CJEM Journal Club

Validity of ultrasonography to diagnosing pneumothorax: a critical appraisal of two meta-analyses

Reviewed by: Aftab Azad, MD*; Saad Al Juma, MD*; Junaid A. Bhatti, PhD†; Jerrald Dankoff, MD*

Clinical question
Can ultrasonography be used in lieu of chest radiography to diagnose pneumothorax?
Articles chosen
1. Ding W, Shen Y, Yang J, et al. Diagnosis of pneumothorax by radiography and ultrasonography: a meta-analysis. *Chest* 2011;140:859-66. [Epub 2011 May 5]
2. Alrajhi K, Woo MY, Vaillancourt C. Test characteristics of ultrasonography for the detection of pneumothorax: a systematic review and meta-analysis. *Chest* 2012; 141:703-8.

Keywords: pneumothorax, sensitivity, ultrasonography

OBJECTIVE OF THE META-ANALYSES

The two meta-analyses aimed to compare the validity of ultrasonography (US) to that of chest radiography (CR) in the diagnosis of pneumothorax (PNX). There have been no published guidelines for the use of US in the diagnosis of PNX in challenging clinical situations (e.g., emergency trauma care or critically ill patients). These meta-analyses review the evidence supporting the superiority of US, suggesting that it is more feasible and radiation free compared to CR and computed tomography (CT).

BACKGROUND

PNX is a frequent diagnosis in emergency departments and intensive care units. CT has become the gold standard for diagnosing PNX, whereas CR is the conventional rapid method used even though studies have shown that CR has low sensitivity when compared to CT. The difficulty in accurately diagnosing PNX in unstable patients has led to the evaluation of alternative methods. US was first used to diagnose PNX in humans in 1987. The characteristic echographic signs commonly used in diagnosing PNX include lung sliding, comet tail artifacts, and the lung point. We present two recent meta-analyses from *Chest* that reviewed the evidence comparing the validity of US and CR in the diagnosis of PNX to the gold standards, either CT scan or air leak on chest tube insertion.

SUMMARY OF METHODS

Although both meta-analyses searched the literature using *MEDLINE* and *EMBASE*, Ding and colleagues also included the *Cochrane Library* database. The inclusion criteria of both reviews were similar, selecting studies that compared the validity of US and CR. The number of included articles varied because of additionally included studies (e.g., prospective studies in Alrajhi and colleagues) and different analytical approaches adopted in each meta-analysis. All articles included in Alrajhi and colleagues' study were also selected in Ding and colleagues'. Ding and colleagues computed pooled sensitivity, specificity, diagnostic odds ratio, and summary receiver operating characteristic (ROC) for the studies comparing US ($n = 15$) and CR ($n = 19$) to the gold standards. Ding and colleagues used Z-statistics based on Cochrane Q to compare the summary ROC curves of US and CR. On

From the *Emergency Medicine Residency Program and †Douglas Hospital Research Centre, McGill University, Montreal, QC.

Correspondence to: Dr. Aftab Azad, Department of Emergency Medicine, Residency Program, McGill Emergency Medicine Residency Program Office, Room A4.62, Royal Victoria Hospital, 687 Pine Avenue West, Montreal, QC H3A 1A1; aftab.azad@mail.mcgill.ca; aftab.azad@hotmail.com.

This article has been peer reviewed.

 CJEM 2015;17(2):199-201 DOI 10.2310/8000.2014.140698

the other hand, Alrajhi and colleagues computed weighted mean differences (WMDs) of the sensitivity and specificity of the studies, such that the findings of both US and CR were compared (*n* = 8). Both studies assessed heterogeneity (as measured by Higgin's I^2) adapted to their analyses. Using meta-regression, Ding and colleagues also explored the possible sources of heterogeneity in detail, for example, operators (radiologists and clinicians) or diagnostic criteria (absence of lung sliding and comet tail sign). Both studies reanalyzed their findings by subgroup. Ding and colleagues assessed the effect of operators other than a radiologist on pooled sensitivity and specificity, whereas Alrajhi and colleagues assessed the effect of trauma patients (*n* = 766) on these parameters. Both meta-analyses used QUADAS to assess the quality of included studies. Ding and colleagues used only 10 items of QUADAS to assess the quality of the studies. Neither of the meta-analyses excluded any study because of low scores on QUADAS.

SUMMARY OF RESULTS

Findings of both meta-analyses on the pooled sensitivity and specificity of US and CR (compared to the same gold standard) in diagnosing PNX were similar (Table 1). They showed that the sensitivity and specificity of US in detecting PNX are approximately 88 to 91% and 98 to 99%, respectively. The sensitivity of CR was low (50–52%), whereas its specificity approached 100%. The WMD analyses of Alrajhi and colleagues showed that US is significantly more sensitive, whereas CR is significantly more specific than US for diagnosing PNX. These differences did not appear to be clinically significant because there were no significant differences in summary ROC curves for both instruments when compared to the gold standards.

The heterogeneity observed in WMD analyses was high for sensitivity and low for specificity. Ding and colleagues also observed a high heterogeneity among studies. Ding and colleagues assessed whether pooled sensitivity and specificity of US differed in studies where operators were nonradiologists compared to overall findings and found no differences. Similarly, Alrajhi and colleagues reanalyzed the study data (*n* = 6) that included only trauma patients and found no differences compared to overall findings. Alrajhi and colleagues discussed some of the included studies that allowed the use of M-mode or power Doppler US, which have slightly different diagnostic criteria, such as seashore sign and power slide sign. They concluded that these effects could not be assessed because of data limitations.

Table 1. Summary of pooled sensitivity and specificity documented in Ding et al (2011) and Alrahjhi et al (2012)

	Ding et al		Alrajhi et al	
	Parameter	95% CI	Parameter	95% CI
Ultrasonography				
N	15		8	
Pooled sensitivity	0.88	0.85–0.91	0.91	0.87–0.94
Heterogeneity (I^2) in %	91.0			
Pooled specificity	0.99	0.98–0.99	0.98	0.97–0.99
Heterogeneity (I^2) in %	75.0			
Chest x-ray				
N	19		8	
Pooled sensitivity	0.52	0.49–0.55	0.50	0.44–0.57
Heterogeneity (I^2) in %	90.6			
Pooled specificity	1.00	1.00–1.00	0.99	0.98–1.00
Heterogeneity (I^2) in %	50.5			
Ultrasonography in specific cases where (sensitivity analyses)	It was performed by clinicians other than radiologist		Only trauma patients were considered	
N	8		6	
Pooled sensitivity	0.90	0.87–0.93	0.90	0.85–0.94
Heterogeneity (I^2) in %	91.2			
Pooled specificity	0.99	0.98–0.99	0.99	0.97–1.00
Heterogeneity (I^2) in %	73.8			

https://doi.org/10.2310/8000.2014.140698 Published online by Cambridge University Press

Table 2. Study inclusion and quality assessment of Ding et al (2011) and Alrajhi et al (2012)

	Ding et al	Alrajhi et al
Databases (*n*)	3	2
Studies included (*n*)	20	8
Studies analyzed	20	8
Overview Quality Assessment Questionnaire		
Reporting of search methods	Poor	Exemplary
Comprehensiveness of search methods	Acceptable	Acceptable
Reporting of inclusion/exclusion criteria	Poor	Good
Avoidance of selection bias	Poor	Acceptable
Reporting of validity criteria	Good	Good
Reporting of methods to combine studies	Good	Good
Assessment of validity	Good	Good
Appropriate combining of findings	Acceptable	Good
Conclusions supported by reported studies	Good	Good
Scientific quality	Acceptable	Good

COMMENTS

Table 2 includes the quality appraisal of both meta-analyses by using the 10-item criteria, the Overview Quality Assessment Questionnaire (OQAQ) by Oxman and Guyatt.[3] Each item is ranked from very poor to exemplary (scoring 1 to 7). We considered a score of 4 as acceptable. Both studies had sufficiently comprehensive search methods, searching two to three databases. Alrajhi and colleagues' reported methods were more detailed than those of Ding and colleagues, who stated that they selected 20 studies without mentioning the number of studies assessed. This affected the ranking of Ding and colleagues' study (poor) for the reporting of inclusion and exclusion criteria and the presence of selection bias. Both studies scored good (score 5) for reporting and assessment of validity, reporting of methods to combine studies, and the conclusions reported by the studies. We ranked the meta-analysis technique of Alrajhi and colleagues better than that of Ding and colleagues as the former used WMD, which included the visualization of effect sizes (for sensitivity and specificity) along with statistical significance. We ranked the overall scientific quality of Ding and colleagues as acceptable and that of Alrajhi and colleagues as good.

CONCLUSION

The two meta-analyses agreed that US can be used for diagnosing PNX in critically ill or trauma patients. Almost all studies made use of supine CR, which might have potential implications for generalizing the results to non–critically ill patients. These studies suggest that present guidelines should include US as an option for diagnosing PNX in critically ill patients. It has not been established what the clinical significance is of the 90% sensitivity of US compared to that of CT. The evidence reviewed supported the ability of operators other than radiologists to diagnose PNX with US. Nonetheless, prudence is necessary regarding training and quality control while implementing such protocols. Further work is necessary to assess the safety, cost-effectiveness, and timeliness of such initiatives.

Competing interests: None declared.

REFERENCES

1. Ding W, Shen Y, Yang J, et al. Diagnosis of pneumothorax by radiography and ultrasonography: a meta-analysis. *Chest* 2011; 140:859-66. [Epub 2011 May 5], doi:10.1378/chest.10-2946.
2. Alrajhi K, Woo MY, Vaillancourt C. Test characteristics of ultrasonography for the detection of pneumothorax: a systematic review and meta-analysis. *Chest* 2012;141:703-8, doi:10.1378/chest.11-0131.
3. Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. *J Clin Epidemiol* 1991;44:1271-8, doi:10.1016/0895-4356(91)90160-B.