





Diagnostic accuracy of point-of-care lung ultrasound for COVID-19: a systematic review and meta-analysis

Ashley Matthies ^{1,2} Michael Trauer ^{2,3} Karl Chopra ^{1,2}
Robert David Jarman ⁴

Handling editor Richard Body

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/emered-2021-212092>).

¹Emergency Department, Homerton University Hospital NHS Foundation Trust, London, UK

²School of Health and Life Sciences, Teesside University, Middlesbrough, UK

³Emergency Department, Guy's and Saint Thomas' NHS Foundation Trust, London, UK

⁴Accident and Emergency Department, Royal Victoria Infirmary, Newcastle upon Tyne, UK

Correspondence to

Dr Ashley Matthies, Emergency Department, Homerton University Hospital NHS Foundation Trust, London, UK; ashley.matthies@nhs.net

Received 18 January 2022
Accepted 31 January 2023
Published Online First
3 March 2023

ABSTRACT

Background Point-of-care (POC) lung ultrasound (LUS) is widely used in the emergency setting and there is an established evidence base across a range of respiratory diseases, including previous viral epidemics. The necessity for rapid testing combined with the limitations of other diagnostic tests has led to the proposal of various potential roles for LUS during the COVID-19 pandemic. This systematic review and meta-analysis focused specifically on the diagnostic accuracy of LUS in adult patients presenting with suspected COVID-19 infection.

Methods Traditional and grey-literature searches were performed on 1 June 2021. Two authors independently carried out the searches, selected studies and completed the Quality Assessment Tool for Diagnostic Test Accuracy Studies (QUADAS-2). Meta-analysis was carried out using established open-source packages in R. We report overall sensitivity, specificity, positive and negative predictive values, and the hierarchical summary receiver operating characteristic curve for LUS. Heterogeneity was determined using the I^2 statistic.

Results Twenty studies were included, published between October 2020 and April 2021, providing data from a total of 4314 patients. The prevalence and admission rates were generally high across all studies. Overall, LUS was found to be 87.2% sensitive (95% CI 83.6 to 90.2) and 69.5% specific (95% CI 62.2 to 72.5) and demonstrated overall positive and negative likelihood ratios of 3.0 (95% CI 2.3 to 4.1) and 0.16 (95% CI 0.12 to 0.22), respectively. Separate analyses for each reference standard revealed similar sensitivities and specificities for LUS. Heterogeneity was found to be high across the studies. Overall, the quality of studies was low with a high risk of selection bias due to convenience sampling. There were also applicability concerns because all studies were undertaken during a period of high prevalence.

Conclusion During a period of high prevalence, LUS had a sensitivity of 87% for the diagnosis of COVID-19 infection. However, more research is required to confirm these results in more generalisable populations, including those less likely to be admitted to hospital.

PROSPERO registration number CRD42021250464

INTRODUCTION

COVID-19 was declared a global pandemic by WHO on 11 March 2020. As of 27 August 2021, there have been more than 214 million confirmed cases and more than 4.4 million deaths.¹

There is now an established evidence base concerning the role of point-of-care (POC) lung ultrasound (LUS) in patients presenting with acute

WHAT IS ALREADY KNOWN ON THIS SUBJECT

- ⇒ There is an established evidence base supporting the use of point-of-care (POC) lung ultrasound (LUS) for many acute respiratory diseases
- ⇒ POC LUS was shown to be useful in previous epidemics and accurate in differentiating bacterial and viral pneumoniae
- ⇒ Chest radiographs have poor diagnostic accuracy for COVID-19 and increase equipment contamination and healthcare provider exposure to COVID-19.

WHAT THIS STUDY ADDS

- ⇒ This systematic review and meta-analysis of data from more than 4000 patients demonstrated that POC lung ultrasound is 87% sensitive for COVID-19 infection.
- ⇒ A quantitative scoring (or grading) system may be more useful for detecting COVID-19 pneumonia with ultrasound than a binary diagnostic threshold.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

- ⇒ POC LUS is a useful adjunct to standard clinical evaluation of patients with suspected COVID-19 infection.
- ⇒ High-quality trials are required to determine the accuracy of POC LUS in more representative populations. This includes populations with reduced disease prevalence, high vaccination rates and varying transmissibility and virulence.

respiratory distress.² First described in 2008, the BLUE (Bedside Lung Ultrasound in Emergency) protocol was found to be greater than 90% accurate for the underlying diagnosis of acute respiratory distress. Since then, there has been a proliferation of research into LUS which has confirmed that it is highly accurate for several respiratory conditions and often more accurate than CXR.^{3,4}

The utility of LUS has also been described during previous viral epidemics. LUS was found to be accurate in differentiating viral and bacterial pneumonia during the influenza epidemic in 2009 and was found to be more accurate than CXR for avian influenza.^{5,6}

Furthermore, CXR and a single initial reverse transcriptase polymerase chain reaction (RT-PCR) test have both been shown to have suboptimal



© Author(s) (or their employer(s)) 2023. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Matthies A, Trauer M, Chopra K, et al. *Emerg Med J* 2023;**40**:407–417.

diagnostic accuracy for COVID-19. In a large meta-analysis, CXR was found to be 81% sensitive and 72% specific.⁷ A single initial RT-PCR test has been found to have a sensitivity ranging between approximately 70% and 95% when compared with serial testing.⁸ While chest CT has been shown to be more accurate than CXR,⁷ this is neither an appropriate nor feasible test to provide for every patient, and is not recommended by the Royal College of Radiologists.⁹

As a result, various roles for LUS have been proposed including triage, diagnosis, prognostication and monitoring of disease progression.¹⁰ Multiple studies have demonstrated that LUS may have superior test characteristics than CXR.^{11–13} Furthermore, LUS is a rapid test that can be performed at the bedside, causes no radiation exposure and has good inter-operator characteristics.¹⁴ It may also have the potential to reduce nosocomial infections and healthcare provider exposure to COVID-19.¹⁵

The sonographic findings of COVID-19 pneumonia (or pneumonitis) are now well described and include pleural line abnormalities, single or confluent B-lines, small areas of peripheral consolidation (also termed sub-pleural) and in severe cases lobar consolidation with characteristic air bronchograms. These features are shared with other viral pneumonias.^{16–18} The constellation of findings is also well understood with abnormalities typically bilateral and patchy,¹⁹ which may help distinguish viral pneumonia from other causes of acute respiratory failure including bacterial pneumonia,⁵ and those causing interstitial syndrome that share individual features but cause a different pattern and/or combination of findings. An example of this would be pulmonary oedema, which also results in B-line formation, but these are typically widespread, symmetrical and occur in the absence of pleural line abnormalities or consolidations.

A scoping review of the diagnostic utility of LUS in COVID-19 was published in August 2020, yielding 33 studies. The methodology of the included studies was generally poor and only five studies reported diagnostic accuracy. However, there was a trend towards high sensitivity and low specificity.²⁰

Since then, we have experienced further waves of infection, and more evidence concerning the role of LUS has accumulated. The purpose of this systematic review was to evaluate the diagnostic accuracy of POC LUS in adult patients presenting with suspected COVID-19 infection compared with three commonly used reference standards used for the diagnosis of COVID-19: (1) RT-PCR, (2) chest CT and (3) aggregate final clinical diagnosis.

METHODS

This systematic review was synthesised according to the Preferred Reporting Items for Systematic Reviews and Meta-analyses,²¹ and the protocol was registered on PROSPERO (CRD42021250464). No funding was sought for the production of this work.

Search strategy

Traditional sources of literature were searched, including Ovid MEDLINE, Embase, SCOPUS, Cochrane Library and Google Scholar. Less traditional sources were also searched including medRxiv, the pre-print server for health sciences.

A search strategy was developed in line with existing guidance²² and searches were performed on 1 June 2021 (see online supplemental file 1).

Eligibility

All prospective and retrospective trials of patients over 16 years of age comparing lung ultrasonography to either RT-PCR, chest CT or a final clinical diagnosis of COVID-19 were included. Case reports, editorials and recommendation/instructional journal papers were excluded. Trials not reported in the English language were excluded from this study.

Study selection

Independent searches for titles and abstracts were carried out independently by two reviewers, both senior emergency physicians (AM, MT). A screening and selection tool (see online supplemental file 2) was agreed on and applied to identify eligible studies from both the abstract and subsequent full-text review. Disagreements reviewed by a third reviewer (RDJ). The reference lists from these studies were also reviewed to identify and further relevant studies.

Definition of the index test/diagnostic threshold

Multiple scanning protocols have been described and the specific threshold of findings necessary to diagnose COVID-19 has also not been agreed on. Studies were included regardless of the number of zones scanned or the diagnostic threshold used.

Definition of the reference standard

Although a positive RT-PCR is recognised as the gold standard diagnostic test for ruling-in COVID-19 infection, it is less reliable at ruling out the disease. While RT-PCR testing is used to detect infection, LUS is used to diagnose viral pneumonia. Chest CT has been described as the reference standard for viral pneumonia²³ but is infrequently used in clinical practice. An alternative to both is the use of a final clinical diagnosis incorporating imaging findings and results of serial RT-PCR tests. We have included studies using all three reference standards, and present combined and individual data.

Data collection and synthesis

Two independent reviewers (AM, MT) extracted the following data (displayed in table 1): study design and setting, exclusion criteria, number of patients, admission rate (used as a measure of disease severity), prevalence of COVID-19, sensitivity, specificity, number of true positive and negative results, number of false positive and negative results, scanning technique and diagnostic threshold (or scoring) for LUS, reference standard and blinding. Disagreements were resolved by a third independent reviewer (RDJ).

Quality appraisal

Two reviewers (AM, MT) independently used the Quality Assessment Tool for Diagnostic Test Accuracy Studies (QUADAS-2) to evaluate the quality of each study and risk of bias. Disagreements were resolved by a third independent reviewer (RDJ).

Data analysis

Data were analysed using a dedicated application, MetaDTA (Complex Reviews Support Unit, National Institute for Health Research, UK), a validated tool for meta-analysis of diagnostic test accuracy studies that utilises open-source packages in R (R Foundation for Statistical Computing, Vienna, Austria).²⁴ Random-effects bivariate binomial models were used to estimate overall sensitivity and

Table 1 Summary of key demographic and epidemiological data as well as results from each of the included studies, arranged by sample size from highest

Author	Publication (submission) dates	Design/enrolment	Setting	Exclusions	n	COVID-19 prevalence (%)	Admission rate (%)	LUS protocol	LUS Dx threshold	Reference standard	Sensitivity LUS (%) (95% CI)	Specificity LUS (%) (95% CI)	TP	FN	FP	TN	Prior training and experience
Volpicelli <i>et al</i>	December 2020 (March 2021)	Prospective, convenience	ED (74%) COVID unit (16%) ICU (10%)	None	1462	70	34	'Complete' (>12 zones)	High or intermediate pattern	PCR	90 (88 to 92)	53 (48 to 57)	921	101	209	231	Documented evidence in performing LUS
Colombi <i>et al</i>	July 2020 (October 2020)	Retrospective, convenience	ED	Those without clinical or epidemiological suspicion	486	70	100	12 zones	>1 zone	PCR	91 (84 to 94)	66 (58 to 74)	311	29	49	97	No details provided
Sorlini <i>et al</i>	June 2020 (October 2020)	Retrospective, convenience	ED	None	384	–	100	12 zones	>2 zones positive	PCR	92 (88 to 95)	65 (55 to 74)	264	23	34	63	LUS accreditation, 1-day LUS course and >150 supervised scans
Bianchi <i>et al</i>	October 2020 (April 2021)	Prospective, convenience	ED	Pregnancy, confirmed COVID-19 infection on PCR	360	39	100	12 zones	Summary assessment	PCR	86 (79 to 91)	71 (66 to 77)	120	20	64	156	LUS accreditation and >50 LUS examinations
Pivetta <i>et al</i>	July 2020 (October 2020)	Prospective, convenience	ED	Intubated patients Confirmed cases Those requiring acute mental health care	228	47	40	12 zones	Operator interpretation	PCR	94 (88 to 97)	95 (90 to 98)	101	6	6	115	LUS accreditation, single learning module and >40 supervised scans
Lieveld <i>et al</i>	July 2020 (October 2020)	Prospective, convenience	ED	Patients with uninterpretable LUS images	187	46	100	12 zones	>2 zones positive	PCR/final Dx	92 (84 to 96) 92 (84 to 96)	71 (61 to 79) 80 (70 to 87)	79	7	29	71	Certified and >20 supervised scans
Brenner <i>et al</i>	December 2020 (March 2021)	Retrospective, convenience	ED Medical ward	None	174	53	77	12 zones	Summary assessment	PCR	86 (77 to 92)	72 (61 to 81)	79	13	23	59	Credentialed in LUS
Schmid <i>et al</i>	July 2020 (December 2020)	Retrospective, convenience	ED	None	135	29	73	12 zones	>2 zones positive	PCR/CT	77 (62 to 87) 65 (43 to 82)	77 (68 to 84) 73 (43 to 90)	30/13	9/7	22/3	74/8	Training in POCUS
Zanforlin <i>et al</i>	June 2020 (December 2020)	Retrospective, convenience	ED	None	111	41	48	20 zones	LUS score >3	Final Dx	85 (72 to 92)	75 (64 to 84)	39	7	16	49	No details provided
Gibbons <i>et al</i>	November 2020 (February 2021)	Prospective, convenience	ED	Pregnancy	110	75	100	8 zones	>2 zones positive	CT	98 (91 to 99)	35 (20 to 53)	79	2	19	10	ACEP credentialled and >25 LUS scans
Tung-Chen <i>et al</i>	October 2020 (January 2021)	Prospective, convenience	ED	None	96	61	77	12 zones	>2 zones positive	PCR	72 (59 to 82) (irregular pleural line) 87 (confluent b-lines)	66 (49 to 79) (irregular pleural line) 33 (confluent b-lines)	39	15	12	23	Experienced sinologists, performing >10 US scans per week
Haak <i>et al</i>	May 2020 (November 2020)	Prospective, convenience	ED	Age <16 CCF No LUS Confirmed COVID	93	29	–	12 zones	>1 zone	PCR/CT	96 (80 to 99) 60 (23 to 88)	59 (48 to 70) 23/3 62 (41 to 79)	23/3	1/2	28/8	41/13	>2.5 years of POCUS experience
Narinx <i>et al</i>	June 2020 (September 2020)	Retrospective, convenience	ED	None	90	17	–	12 zones	>1 zone	PCR	93 (70 to 99)	12 (14 to 32)	14	1	59	16	Board certified for LUS and >5 years of experience in US
Karagöz <i>et al</i>	May 2020 (December 2020)	Prospective, convenience	ED	Age <18 Chest pain Pregnant Hypotensive Previous thoracic surgery if unable to obtain optimal images	72	48	100	12 zones	Not defined	CT	97 (85 to 99)	92 (80 to 97)	32	1	3	36	5 years of experience, 3-hour theory module, 20 live demonstrations
Jalil <i>et al</i>	September 2020 (October 2020)	Retrospective, convenience	Mixed secondary care setting (LUS performed <24 hours from presentation)	None	69	52	100	8 zones	LUS-COV criteria (summary assessment of images)	PCR	86 (71 to 94)	76 (76 to 97)	31	5	3	30	No details provided, but all 3 operators were intensive care physicians
Fonsi <i>et al</i>	May 2020 (September 2020)	Prospective, convenience	ED/COVID unit	None	63	70	100	12 zones	Not defined	PCR	68 (53 to 80)	79 (57 to 91)	30	14	4	15	No details provided
Bosso <i>et al</i>	May 2020 (October 2020)	Prospective, convenience	ED/COVID unit	None	53	49	100	12 zones	Severity score used	PCR	73 (54 to 86) (optimal cut-off >12.5)	88 (72 to 96) (optimal cut-off >12.5)	19	7	3	24	No details provided

Continued

Table 1 Continued

Author	Publication (submission) dates	Design/enrolment	Setting	Exclusions	n	COVID-19 prevalence (%)	Admission rate (%)	LUS protocol	LUS Dx threshold	Reference standard	Sensitivity LUS (%) (95% CI)	Specificity LUS (%) (95% CI)	TP	FN	FP	TN	Prior training and experience
Spiedel <i>et al</i>	September 2020 (December 2020)	Prospective, convenience	Admission ward (LUS performed <24 hours from presentation)	If PCR result already available ICU admissions	49	22	100	12 zones	Severity score used	Final Dx	91 (62 to 98) (optimal cut-off score >8)	76 (61 to 87) (optimal cut-off score >8)	10	1	9	29	Specific POCUS qualification, >3 years of experience in US, >1 year of experience in LUS
Walsh <i>et al</i>	July 2020 (September 2020)	Retrospective, convenience	ED	History of Congestive cardiac failure or other pathology likely to affect LUS	49	12	39	Continuous posterior assessment+single anterior and axillary zones	Operator interpretation	CT	100 (76 to 100)	80 (49 to 94)	12	0	2	8	Facile in LUS, but no specific details provided
Pare <i>et al</i>	April 2020 (June 2020)	Retrospective, convenience	ED	If there was no CXR, LUS or PCR test If the patient had an underlying condition predisposing them to B-lines on LUS	43	70	72	Unclear (median number of saved images=6)	>1 B-line	PCR	89 (72 to 96)	56 (33 to 77)	24	3	7	9	No details provided. Images overinterpreted by 3 fellowship trained physicians

Dx, diagnosis; FN, false negative; FP, false positive; ICU, intensive care unit; LUS, lung ultrasound; POCUS, point-of-care ultrasound; TN, true negative; TP, true positive.

specificity and provide the hierarchical summary receiver operating characteristic (HSROC) curve for LUS. The open-source package *Metafor* in R was used to determine heterogeneity for univariate sensitivity and specificity analysis using a DerSimonian and Laird random-effects model, providing the I^2 statistic.^{25 26} Combined and separate analyses were performed for each of the reference standards.

RESULTS

Our search identified 3041 studies from traditional databases and registers, and 221 from other sources. Following removal of duplicates, eligibility screening by title/abstract and then full-text review, a total of 20 studies were included. A flow chart was produced in line with the PRISMA guidelines (figure 1).

Quality appraisal

The results of the QUADAS-2 assessment are shown in figure 2.

Patient selection

We determined all studies to be at high risk of selection bias. All used convenience sampling and six had exclusion criteria that introduced a risk of spectrum bias (sampling methods and exclusion criteria for each study are shown in table 1). In two studies, patients with indeterminate LUS examinations were excluded, further increasing the risk of spectrum bias. In addition, high prevalence, high rates of severe disease and hospital admission across all the studies introduced both a risk of referral bias and applicability concerns.

Index test

Overall, there was significant variation in the LUS protocols used as well as the diagnostic threshold for a positive scan. Furthermore, prior experience and training in ultrasound varied between studies. In five studies, it was not specified whether the operator was blinded to clinical information; therefore, the risk of bias was deemed unclear. One study was deemed to have applicability concerns due to the use of a low diagnostic threshold (a single B-line), which in the context of high disease prevalence may have overestimated diagnostic test performance. There were unclear applicability concerns in three studies, in which minimal or no information was given regarding prior training and experience of the clinicians performing the index test.

Reference standard

The two studies that used the final clinical diagnosis as their only reference standard were deemed to have a high risk of incorporation bias. A third study also comparing LUS with the final clinical diagnosis also provided separate comparison to RT-PCR and as such was not deemed high risk of incorporation overall.

Summary of results

An overview of the characteristics of all the included studies is shown in table 1. A total of 4314 patients were included. There was an even mix of prospective (n=11) and retrospective (n=9) studies, and all were convenience samples. The prevalence of COVID-19 infection ranged from 12% to 75%. In 16 of the studies, patients were recruited in the ED. In 4 of the studies, patients were recruited from mixed secondary care settings (ED, medical admissions ward, intensive care unit). Importantly, in these studies the LUS

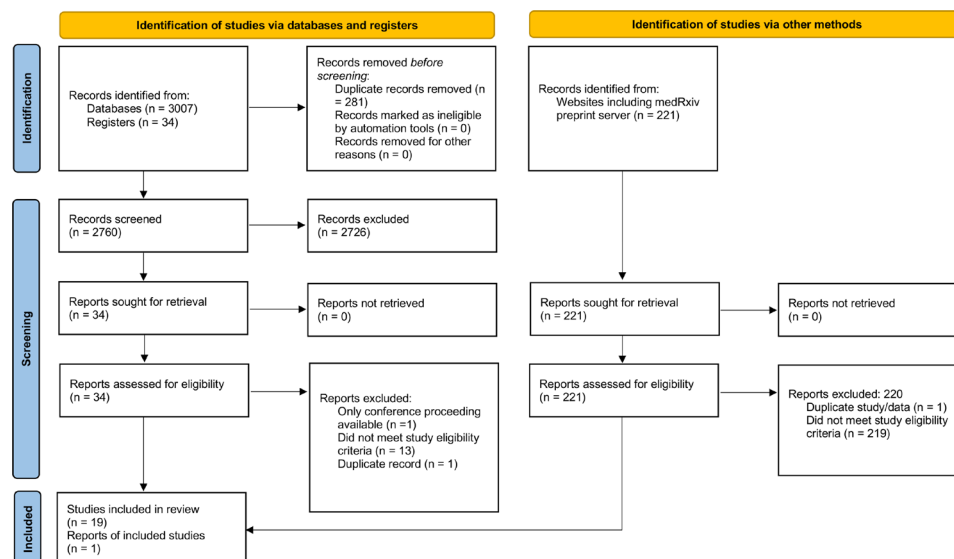


Figure 1 PRISMA flow diagram.

was performed within 24 hours of presentation and prior to results of the RT-PCR test being available.

Fifteen studies used RT-PCR as the reference standard, while five used chest CT and three the final clinical diagnosis. The smallest study included 43 patients while the largest included 1462. Sensitivity of LUS ranged between 68% and 100%, and specificity ranged between 21% and 97%.^{11–14 27–42}

Comparison with RT-PCR testing

Fifteen studies used RT-PCR as the reference standard. In this group, sensitivity ranged between 68% and 96% while specificity ranged between 21% and 91%.^{11 12 14 27–38}

This included the largest published study which included 1462 patients and classified patients according to their clinical phenotype: 1) mild (without dyspnoea or respiratory failure, $n=402$); 2) severe (with dyspnoea or respiratory failure, $n=725$); 3) mixed (patients with cardiorespiratory comorbidities, $n=335$). Furthermore, LUS studies were classified as either low, intermediate or high probability of COVID, or suggestive of an alternate diagnosis. The authors describe findings such as isolated large consolidations, large pleural effusions and a homogeneous diffuse B-line pattern as being more consistent with an alternate diagnosis.¹⁴

In the overall population, a high probability LUS demonstrated a modest sensitivity of 60% (95% CI 57 to 63) but a relatively strong specificity of 89% (95% CI 86 to 92), while combined high/intermediate LUS studies predictably demonstrated a higher sensitivity of 90% (95% CI 88 to 92) but lower specificity of 53% (95% CI 48 to 57). In patients with an 'alternate' diagnosis (or disease pattern) on LUS, the majority tested negative on RT-PCR (85% and 84%, respectively, for patients classified as the severe or mixed phenotype, respectively). In patients with the mild phenotype, a high probability LUS demonstrated a sensitivity of 32% (95% CI 26 to 38) and a specificity of 94% (95% CI 90 to 97). In the same group of patients, combined high/intermediate LUS demonstrated a sensitivity of 68% (95% CI 61 to 74) and specificity of 64% (95% CI 56 to 71). In those classified as having a mixed phenotype, a high probability LUS demonstrated an improved sensitivity of 68% (95% CI 61 to 74) and a specificity of 88% (95% CI 81 to 93).

In this group of patients, combined high/intermediate LUS demonstrated a sensitivity of 95% (95% CI 90 to 97) and specificity of 50% (95% CI 41 to 59). In patients with the severe phenotype, a high probability LUS demonstrated an improved sensitivity of 69% (95% CI 65 to 72) and a specificity of 82% (95% CI 74 to 88). In this group of patients, combined high/intermediate LUS demonstrated a sensitivity of 97% (95% CI 96 to 98) and specificity of 39% (95% CI 31 to 48).¹⁴

Three studies compared both LUS and chest CT to the reference standard RT-PCR result. In each study, LUS demonstrated sensitivity greater than 90%, whereas the sensitivity of chest CT ranged between 69% and 90%. Both LUS and CT were less specific than RT-PCR. Specificity of LUS ranged between 21% and 66% and for CT between 43% and 80%.^{31 34 35}

In one study, the authors compared traditional clinical evaluation to an integrated assessment incorporating POC LUS. Clinical evaluation was found to be 81% (95% CI 73 to 88) sensitive and 64% (95% CI 54 to 72) specific whereas integrated assessment with LUS was 94% (95% CI 88 to 98) sensitive and 95% (95% CI 90 to 98) specific.

Two studies adopted scoring systems for the LUS findings. Unsurprisingly, those who tested positive on RT-PCR had higher scores (optimal cut-off scores are shown in table 1).^{28 29} It was also demonstrated that complete (12-zone) studies were 5% more sensitive and 4% more specific than less thorough studies.²⁹

Comparison with chest CT

Five studies (a total of 282 patients) used chest CT as the reference standard. Sensitivity ranged between 65% and 97% and specificity between 59% and 97%.^{13 30 37 39 40}

Comparison with final clinical diagnosis

Three studies used a final clinical diagnosis as the reference standard (a total of 347 patients). Sensitivity of LUS ranged between 85% and 91% and specificity between 75% and 80%.^{34 41 42}

Two studies adopted scoring systems for LUS and optimal cut-off scores demonstrating sensitivities above 90% and specificities above 75%.^{41 42}

Study	Risk of Bias				Applicability concerns		
	Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
Volpicelli	⊗	☺	☺	☺	⊗	☺	☺
Colombi	⊗	☺	☺	☺	⊗	?	☺
Sorlini	⊗	☺	☺	☺	⊗	☺	☺
Bianchi	⊗	☺	☺	☺	⊗	☺	☺
Pivetta	⊗	☺	☺	☺	⊗	☺	☺
Lieveld	⊗	☺	☺	☺	⊗	☺	☺
Brenner	⊗	?	☺	☺	⊗	☺	☺
Schmid	⊗	☺	☺	☺	⊗	☺	☺
Zanforlin	⊗	☺	⊗	☺	⊗	?	⊗
Gibbons	⊗	?	☺	☺	⊗	☺	☺
Tung-Chen	⊗	☺	☺	☺	⊗	☺	☺
Haak	⊗	☺	☺	☺	⊗	☺	☺
Narinx	⊗	☺	☺	☺	⊗	☺	☺
Karagoz	⊗	?	☺	☺	⊗	☺	☺
Jalil	⊗	☺	☺	☺	⊗	?	☺
Fonsi	⊗	?	☺	☺	⊗	☺	☺
Bosso	⊗	☺	☺	☺	⊗	☺	☺
Spiedel	⊗	☺	⊗	☺	⊗	☺	⊗
Walsh	⊗	?	☺	☺	⊗	☺	☺
Pare	⊗	☺	☺	☺	⊗	⊗	☺

Figure 2 QUADAS-2 assessment (☺ low risk, ⊗ high risk, ? unclear risk). Studies arranged by sample size, from largest.

META-ANALYSIS

Combining all studies, LUS had an overall sensitivity of 87.2% (95% CI 83.6 to 90.2) and specificity of 69.5% (62.2 to 72.5). Forest plots for each are shown in [figure 3A](#). Positive and negative likelihood ratios were 3.04 (2.27 to 4.06) and 0.16 (0.12 to 0.22), respectively. The HSROC curve is shown in [figure 4](#). Heterogeneity between all studies was high, with I^2 values of 71.6% (95% CI 58.7 to 91.5) and 86.8% (84.4 to 96.0) for sensitivity and specificity, respectively.

When only studies using RT-PCR as the reference standard were included, sensitivity and specificity remained similar at 87.2% (83.4 to 90.3) and 68.7% (60.2 to 75.3), respectively. Forest plots are shown in [figure 3B](#). Positive and negative likelihood ratios were 2.94 (2.06 to 4.19) and 0.17 (0.12 to 0.24), respectively. Heterogeneity remained high with I^2 values of

73.5% (55.7 to 92.1) and 89.4% (87.6 to 97.6) for sensitivity and specificity, respectively.

When only the five studies using chest CT as the reference standard were analysed, sensitivity increased to 89.7% (64.1 to 97.7) and specificity to 70.5% (43.9 to 87.9). Forest plots are shown in [figure 3C](#). Positive and negative likelihood ratios were 3.24 (1.56 to 6.73) and 0.10 (0.02 to 0.47), respectively. Again, heterogeneity was high, with I^2 values of 79.0% (33.9 to 97.0) and 81.1% (42.4 to 97.3) for sensitivity and specificity, respectively.

When only the studies comparing LUS with a final clinical diagnosis were examined, sensitivity remained similar at 88.0% (81.9 to 92.2); however, specificity increased to 77.6% (71.1 to 83.0). Forest plots are shown in [figure 3D](#). Heterogeneity between these three studies was low with I^2 values of 0% (0.0 to 88.1) and 0% (0.0 to 0.81).

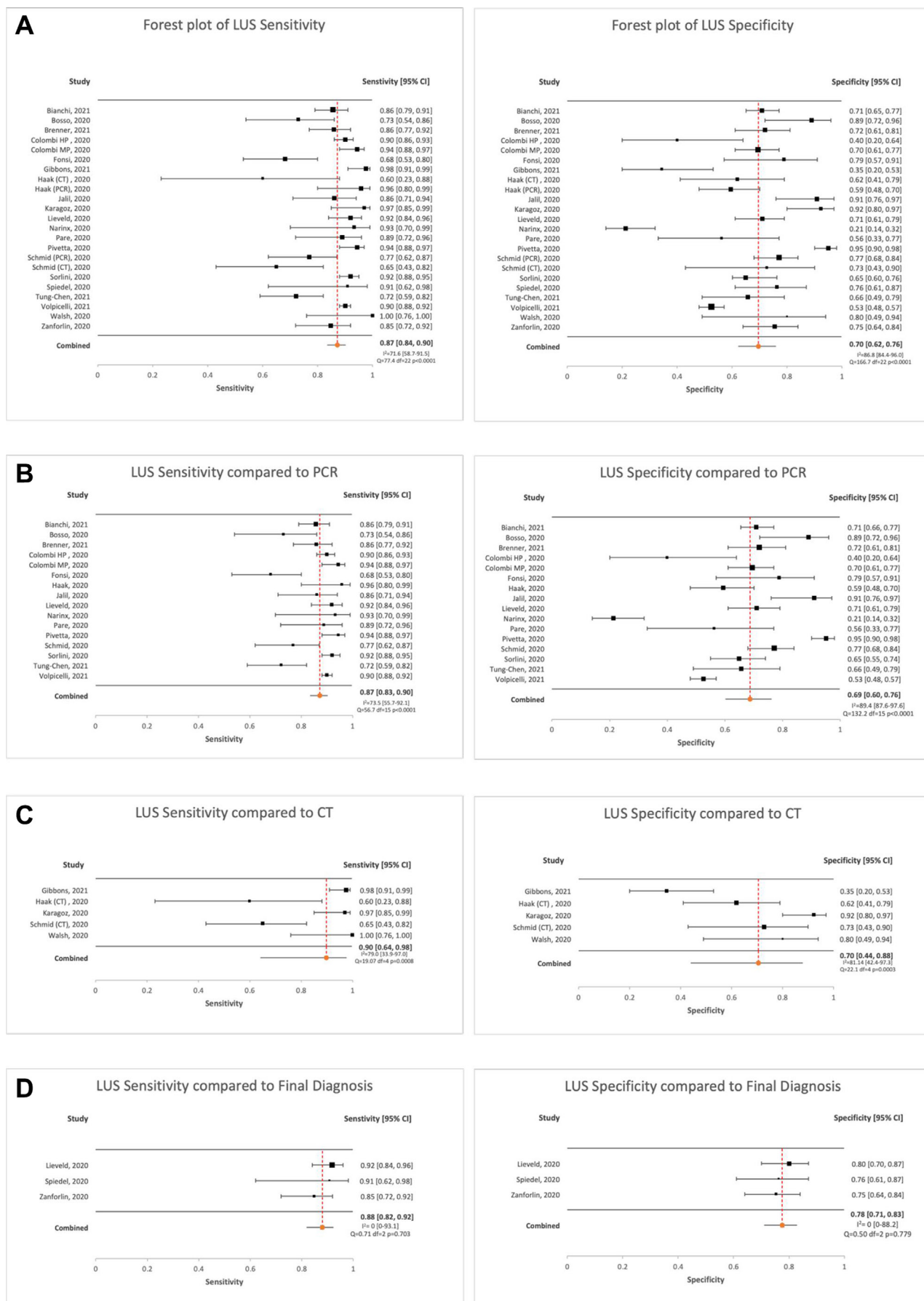


Figure 3 Forest plots demonstrating sensitivity and specificity of LUS overall (A) and compared with each of the three reference standards (RT-PCR (B), chest CT (C) and final clinical diagnosis (D)). LUS, lung ultrasound.

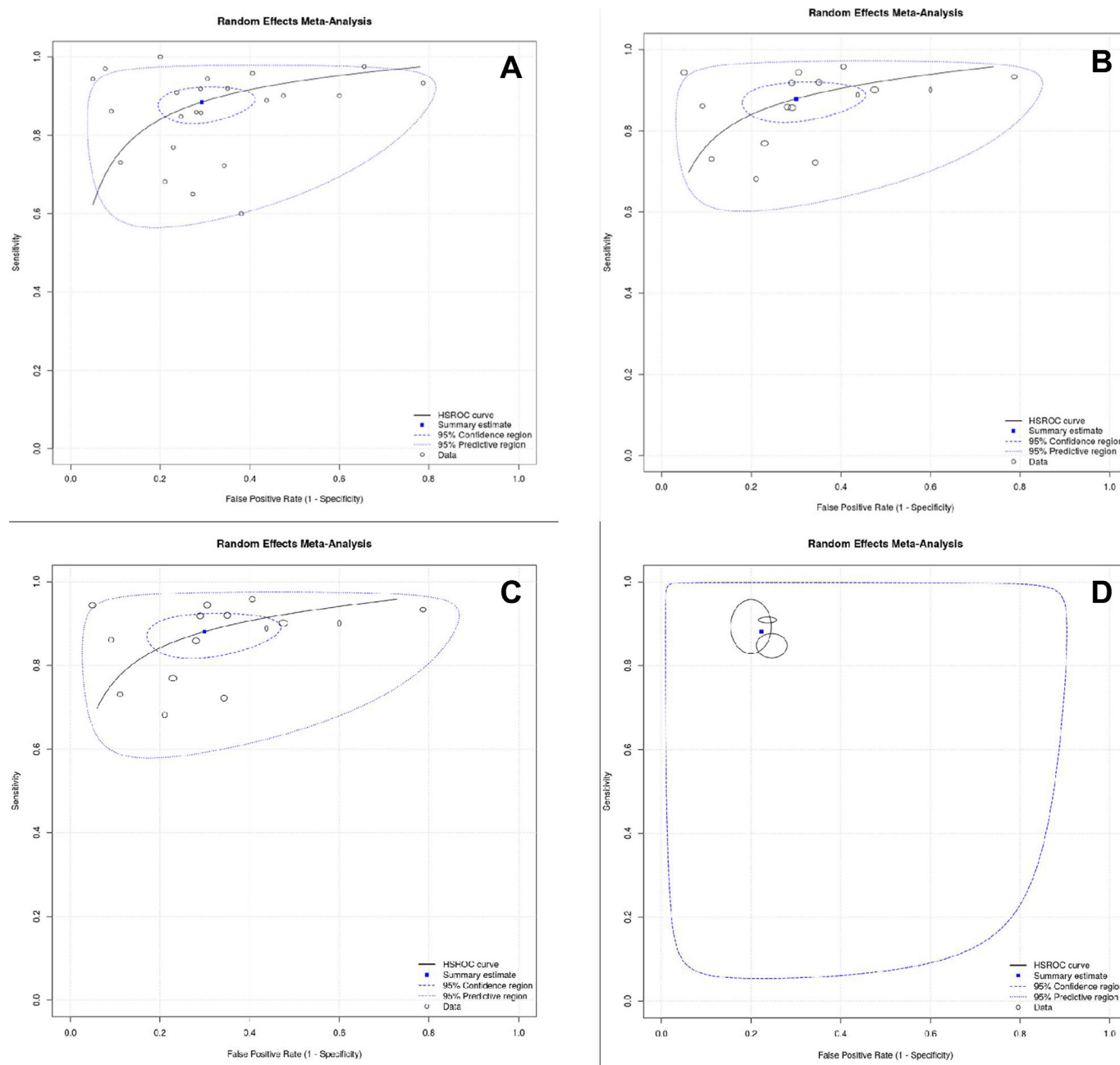


Figure 4 Hierarchical summary receiver operating characteristic (HSROC) curves for LUS overall (A) and compared with each of the three reference standards (RT-PCR (B), chest CT (C) and final clinical diagnosis (D)). LUS, lung ultrasound.

HSROC curves for LUS compared with each of the reference standards are shown in [figure 4B–D](#). These plot the true positive rates (sensitivity) against the false positive rates (1 – specificity) in each study, providing a graphical representation of the relationship between sensitivity and specificity, and a summary estimate for test performance.

DISCUSSION

In this systematic review, we aimed to evaluate the diagnostic accuracy of LUS for COVID-19 infection, diagnosed either by RT-PCR testing, chest CT scan or an aggregate final clinical diagnosis. We presented data from 20 studies consisting of more than 4000 patients. The majority of studies were conducted in high-prevalence populations with high rates of hospital admission. Overall, LUS was found to be 87% sensitive but generally less

specific for COVID-19 infection. The negative likelihood ratio (for ruling out disease) was reasonable; however, the positive likelihood ratio was less useful.

We reported test characteristics for LUS combining all studies and separately for the three reference standards commonly reported in the literature. However, it may be argued that in the population studied, all three are in fact comparable. The rate of hospital admission was high across the studies and therefore it would be expected that the majority of patients included presented with more severe disease and secondary complications of COVID-19, including viral pneumonia.⁴³ Chest CT is the recognised gold standard test for viral pneumonia²³ and therefore may be considered broadly equivalent to RT-PCR testing in this population. This is supported by separate analyses for each

of the reference standards demonstrating similar pooled sensitivities (87–90%) and specificities (69%–78%).

It is important to remember that LUS only tests for the pulmonary manifestations of COVID-19 and its sensitivity may therefore be significantly lower in non-admitted patients with mild or non-respiratory symptoms. The choice of reference standard for comparison to LUS is therefore crucial and we recommend future research examining LUS as a screening or diagnostic test for COVID-19 infection use RT-PCR testing, whereas studies examining the accuracy of LUS at diagnosing pulmonary manifestations should use chest CT.

Although in this study we reported sensitivity and specificity, examining the positive and negative predictive values (PPV and NPV), which incorporate the effect of disease prevalence, can be useful measures of the performance of a diagnostic test. This is an important consideration given the overall high prevalence rates reported in the studies. In an era of relatively lower prevalence, the PPV of LUS may be significantly lower than has been estimated. Caution should therefore be taken when considering the application of LUS for early diagnosis, screening and cohorting patients in the setting of lower prevalence.

The included studies were largely undertaken during periods of high disease prevalence and low or no vaccination, with data collection ending in June 2020. Since this time, WHO have designated several further variants of concern, including the currently circulating Omicron variants. It is widely appreciated that both virulence and transmissibility differ between variants and as such, the emergence of the Omicron variant has resulted in reduced rates of severe disease and hospitalisation.^{44–46}

While the rates of severe disease may vary between strains, the pulmonary complications and their radiological manifestations appear to be consistent.⁴⁷ Also, these are indeed shared by other causes of viral pneumonia.^{5 6} In one study comparing chest CT findings in patients infected with different variants of COVID-19, while there was a difference in the severity of radiological findings, the nature and distribution of changes was similar.⁴⁷ Given that LUS findings correspond well with those on chest CT,^{31 48} it is likely to be useful at detecting viral pneumonia caused by all strains of the disease. However, in populations with fewer manifestations of severe disease (including viral pneumonia), the sensitivity of LUS as a screening test is likely to be reduced. This should be confirmed with ongoing work during periods of reduced prevalence, high vaccination rates and changing variants.

There is a rapidly expanding evidence base demonstrating favourable characteristics for molecular and other non-RT-PCR tests for COVID-19 that are much easier and faster to perform and can turn over rapid results.⁸ The availability of these tests will likely impact on the value of LUS as an early bedside clinical tool.

Where there is access to rapid bedside molecular testing, LUS use is likely to be most useful in certain circumstances, for instance, patients presenting with acute respiratory distress where viral pneumonia secondary to COVID-19 infection is part of a wide differential. These patients may have incidental COVID-19 infection alongside an alternative primary respiratory illness, and this is made more likely by the dominance of less virulent strains and vaccination. In these patients, the pattern of abnormal findings on LUS can be used to accurately differentiate the underlying cause of respiratory disease.^{3 49}

Integrated LUS, used in combination with other clinical data, may also improve early diagnostic accuracy in the ED. Walsh *et al* showed that both sensitivity and specificity were increased when LUS was integrated with clinical findings as opposed to

interpreted in isolation.¹³ The potential to integrate bedside imaging findings is unique to point-of-care ultrasound (POCUS) and may enhance diagnostic accuracy. This concept was exemplified by Pivetta *et al* when they showed that an integrated LUS approach was highly accurate in patients with acute decompensated heart failure, more so than CXR or B-natriuretic peptide levels.⁴

All but three studies adopted a comprehensive 12-zone LUS technique. Given that the changes observed in viral pneumonia are bilateral and patchy, scanning technique is likely to affect sensitivity. Various scanning techniques have been described ranging from only limited views of the anterior chest to a more thorough examination of each intercostal space both anteriorly and posteriorly. Brenner *et al* showed that when only complete 12-zone studies were included in their analysis, both sensitivity and specificity of LUS were improved.²⁹

The heterogeneous criteria are used to determine whether a positive LUS may also have affected the results of the included studies. Various thresholds were used for ruling in this diagnosis, ranging from any B-lines to a minimum number of affected zones to a minimum severity score. One study used a low diagnostic threshold for LUS, where a single B-line was deemed positive. In a population of high disease prevalence, this is likely to have overestimated diagnostic performance.¹²

Lower diagnostic thresholds generally favoured improved sensitivity while higher thresholds favoured specificity. While relatively common in COVID-19 pneumonia, B-lines are also commonly seen in multiple respiratory diseases.² Intuitively, a low diagnostic threshold (eg, the presence of any B-lines) is likely to be sensitive but non-specific. Data presented by Volpelli *et al* suggest that a diagnostic score weighted to the specific sonographic findings is likely to be more useful than a binary cut-off.¹⁴ We therefore propose the development of a diagnostic (as opposed to severity) scoring system. Such a system could quantify the likelihood of COVID-19 infection or pneumonia depending on how typical the pattern of LUS features. A resulting estimated likelihood ratio could then be integrated with a pre-test probability to provide the treating clinician with a post-test probability.

STUDY LIMITATIONS

Included studies were limited to those published in the English language. In the context of an international pandemic, this may have excluded some pertinent studies. Second, there may have been a risk of publication bias and indeed, no dedicated assessment of this was made.

For both sensitivity and specificity analyses, heterogeneity was found to be high and as such the results should be interpreted with caution. This may in part be attributable to the use of broad inclusion criteria with respect to the index test. There was significant variation in the LUS protocols used as well as the diagnostic threshold for a positive scan. Furthermore, there was some variability regarding the prior experience and training of the operators. POCUS is an operator-dependent technique where a single provider is responsible for both image acquisition and interpretation. Therefore, both the specific training received, and prior experience of the scanning physicians is likely to influence diagnostic accuracy.

Searches were carried out on 1 June 2021, and the included studies reported data up until June 2020, 3 months after the declaration of the COVID-19 pandemic by WHO. Therefore, studies were limited to a period of high disease prevalence, during which admission rates with pulmonary complications

and severe disease were high. As a result, the conclusions of this meta-analysis may not be generalisable to current populations, in whom the emergence of less virulent strains of disease (such as the Omicron variant) and effective vaccines have reduced both prevalence and the likelihood of severe disease.

This meta-analysis investigated the role of LUS as a diagnostic test for COVID-19 infection when rates of severe disease (including pneumonia) were high, and during a period where alternative rapid point-of-care testing was not readily available. However, our results may not be applicable to current practice where the role of LUS is less clear and more likely to be useful in diagnosing viral pneumonia than COVID-19 infection itself.

FUTURE RESEARCH

Future research should focus on wider populations, including mildly and asymptomatic patients, during periods of lower disease prevalence to define the role and setting in which LUS is most useful. As previously described, future work is required to define the optimal diagnostic threshold for LUS. Finally, further work is required to understand the training and experience required to gain proficiency in LUS.

RECOMMENDATIONS/CONCLUSIONS

LUS was found to be 87% sensitive for the diagnosis of COVID-19 infection in a population of both high prevalence and mostly admitted patients, and may improve detection of COVID-19 pneumonia in this group compared with CXR. Indeed, four of the studies included in this meta-analysis also reported data for CXR each demonstrating LUS to be more sensitive^{11 12 38 39}. In patients requiring admission, a normal LUS should prompt consideration of an alternative diagnosis. Although overall specificity was less optimal, the presence of a full spectrum or typical pattern of LUS findings is likely to be more specific, and as such, a quantitative diagnostic LUS scoring system is likely to be more useful than a binary threshold.

Twitter Ashley Matthies @AshleyMatthies and Michael Trauer @michael_trauer

Contributors All four authors contributed significantly to the design of the study and manuscript preparation. Three authors performed the literature searches, selected studies and extracted data (AM, MT, RDJ). Two authors completed the meta-analysis (AM, KC). AM is the guarantor of the study.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as online supplemental information.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

This article is made freely available for personal use in accordance with BMJ's website terms and conditions for the duration of the covid-19 pandemic or until otherwise determined by BMJ. You may download and print the article for any lawful, non-commercial purpose (including text and data mining) provided that all copyright notices and trade marks are retained.

ORCID iDs

Ashley Matthies <http://orcid.org/0000-0003-1625-2205>

Michael Trauer <http://orcid.org/0000-0002-7631-8061>

Karl Chopra <http://orcid.org/0000-0003-2883-9185>

Robert David Jarman <http://orcid.org/0000-0003-4557-9274>

REFERENCES

- 1 WHO. Coronavirus disease (COVID-19) dashboard 2021. n.d. Available: <https://covid19.who.int/>
- 2 Volpicelli G, Elbarbary M, Blaivas M, et al. International evidence-based recommendations for point-of-care lung ultrasound. *Intensive Care Med* 2012;38:577–91.
- 3 Lichtenstein DA, Mezière GA. Relevance of lung ultrasound in the diagnosis of acute respiratory failure: the BLUE protocol. *Chest* 2008;134:117–25.
- 4 Pivetta E, Goffi A, Nazerian P, et al. Lung ultrasound integrated with clinical assessment for the diagnosis of acute decompensated heart failure in the emergency department: a randomized controlled trial. *Eur J Heart Fail* 2019;21:754–66.
- 5 Tsung JW, Kessler DO, Shah VP. Prospective application of clinician-performed lung ultrasonography during the 2009 H1N1 influenza A pandemic: distinguishing viral from bacterial pneumonia. *Crit Ultrasound J* 2012;4:16.
- 6 Zhang Y-K, Li J, Yang J-P, et al. Lung ultrasonography for the diagnosis of 11 patients with acute respiratory distress syndrome due to bird flu H7N9 infection. *Viral J* 2015;12:176.
- 7 Islam N, Ebrahimzadeh S, Salameh J-P, et al. Thoracic imaging tests for the diagnosis of COVID-19. *Cochrane Database Syst Rev* 2021;3:CD013639.
- 8 Axell-House DB, Lavingia R, Rafferty M, et al. The estimation of diagnostic accuracy of tests for COVID-19: a scoping review. *J Infect* 2020;81:681–97.
- 9 RCoR. The role of CT in patients with COVID-19 infection, statement of the Royal College of Radiologists. 2020. Available: <https://www.rcr.ac.uk/posts/role-ct-patients-suspected-covid-19-infection-12-march-2020>
- 10 Soldati G, Smargiassi A, Inchingolo R, et al. Is there a role for lung ultrasound during the COVID-19 pandemic? *J Ultrasound Med* 2020;39:1459–62.
- 11 Sorlini C, Femia M, Nattino G, et al. The role of lung ultrasound as a frontline diagnostic tool in the era of COVID-19 outbreak. *Intern Emerg Med* 2021;16:749–56.
- 12 Pare JR, Camelo I, Mayo KC, et al. Point-of-care lung ultrasound is more sensitive than chest radiograph for evaluation of COVID-19. *West J Emerg Med* 2020;21:771–8.
- 13 Walsh P, Hankins A, Bang H. Point-of-care lung ultrasound is useful to evaluate emergency department patients for COVID-19. *West J Emerg Med* 2020;21:24–31.
- 14 Volpicelli G, Gargani L, Perlini S, et al. Lung ultrasound for the early diagnosis of COVID-19 pneumonia: an international multicenter study. *Intensive Care Med* 2021;47:444–54.
- 15 Mongodi S, Orlando A, Arisi E, et al. Lung ultrasound in patients with acute respiratory failure reduces conventional imaging and health care provider exposure to COVID-19. *Ultrasound Med Biol* 2020;46:2090–3.
- 16 Volpicelli G, Frascisco MF. Sonographic detection of radio-occult interstitial lung involvement in measles pneumonitis. *Am J Emerg Med* 2009;27:128.
- 17 Tsai NW, Ngai CW, Mok KL, et al. Lung ultrasound imaging in avian influenza A (H7N9) respiratory failure. *Crit Ultrasound J* 2014;6:6.
- 18 Testa A, Soldati G, Copetti R, et al. Early recognition of the 2009 pandemic influenza A (H1N1) pneumonia by chest ultrasound. *Crit Care* 2012;16:R30.
- 19 Volpicelli G, Gargani L. Sonographic signs and patterns of COVID-19 pneumonia. *Ultrasound J* 2020;12:22.
- 20 Trauer MM, Matthies A, Mani N, et al. The utility of lung ultrasound in COVID-19: a systematic scoping review. *Ultrasound* 2020;28:208–22.
- 21 Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71.
- 22 McGowan J, Sampson M, Salzwedel DM, et al. PRESS peer review of electronic search strategies: 2015 guideline statement. *J Clin Epidemiol* 2016;75:40–6.
- 23 Winkler MH, Touw HR, van de Ven PM, et al. Diagnostic accuracy of chest radiograph, and when concomitantly studied lung ultrasound, in critically ill patients with respiratory symptoms: a systematic review and meta-analysis. *Crit Care Med* 2018;46:e707–14.
- 24 Freeman SC, Kerby CR, Patel A, et al. Development of an interactive web-based tool to conduct and interrogate meta-analysis of diagnostic test accuracy studies: metadta. *BMC Med Res Methodol* 2019;19:81.
- 25 Shim SR, Kim S-J, Lee J. Diagnostic test accuracy: application and practice using R software. *Epidemiol Health* 2019;41:e2019007.
- 26 W.V. conducting meta-analyses in R with the metafor package. *J Stat Softw* 2010;36:1–48.
- 27 Bianchi S, Savinelli C, Paolucci E, et al. Point-of-care ultrasound (POCUS) in the early diagnosis of novel coronavirus 2019 disease (COVID-19) in a first-level emergency department during a SARS-cov-2 outbreak in Italy: a real-life analysis. *Intern Emerg Med* 2022;17:193–204.
- 28 Bosso G, Allegorico E, Pagano A, et al. Lung ultrasound as diagnostic tool for SARS-cov-2 infection. *Intern Emerg Med* 2021;16:471–6.
- 29 Brenner DS, Liu GY, Omron R, et al. Diagnostic accuracy of lung ultrasound for SARS-cov-2: a retrospective cohort study. *Ultrasound J* 2021;13:12.

- 30 Haak SL, Renken IJ, Jager LC, *et al.* Diagnostic accuracy of point-of-care lung ultrasound in COVID-19. *Emerg Med J* 2021;38:94–9.
- 31 Colombi D, Petrini M, Maffi G, *et al.* Comparison of admission chest computed tomography and lung ultrasound performance for diagnosis of COVID-19 pneumonia in populations with different disease prevalence. *Eur J Radiol* 2020;133:109344.
- 32 Fonsi GB, Sapienza P, Brachini G, *et al.* Is lung ultrasound imaging a worthwhile procedure for severe acute respiratory syndrome coronavirus 2 pneumonia detection? *J Ultrasound Med* 2021;40:1113–23.
- 33 Jalil BA, Khan A, Kugasia IR, *et al.* Lung ultrasound in early SARS-cov-2 pneumonia and the LUS-cov criteria. *Proc (Bayl Univ Med Cent)* 2020;34:1–4.
- 34 Lieveld AWE, Kok B, Schuit FH, *et al.* Diagnosing COVID-19 pneumonia in a pandemic setting: lung ultrasound versus CT (LUVCT)-a multicentre, prospective, observational study. *ERJ Open Res* 2020;6:00539-2020.
- 35 Narinx N, Smismans A, Symons R, *et al.* Feasibility of using point-of-care lung ultrasound for early triage of COVID-19 patients in the emergency room. *Emerg Radiol* 2020;27:663–70.
- 36 Pivetta E, Goffi A, Tizzani M, *et al.* Lung ultrasonography for the diagnosis of SARS-cov-2 pneumonia in the emergency department. *Ann Emerg Med* 2021;77:385–94.
- 37 Schmid B, Feuerstein D, Lang CN, *et al.* Lung ultrasound in the emergency department-a valuable tool in the management of patients presenting with respiratory symptoms during the SARS-cov-2 pandemic. *BMC Emerg Med* 2020;20:96.
- 38 Tung-Chen Y, Algora-Martín A, Llamas-Fuentes R, *et al.* Point-of-care ultrasonography in the initial characterization of patients with COVID-19. *Med Clin (Barc)* 2021;156:477–84.
- 39 Gibbons RC, Magee M, Goett H, *et al.* Lung ultrasound vs. chest X-ray study for the radiographic diagnosis of COVID-19 pneumonia in a high-prevalence population. *J Emerg Med* 2021;60:615–25.
- 40 Karagöz A, Sağlam C, Demirbaş HB, *et al.* Accuracy of bedside lung ultrasound as a rapid triage tool for suspected covid-19 cases. *Ultrasound Q* 2020;36:339–44.
- 41 Speidel V, Conen A, Gisler V, *et al.* Lung assessment with point-of-care ultrasound in respiratory coronavirus disease (COVID-19): a prospective cohort study. *Ultrasound Med Biol* 2021;47:896–901.
- 42 Zanforlin A, Strapazzon G, Falk M, *et al.* Lung ultrasound in the emergency department for early identification of COVID-19 pneumonia. *Respiration* 2021;100:145–53.
- 43 Ai T, Yang Z, Hou H, *et al.* Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in china: a report of 1014 cases. *Radiology* 2020;296:E32–40.
- 44 Ong SWX, Chiew CJ, Ang LW, *et al.* n.d. Clinical and virological features of SARS-cov-2 variants of concern: a retrospective cohort study comparing B.1.1.7 (alpha), B.1.315 (beta), and B.1.617.2 (delta). *SSRN Journal*
- 45 Twohig KA, Nyberg T, Zaidi A, *et al.* Hospital admission and emergency care attendance risk for SARS-cov-2 delta (B.1.617.2) compared with alpha (B.1.1.7) variants of concern: a cohort study. *Lancet Infect Dis* 2022;22:35–42.
- 46 Duong BV, Larpruenrudee P, Fang T, *et al.* Is the SARS cov-2 omicron variant deadlier and more transmissible than delta variant? *Int J Environ Res Public Health* 2022;19:4586.
- 47 Brakohiapa EKK, Sarkodie BD, Botwe BO, *et al.* Comparing radiological presentations of first and second strains of COVID-19 infections in a low-resource country. *Heliyon* 2021;7:e07818.
- 48 Benchoufi M, Bokobza J, Chauvin A, *et al.* Lung injury in patients with or suspected COVID-19 : a comparison between lung ultrasound and chest CT-scanner severity assessments, an observational study. *Radiology and Imaging [Preprint]* 2020.
- 49 Laursen CB, Sloth E, Lassen AT, *et al.* Point-of-care ultrasonography in patients admitted with respiratory symptoms: a single-blind, randomised controlled trial. *Lancet Respir Med* 2014;2:638–46.