Lung Ultrasonography: A Viable Alternative to Chest Radiography in Children with Suspected Pneumonia?

Lilliam Ambroggio, PhD, MPH[1,2,3], Heidi Sucharew, PhD[2,3], Mantosh S. Rattan, MD[3,4], Sara M. O'Hara, MD[3,4], Diane S. Babcock, MD[3,4], Caitlin Clohessy, BA[1], Mark C. Steinhoff, MD[3,5], Maurizio Macaluso, MD, DrPH[2,3], Samir S. Shah, MD, MSCE[1,3,6], and Brian D. Coley, MD[3,4]

Objective To determine the interrater reliability (IRR) of lung ultrasonography (LUS) and chest radiography (CXR) and evaluate the accuracy of LUS compared with CXR for detecting pediatric pneumonia compared with chest computed tomography (CT) scan.

Study design This was a prospective cohort study of children aged 3 months to 18 years with a CXR and LUS performed between May 1, 2012, and January 31, 2014 with or without a clinical diagnosis of pneumonia. Four pediatric radiologists blinded to clinical information reported findings for the CXR and LUS images. IRR was estimated for 50 LUS and CXR images. The main outcome was the finding from CT ordered clinically or the probability of the CT finding for patients clinically requiring CT. Two radiologists reviewed CT scans to determine an overall finding. Latent class analysis was used to evaluate the sensitivity and specificity for findings (eg, consolidation) for LUS and CXR compared with CT.

Results Of the 132 patients in the cohort, 36 (27%) had CT performed for a clinical reason. Pneumonia was clinically documented in 47 patients (36%). The IRR for lung consolidation was 0.55 (95% CI, 0.40-0.70) for LUS and 0.36 (95% CI, 0.21-0.51) for CXR. The sensitivity for detecting consolidation, interstitial disease, and pleural effusion was statistically similar for LUS and CXR compared with CT; however, specificity was higher for CXR. The negative predictive value was similar for CXR and LUS.

Conclusions LUS has a sufficiently high IRR for detection of consolidation. Compared with CT, LUS and CXR have similar sensitivity, but CXR is more specific for findings indicating pneumonia. *(J Pediatr 2016;176:93-8)*.

Community-acquired pneumonia is a leading cause of morbidity and mortality in children worldwide. Chest radiography (CXR) facilitates the diagnosis of pneumonia as well as pneumonia-related complications[1]; however, CXR is an imperfect diagnostic test for pneumonia. It exposes patients to ionizing radiation, and ill children with suspected pneumonia may receive multiple CXRs, posing a small increased risk of cancer later in life.[2-4]

Lung ultrasonography (LUS) has several potential advantages over CXR. Compared with CXR, LUS does not expose the child to ionizing radiation, when used to address specific diagnostic questions requires minimal training for the provider, and can be performed at the point of care.[5,6] Compared with chest computed tomography (CT) scans in adults, LUS correctly identified pneumonia in >90% of cases and correctly identified nonpneumonia cases in approximately 95% of cases.[7]

The objectives of this study were to determine the interrater reliability (IRR) of LUS and CXR and to determine whether chest LUS is as accurate as CXR for detecting disease and thus can provide a viable alternative for assessing pneumonia in children.

Methods

This prospective cohort study enrolled patients seeking medical care at Cincinnati Children's Hospital Medical Center (CCHMC) between May 1, 2012, and January 31, 2014. Patients' guardians gave written consent, and patients age ≥11 years gave assent to participate. This study was approved by the Institutional Review Board at CCHMC.

CCC	Complex chronic condition
CCHMC	Cincinnati Children's Hospital Medical Center
CT	Computed tomography
CXR	Chest radiography
IRR	Interrater reliability
LCM	Latent class model
LUS	Lung ultrasonography
WHO	World Health Organization

From the Divisions of [1]Hospital Medicine and [2]Biostatistics and Epidemiology, Cincinnati Children's Hospital Medical Center; Departments of [3]Pediatrics and [4]Radiology, University of Cincinnati College of Medicine; and the [5]Global Health Center and [6]Division of Infectious Diseases, Cincinnati Children's Hospital Medical Center, Cincinnati, OH

Funded by the Thrasher Research Fund and the Bureau of Health Professions, Health Resources and Services Administration, and the Department of Health and Human Services (NRSA T32HP10027-14-00 [to L.A.]). S.O. is on the Toshiba America medical advisory board for ultrasonography equipment, which had no involvement in the design, execution, or conduct of this study. The other authors declare no conflicts of interest.

http://dx.doi.org/10.1016/j.jpeds.2016.05.033

Patients aged 3 months to 18 years who had a CT scan ordered for a clinical reason (eg, pneumonia, tumor) or who were hospitalized with a respiratory diagnosis (ie, pneumonia, wheezing, asthma, bronchiolitis, pleural effusion, parapneumonic effusion) were eligible. Children aged <3 months were excluded, because CT for acute respiratory illness is rarely performed in this population. Children imaged by portable CXR were excluded, because often only a single view of the chest was obtained, which can be associated with less certainty for diagnosing pneumonia. Imaging studies obtained within 36 hours of each other for individual patients were entered, to capture a similar disease state within that patient.

The imaging studies were deidentified, assigned a research study number, and then placed into a dedicated research picture archiving and retrieval system for image interpretation. The unique study identification numbers were generated by a random number table, thereby eliminating the possibility of reviewers linking the LUS with the CXR of an individual patient. Additional demographic and clinical information (eg, age, sex, complex chronic condition [CCC],[8] length of stay, underlying condition of asthma) from the electronic medical record were obtained using a standard case report form and entered into a separate secure electronic database.

LUS and CXR

All study participants underwent LUS performed by 1 of 5 sonographers trained to perform the study protocol (**Appendix 1**; available at www.jpeds.com), using an Aplio XG ultrasound machine (Toshiba America Medical Systems, Tustin, California) with 2- to 6-MHz convex and 5- to 12-MHz linear array transducers. For smaller children, 4- to 10-MHz curved and 5- and 12-MHz linear array transducers were used. One radiologist (B.C.) provided a 1-hour training session, which included examples of a normal lung, presence of consolidation (with and without air bronchograms), interstitial disease, pleural effusion, and lung abscess. The LUS reporting tool was developed and modified by the group of radiologists using 40 retrospectively obtained LUS studies from patients not included in this study (**Appendix 2**; available at www.jpeds.com).

LUS findings were categorized as normal, lobar or patchy consolidation, interstitial disease, pleural effusion, or other. Standardized definitions were developed based on previous literature and agreement by the study radiologists (**Appendix 3**; available at www.jpeds.com). "Normal" was defined as the presence of A-lines, normal lung sliding, and the absence of other findings. "Interstitial disease" was defined as the presence of ≥3 B-lines per imaging field.[9] "Consolidation, either patchy or lobar" was defined as nonaerated lung with or without air bronchograms. "Pleural effusion" was defined as the presence of fluid >3 mm in width within the pleural space. "Other" was defined as an abnormality that did not clearly fit into any specified category (**Appendix 3**).

Anteroposterior and lateral chest radiographs were obtained for each patient, and were interpreted using the World Health Organization (WHO) classification scheme for radiographic pneumonia.[10] Main findings for the CXR were categorized as normal, lobar or patchy consolidation, interstitial disease, pleural effusion, or other. All images were reviewed by board-certified radiologists with a pediatric radiology certificate of added qualifications at CCHMC using a standardized reporting tool (**Appendix 2**).

The radiologists were not provided with any patient clinical information or physical findings when interpreting either the CXR or LUS.

Outcome Assessment

The primary outcome was the pulmonary CT diagnosis as independently classified by 2 radiologists as normal, consolidation, interstitial disease, pleural effusion, or other. There was complete agreement by both radiologists on all of the CT findings, and thus a third radiologist was not needed for additional review. For ethical reasons, CTs were obtained for patients only if ordered as part of their clinical care; therefore, for patients without a CT ordered clinically, the probability of the CT diagnosis with a finding of normal, consolidation, interstitial disease, pleural effusion, or other served as the outcome, with the probabilities obtained from a latent class model (LCM).

Statistical Analyses

Patient demographic data (eg, age) and clinical characteristics (eg, CCC) were described using median and IQR for continuous variables, which were non-normally distributed, and a count and percentage for categorical variables. All LUS and CXR images were interpreted by 2 of the 4 study radiologists. In addition, 50 of the LUS and CXR images were selected at random and read by all 4 radiologists to calculate the IRR.

The IRR was calculated using a free- rather than fixed-marginal multirater κ statistic for each binary imaging finding (eg, presence or absence of consolidation) for each type of imaging modality (ie, LUS and CXR). Fixed-marginal multirater κ statistics (eg, Fleiss κ) depend on the observed marginal prevalence and symmetry of the raters' findings. They are suitable when the marginal distributions of the findings are known to the raters beforehand (eg, number of patients with consolidation). The free-marginal multirater κ statistic is an appropriate measure of agreement when the number of findings is not known by the raters a priori, as in this study, such that the raters are not restricted in the number of findings that can be assigned, and the raters hold similar expertise.[11] We classified the strength of agreement measured by the κ statistic as poor (<0.0), slight (0.0-0.2), fair (0.2-0.4), moderate (0.4-0.6), substantial (0.6-0.8), or almost perfect (0.8-1.0).[12] We calculated the 95% CI for each κ estimate using a bootstrap resampling methodology.

CT was considered the gold standard for calculating estimates of sensitivity and specificity. LCMs are used predominately in the field of infectious disease diagnostics when a gold standard exists but is not available for all individuals.[13] We used 2 LCMs to estimate and compare the sensitivity