TRANSFORM HEALTHCARE THROUGH A

Dedication to Academics

“

At Vituity, I get to build the future of medicine. As a Fellowship director, I help physicians become true experts and leaders in their field.

KIMBERLY SOKOL, MD, MS, MACM
Co-Director of Clinical Teaching and Simulation Fellowship, Vituity

Vituity's established residency programs and specialized fellowships prepare the next generation of clinicians to face the unpredictability of a normal day.

Learn More

vituity®

Received: 21 August 2022 | Revised: 27 December 2022 | Accepted: 3 January 2023

DOI: 10.1111/acem.14660

Academic Emergency Medicine

SYSTEMATIC REVIEWS

Test characteristics of ultrasound for the diagnosis of peritonsillar abscess: A systematic review and meta-analysis

Daniel J. Kim MD[1,2] | Justin E. Burton MD, MSc[1,3] | Abdullah Hammad MBBS[4] | Vikram Sabhaney MD[1,5] | Jason Freder MDCM, MBA[6,7] | Jeffrey N. Bone MSc[8] | Justin S. Ahn MD[1,9]

[1]Department of Emergency Medicine, University of British Columbia, Vancouver, British Columbia, Canada

[2]Department of Emergency Medicine, Vancouver General Hospital, Vancouver, British Columbia, Canada

[3]Department of Emergency Medicine, Langley Memorial Hospital, Langley, British Columbia, Canada

[4]Department of Pediatric Emergency Medicine, King Saud bin Abdulaziz University for Health Sciences, Jeddah, Kingdom of Saudi Arabia

[5]Department of Pediatrics, University of British Columbia, Vancouver, British Columbia, Canada

[6]Department of Emergency Medicine, St. Mary's Hospital, Montreal, Quebec, Canada

[7]Department of Emergency Medicine, McGill University, Montreal, Quebec, Canada

[8]Biostatistics, Research Informatics, BC Children's Hospital Research Institute, Vancouver, British Columbia, Canada

[9]Department of Emergency Medicine, Royal Columbian Hospital, New Westminster, British Columbia, Canada

Correspondence
Daniel J. Kim, Department of Emergency Medicine, University of British Columbia, Vancouver, BC V6T 2B5, Canada.
Email: dkim000@gmail.com

Abstract

Background: Distinguishing peritonsillar abscess (PTA) from peritonsillar cellulitis using clinical assessment is challenging as many features overlap for both conditions, and physical examination is only about 75% sensitive and 50% specific for diagnosing PTA. The primary objective of this systematic review was to determine the test characteristics of ultrasound for diagnosing PTA when compared to a reference standard of computed tomography or acquisition of pus via needle aspiration or incision and drainage.

Methods: This systematic review was performed in accordance with the Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy (PRISMA-DTA) guidelines. We searched seven databases from 1960 to November 2022. Two independent reviewers completed study selection, data extraction, and QUADAS-2 risk-of-bias assessment. We used a bivariate random-effects model to calculate pooled sensitivity, specificity, positive likelihood ratio (LR+), and negative likelihood ratio (LR–). We also conducted subgroup analyses on radiology ultrasound compared to point-of-care ultrasound (POCUS) and intraoral compared to transcervical scanning techniques.

Results: From 339 citations, we identified 18 studies for inclusion. Because one study only reported positive cases of PTA (thereby preventing the calculation of specificity), it was excluded from the analysis, so the analysis included a total of 17 studies with 812 patients, of whom 541 had PTA. Pooled bivariate sensitivity was 86% (95% confidence interval [CI] 78%–91%), specificity 76% (95% CI 67%–82%), LR+ 3.51 (95% CI 2.59–4.89), and LR– 0.19 (95% CI 0.12–0.30). On subgroup analysis, radiology-performed ultrasound had a sensitivity and specificity of 89% and 71%, compared to POCUS, which had a sensitivity and specificity of 74% and 79%. Comparing the two different techniques, intraoral had a sensitivity and specificity of 91% and 75% while transcervical had a sensitivity and specificity of 80% and 81%.

Supervising Editor: Dr. Christopher Carpenter.

Conclusions: Ultrasound demonstrates high sensitivity for ruling out PTA, but it only has moderate specificity for ruling in the diagnosis.

INTRODUCTION

Peritonsillar abscess (PTA) is one of the most common infections of the head and neck with an estimated incidence between 10 and 40 cases per 100,000 annually.[1] There are about 63,000 emergency department (ED) visits annually in the United States for PTA.[2] PTA is a purulent collection between the palatine tonsil and its capsule and it usually develops from an initial tonsillitis or pharyngitis. Complications from this infection can include airway obstruction, aspiration pneumonitis, mediastinal extension, and carotid sheath erosion.[3]

Distinguishing PTA from peritonsillar cellulitis (PTC) using clinical assessment is challenging as many features including fever, sore throat, dysphagia, trismus, and "hot potato" voice overlap for both conditions. Physical examination is only about 75% sensitive and 50% specific for the diagnosis of PTA.[4] The definitive diagnosis of PTA traditionally relies on needle aspiration or incision drainage of the purulent fluid. As this is invasive, there is increasing interest in the use of imaging modalities like computed tomography (CT) and ultrasound to establish the diagnosis. National databases report the use of CT in 5%–20% of ED patients presenting with PTA.[2]

While cross-sectional imaging with CT is highly accurate, it is associated with increased cost, resource utilization, and radiation exposure.[5,6] Additionally, the use of CT imaging carries the risks of overtesting, overdiagnosis, and overtreatment,[7] especially when it is unclear if CT imaging improves outcomes in patients with PTA.[8] Due to these considerations, ultrasound has been suggested by both emergency medicine (EM) and ear, nose, and throat (ENT) specialists as an alternative imaging modality to avoid radiation exposure and to reduce costs.[9,10] There is a paucity of clinical practice guidelines on the diagnostic imaging approach to suspected PTA. The American College of Emergency Physicians (ACEP),[11] Society for Academic Emergency Medicine (SAEM),[12] and American Academy of Otolaryngology–Head and Neck Surgery (AAO-HNS)[13] do not have policy statements or guidelines for approaching patients presenting with suspected PTA. Furthermore, the American College of Radiology (ACR)[14] does not have any appropriateness criteria for imaging in patients presenting with sore throat.

The primary objective of this systematic review was to determine the accuracy of ultrasound for diagnosing PTA when compared to a reference standard of CT or acquisition of pus via needle aspiration or incision drainage. We also planned a priori subgroup analyses of the primary objective based on radiology-performed ultrasound compared to point-of-care ultrasound (POCUS), intraoral compared to transcervical scanning technique, and adult compared to pediatric patient populations.

METHODS

This systematic review was performed in accordance with the Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy (PRISMA-DTA) guidelines.[15] We registered this review with PROSPERO (CRD42021224241).

We developed our search strategy (Appendix S1) with an experienced medical librarian and searched the following databases: Ovid MEDLINE, Embase, Cumulative Index to Nursing and Allied Health Literature (CINAHL), the Cochrane Central Register of Controlled Trials, Latin American and Caribbean Health Sciences Literature database (LILACS), and Web of Science from 1960 to December 2020. A repeat search was completed in May 2021 and November 2022 to ensure no new data had been published.

Study selection and definitions

Studies were eligible if they evaluated the test characteristics of ultrasound in diagnosing PTA compared to a reference standard of CT or the aspiration or drainage of pus. We included retrospective, prospective, and randomized control trials. There were no date or age restrictions. Each study required a 2×2 table of true-positive, false-negative, true-negative, and false-positive counts, either extracted from the original paper or calculated from the reported results. We excluded case reports, case series with 10 or fewer patients, protocol papers, and narrative reviews. Studies were also excluded if they were not in English.

Two investigators (A.H., J.F.) independently performed the search and independently assessed the studies for eligibility based on the above criteria. They identified all potentially relevant studies by screening titles and abstracts. They then independently reviewed the full text of the selected articles. Disagreement was resolved through discussion and, when necessary, a third party (D.K.).

Data extraction

One investigator (A.H.) collected the following variables from the included studies: author information; country of publication; study design; number of patients included; patient age characteristics; pediatric specific data (if possible); clinical follow-up; clinical setting; ultrasound machine; ultrasound transducer; scanning technique; background of sonographers; training or qualification of sonographers; blinding; reference standard; and true-positive, false-negative, true-negative, and false-positive counts. Study investigators were contacted when data were missing or unreported, and

variables were left as missing if the corresponding author did not respond after two attempts. A second investigator (J.A.) reviewed the extracted data to ensure its accuracy. Disagreements were resolved through discussion, with conflicts adjudicated by a third party (D.K.). The third reviewer (D.K.) independently verified the extracted 2×2 counts including performing back-calculations when necessary to ensure accuracy before statistical analysis.

Quality appraisal

Two investigators (J.E.B., J.F.) independently assessed study quality using the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool.[16] Discrepancies were resolved by consensus when needed. These authors used several a priori conditions to evaluate each individual study's risk of bias and degree of applicability.

- For patient selection, a study was deemed to have low concerns for risk of bias if: (1) a consecutive or random sample of patients was enrolled, (2) a case–control design was avoided, and (3) the study avoided inappropriate exclusions.
- An inappropriate exclusion was defined as any participant who was included in the initial sample who was then taken out of the final analysis or where the enrollment of patients was unclear. Additionally, if only retrospective data were used or if patient selection was solely based on the availability of sonographers in the clinical setting, this was also defined as an inappropriate exclusion.
- The index test risk of bias was assessed by whether or not the index test results were interpreted without knowledge of the results of the reference standard and whether or not a threshold was prespecified. If there was no prespecified description of the ultrasound technique and ultrasound definition of PTA, the study was deemed to not have a prespecified threshold, and conduct of the index test was deemed to be at high risk of bias.
- Concerns regarding applicability of the index test were based on a typical single encounter in the clinical setting. For example, if the index test relied on the patient needing to be reassessed and rescanned, applicability was questioned.
- The ideal index test was ultrasound performed by an emergency physician, ENT specialist, or radiologist (intraoral or transcervical) with comparison to a reference standard (CT or acquisition of pus via needle aspiration or incision and drainage) with no prior knowledge of one from the other.
- A positive index test was one in which a distinct anechoic area was detected on ultrasound. A positive reference standard was CT or acquisition of pus via needle aspiration or incision and drainage.
- The reference standard introduced bias if it was interpreted with prior knowledge of the index test.
- The time between index test and reference standard introduced bias if it was beyond acceptable current clinical care (>24 h) or if a confounding variable such as antibiotics was administered in the time frame between index test and reference standard.

Data analysis

We used a bivariate random-effects models to calculate pooled sensitivity, specificity, positive likelihood ratio (LR+), and negative likelihood ratio (LR–) with associated 95% confidence intervals (CIs). We additionally used a hierarchical summary receiver operating characteristic (SROC) curve with a 95% CI region to further assess the diagnostic properties of the two tests. Heterogeneity was summarized with I^2 statistics for the univariate analyses of sensitivity and specificity was estimated. Publication bias for these measures was assessed graphically with funnel plots.

We tested for moderation in bivariate models and conducted possible subgroup analyses accounting for the setting of the ultrasound study (radiology ultrasound vs. POCUS) and scanning technique (intraoral vs. transcervical). In these cases, we provide sensitivity, specificity, LR+, and LR–. Statistical analyses were conducted using the Meta-Analysis of Diagnostic Accuracy (mada) package (version 0.5.10) in R statistical software version 4.0.3.

RESULTS

Of the 339 studies identified through our search, 192 abstracts were screened after removing duplicates, and 33 studies were selected for full-text review. A total of 18 studies met criteria for inclusion in the meta-analysis (Figure 1).

Table 1 describes the characteristics of the included studies. The studies were published from 1992 to 2021, all were published in English, 17 studies were published as full-text papers, and one study[23] was published only as an abstract. Of the 18 included studies, one was an RCT, 14 were prospective cohort studies, and three were retrospective cohort studies. All included studies utilized drainage or surgical intervention as a possible reference standard, and five studies additionally used CT as an option for reference standard. Pediatric patients were enrolled in 10 studies, but the reported data were not granular enough to calculate test characteristics of ultrasound in this population. Because one study[22] only reported positive cases of PTA (thereby preventing the calculation of specificity), it was excluded from the analysis, so the analysis included a total of 812 patients of whom 541 had PTA from 17 studies. Of note, the two largest studies[25,30] included 430 patients (53% of the total) of whom 305 had PTA.

Transcervical scanning was performed in eight studies and intraoral scanning in 13 studies; three studies assessed both scanning techniques. Radiology ultrasound was used in 12 studies and POCUS was utilized in six studies. Of these POCUS studies, four studies assessed emergency physicians, and two studies assessed ENT operators. The training or qualifications of the sonographers were not specified in 13 studies. The remaining five studies all described differing levels of training or qualification: (a) ultrasound credentialed

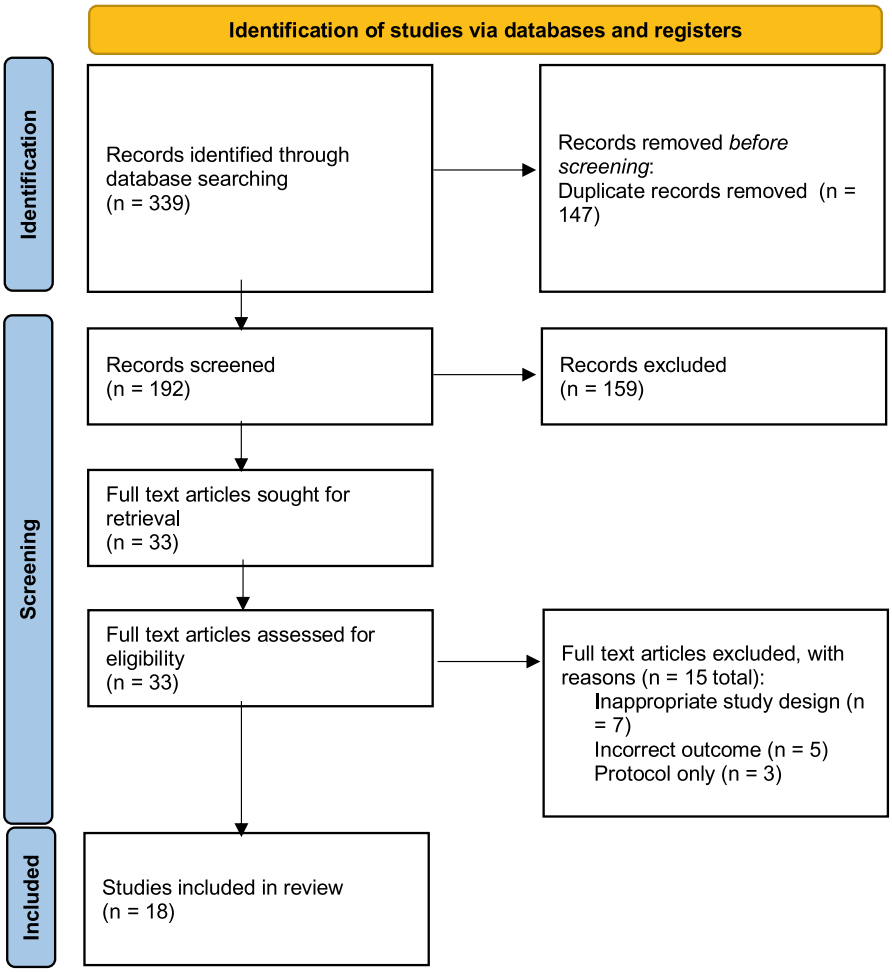


FIGURE 1 PRISMA flow chart illustrating evidence search and study selection. PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

at their hospital,[26] (b) met the ACEP emergency ultrasound training guidelines,[20] (c) a 30-min didactic and hands-on session,[21] (d) certification by the German Society of Ultrasound in Medicine,[30] and (e) formal training with >300 prior head and neck ultrasound scans.[32]

Figure 2 summarizes the quality assessment using the QUADAS-2 tool. One study[23] was only published as a conference abstract and not as a full-text manuscript, so it could not be assessed using QUADAS-2. All included studies were found to be of high or unclear risk of bias. No studies were found to be of low risk of bias. In general, the studies were at low risk with regard to applicability concerns. Funnel plots for the primary outcome measure were symmetrical for the pooled estimate of specificity. However, there was some asymmetry around the pooled estimate of sensitivity, graphically suggestive of possible publication bias for this outcome measure.

Pooled bivariate sensitivity and specificity were 86% (95% CI 78%–91%) and 76% (95% CI 67%–82%), respectively. Diagnostic odds ratio was 19.5 (95% CI 9.63–35.4). Pooled bivariate LR+ and LR− were 3.51 (95% CI 2.59–4.89) and 0.19 (95% CI 0.12–0.30), respectively. I^2 was not calculated for our bivariate model, but I^2 from our univariate model for sensitivity was 35% and for specificity was 28%. Figure 3 demonstrates descriptive forest plots of sensitivity and specificity of ultrasound for the diagnosis of PTA. Figure 4 illustrates the area under the receiver operator curve (AUROC) as 0.87. When the two largest contributing studies[25,30] were excluded from the analysis, the pooled diagnostic test characteristics were similar: sensitivity 88% (95% CI 79%–93%), specificity 73% (95% CI 65%–80%), and AUROC 0.83.

Table 2 details the subgroup analyses comparing radiology ultrasound to POCUS as well as intraoral compared to transcervical scanning techniques. The diagnostic test characteristics for both subgroup analyses were comparable.

DISCUSSION

To our knowledge, this is the first systematic review and meta-analysis to characterize the diagnostic test characteristics of ultrasound for PTA. PTA is a common head and neck infection encountered in the ED,[2] and accurate diagnosis is important to institute appropriate treatment as inaccurate diagnosis of PTA exposes patients to unnecessary surgical intervention without clinical benefit.[20] Inadequately treated PTA can lead to significant morbidity either through extension of infection or through airway obstruction.[3] Although CT is accurate, it exposes patients to substantial ionizing radiation and intravenous (IV) contrast.[5,6] Based on this systematic review, ultrasound has reasonably high sensitivity for ruling out PTA, but it only demonstrates moderate specificity for establishing the diagnosis.

TABLE 1 Summary of included studies

Author and year	Country	Study design	Patients	Setting	Scanning technique	Operator	Ultrasound	Reference standard
Ahmed 1994[10]	UK	Prospective	27	Radiology	Transcervical	Radiology	Aloka Echo Camera SSD-650, curvilinear 7.5 MHz	Needle aspiration
Araujo Fiho 2006[17]	Brazil	Prospective	39 (35 scanned intraoral, 39 scanned transcervical)	Radiology	Intraoral and transcervical	Radiology	General Electric 500, endocavitary 6.5 MHz, linear 7.5 MHz	Needle aspiration
Boesen 1992[18]	Denmark	Prospective	27	Radiology	Transcervical	Radiology	Machine not specified, linear 5–7 MHz	Needle aspiration, I&D, or tonsillectomy
Buckley 1994[19]	Canada	Prospective	17	Radiology	Intraoral and transcervical	Radiology	Siemens ACUSON 128, endocavitary 5 MHz, linear 5 MHz	Needle aspiration or I&D
Costantino 2012[20]	USA	Prospective RCT	14	ED	Intraoral	Emergency medicine	SonoSite Micromaxx, endocavitary 5–8 MHz	Needle aspiration
Fordham 2015[21]	USA	Prospective	43	Radiology	Transcervical	Radiology	Machine not specified, linear	Needle aspiration or I&D
Haeggström 1993[22]	Sweden	Prospective	12	Radiology	Intraoral	Radiology	Bruel and Kjaer 1846, endocavitary 7 MHz	Needle aspiration or tonsillectomy
Kelley 2017[23]	USA	Prospective	46	ED	Intraoral	Emergency medicine	Machine not specified, endocavitary 5–9 MHz	I&D or CT
Kew 1998[24]	Hong Kong	Prospective	14	Radiology	Intraoral	Radiology	Philips ATL, endocavitary 7–9 MHz	CT or surgical intervention
Klimek 1993[25]	Germany	Prospective	118	Radiology	Transcervical	Radiology	Hitachi PICKER 7000, curvilinear 5 MHz	Tonsillectomy
Lyon 2005[26]	USA	Retrospective	43	ED	Intraoral	Emergency medicine	Philips ATL HDI 4000, endocavitary 4–8 MHz	Needle aspiration
Miziara 2001[27]	Brazil	Prospective	21	Radiology	Intraoral and transcervical	Radiology	Machine not specified, endocavitary 7 MHz, linear 5 MHz	Needle aspiration or I&D
Nogan 2015[28]	USA	Prospective	21	ED	Intraoral	Emergency medicine	SonoSite (model not specified), endocavitary 5–9 MHz	Needle aspiration or I&D
Salihoglu 2013[29]	Turkey	Retrospective	26	Radiology	Intraoral	Radiology	6000 Povervision Toshiba SSA- 370A, endocavitary 6 MHz	CT, needle aspiration, or I&D
Scott 1999[4]	Hong Kong	Prospective	13	Radiology	Intraoral	Radiology	Philips ATL HDI, endocavitary 10 MHz	CT, I&D, or tonsillectomy
Sievert 2021[30]	Germany	Retrospective	312	ENT	Transcervical	ENT	Siemens ACUSON S2000 or S3000, linear 4–10 MHz	Needle aspiration, I&D, or tonsillectomy
Strong 1995[31]	USA	Prospective	15	Radiology	Intraoral	Radiology	Machine not specified, endocavitary 5 MHz	CT, MRI, needle aspiration, I&D, or tonsillectomy
Todsen 2021[32]	Denmark	Prospective	16	ENT	Intraoral	ENT	BK Medical Flex Focus 800, burr-hole 3.8–10 MHz, hockey stick 6–15 MHz	I&D

Abbreviations: ENT, ear, nose, and throat; I&D, incision and drainage; RCT, randomized controlled trial.

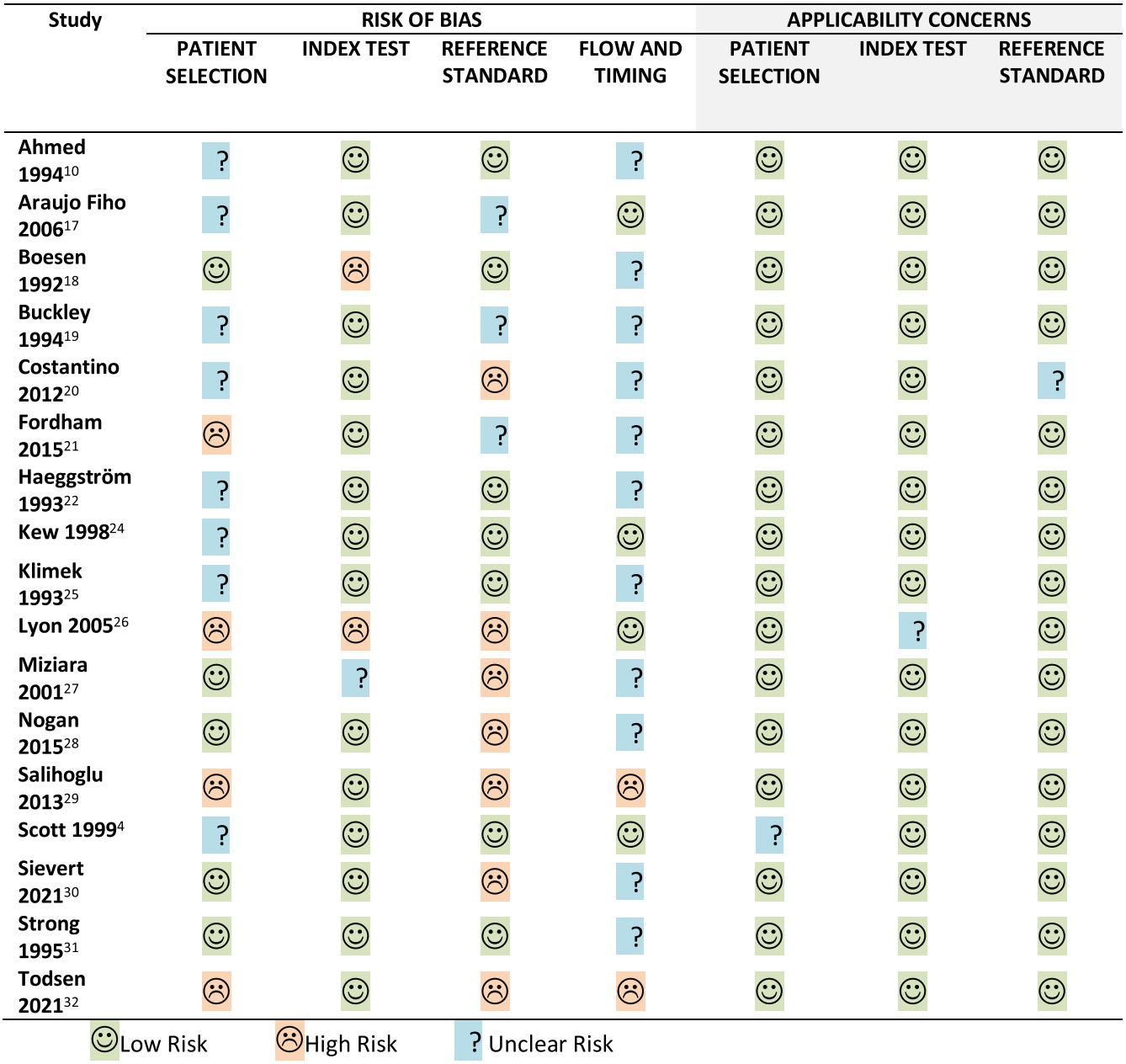


FIGURE 2 QUADAS-2 risk of bias assessment of the included studies. Although a total of 18 studies are included in this systematic review, the study by Kelley[23] was only published as a conference abstract and never as a full-text manuscript. As such, the authors were unable to assess its quality using QUADAS-2, so this figure only includes 17 studies. QUADAS, Quality Assessment of Diagnostic Accuracy Studies.

Subgroup analysis did not reveal a statistically significant difference between intraoral and transcervical scanning techniques. However, the intraoral technique trended toward a higher sensitivity (91% vs. 80%) and lower specificity (75% vs. 81%) compared to the transcervical technique. These diagnostic test characteristics are similar to those of the published literature. A narrative review on ultrasound for PTA found the intraoral technique to have a sensitivity in the range of 89%–95% and a specificity in the range of 79%–100%. This same review described a sensitivity in the range of 83%–91% and specificity in the range of 80%–93% for the transcervical technique.[33] Our systematic review provides a more robust estimate of the diagnostic test characteristics of the two different techniques compared to this previously published narrative review.[33] The intraoral technique involves inserting the endocavitary probe with a protective cover into the patient's open mouth directly over the tonsils (Figure 5A–C). This approach is more intuitive but faces limitations such as access to an endocavitary probe, appropriate cleaning and disinfection of the probe, issues with patient tolerance in cases where significant trismus and pain are present, and issues with patient cooperation in pediatric populations. As an alternative, the transcervical technique involves placing either a linear or microconvex probe over the patient's submandibular space just inferior and medial to the

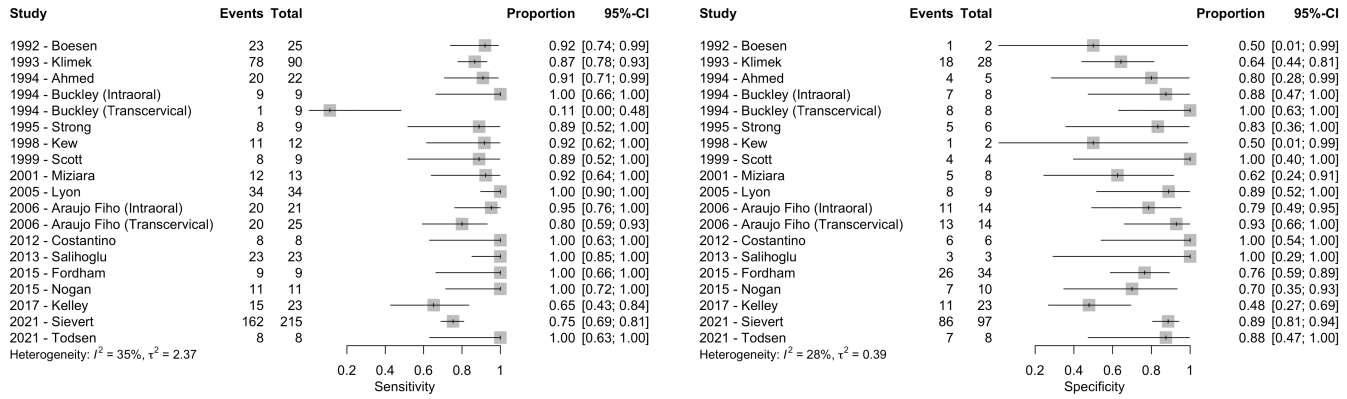
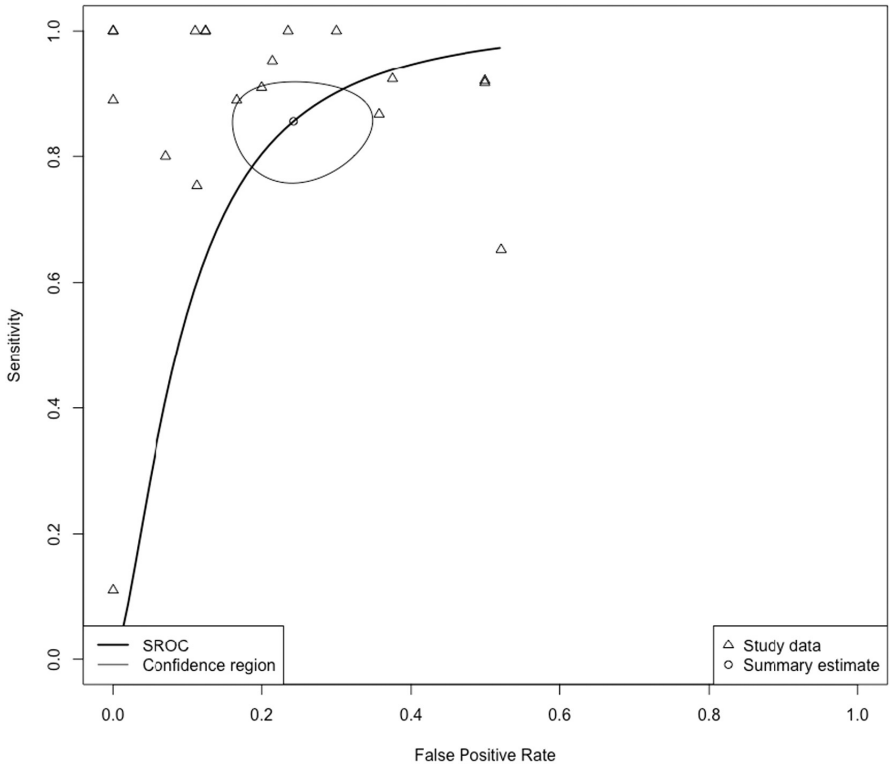


FIGURE 3 Descriptive forest plots of the sensitivity (on the left) and specificity (on the right) of ultrasound for the diagnosis of peritonsillar abscess.

FIGURE 4 Plot of the summary receiver operating curve for the ultrasound diagnosis of peritonsillar abscess. The summary estimate is presented as a circle, each individual study is presented as a triangle, and the 95% confidence region is presented as light gray line. SROC, summary receiver operating characteristic.



angle of the mandible to identify the submandibular gland and tonsillar tissue (Figure 5D–F). A more detailed description of the technique can be found at https://www.youtube.com/watch?v=JkIYOhKCwel.

A further subgroup analysis shows that radiology performed scans have greater sensitivity (89%) than POCUS (74%). Only six studies assessed the use of POCUS, and the majority of these were published more recently since 2012. The studies assessing POCUS had a heterogenous group of operators, including both EM and ENT residents and specialists, and the amount and level of ultrasound training was highly variable.

It was not possible to perform a subgroup analysis of the test characteristics of ultrasound for PTA in pediatric populations as the included studies did not present individual level data for children except in one study. We did not specifically assess the effect of ultrasound on the patient's clinical management or clinical course. However, some studies have demonstrated relatively high success with the conservative management of PTA using IV antibiotics and corticosteroids, comparable to surgical intervention.[34,35] Conservative management seems to be more successful in smaller PTAs.[35] Given this, the diagnostic test characteristics of ultrasound for PTA may be sufficiently acceptable for diagnosing clinically important PTA requiring either aspiration or drainage. Within the current limitations of evidence, the benefits of ultrasound, including lack of ionizing radiation and ability to perform at the bedside, must be weighed against the limitations, including inaccurate diagnosis and patient discomfort. Ultimately, these options should be provided to the patient and the most appropriate investigation can be selected through shared decision making.[36]

There are many areas to explore with future research, and the lack of clinical guidelines and policy statements on the use of ultrasound in

TABLE 2 Subgroup analyses of diagnostic test characteristics for ultrasound to diagnose peritonsillar abscess based on operator type (radiology vs. POCUS) and scanning technique (intraoral vs. transcervical)

	Sensitivity (95% CI)	Specificity (95% CI)	DOR (95% CI)	LR+ (95% CI)	LR− (95% CI)
Radiology vs. POCUS ($p = 0.60$)					
Radiology ultrasound (reference standard: drainage or CT)	89% (84%–92%)	71% (63%–79%)	19.5 (11.0–34.5)	3.10 (2.35–4.20)	0.16 (0.11–0.23)
POCUS (reference standard: drainage or CT)	74% (67%–80%)	79% (61%–90%)	10.9 (3.7–32.5)	3.54 (1.80–7.83)	0.34 (0.23–0.50)
Intraoral vs. transcervical ($p = 0.42$)					
Intraoral (reference standard: drainage or CT)	91% (82%–95%)	75% (63%–84%)	29.4 (9.5–90.9)	3.62 (2.32–5.92)	0.12 (0.06–0.26)
Transcervical (reference standard: drainage or CT)	80% (67%–89%)	81% (66%–91%)	18.4 (10.9–29.4)	4.29 (2.53–7.77)	0.25 (0.16–0.37)

Abbreviations: DOR, diagnostic odds ratio; LR, likelihood ratio; POCUS, point-of-care ultrasound.

the workup of suspected PTA from major EM, ENT, and radiology societies[11–14] may reflect a lack of high-quality evidence. When it comes to POCUS, there is a need for larger studies with standardized training strategies to determine the diagnostic test characteristics of POCUS for PTA when performed by a skilled EM operator. Future studies will need to determine the learning curve of this skill and to assess the minimum amount of training necessary for a nonexpert point-of-care operator to perform the scan with sufficient diagnostic accuracy. The cost-effectiveness of a POCUS-based approach should be explored, especially compared to an approach depending more heavily on CT use. Pediatric populations have been inadequately studied, so future studies should attempt to elucidate the diagnostic test characteristics of ultrasound for PTA in this patient population. Finally, more study is needed to determine the effect of the use of ultrasound on the patient's clinical course and outcome. An ideal study would randomize a large number of consecutive patients with suspected PTA to a POCUS group versus a non-POCUS group and would follow them out to 30 days to determine treatment success, recurrence, complications, and return visits. Once data from high-quality studies are available, it could provide the impetus for the derivation and validation of a clinical decision rule to assist clinicians with selecting the most appropriate diagnostic imaging test for investigating PTA as described by the 2015 *Academic Emergency Medicine* Consensus Conference.[37]

LIMITATIONS

The majority of included studies had very small sample sizes, and a significant number of patients in the analysis were from two larger studies. It was not possible to determine whether 17 of the 18 included studies were adequately powered as sample size calculations were not described in the articles. The one study that provided a sample size estimate achieved its enrollment target.[20] Additionally, of the eight studies published since 2003, only one study[21] had >75% compliance with the Standards for Reporting of Diagnostic Accuracy (STARD) reporting standards.[38] This raises concerns about the transparency and reproducibility of these studies.[39] Overall, the included studies were found to be of high or unclear risk of bias using the QUADAS-2 tool. The majority of studies enrolled a convenience sample of patients, which introduces the possibility of selection bias. Studies were conducted in a variety of settings, and some included a higher proportion of participants with confirmed PTA than would be expected (four studies reported PTA rates >80%[10,18,22,29]). High rates of PTA suggest the population in these studies may have been sicker, introducing the risk of spectrum bias and overestimating the test characteristics of ultrasound for the diagnosis of PTA.[40]

Using CT and needle aspiration as reference standards has some limitations and introduces imperfect criterion standard bias.[40] While CT has near perfect sensitivity, it has a reported specificity of 75%.[4] It is possible that some negative ultrasounds were misclassified by CT as PTA, which would result in a falsely lowered sensitivity for ultrasound. Conversely, it has been suggested that needle aspiration is not perfectly sensitive, as puncture may be performed in the incorrect location or depth resulting in a negative needle aspiration.[22] In these cases, a positive ultrasound may have correctly diagnosed PTA resulting in a falsely lowered specificity for ultrasound.

There were also issues with some of the studies related to a lack of consistent comparison to a reference standard of CT or drainage of pus, as in some studies, a patient with a negative ultrasound was clinically presumed to not have PTA despite not undergoing CT or attempted drainage. Though this may have introduced the possibility of incorporation bias if the final clinical diagnosis integrated ultrasound results, it reduced the risk of creating partial verification bias by including those who did not undergo CT or drainage. Differential verification bias may also have influenced reported test characteristics for some studies, as those with a false-negative ultrasound may have had resolution of their PTA with medical treatment only. While this would falsely raise the sensitivity of ultrasound, this concern

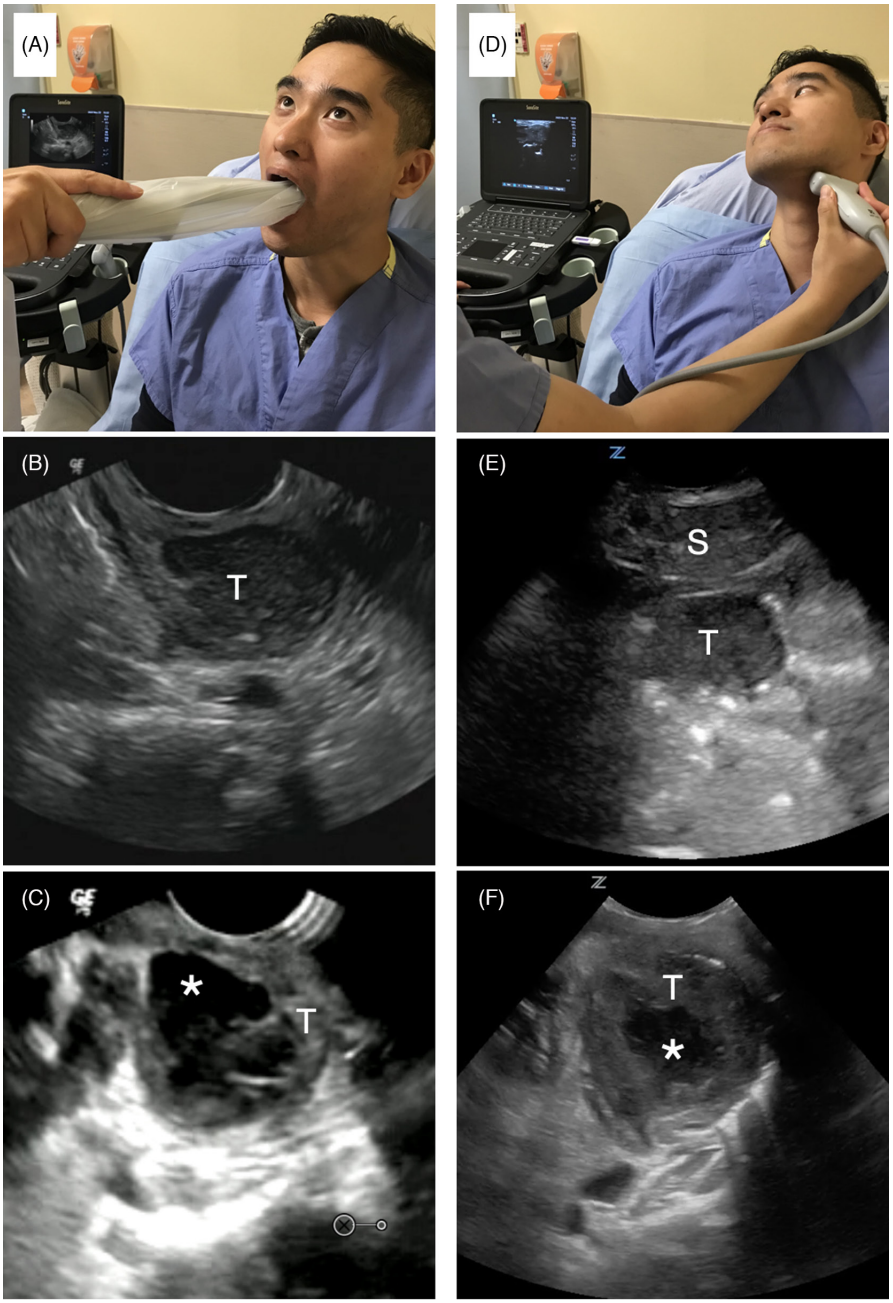


FIGURE 5 . (A) Intraoral approach where the endocavitary probe is placed intraorally with a protective cover over the patient's tonsils. (B) Intraoral view demonstrating normal tonsil (T) with homogenous echotexture. (C) Intraoral view demonstrating abnormal tonsil (T) with loculated anechoic fluid collections (*) indicating peritonsillar abscess. (D) Transcervical approach where the linear or microconvex probe is placed in the submandibular space inferior to the angle of the mandible. (E) Transcervical view demonstrating normal tonsil (T) with homogenous echotexture deep to the submandibular gland (S). (F) Transcervical view demonstrating abnormal tonsil (T) with anechoic fluid collection (*) indicating peritonsillar abscess. Credit for (F) goes to Mark Tessaro, MD, Division of Pediatric Emergency Medicine, Hospital for Sick Children, Toronto, Ontario, Canada.

may be of little clinical importance in these cases as the false-negative ultrasound result would not meaningfully impact patient outcomes.[40] However, it is uncertain how common this scenario was as the majority of studies did not report on the patient's clinical course. It is impossible to know how the accuracy of the ultrasound scan ultimately affected the patient's clinical outcome. The majority of included studies did not describe the level of training or qualifications of the operators performing the ultrasound scans.

CONCLUSIONS

Ultrasound is an acceptable alternative to computed tomography and interventional aspiration or drainage for establishing the diagnosis of peritonsillar abscess. It has high sensitivity for ruling out peritonsillar abscess, but it demonstrates only moderate specificity for ruling in the diagnosis. Future studies are needed to determine the amount of training required to learn the skill, the diagnostic test characteristics of point-of-care ultrasound, and the impact of incorporating the results of an ultrasound scan on the patient's clinical course.

AUTHOR CONTRIBUTIONS

Daniel J. Kim, Vikram Sabhaney, and Justin S. Ahn conceived and designed the study with input during the process by Abdullah Hammad and Jeffrey N. Bone. Abdullah Hammad and Jason Freder independently performed the search and assessed the studies for eligibility. Abdullah Hammad extracted the data from the included studies and

Justin S. Ahn reviewed the extracted data to ensure its accuracy. Justin E. Burton and Jason Freder independently assessed study quality using the QUADAS-2 tool. Jeffrey N. Bone analyzed the data. Justin E. Burton and Daniel J. Kim drafted the manuscript. All authors contributed substantially to its revision. Daniel J. Kim takes responsibility for the paper as a whole.

ACKNOWLEDGMENTS

The authors thank Colleen Pawliuk, MLIS, BC Children's Hospital Research Institute, Vancouver, British Columbia, Canada, for providing assistance in the development of our literature search strategy. We also thank Mark Tessaro, MD, Division of Pediatric Emergency Medicine, Hospital for Sick Children, Toronto, Ontario, Canada, for providing an image and video of the transcervical ultrasound approach for PTA.

CONFLICT OF INTEREST

Daniel J. Kim provides consultant services to Fujifilm Sonosite. The rest of the authorship team declare no potential conflict of interest.

ORCID

Daniel J. Kim https://orcid.org/0000-0003-0470-8279

REFERENCES

1. Klug TE. Incidence and microbiology of peritonsillar abscess: the influence of season, age, and gender. *Eur J Clin Microbiol Infect Dis*. 2014;33(7):1163-1167.
2. Johnson RF. Emergency department visits, hospitalizations, and readmissions of patients with a peritonsillar abscess. *Laryngoscope*. 2017;127(Suppl 5):S1-S9.
3. Powell J, Wilson JA. An evidence-based review of peritonsillar abscess. *Clin Otolaryngol*. 2012;37(2):136-145.
4. Scott PM, Loftus WK, Kew J, Ahuja A, Yue V, van Hasselt CA. Diagnosis of peritonsillar infections: a prospective study of ultrasound, computerized tomography and clinical diagnosis. *J Laryngol Otol*. 1999;113(3):229-232.
5. Patel KS, Ahmad S, O'Leary G, Michel M. The role of computed tomography in the management of peritonsillar abscess. *Otolaryngol Head Neck Surg*. 1992;107(6 Pt 1):727-732.
6. Brenner DJ, Hall EJ. Computed tomography—an increasing source of radiation exposure. *N Engl J Med*. 2007;357(22):2277-2284.
7. Carpenter CR, Raja AS, Brown MD. Overtesting and the downstream consequences of overtreatment: implications of "preventing overdiagnosis" for emergency medicine. *Acad Emerg Med*. 2015;22(12):1484-1492.
8. Grant MC, Guarisco JL. Association between computed tomographic scan and timing and treatment of peritonsillar abscess in children. *JAMA Otolaryngol Head Neck Surg*. 2016;142(11):1051-1055.
9. Secko M, Sivitz A. Think ultrasound first for peritonsillar swelling. *Am J Emerg Med*. 2015;33(4):569-572.
10. Ahmed K, Jones AS, Shah K, Smethurst A. The role of ultrasound in the management of peritonsillar abscess. *J Laryngol Otol*. 1994;108(7):610-612.
11. Clinical policies. American College of Emergency Physicians. 2023. Accessed November 17, 2022. https://www.acep.org/patient-care/clinical-policies/
12. Guidelines for Reasonable and Appropriate Care in the Emergency Department. Society for Academic Emergency Medicine. 2022. Accessed November 17, 2022. https://www.saem.org/publications/grace
13. Clinical Practice Guidelines. American Academy of Otolaryngology–Head and Neck Surgery (AAO-HNS). Accessed November 17, 2022. https://www.entnet.org/quality-practice/quality-products/clinical-practice-guidelines/
14. American College of Radiology. Appropriateness criteria. Accessed November 17, 2022. https://acsearch.acr.org/list
15. 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(4):388-396.
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(8):529-536.
17. Araujo Filho BC, Sakae FA, Sennes LU, Imamura R, de Menezes MR. Intraoral and transcutaneous cervical ultrasound in the differential diagnosis of peritonsillar cellulitis and abscesses. *Braz J Otorhinolaryngol*. 2006;72(3):377-381.
18. Boesen T, Jensen F. Preoperative ultrasonographic verification of peritonsillar abscesses in patients with severe tonsillitis. *Eur Arch Otorhinolaryngol*. 1992;249(3):131-133.
19. Buckley AR, Moss EH, Blokmanis A. Diagnosis of peritonsillar abscess: value of intraoral sonography. *Am J Roentgenol*. 1994;162(4):961-964.
20. Costantino TG, Satz WA, Dehnkamp W, Goett H. Randomized trial comparing intraoral ultrasound to landmark-based needle aspiration in patients with suspected peritonsillar abscess. *Acad Emerg Med*. 2012;19(6):626-631.
21. Fordham MT, Rock AN, Bandarkar A, et al. Transcervical ultrasonography in the diagnosis of pediatric peritonsillar abscess. *Laryngoscope*. 2015;125(12):2799-2804.
22. Haeggström A, Gustafsson O, Engquist S, Engström CF. Intraoral ultrasonography in the diagnosis of peritonsillar abscess. *Otolaryngol Head Neck Surg*. 1993;108(3):243-247.
23. Kelley K, Dean T. 367 point-of-care ultrasound for evaluation of Peritonsillar abscess. *Ann Emerg Med*. 2017;70(4):S144.
24. Kew J, Ahuja A, Loftus WK, Scott PM, Metreweli C. Peritonsillar abscess appearance on intra-oral ultrasonography. *Clin Radiol*. 1998;53(2):143-146.
25. Klimek L, Mösges R, Wein B, Schmelzer B, Kuth G. Surgical or medical treatment of questionable peritonsillar abscess? The use of B-mode ultrasonography. *Acta Otorhinolaryngol Belg*. 1993;47(4):439-442.
26. Lyon M, Blaivas M. Intraoral ultrasound in the diagnosis and treatment of suspected peritonsillar abscess in the emergency department. *Acad Emerg Med*. 2005;12(1):85-88.
27. Miziara ID, Koishi HU, Zonato AI, Valentini M, Miniti A, De Menezes MR. The use of ultrasound evaluation in the diagnosis of peritonsillar abscess. *Rev Laryngol Otol Rhinol (Bord)*. 2001;122(3):201-203.
28. Nogan S, Jandali D, Cipolla M, DeSilva B. The use of ultrasound imaging in evaluation of peritonsillar infections. *Laryngoscope*. 2015;125(11):2604-2607.
29. Salihoglu M, Eroglu M, Yildirim AO, Cakmak A, Hardal U, Kara K. Transoral ultrasonography in the diagnosis and treatment of peritonsillar abscess. *Clin Imaging*. 2013;37(3):465-467.
30. Sievert M, Miksch M, Mantsopoulos K, et al. The value of transcutaneous ultrasound in the diagnosis of tonsillar abscess: a retrospective analysis. *Auris Nasus Larynx*. 2021;48(6):1120-1125.
31. Strong EB, Woodward PJ, Johnson LP. Intraoral ultrasound evaluation of peritonsillar abscess. *Laryngoscope*. 1995;105(8 Pt 1):779-782.
32. Todsen T, Stage MG, Hahn CH. Surgeon-performed transoral ultrasound-guided aspiration of peritonsillar abscess. *Laryngoscope*. 2021;131(10):2241-2245.
33. Froehlich MH, Huang Z, Reilly BK. Utilization of ultrasound for diagnostic evaluation and management of peritonsillar abscesses. *Curr Opin Otolaryngol Head Neck Surg*. 2017;25(2):163-168.

34. Forner D, Curry DE, Hancock K, et al. Medical intervention alone vs. surgical drainage for treatment of peritonsillar abscess: a systematic review and meta-analysis. *Otolaryngol Head Neck Surg*. 2020;163(5):915-922.
35. Urban MJ, Masliah J, Heyd C, Patel TR, Nielsen T. Peritonsillar abscess size as a predictor of medical therapy success. *Ann Otol Rhinol Laryngol*. 2021;131:211-218.
36. Hess EP, Grudzen CR, Thomson R, Raja AS, Carpenter CR. Shared decision-making in the emergency department: respecting patient autonomy when seconds count. *Acad Emerg Med*. 2015;22(7):856-864.
37. Finnerty NM, Rodriguez RM, Carpenter CR, et al. Clinical decision rules for diagnostic imaging in the emergency department: a research agenda. *Acad Emerg Med*. 2015;22(12):1406-1416.
38. Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ*. 2003;326(7379):41-44.
39. Gallo L, Hua N, Mercuri M, Silveira A, Worster A, Best Evidence in Emergency Medicine (BEEM; beem.ca). Adherence to standards for reporting diagnostic accuracy in emergency medicine research. *Acad Emerg Med*. 2017;24(8):914-919.
40. Kohn MA, Carpenter CR, Newman TB. Understanding the direction of bias in studies of diagnostic test accuracy. *Acad Emerg Med*. 2013;20(11):1194-1206.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Kim DJ, Burton JE, Hammad A, et al. Test characteristics of ultrasound for the diagnosis of peritonsillar abscess: A systematic review and meta-analysis. *Acad Emerg Med*. 2023;30:859-869. doi:10.1111/acem.14660

Penn State Health Emergency Medicine

About Us:

Penn State Health is a multi-hospital health system serving patients and communities across central Pennsylvania. We are the only medical facility in Pennsylvania to be accredited as a Level I pediatric trauma center and Level I adult trauma center. The system includes Penn State Health Milton S. Hershey Medical Center, Penn State Health Children's Hospital, and Penn State Cancer Institute based in Hershey, Pa.; Penn State Health Hampden Medical Center in Enola, Pa.; Penn State Health Holy Spirit Medical Center in Camp Hill, Pa.; Penn State Health St. Joseph Medical Center in Reading, Pa.; Penn State Health Lancaster Pediatric Center in Lancaster, Pa.; Penn State Health Lancaster Medical Center (opening fall 2022); and more than 3,000 physicians and direct care providers at more than 126 outpatient practices in 94 locations. Additionally, the system jointly operates various health care providers, including Penn State Health Rehabilitation Hospital, Hershey Outpatient Surgery Center, Hershey Endoscopy Center, Horizon Home Healthcare and the Pennsylvania Psychiatric Institute.

We foster a collaborative environment rich with diversity, share a passion for patient care, and have a space for those who share our spark of innovative research interests. Our health system is expanding and we have opportunities in both academic hospital as well community hospital settings.

Benefit highlights include:

- Competitive salary with sign-on bonus
- Comprehensive benefits and retirement package
- Relocation assistance & CME allowance
- Attractive neighborhoods in scenic central Pennsylvania

PennState Health

FOR MORE INFORMATION PLEASE CONTACT:

Heather Peffley, PHR CPRP - Penn State Health Lead Physician Recruiter
hpeffley@pennstatehealth.psu.edu