



Published in final edited form as:

Ultrasound Med Biol. 2023 May ; 49(5): 1194–1201. doi:10.1016/j.ultrasmedbio.2023.01.005.

Lung ultrasound protocol and quality control of image interpretation using an adjudication panel in the Household Air Pollution Intervention Network (HAPIN) trial

Suzanne M. Simkovich^{1,2,3,4}, Shakir Hossen^{1,2}, Eric D. McCollum^{5,6}, Ashley K. Toenjes⁷, John P. McCracken⁸, Lisa M. Thompson⁹, Adly Castañaza¹⁰, Anaite Diaz¹⁰, Ghislaine Rosa¹¹, Miles A. Kirby¹², Alexie Mukeshimana¹³, Rachel Myers⁷, Patricia M. Lenzen⁷, Rachel Craik¹⁴, Shirin Jabbarzadeh¹⁵, Lisa Elon¹⁵, Sarada S Garg¹⁶, Kalpana Balakrishnan¹⁶, Gurusamy Thangavel¹⁶, Jennifer L. Peel¹⁷, Thomas F. Clasen¹⁸, Victor G. Dávila-Román⁷, Aris T. Papageorghiou¹⁴, Lisa de las Fuentes⁷, William Checkley, MD, PhD^{1,2},

HAPIN Investigators

¹Division of Pulmonary and Critical Care, Johns Hopkins University School of Medicine, Baltimore, USA.

²Centre for Global Non-Communicable Disease Research and Training, School of Medicine, Johns Hopkins University, Baltimore, USA.

³ Division of Healthcare Delivery, MedStar Health Research Institute, Hyattsville, USA

Correspondence: William Checkley, MD, PhD, Division of Pulmonary and Critical Care, Johns Hopkins University, 1830 E. Monument Street Room 555, Baltimore, MD 21287, wcheckl1@jhmi.edu, Telephone: (443) 287-4587.

HAPIN Investigators: Vigneswari Aravindalochanan, Kalpana Balakrishnan, Dana Boyd Barr, Vanessa Burrowes, Devan Campbell, Julia McPeck Campbell, Eduardo Canuz, Adly Castañaza, Howard Chang, William Checkley, Yunyun Chen, Marilú Chiang, Maggie L. Clark, Thomas Clasen, Rachel Craik, Mary Crocker, Victor G. Dávila-Román, Lisa de las Fuentes, Oscar De León, Anaite Diaz-Artiga, Ephrem Dusabimana, Lisa Elon, Juan Gabriel Espinoza, Irma Sayury Pineda Fuentes, Sarada Garg, Dina Goodman, Savannah Gupton, Meghan Hardison, Stella Hartinger, Steven A. Harvey, Mayari Hengstermann, Phabiola Herrera, Shakir Hossen, Penelope Howards, Lindsay Jaacks, Shirin Jabbarzadeh, Michael A. Johnson, Abigail Jones, Katherine Kearns, Miles Kirby, Jacob Kremer, Margaret Laws, Patricia M. Lenzen, Jiawen Liao, Amy Lovvorn, Fiona Majorin, Eric McCollum, John P. McCracken, Rachel M. Meyers, J. Jaime Miranda, Erick Mollinedo, Lawrence Moulton, Krishnendu Mukhopadhyay, Laura Nicolaou, Luke Naeher, Abidan Nambajimana, Florian Ndagijimana, Azhar Nizam, Jean de Dieu Ntivuguruzwa, Aris Papageorghiou, Jennifer Peel, Ricardo Piedrahita, Ajay Pillarisetti, Naveen Puttaswamy, Elisa Puzolo, Ashlinn Quinn, Sarah Rajkumar, Usha Ramakrishnan, Davis Reardon, Ghislaine Rosa, Joshua Rosenthal, P. Barry Ryan, Zoe Sakas, Sankar Sambandam, Jeremy Sarnat, Suzanne Simkovich, Sheela Sinharoy, Kirk R. Smith, Kyle Steenland, Damien Swearing, Gurusamy Thangavel, Lisa M. Thompson, Ashley K. Toenjes, Lindsay Underhill, Jean Damascene Uwizeyimana, Viviane Valdes, Amit Verma, Lance Waller, Megan Warnock, Kendra Williams, Wenlu Ye, Bonnie Young.

HAPIN sonographers: Libny Yadira Monroy Alarcón, Adly Castañaza Gonzalez de Durante, Claudia López Ortega, Irma sayury Pineda, Maria Fernanda Gonzalez, Lakshminarayanan Sowrirajan, Shanthi Priya Paramanandam, K Shanmugavadivu, Sudharsanan V., Suresh Seshadri, Vigneswari Aravindalochanan, Danielle Mendoza Apaza, Adhemir Edwing Yupanqui Fredes, Mario Hanco Gómez, Ronald Apaza R, Juan Felix Persivale Calle, Elizabeth Quispe, Carlos Leon Ponce, Victor Villar Gonzales, Rebeca Andrade Salas, Jhon Elmer Herrera, Luzdelia Ramos Mamani, Yessica Lopez, Giovanna Quiza, Yadel Hinojosa, Madeluz Gomez Quispe, Gery Frisancho Parada, Danielle Isabel Mendoza Apaza, Luz Roxana Quispe Flores, Alexie Mukeshimana, Niyitegeka Francois Xavier, Grace Ufimana, Elie Tuzayisenga, Valens Nkurunziza.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Administrative: The trial is registered in clinicaltrials.gov (NCT02944682 Checkley, Clasen, Peel).

Conflict of Interest: There are no conflicts of interest by any of the authors.

- ⁴Division of Pulmonary and Critical Care Medicine, Georgetown University School of Medicine, Washington, USA
- ⁵Eudowood Division of Pediatric Respiratory Sciences, Department of Pediatrics, School of Medicine, Johns Hopkins University, Baltimore, USA.
- ⁶Global Program on Pediatric Respiratory Sciences, Department of Pediatrics, School of Medicine, Johns Hopkins University, Baltimore, USA.
- ⁷Cardiovascular Division, Department of Medicine, Washington University in St. Louis, St. Louis, USA.
- ⁸Global Health Institute, Department of Epidemiology and Biostatistics, College of Public Health, University of Georgia
- ⁹Nell Hodgson Woodruff School of Nursing, Emory University, Atlanta, USA.
- ¹⁰Center for Health Studies, Universidad del Valle de Guatemala, Guatemala City, Guatemala.
- ¹¹Department of Disease Control, London School of Hygiene and Tropical Medicine, London, United Kingdom.
- ¹²Department of Global Health and Population, Harvard T.H. Chan School of Public Health, Harvard University, Boston, USA
- ¹³Eagle Research Center, Kigali, Rwanda.
- ¹⁴Nuffield Department of Women's and Reproductive Health, University of Oxford, Oxford, UK
- ¹⁵Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University USA
- ¹⁶Department of Environmental Health Engineering, ICMR Center for Advanced Research on Air Quality, Climate and Health, Sri Ramachandra Institute for Higher Education and Research (Deemed University), Chennai, India
- ¹⁷Department of Environmental and Radiological Health Sciences, Colorado State University, Fort Collins CO, USA.
- ¹⁸Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, USA.

Abstract

Objective: Lung ultrasound (LUS) is an alternative to chest radiography to confirm a diagnosis of pneumonia. For research and disease surveillance, methods to use LUS to diagnose pneumonia are needed.

Methods: In the Household Air Pollution Intervention Network (HAPIN) trial, LUS was used to confirm a clinical diagnosis of severe pneumonia in infants. We developed a standardized definition of pneumonia, protocols for recruitment and training of sonographers, along with LUS image acquisition and interpretation. We use a blinded panel approach to interpretation with LUS cine-loops randomized to non-scanning sonographers with expert review.

Discussion: We obtained 357 lung ultrasound scans: 159, 8 and 190 scans were collected in Guatemala, Peru and Rwanda, respectively. The diagnosis of primary endpoint pneumonia (PEP) required an expert tie breaker in 181 scans (39%). PEP was diagnosed in 141 scans (40%), not diagnosed in 213 (60%), with 3 scans (<1%) deemed uninterpretable. Agreement among the two blinded sonographers and the expert reader in Guatemala, Peru and Rwanda was 65%, 62%, and 67%, with a prevalence-and-bias-corrected kappa of 0.30, 0.24 and 0.33, respectively.

Conclusion: In the HAPIN study, the use of standardized imaging protocols, training and an adjudication panel resulted in high confidence for the diagnosis of pneumonia using LUS.

Keywords

Respiratory Tract Infections; Infant; Developing Countries; Ultrasound; Pneumonia; Imaging; research protocol

INTRODUCTION

Pneumonia remains the leading infectious cause of death in children under the age of five in low- and middle-income countries (LMICs), despite a significant reduction in the number of deaths from pneumonia over the past fifteen years.¹ In LMICs, clinical definitions based on signs and symptoms are used to make a diagnosis of pneumonia, with the World Health Organization Integrated Management of Childhood Illness (WHO IMCI) definition being the most commonly used.² Because a clinical definition is used primarily for case management rather than surveillance, the algorithm intentionally favors sensitivity over specificity, resulting in a high proportion of false positive cases.^{2,3} Imaging of children with clinical pneumonia could improve specificity in confirming a diagnosis.³ Chest radiography is currently the most widely used chest imaging technique for clinical purposes and is the gold standard imaging technique used in epidemiologic studies and vaccine trials. Standardized definitions of pneumonia by chest radiography interpretation are well validated.⁴ Although widely accepted, radiography equipment requires dedicated personnel, regular maintenance, and adequate infrastructure to support its use.⁴ Cost and lack of portability of radiography equipment also limits access for clinical practice and in research studies in LMICs.^{4,5} Furthermore, radiography exposes children to carcinogenic ionizing radiation.⁶

Lung ultrasound (LUS) is an alternative chest imaging modality that offers improved, sensitivity and specificity compared to chest radiography in detecting pneumonia in neonates and infants.⁷⁻¹³ Modern ultrasound systems are portable, less expensive than radiography equipment, do not emit radiation, and clinicians without specialized training can be taught to perform ultrasound to facilitate a rapid diagnosis.¹⁴⁻¹⁶ For research purposes and disease surveillance, standardized methods to utilize LUS need to be established. In the Household Air Pollution Intervention Network (HAPIN) trial, LUS was used as a method to confirm a clinical diagnosis of severe pneumonia in children less than one year of age.¹⁷ In this manuscript, we describe the protocols and procedures for utilizing LUS in our trial and the results of a panel interpretation approach for LUS. We present the quality control assessments used during the trial to develop strengthened protocols and training for future use of LUS in research studies or disease surveillance.

MATERIALS AND METHODS

Study design and setting

The HAPIN trial is a randomized controlled trial of a liquefied petroleum gas (LPG) stove and continuous, free fuel distribution intervention combined with behavioral change reinforcement to encourage LPG use in 3,200 households across four LMICs: Guatemala, India, Peru, and Rwanda. Each International Research Centre (IRC) recruited 800 pregnant women (18–34 years of age, 9–19 weeks gestation) and followed infants born to these women for their first year of life to determine the incidence of severe pneumonia. Full details of the main trial are described elsewhere.^{18,19} As detailed elsewhere, the HAPIN trial created a protocol and uses an adaptation of the most recent WHO IMCI case definition of severe pneumonia.¹⁷ As shown in Table 1, one component of our definition of pneumonia is a positive finding on LUS, which we refer to as primary endpoint pneumonia (PEP).¹⁷ To determine if PEP was present, local clinicians were trained in Guatemala, Peru, and Rwanda to perform a LUS when a suspected case of pneumonia presented to a hospital or health center where surveillance was being performed. Due to local regulations limiting ultrasound use, it was not feasible to perform LUS in India; chest radiography was used for imaging confirmation.¹⁷

Performing an ultrasound

A LUS was performed on all participating children aged 12 months who presented to a hospital or health center with suspected pneumonia. Ultrasound equipment was provided to each site (SonoSite Edge, Sonosite-FujiFilm, Bothell, QA). When a participant child presented with suspected pneumonia, a trial certified sonographer performed a LUS on the child within twenty-four hours of the initial assessment by the local field team.¹⁷ To meet the definition of PEP, the scan had to meet minimal quality standards of high or medium quality and have findings consistent with the definition of pneumonia (Table 1 and Figure 1).¹⁶

Storage and interpretation of scans

We used TRICE (TRICE Imaging, Del Mar, California, USA) to store and read images.²⁰ TRICE customized features for storage and blinded interpretation of LUS videos for our trial. Specifically, TRICE developed a dashboard that allowed a coordinator to assign and notify sonographers to read images. This feature in TRICE was important to our interpretation strategy, which required each LUS video to be interpreted by up to four different sonographers (scanning sonographer, Reader A, Reader B, and expert tie-breaker sonographer). Still images captured from ultrasound cine-loops were used to measure the largest vertical and horizontal dimensions of a consolidation. Annotations were saved on the system but could only be seen by the sonographer who placed the annotations to maintain blinding of the interpretation to other sonographers.

Sonographer training

We recruited local doctors and nurses at each IRC to become certified sonographers in LUS. We sought to have three sonographers at each IRC. The experience of these providers

ranged from no experience to prior experience with ultrasound. Sonographers underwent a standardized training process to become certified and scan participants for the trial.¹⁶ First, each trainee completed web-based didactic sessions and later attended a centralized training workshop at Washington University School of Medicine in St. Louis, Missouri, USA. Details of this workshop are described elsewhere.²⁰ An expert sonographer, who had significant experience in performing LUS in field trials, then traveled to each IRC and provided additional didactic and hands-on training. Local sonographers were asked to complete a series of scans and supervised interpretations and had to meet quality standards to become certified. If a certified sonographer was unable to continue, then another sonographer was recruited and trained as a replacement. The training of the replacement sonographer occurred via a virtual platform secondary to the COVID-19 pandemic.²¹ Details of training are available in the Online Supplement. A total of 17 sonographers (5 from Guatemala, 7 from Peru, and 5 from Rwanda) were trained during the HAPIN trial of which 14 were trained in person and three were trained virtually. The number of sonographers varied during the trial.

Image interpretation

We used a panel approach to interpret and adjudicate all LUS images. The design of the panel accomplished two objectives (Figure 2): panel interpretation of LUS to reach consensus on a diagnosis; and continuous preliminary reporting on quality and interpretation of each scan. Scanning sonographers uploaded the LUS images to TRICE and performed the initial quality control and interpretation. This was done to ensure that the required 24 LUS cine-loops were obtained during scanning. Scanning sonographers were asked to provide an initial interpretation to maintain interpretive skills. While the imaging sonographer's interpretation was valuable as a quality assurance strategy to improve image quality and to identify errors in image acquisition, labeling, or upload, their interpretation was not considered in the final adjudication of PEP, as they were not blinded to the clinical presentation of the child or randomization in the trial.

All certified sonographers were pooled to serve as panelists for adjudication. When a scan was uploaded to the system, the coordinator copied the cine-loop in its original form and assigned two sonographers (Reader A and Reader B) who were blinded to clinical presentation to interpret the cine-loop. Reader A and Reader B were not from the same IRC as the scan was conducted to minimize bias. Reader A and Reader B were titles assigned for tracking to ensure each image received two different interpreting sonographers. Panelists were blinded to the clinical presentation of the child, randomization in the trial, and interpretation from other sonographers. If Readers A and B agreed on the diagnosis, then the diagnosis was final. If they disagreed, then an expert sonographer — who was also blinded to the clinical presentation, randomization assignment and to the interpretation of other sonographers — was used as a tiebreaker to reach agreement. The expert was a physician who was previously trained by our team for another study and had interpreted more than 5,000 LUS cine-loops.

To provide monthly feedback to the IRCs and funders on progress in data collection, all scans were interpreted by the expert sonographer. This served three purposes: it provided a

preliminary interpretation of the scan to give feedback to IRCs; it was used as tie breaker if the two sonographers disagreed on the diagnosis; and the expert sonographer read was used in quality assurance.

Quality Control of Image Collection and Interpretation

Interpretations by the sonographers were compared to the expert sonographer on a quarterly basis by investigators. If a trend of poor agreement was noticed (agreement <60%) then the sonographer was asked to stop interpreting images and undergo retraining with the expert sonographer. The sonographer and the expert assessed the areas of error in conduct and interpretation of scans and developed a retraining plan to strengthen the standardization of interpretation of scans.

Statistical methods

We calculated agreement between our interpreting sonographers (Reader A and Reader B) and expert (Reader A and expert, Reader B and expert) after categorizing the image as high or medium quality and PEP as present or absent. To account for agreement by chance, we calculated the unadjusted Cohen's kappa (κ) statistic to measure strength of agreement between the sonographer and expert. We used a kappa statistic, both unadjusted and adjusted for prevalence and bias (PABAK) to estimate reader agreement not expected by chance.²² If an interpreting sonographer performed worse than chance, then the panel interpretations by this sonographer were not used and the expert replaced the interpreting sonographer. All analyses were conducted using R software version 4.1.2 (Bird Hippie) using the packages tidyverse and epiR.

Ethical Approval and Dissemination: The study protocol has been reviewed and approved by institutional review boards (IRBs) or Ethics Committees at Emory University (00089799), Johns Hopkins University (00007403), Sri Ramachandra Institute of Higher Education and Research (IEC-N1/16/JUL/54/49) and the Indian Council of Medical Research – Health Ministry Screening Committee (5/8/4–30/(Env)/Indo-US/2016-NCD-I), Universidad del Valle de Guatemala (146–08-2016/11–2016) and Guatemalan Ministry of Health National Ethics Committee (11–2016), A.B. PRISMA, the London School of Hygiene and Tropical Medicine (11664–5) and the Rwandan National Ethics Committee (No.357/RNEC/2018), and Washington University in St. Louis (201611159). The study results will be disseminated to the appropriate stakeholders through presentations, conferences, and peer-reviewed journals.

RESULTS

Conduct, randomization of scans and identification of PEP on lung ultrasound

A total of 357 LUS scans were conducted between January 17, 2019, and April 24, 2021 (Figure 3). Rwanda conducted 190 scans, Guatemala conducted 159, and Peru conducted eight. Reader A and Reader B were randomly assigned to interpret these scans for a total of 714 interpretations. The expert reader interpreted all 357 scans. The interpretations for one sonographer were removed from the panel because of poor performance in interpreting within the panel without improvement after retraining (n=82). She conducted one scan of a

child during the trial. For these 82 scans, we used Reader's A (or B) interpretation and the expert sonographer.

Reader A and Reader B agreed on PEP for 67 scans and agreed on absence of PEP on 106 scans. Three scans (<1%) were uninterpretable and excluded. PEP was identified in 141 (40%) scans whereas PEP was not found in 213 (60%) scans after adjudication. A total of 98 scans met PEP criteria through the identification of consolidation defined as >1 intercostal space in size or 1 cm in width or height. The remaining scans met PEP through more than one criterion, such as having a consolidation that was >1 intercostal space in size or 1 cm in width or height and pleural effusion and presence of 3 B-lines (Table 2).

Identification of sonographer errors

By comparing reads of the panel sonographers to the reads of the expert, key trends were found where sonographers had difficulty identifying pathology. There was >78% agreement in findings with the expert for identifying air or fluid bronchograms and shred sign in all 12 regions of the lung. The agreement between sonographers and the expert on identification of consolidation occurred in >71% of scans. Sonographers had difficulty identifying 3 B-lines present (agreement 46–56%) and pleural abnormalities (25–35%). These trends occurred throughout all scans and were not specific to a particular lung field. See E-Tables 1 and 2 for further details.

Agreement among ultrasound interpretation panel sonographers with Ultrasound Core Lab expert sonographer

Inter-reader agreement among Readers A and B, and Reader A or B and the expert was 62–64% with an unadjusted kappa 0.23–0.27 and a PABAK of 0.24–0.27 (Table 3). Sonographers in Rwanda had the highest agreement with the expert reader (66.7%) with the highest kappa (0.33) and PABAK (0.33) along with the highest volume of scans conducted (n=190). Peru sonographers had the lowest agreement (62.2%), kappa (0.26) and PABAK (0.24) along with the lowest volume of scans conducted (n=8). Individual sonographer's agreement with the expert was directly associated with volumes of scans interpreted (E-Table 3).

DISCUSSION

We developed a comprehensive study protocol for the use of LUS for disease surveillance and research that included both quality control and panel adjudication for interpretation. Our protocol and panel approach to interpretation were successful in confirming a diagnosis of severe pneumonia in children as part of the HAPIN trial.¹⁷ Through the analysis of quality control data, we found that sonographer performance was associated with the volume of scans conducted and interpreted. This was further seen that all of our sonographers needed to undergo retraining following a gap in conducting scans during the COVID-19 lockdowns. Furthermore, through evaluating performance of sonographers, we identified key areas to strengthen during training programs and provide recommendations for using LUS in field trials.

Our findings suggest that performance by our LUS panel was associated with the number of scans that were conducted by each center.¹⁶ There is a dearth of literature on the use of LUS in field trials for children, thus not allowing us to compare performance in our studies with others. Drawing on the performance of the use of chest x-ray in children in field trials or observational studies, our agreement and kappa statistics are similar or lower than that of reading panels utilized in trials assessing chest x-ray agreement. The Pneumonia Etiology Research for Child Health (PERCH), a case-control study on child pneumonia etiology from seven LMIC settings utilized a reading panel to interpret chest x-rays to confirm findings of pneumonia. PERCH found readers agreed on the presence or absence of PEP in 77.8% of 3497 interpretable images with an unadjusted kappa 0.50 and a prevalence and bias-adjusted kappa of 0.56.²³ In a study of WHO-defined community acquired pneumonia in five Indian states, resulting in 2,829 interpretable x-rays, agreement was 86% and kappa ranged from 0.31–0.46 among the reading panel.²⁴ In a study to evaluate vaccine effectiveness of the pneumococcal 10-valent conjugate vaccine rollout in Bangladesh, interobserver agreement for 9,723 images was 79.0% with an unadjusted kappa 0.35 and a PABAK of 0.58.²⁵ The same study in Bangladesh obtained 9,051 LUS on children with clinical pneumonia and a panel interpretation of the images showed an agreement of 91% and a kappa of 0.86.¹⁶ Our sonographer panel performed with 61.2% agreement with a kappa 0.23 and a PABAK or 0.24. We suspect the performance of our panel is not as strong for LUS compared to other studies given the lower volumes of scans in our study.²⁶ Volume of children presenting for clinical evaluation was likely lower secondary to the COVID-19 pandemic. Moreover, LUS use in field trials is early in development. Protocols and training for the use of LUS need to be informed by data, such as ours, to strengthen the interpretation approach, quality assessment and control.

Through this process, we found key areas for improvement when using LUS in future studies or trials. Training for LUS was more feasibly conducted in-person than virtually. As the COVID-19 pandemic spread, we had to transition from in-person training to a virtual platform. This created difficulty with scheduling training sessions secondary to time zone differences, connectivity issues in LMIC settings, and then a translator needed to be scheduled concurrently to assist with the training sessions. We attribute this factor to the poor performance of one of our sonographers, whose interpretations had to be excluded, as we did not feel as though the virtual training platform was ideal for this sonographer. We recommend in-person training when feasible, but this is more costly than virtual training. Based on our review of the scans, enhanced training on identifying pleural abnormalities and B-lines is needed to strengthen sonographer skills. We recommend having sonographers conduct a higher volume of scans and interpret scans regularly without a prolonged delay between scans. A higher volume of scans keeps the sonographers engaged, although may contribute to additional study costs. One alternative, to keep costs low, is to create an open-source bank of scans be created for routine interpretation by sonographers to ensure proficiency.

To set up an approach that minimizes bias and maximizes information, protocols and procedures were closely followed during the conduct of our trial. The standardized protocols we developed can be applied for the field use of LUS for other randomized controlled trials, observational trials, and routine disease surveillance in children and adults. Studies have

shown that based on the ease of use of LUS, local clinicians were easily trained to perform LUS in a standardized fashion. In our study, we found that remote training using a virtual platform posed challenges that may have resulted in less effective training compared to in-person training. As new lower-cost, portable ultrasound systems are introduced and scaled up, the routine use of ultrasound in low-resource settings will become more feasible. Further scaling-up quality assessment and control procedures will be necessary to ensure high quality clinical assessments are being conducted. Given the ongoing COVID-19 pandemic and the threat of future outbreaks of respiratory infections, there is an urgent need for integrating resources that offer rapid confirmation of a clinical diagnosis of pneumonia. The HAPIN trial training program and protocols may serve as a resource to build local capacity in image acquisition and interpretation. The use of an adjudication panel for interpretation lessens the concerns around inter-rater variability through reaching consensus.

Our strategy for interpretation has several strengths. The expert sonographer provided a preliminary interpretation for feedback and then served as an arbitrator when there were disagreements between sonographers. Furthermore, the expert sonographer provided quality control checks and timely feedback to sonographers. One of the costs of this approach is that the expert sonographer needed to review and interpret all images. We believed the benefits outweighed the costs as each scan took approximately ten minutes to interpret. Having two sonographers, each blinded to the clinical presentation, that agree on interpretation minimizes bias and improves the quality of the information the ultrasound provides. Customizing a cloud-based secured software with tools for image tracking, measuring, and labelling videos created ease in the data analysis process. Training local clinicians to perform LUS builds local capacity. We hope that the training these health care professionals have received will be carried forward to their clinical practice to diagnose and triage patients with pneumonia in local health facilities where ultrasound is available.

However, there are potential limitations to our approach. First, the coordination of three sonographers at each international center and an expert reader demanded substantial time and financial resources. Our approach also required an expert reader to adjudicate cases. Second, use of ultrasound may not be possible in all sites secondary to legal regulation. Third, we also found that performance of conducting and interpreting scans was based on the volume of scans obtained. Indeed, agreement and inter-rater variability was higher in Guatemala and Rwanda when compared to Peru. Based on this finding, we recommend that sonographers conduct scans more routinely to keep up their acquisition and interpretation skills. Fourth, due to the COVID-19 pandemic, in-person training and retraining was not always possible. Instead, we relied on virtual platforms that created barriers for scheduling mutual times for training given different time zones, connection difficulties in LMIC settings, difficulty with language barriers, and lack of hands-on training with live-subjects. This resulted in poor performance by one sonographer and inability to utilize this sonographer's interpretations. However, this experience presents itself as an opportunity to understand the facilitators and barriers to online training for LUS and could potentially be a strategy for training in low resource settings with refinement. Further refinement of approaches for virtual training needs to be conducted to make this component more successful.

CONCLUSION

Portable LUS is emerging as a new imaging modality for diagnosing pneumonia in LMICs. Standardized protocols for conducting field-based research studies, such as the one we present in this manuscript, are needed to evaluate the use of this technology in confirming a clinical diagnosis of pneumonia. In the HAPIN trial, we successfully utilized LUS to confirm a diagnosis of clinical pneumonia in children for research purposes with high confidence through the use of standardized imaging protocols, training and adjudication. The volume of scans acquired and interpreted affects performance; as such, we recommend sonographers regularly acquire and interpret scans to maintain skill level. We recommend a strategy that includes intensive in-person training with continued practice, an adjudication panel, and interpretation of images by an expert to provide quality control and arbitration in interpretation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We would like to thank our pneumonia experts who participated in our workshops: Prof. Heather Zar of the University of Cape Town (Cape Town, South Africa), Prof. Harry Campbell of the University of Edinburgh (Edinburgh, United Kingdom), Prof Claudio Lanata of the Instituto De Investigación Nutricional (Lima, Peru) and Dr. Carina King (Karolinska Institute, Sweden), and Dr. Laura Hammitt (Johns Hopkins University, USA). We additionally would like to thank Katerina Lescouflair MSPH and Delaney Connolly both of Johns Hopkins University (Baltimore, Maryland) for assistance with preparation of the protocols and development of materials. A multidisciplinary, independent Data and Safety Monitoring Board (DSMB) appointed by the National Heart, Lung, and Blood Institute (NHLBI) monitors the quality of the data and protects the safety of patients enrolled in the HAPIN trial. NHLBI DSMB: Nancy R. Cook, Sc.D.; Stephen Hecht, Ph.D.; Catherine Karr, M.D., Ph.D.; Katie H. Kavounis, M.P.H.; Dong-Yun Kim, Ph.D.; Joseph Millum, Ph.D.; Lora A. Reineck, M.D., M.S.; Nalini Sathiakumar, M.D., Dr.P.H.; Paul K. Whelton, M.D.; Gail G. Weinmann, M.D.

Program Coordination: Gail Rodgers, M.D., Bill & Melinda Gates Foundation; Claudia L. Thompson, Ph.D., National Institute of Environmental Health Science (NIEHS); Mark J. Parascandola, Ph.D., M.P.H., National Cancer Institute (NCI); Danuta M. Krotoski, Ph.D., Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD); Joshua P. Rosenthal, Ph.D., Fogarty International Center (FIC); Conception R. Nierras, Ph.D., NIH Office of Strategic Coordination Common Fund; Antonello Punturieri, M.D., Ph.D. and Barry S. Schmetter, B.S., National Heart, Lung, and Blood Institute (NHLBI).

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the U.S. National Institutes of Health or Department of Health and Human Services.

Funding:

This study is funded by the U.S. National Institutes of Health (cooperative agreement 1UM1HL134590) in collaboration with the Bill & Melinda Gates Foundation (OPP1131279). Participating NIH organizations include the National Heart, Lung and Blood Institute, National Institute of Environmental Health Sciences, National Cancer Institute, National Institute of Child Health and Human Development, Fogarty International Center, and the NIH Common Fund. Suzanne M. Simkovich was supported by funding from the National Heart, Lung and Blood Institute T32 HL007534–36, National Heart, Lung, and Blood Institute 1F32HL143909–01, and National Heart, Lung, and Blood Institute K12HL137942.

Data Sharing Statement:

A de-identified data set will be available upon request to the authors.

References

1. McAllister DA, Liu L, Shi T, et al. Global, regional, and national estimates of pneumonia morbidity and mortality in children younger than 5 years between 2000 and 2015: a systematic analysis. *Lancet Glob Health*. 2019;7(1):e47–e57. [PubMed: 30497986]
2. Organization WH. Pocket book of hospital care for children: Second edition. Geneva: World Health Organization 2013.
3. Goodman D, Crocker ME, Pervaiz F, et al. Challenges in the diagnosis of paediatric pneumonia in intervention field trials: recommendations from a pneumonia field trial working group. *Lancet Resp Med*. 2019; 7(12):1068–108.
4. Mahomed N, Fancourt N, de Campo J, et al. Preliminary report from the World Health Organisation Chest Radiography in Epidemiological Studies project. *Pediatr Radiol*. 2017;47(11):1399–1404. [PubMed: 29043423]
5. Smith KR, McCracken JP, Weber MW, et al. Effect of reduction in household air pollution on childhood pneumonia in Guatemala (RESPIRE): a randomised controlled trial. *Lancet*. 2011;378(9804):1717–1726. [PubMed: 22078686]
6. Komatsu CV, Silva CC, Souza LRMFd, Gonçalves LF. Excess Radiation to Newborns Hospitalized in the Intensive Care Unit. *Radiat Prot Dosimetry*. 2017;177(3):331–341. [PubMed: 28444292]
7. Ellington LE, Gilman RH, Chavez MA, et al. Lung ultrasound as a diagnostic tool for radiographically-confirmed pneumonia in low resource settings. *Respir Med*. 2017;128:57–64. [PubMed: 28610670]
8. Iorio G, Capasso M, Prisco S, et al. Lung Ultrasound Findings Undetectable by Chest Radiography in Children with Community-Acquired Pneumonia. *Ultrasound Med Bio*. 2018;44(8):1687–1693. [PubMed: 29759424]
9. Miller LE, Stoller JZ, Fraga MV. Point-of-care ultrasound in the neonatal ICU. *Curr Opin Pediatrics*. 2020;32(2):216–227.
10. Orso D, Ban A, Guglielmo N. Lung ultrasound in diagnosing pneumonia in childhood: a systematic review and meta-analysis. *Journal of ultrasound*. 2018;21(3):183–195. [PubMed: 29931473]
11. Balk DS, Lee C, Schafer J, et al. Lung ultrasound compared to chest X-ray for diagnosis of pediatric pneumonia: A meta-analysis. *Pediatr Pulmonol*. 2018;53(8):1130–1139. [PubMed: 29696826]
12. Pereda MA, Chavez MA, Hooper-Miele CC, et al. Lung ultrasound for the diagnosis of pneumonia in children: a meta-analysis. *Pediatrics*. 2015;135(4):714–722. [PubMed: 25780071]
13. Tsou PY, Chen KP, Wang YH, et al. Diagnostic Accuracy of Lung Ultrasound Performed by Novice Versus Advanced Sonographers for Pneumonia in Children: A Systematic Review and Meta-analysis. *Acad Emerg Med*. 2019;26(9):1074–1088. [PubMed: 31211896]
14. Marini TJ, Castaneda B, Baran T, et al. Lung Ultrasound Volume Sweep Imaging for Pneumonia Detection in Rural Areas: Piloting Training in Rural Peru. *Clin Imaging Sci*. 2019;9:35.
15. Nadimpalli A, Tsung JW, Sanchez R, et al. Feasibility of Training Clinical Officers in Point-of-Care Ultrasound for Pediatric Respiratory Diseases in Aweil, South Sudan. *Am J Trop MedHyg*. 2019;101(3):689–695.
16. Pervaiz F, Hossen S, Chavez MA, et al. Training and standardization of general practitioners in the use of lung ultrasound for the diagnosis of pediatric pneumonia. *Pediat Pulmonol*. 2019;54(11):1753–1759.
17. Simkovich SM, Underhill LJ, Kirby MA, et al. Design and conduct of facility-based surveillance for severe childhood pneumonia in the Household Air Pollution Intervention Network (HAPIN) trial. *ERJ Open Res*. 2020;6(1).
18. Barr DB, Puttaswamy N, Jaacks LM, et al. Design and Rationale of the Biomarker Center of the Household Air Pollution Intervention Network (HAPIN) Trial. *Environ Health Perspect*. 2020;128(4):47010. [PubMed: 32347765]
19. Clasen T, Checkley W, Peel JL, et al. Design and Rationale of the HAPIN Study: A Multicountry Randomized Controlled Trial to Assess the Effect of Liquefied Petroleum Gas Stove and Continuous Fuel Distribution. *Environ Health Perspect*. 2020;128(4):47008. [PubMed: 32347766]

20. Dávila-Román VG, Toenjes AK, Meyers RM, et al. Ultrasound Core Laboratory for the Household Air Pollution Intervention Network Trial: Standardized Training and Image Management for Field Studies Using Portable Ultrasound in Fetal, Lung, and Vascular Evaluations. *Ultrasound Med Biol.* 2021;47(6):1506–1513. [PubMed: 33812692]
21. Simkovich SM, Thompson LM, Clark ML, et al. A risk assessment tool for resumption of research activities during the COVID-19 pandemic for field trials in low resource settings. *BMC Med Res Methodol.* 2021;21(1):68. [PubMed: 33845785]
22. Sim J, Wright CC. The Kappa Statistic in Reliability Studies: Use, Interpretation, and Sample Size Requirements. *Phys Ther.* 2005;85(3):257–268. [PubMed: 15733050]
23. Fancourt N, Deloria Knoll M, Barger-Kamate B, et al. Standardized Interpretation of Chest Radiographs in Cases of Pediatric Pneumonia From the PERCH Study. *Clin Infect Dis.* 2017;64(suppl_3):S253–s261. [PubMed: 28575359]
24. McCollum ED, Ahmed S, Chowdhury NH, et al. Chest radiograph reading panel performance in a Bangladesh pneumococcal vaccine effectiveness study. *BMJ Open Respir Res.* 2019;6(1):e000393.
25. Awasthi S, Rastogi T, Mishra N, et al. Chest radiograph findings in children aged 2–59 months hospitalised with community-acquired pneumonia, prior to the introduction of pneumococcal conjugate vaccine in India: a prospective multisite observational study. *BMJ Open.* 2020;10(5):e034066.
26. Gray DM, Davies MA, Githinji L, et al. COVID-19 and Pediatric Lung Disease: A South African Tertiary Center Experience. *Front Pediatr.* 2020;8:614076. [PubMed: 33553073]

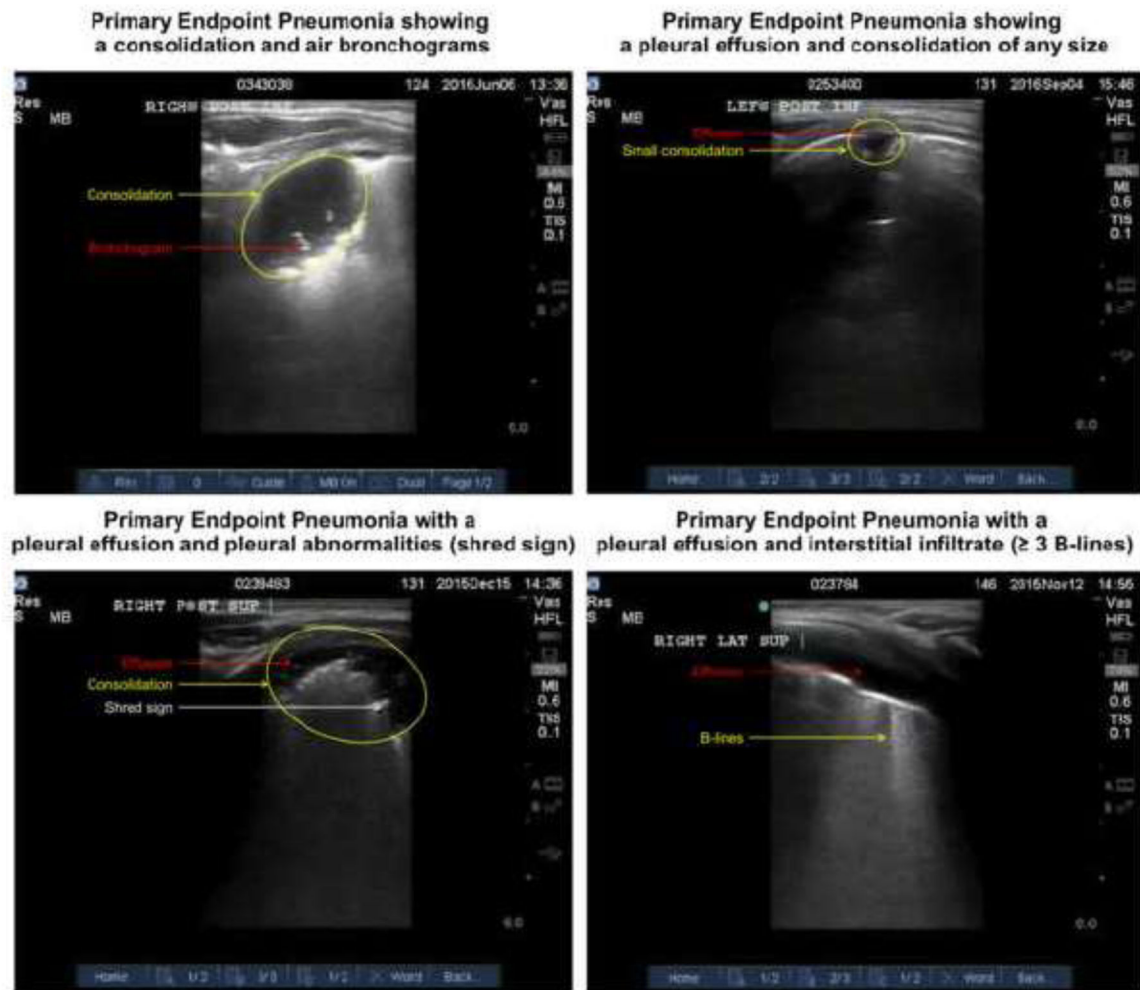


Figure 1. Examples of Primary Endpoint Pneumonia (PEP).

Left Upper Corner: This image shows the presence of a consolidation with air bronchograms indicating PEP. Right Upper Corner: This image shows the presence of a pleural effusion associated with a consolidation indicating PEP. Left Lower Corner: This image shows a pleural effusion with pleural abnormalities (i.e. shred sign). Right Lower Corner: This image shows a pleural effusion with interstitial infiltrates defined by B-lines.

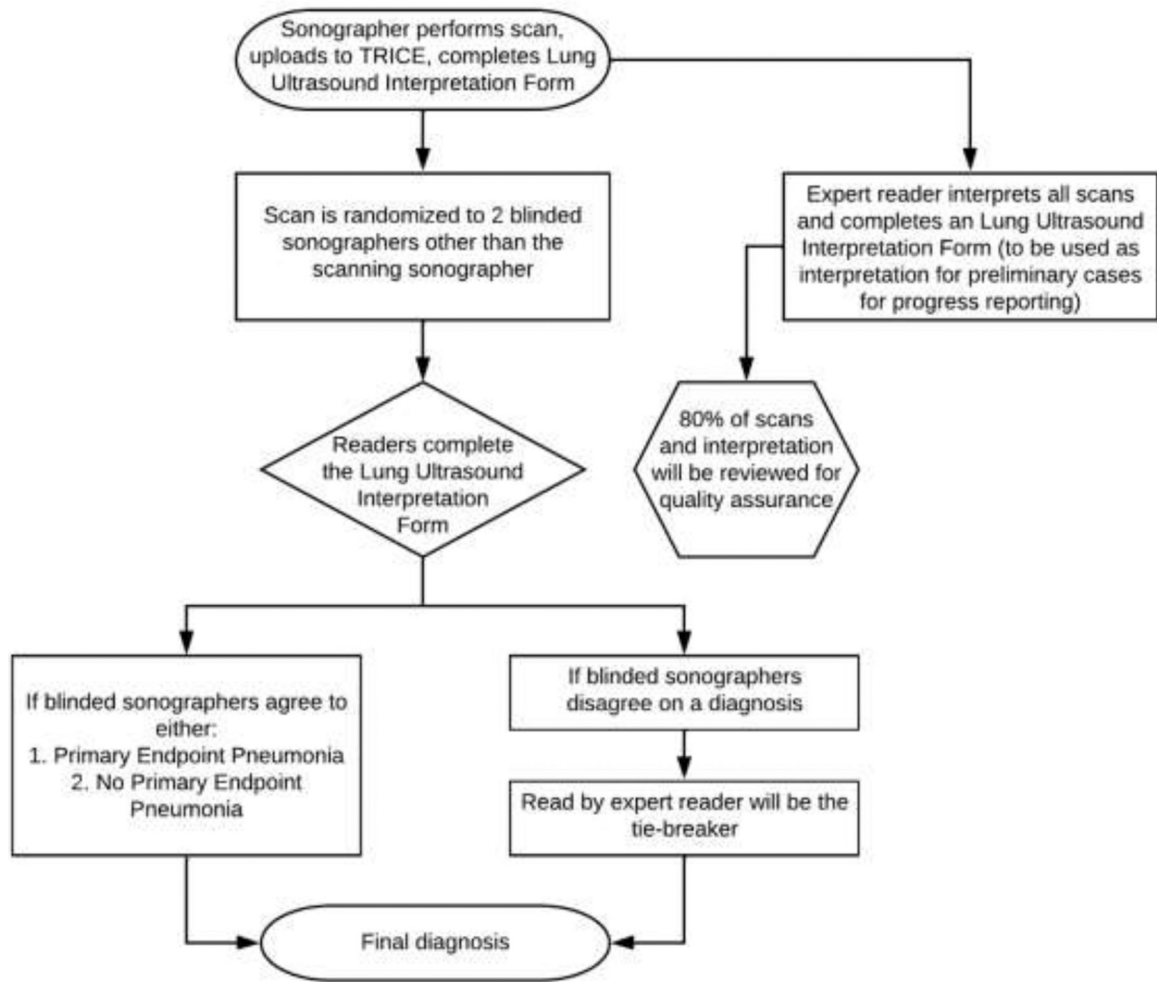


Figure 2. Workflow for Image Interpretation.

This flow chart demonstrates the step-by-step process for the acquisition of the cine-loop image by the scanning sonographer, randomization to the two blinded sonographers, expert quality control, and determination of the final diagnosis.

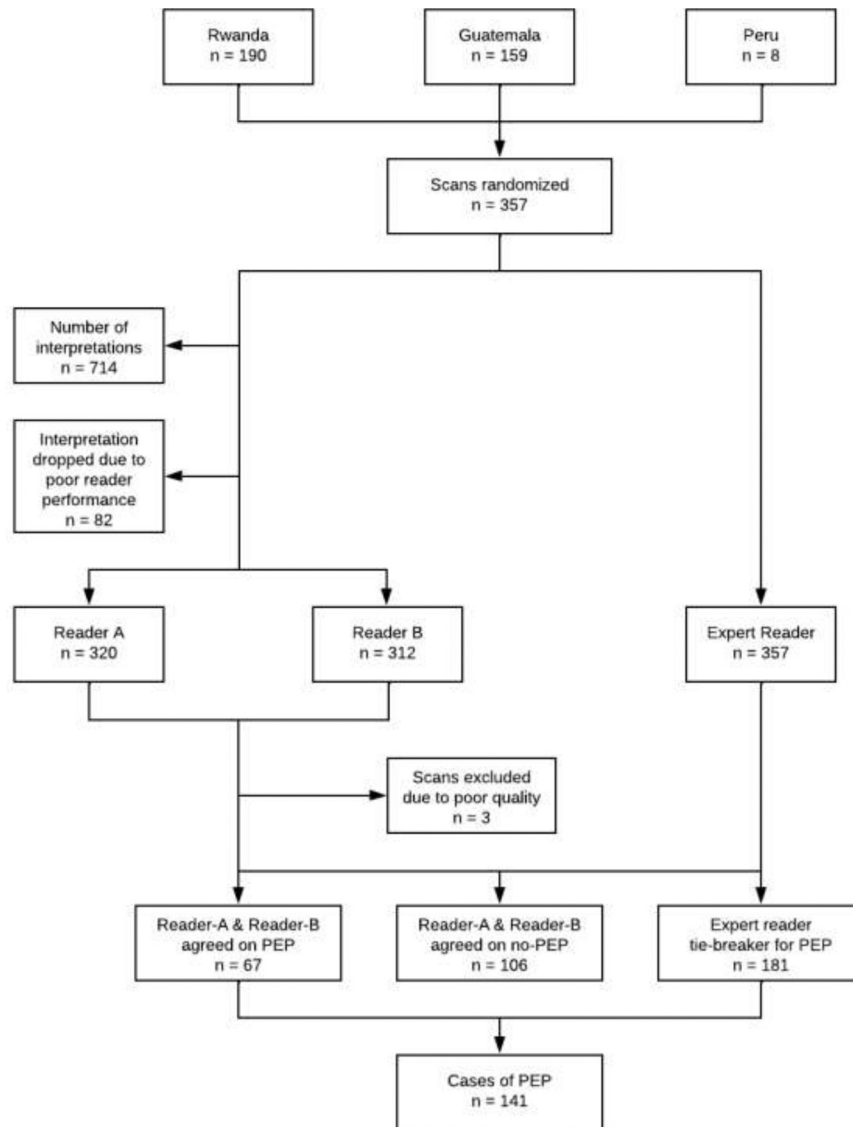


Figure 3. LUS Adjudication Diagram.

This consort diagram demonstrates the number of ultrasounds acquired by IRC, randomization, interpretation and determination of final diagnosis within our panel.

Table 1.

Definition of Primary Endpoint Pneumonia (PEP) on lung ultrasound

Quality Standards	Definition of Primary Endpoint Pneumonia
High Quality (all criteria must be met): <ul style="list-style-type: none"> - Depth 6 cm - Gain allows visibility of 6 cm - Good visualization of the lung parenchyma - 24 complete views, both sagittal and longitudinal images are obtained 	1) Consolidation that is either >1 intercostal space in size or 1 cm in width or height with or without the presence of air bronchograms or shred sign or
Medium Quality (all criteria must be met): <ul style="list-style-type: none"> - Depth 6 cm - Gain allows visibility of 6 cm - Good visualization of the lung parenchyma - 24 regions recorded with 1 or 2 zones not interpretable. 	2) Pleural effusion of any size that has any of the following spatially located findings: <ul style="list-style-type: none"> - Consolidation of any size - Pleural abnormalities - Interstitial infiltrate (3 B-lines)
Low Quality: <ul style="list-style-type: none"> - The criteria for good or medium quality are not met. 	
Appearance of findings on lung ultrasound	
Consolidation: Appears as a hypoechoic area or tissue pattern with loss or attenuation of distinct pleural lines.	
Air bronchogram: Appears as a dot-like or a linear hyperechoic artifact within the consolidation.	
Shred sign: Appears as an irregular border that interrupts the pleura and goes beneath the pleura and is a consolidation that is limited in depth by an irregular border.	
Pleural abnormalities: Appears as an irregular border that interrupts the pleura but is still along the pleura.	
Pleural effusion: Appears as an anechoic (black area with no a-lines or air bronchograms) and represents fluid.	
Lung sliding: Appears as rhythmic movements of the pleural line with respiration.	
A-lines: Equally spaced horizontal lines parallel to the pleural line, representing artifact generated by subpleural air.	
B-lines: Appear as a line that starts at the pleura and reaches the bottom of the screen and erase A-lines when the lung is sliding.	
Interstitial infiltrate: Appears as 3 B-lines	

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2.

Primary Endpoint Pneumonia (PEP) determination by lung ultrasound finding. This includes the scans that will be used in the primary analysis (Agreement on Reader A & Reader B or tie breaker). Includes the number of scans who are considered primary endpoint pneumonia and how the diagnosis of PEP was determined; and is not mutually exclusive; therefore, a scan may meet the qualification under one or more categories.

	Consolidation >1 intercostal space or 1cm in width or height <i>with or without air bronchograms or shred sign</i>	Pleural effusion + consolidation of any size	Pleural Effusion + pleural abnormalities	Pleural Effusion + Interstitial Abnormalities (3 B-lines)
Number of scans diagnosed as PEP	98	65	44	55

Table 3.

Agreement among reading panel members.

Readers	Agreement	Unadjusted-kappa	Prevalence and bias adjusted kappa
Reader A- Reader B	62%	0.23	0.24
Reader A- Expert Reader	64%	0.26	0.27
Reader B- Expert Reader	62%	0.24	0.24
Guatemala Reader- Expert Reader	65%	0.29	0.30
Peru Readers- Expert Readers	62%	0.26	0.24
Rwanda Readers- Expert Reader	67%	0.33	0.33

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript