the determination of the patient's "pre-test probability" (in other words, identifying those that are "high risk") is based upon the clinician's assessment of history and physical examination findings, which are demonstrably scarce in evidence. Defining the appropriate population in which to apply the

LRINEC or CT is challenging, as reflected by the varying study definitions (and resultant variation in NSTI prevalence across studies) in this review.

CONCLUSIONS

Our systematic review found that individual physical examination signs (fever, hemorrhagic bullae, and hypotension) were poorly sensitive for diagnosis of NSTI. CT had superior sensitivity and specificity to plain radiography in diagnosing NSTI, but may not be readily available in all centers, and may not be suitable for unstable patients. Finally, the LRINEC score was poorly sensitive for diagnosis of NSTI, suggesting that a low score is not sufficient to rule out the diagnosis.

ACKNOWLEDGMENTS

The authors thank Dr. Stefan Leichtle (Virginia Commonwealth University) for providing data for our meta-analysis.

REFERENCES

1. Stevens DL, Bryant AE. Necrotizing soft-tissue infections. *N Engl J Med*. 2017;377:2253–2265.
2. Bonne SL, Kadri SS. Evaluation and management of necrotizing soft tissue infections. *Infect Dis Clin North Am*. 2017;31:497–511.
3. Miller AT, Saadai P, Greenstein A, et al. Postprocedural necrotizing fasciitis: a 10-year retrospective review. *Am Surg*. 2008;74:405–409.
4. Naseer U, Steinbakk M, Blystad H, et al. Epidemiology of invasive group A streptococcal infections in Norway 2010-2014: a retrospective cohort study. *Eur J Clin Microbiol Infect Dis*. 2016;35:1639–1648.
5. Elliott DC, Kufera JA, Myers RA. Necrotizing soft tissue infections. Risk factors for mortality and strategies for management. *Ann Surg*. 1996;224:672–683.
6. Bocking N, Matsumoto CL, Loewen K, et al. High Incidence of invasive group A streptococcal infections in remote indigenous communities in Northwestern Ontario, Canada. *Open Forum Infect Dis*. 2017;4:ofw243.
7. Khamnuan P, Chongruksut W, Jearwattanakanok K, et al. Necrotizing fasciitis: epidemiology and clinical predictors for amputation. *Int J Gen Med*. 2015;8:195–202.
8. McHenry CR, Piotrowski JJ, Petrinic D, et al. Determinants of mortality for necrotizing soft-tissue infections. *Ann Surg*. 1995;221:558–563. discussion 563–555.
9. Anaya DA, Dellinger EP. Necrotizing soft-tissue infection: diagnosis and management. *Clin Infect Dis*. 2007;44:705–710.
10. Wong CH, Khin LW, Heng KS, et al. The LRINEC (Laboratory Risk Indicator for Necrotizing Fasciitis) score: a tool for distinguishing necrotizing fasciitis from other soft tissue infections. *Crit Care Med*. 2004;32:1535–1541.
11. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ*. 2009;339:b2535.
12. McInnes MDF, Moher D, Thombs BD, et al. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: The PRISMA-DTA Statement. *JAMA*. 2018;319:388–396.
13. Deeks JJ, Bossuyt PM, Gatsonis C. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy, Version 1.0.0. 2013; http://srdta.cochrane.org/.
14. Leeflang MM, Deeks JJ, Gatsonis C, et al. Systematic reviews of diagnostic test accuracy. *Ann Intern Med*. 2008;149:889–897.
15. Sampson M, Shojania KG, McGowan J, et al. Surveillance search techniques identified the need to update systematic reviews. *J Clin Epidemiol*. 2008;61:755–762.
16. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155:529–536.
17. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med*. 2001;20:2865–2884.
18. Takwoingi Y. Deeks JJ: MetaDAS: a SAS Macro for Meta-analysis of Diagnostic Accuracy Studies. User Guide Version 1.3. In Released 30 July 2010.
19. Fernando SM, Tran A, Taljaard M, et al. Prognostic accuracy of the quick sequential organ failure assessment for mortality in patients with suspected infection: a systematic review and meta-analysis. *Ann Intern Med*. 2018;168:266–275.
20. Balshem H, Helfand M, Schunemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol*. 2011;64:401–406.
21. Alayed KA, Tan C, Daneman N. Red flags for necrotizing fasciitis: a case control study. *Int J Infect Dis*. 2015;36:15–20.
22. Borschitz T, Schlicht S, Siegel E, et al. Improvement of a clinical score for necrotizing fasciitis: 'pain out of proportion' and high CRP levels aid the diagnosis. *PLoS One*. 2015;10:e0132775.
23. Carbonetti F, Cremona A, Carusi V, et al. The role of contrast enhanced computed tomography in the diagnosis of necrotizing fasciitis and comparison with the laboratory risk indicator for necrotizing fasciitis (LRINEC). *Radiol Med*. 2016;121:106–121.
24. Chao WN, Tsai SJ, Tsai CF, et al. The laboratory risk indicator for necrotizing fasciitis score for discernment of necrotizing fasciitis originated from *Vibrio vulnificus* infections. *J Trauma Acute Care Surg*. 2012;73:1576–1582.
25. Cranendonk DR, van Vught LA, Wiewel MA, et al. Clinical characteristics and outcomes of patients with cellulitis requiring intensive care. *JAMA Dermatol*. 2017;153:578–582.
26. Holland MJ. Application of the Laboratory Risk Indicator in Necrotising Fasciitis (LRINEC) score to patients in a tropical tertiary referral centre. *Anaesth Intensive Care*. 2009;37:588–592.
27. Kim KT, Kim YJ, Won Lee J, et al. Can necrotizing infectious fasciitis be differentiated from nonnecrotizing infectious fasciitis with MR imaging? *Radiology*. 2011;259:816–824.
28. Lam SK, Wong HT. The LRINEC (Laboratory Risk Indicator for Necrotizing Fasciitis) Score: a review on the tool for distinguishing necrotizing fasciitis from other soft tissue Infections. *Hong Kong J Emerg Med*. 2010;17:419.
29. Leichtle SW, Tung L, Khan M, et al. The role of radiologic evaluation in necrotizing soft tissue infections. *J Trauma Acute Care Surg*. 2016;81:921–924.
30. Liao CI, Lee YK, Su YC, et al. Validation of the laboratory risk indicator for necrotizing fasciitis (LRINEC) score for early diagnosis of necrotizing fasciitis. *Tzu Chi Med J*. 2012;24:73–76.
31. Martinez M, Peponis T, Hage A, et al. The role of computed tomography in the diagnosis of necrotizing soft tissue Infections. *World J Surg*. 2018;42:82–87.
32. McGillicuddy EA, Lischuk AW, Schuster KM, et al. Development of a computed tomography-based scoring system for necrotizing soft-tissue infections. *J Trauma*. 2011;70:894–899.
33. Narasimhan V, Ooi G, Weidlich S, et al. Laboratory risk indicator for necrotizing fasciitis score for early diagnosis of necrotizing fasciitis in Darwin. *ANZ J Surg*. 2017.
34. Neeki MM, Dong F, Au C, et al. Evaluating the laboratory risk indicator to differentiate cellulitis from necrotizing fasciitis in the emergency department. *West J Emerg Med*. 2017;18:684–689.
35. Sandner A, Moritz S, Unverzagt S, et al. Cervical necrotizing fasciitis—the value of the laboratory risk indicator for necrotizing fasciitis score as an indicative parameter. *J Oral Maxillofac Surg*. 2015;73:2319–2333.
36. Syed A, Alvin T, Fazrina A, et al. Determining if positive predictive value using laboratory risk indicator for necrotising fasciitis is applicable in malaysian patients with necrotising fasciitis. *Malays Orthop J*. 2017;11:36–39.
37. Thomas AJ, Meyer TK. Retrospective evaluation of laboratory-based diagnostic tools for cervical necrotizing fasciitis. *Laryngoscope*. 2012;122:2683–2687.
38. Wall DB, de Virgilio C, Black S, et al. Objective criteria may assist in distinguishing necrotizing fasciitis from nonnecrotizing soft tissue infection. *Am J Surg*. 2000;179:17–21.
39. Wall DB, Klein SR, Black S, et al. A simple model to help distinguish necrotizing fasciitis from nonnecrotizing soft tissue infection. *J Am Coll Surg*. 2000;191:227–231.
40. Wang TL, Hung CR. Role of tissue oxygen saturation monitoring in diagnosing necrotizing fasciitis of the lower limbs. *Ann Emerg Med*. 2004;44:222–228.
41. Zacharias N, Velmahos GC, Salama A, et al. Diagnosis of necrotizing soft tissue infections by computed tomography. *Arch Surg*. 2010;145:452–455.
42. Zemplenyi K, Lopez B, Sardesai M, et al. Can progression of odontogenic infections to cervical necrotizing soft tissue infections be predicted? *Int J Oral Maxillofac Surg*. 2017;46:181–188.
43. Kiat HJ, En Natalie YH, Fatimah L. Necrotizing fasciitis: how reliable are the cutaneous signs? *J Emerg Trauma Shock*. 2017;10:205–210.
44. Burnham JP, Kirby JP, Kollef MH. Diagnosis and management of skin and soft tissue infections in the intensive care unit: a review. *Intensive Care Med*. 2016;42:1899–1911.
45. Kehrl T. Point-of-care ultrasound diagnosis of necrotizing fasciitis missed by computed tomography and magnetic resonance imaging. *J Emerg Med*. 2014;47:172–175.