Computer-aided diagnosis for World Health Organization-defined chest radiograph primary-endpoint pneumonia in children

Nasreen Mahomed, Bram van Ginneken, Rick H. H. M. Philipsen, Jaime Melendez, David P. Moore, Halvani Moodley, Tanusha Sewchuran, et

Pediatric Radiology

ISSN 0301-0449

Pediatr Radiol
DOI 10.1007/s00247-019-04593-0

Springer

Pediatric Radiology
https://doi.org/10.1007/s00247-019-04593-0

ORIGINAL ARTICLE

Computer-aided diagnosis for World Health Organization-defined chest radiograph primary-endpoint pneumonia in children

Nasreen Mahomed[1,2] · Bram van Ginneken[3,4] · Rick H. H. M. Philipsen[3,4] · Jaime Melendez[3] · David P. Moore[2,5,6] · Halvani Moodley[1] · Tanusha Sewchuran[1] · Denny Mathew[1] · Shabir A. Madhi[2,6]

Received: 4 June 2019 / Revised: 26 September 2019 / Accepted: 28 November 2019

Abstract

Background The chest radiograph is the most common imaging modality to assess childhood pneumonia. It has been used in epidemiological and vaccine efficacy/effectiveness studies on childhood pneumonia.

Objective To develop computer-aided diagnosis (CAD4Kids) for chest radiography in children and to evaluate its accuracy in identifying World Health Organization (WHO)-defined chest radiograph primary-endpoint pneumonia compared to a consensus interpretation.

Materials and methods Chest radiographs were independently evaluated by three radiologists based on WHO criteria. Automatic lung field segmentation was followed by manual inspection and correction, training, feature extraction and classification. Radiographs were filtered with Gaussian derivatives on multiple scales, extracting texture features to classify each pixel in the lung region. To obtain an image score, the 95th percentile score of the pixels was used. Training and testing were done in 10-fold cross validation.

Results The radiologist majority consensus reading of 858 interpretable chest radiographs included 333 (39%) categorised as primary-endpoint pneumonia, 208 (24%) as other infiltrate only and 317 (37%) as no primary-endpoint pneumonia or other infiltrate. Compared to the reference radiologist consensus reading, CAD4Kids had an area under the receiver operator characteristic (ROC) curve of 0.850 (95% confidence interval [CI] 0.823–0.876), with a sensitivity of 76% and specificity of 80% for identifying primary-endpoint pneumonia on chest radiograph. Furthermore, the ROC curve was 0.810 (95% CI 0.772–0.846) for CAD4Kids identifying primary-endpoint pneumonia compared to other infiltrate only.

Conclusion Further development of the CAD4Kids software and validation in multicentre studies are important for future research on computer-aided diagnosis and artificial intelligence in paediatric radiology.

Keywords Accuracy · Children · Computer-aided diagnosis · Pneumonia · Primary-endpoint · Radiography · World Health Organization

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00247-019-04593-0) contains supplementary material, which is available to authorized users.

✉ Nasreen Mahomed
Nasreen.Mahomed@wits.ac.za

[1] Department of Radiology, Faculty of Health Sciences, University of Witwatersrand, Johannesburg 2000, South Africa

[2] Medical Research Council, Respiratory and Meningeal Pathogens Research Unit, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

[3] Diagnostic Image Analysis Group, Department of Radiology and Nuclear Medicine, Radboudumc, Nijmegen, the Netherlands

[4] Radboud UMC, Thirona, Nijmegen, the Netherlands

[5] Departments of Paediatrics and Child Health, Chris Hani Baragwanath Academic Hospital, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

[6] Departments of Science/National Research Foundation: Vaccine Preventable Diseases, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

Published online: 13 January 2020

Introduction

Pneumonia is the leading infectious cause of morbidity and mortality in children younger than 5 years of age globally. In developing countries, childhood pneumonia accounted for 1.5 million deaths per year in 2008 [1]. *Streptococcus pneumoniae* (pneumococcus) and *Haemophilus influenzae* type b (Hib) are the most significant causes of vaccine-preventable deaths in children younger than 5 years of age, in the absence of vaccination against these pathogens [2, 3]. The chest radiograph remains the most common and readily available, low-cost imaging modality for the assessment and specific diagnosis of childhood pneumonia [4, 5].

Pneumonia defined by chest radiography is used as an important outcome measure in pneumonia epidemiological studies and vaccine efficacy trials. In 1997, the World Health Organization (WHO) convened a Radiology Working Group to provide a consensus method for interpreting chest radiographs in the context of it being used to evaluate endpoints in pneumococcal conjugate vaccine efficacy trials [6]. These definitions were not designed for use in individual patient clinical management because of their emphasis on specificity at the expense of sensitivity [7].

Chest radiographs are a two-dimensional (2-D) representation of a three-dimensional (3-D) structure and superimposition of anatomical structures overlapping with regions of interest make detecting abnormalities difficult, even for experienced radiologists [8]. The correct interpretation of chest radiographs is complex and subjective, with different readers being influenced by expertise, professional training and momentary factors like fatigue, distraction and focus [9]. In contrast, the use of computer-aided diagnosis (CAD)-based software in the automated reading of chest radiographs has the potential to be devoid of inter- and intra-observer variability and may reduce detection errors [8–10] Most of the existing CAD systems on chest radiography are aimed at adults and the early detection of cancer, with only a small number of studies devoted to detecting other pathologies [5].

There are a limited number of trained radiologists and paediatricians in low-income countries with experience and certification in WHO standardized chest radiograph interpretation. The application of CAD in radiography of children with pneumonia may be an alternative and supplement to human reading [4, 8, 11]. There is limited literature on the use of CAD on chest radiography in children with pulmonary pathology [8]. A study by Oliveira et al. [12] that investigated the use of CAD on chest radiographic pneumonia in children was limited by a small sample size and the use of photos of chest radiographs captured by a digital camera, whose integrity was compromised by poor image quality [6].

In order to develop artificial intelligence algorithms involving chest radiographs to predict clinical outcome, the clinical setting and disease burden profile including the HIV prevalence needs to be considered. We aimed to develop computer-aided diagnosis (CAD4Kids) software to interpret chest radiographs for WHO-defined primary-endpoint pneumonia and to evaluate the sensitivity and specificity of CAD4Kids to a majority consensus radiologist interpretation as the reference standard.

Materials and methods

Ethics approval for this study was obtained from the University of the Witwatersrand Human Research Ethics Committee. Parents/legal guardians consented to their children's participation in the Pneumonia Etiology Research in Child Health (PERCH) study.

This study was nested within the PERCH study [13, 14] at the South African site based at Chris Hani Baragwanath Academic Hospital, Johannesburg. Enrollment of cases occurred over a 2-year period (August 2011 to August 2013), during which 920 children younger than 5 years old with WHO-defined clinically severe or very severe pneumonia were enrolled; 885 (96%) of them had chest radiographs done. WHO clinical criteria for severe and very severe pneumonia are defined as the presence of cough and/or difficulty breathing with the current illness and lower chest wall in-drawing and/or danger signs [15].

The software CAD4Kids was developed from 2015 to 2017 with the PERCH South African team and Diagnostic Image Analysis Group, Department of Radiology, the Netherlands, who have done extensive work on CAD on chest radiography in adults, but not in children younger than 5 years. The paediatric radiology unit has two digital radiograph units (Philips and Second Opinion, Johannesburg, South Africa) and one portable radiograph unit (Shimadzu, Tokyo, Japan). Chest radiographs were routinely printed and manually scanned (using an Epson Expression 10000 XL flatbed scanner, Los Angeles, CA) due to the absence of software to upload and store the digital radiographs directly.

Chest radiographic findings have been briefly reported [14]. Chest radiographs were read independently by 3 radiologists (N.M. with 8 years of clinical experience, H.M. with 4 years, T.S. with 3 years) using WHO standardized chest radiographic interpretation methodology in a binary classification [6]. All three radiologists have experience in interpreting chest radiographs in a high HIV prevalence setting and similar scope of clinical practice and they interpreted the chest radiographs independently. A majority consensus reading (2/3 or 3/3 readers) of the chest radiographic findings was used during the data analysis phase.

WHO chest radiographic primary-endpoint pneumonia was defined as a "dense or fluffy opacity that occupies either a portion, lobe or entire lung, with or without air bronchograms." Other infiltrate (non-endpoint) was defined as "linear, patchy densities (interstitial infiltrate) in a lacy

pattern associated with peribronchial thickening and atelectasis not sufficient enough to call endpoint consolidation." The primary-endpoint pneumonia was defined as "endpoint consolidation or pleural effusion involving the lateral pleural space" that is "spatially associated with pulmonary parenchymal infiltrate (including other infiltrate) or if the effusion obliterates the hemi-thorax to obscure an opacity" [6].

Of the 858 interpretable chest radiographs, 333 (39%) had WHO-defined primary-endpoint pneumonia, 208 (24%) had other infiltrate only and 317 (37%) had no primary-endpoint pneumonia or other infiltrate only.

The entire data set of South African PERCH chest radiographs were stored in a JPEG (Joint Photographic Experts Group) digital format with 256 grey levels (8 bit). Using the majority consensus reading, the principal investigator used the program Paint.NET (https://www.getPaint.NET/) to manually outline the areas of primary-endpoint pneumonia and other infiltrate only on the chest radiographs

Automatic lung segmentation

The first step of the computer analysis involved automatic lung field segmentation, as the structures of interest are the lung fields. From the 858 interpretable chest radiographs, the lung fields were manually outlined in 25% (214/858) randomly selected chest radiographs by the principal investigator. This was used to train the CAD4Kids software to automatically segment the lung fields of the remaining chest radiographs following the approach described by van Ginneken et al. [16]. Briefly, this approach computes a set of filtered versions of the radiograph from convolutions with the 2-D Gaussian function and its derivatives up to second order, at 5 scales. The output of these filters at each pixel is used as a feature vector for that pixel, augmented with the spatial coordinates of that pixel as two additional features. These feature vectors are used to train a k-nearest neighbour classifier for each pixel as belonging to the object, i.e. right lung, left lung or background. Once the classifier had been trained on the 214 randomly selected chest radiographs, it was applied to the remaining 644 chest radiographs and provided a probability for each pixel as belonging to right lung, left lung or background. These lung probability maps were smoothed and post-processed with mathematical morphology filters [16].

The automated lung segmentations were manually inspected and if deemed inadequate, they were adjusted manually. From the chest radiographs with automated lung segmentation (n=644), manual inspection demonstrated 237 (37%) of chest radiographs to have inadequate automatic lung field segmentation, hence these chest radiographs were manually adjusted. The required adjustments were minor involving the superior mediastinum in most cases, where the software had misinterpreted the superior mediastinum for upper lobe lung fields on chest radiography. The rationale for manually correcting the outlines was to analyse the performance of CAD4Kids software texture analysis, undisturbed by errors in the lung field extraction.

There are greater technical challenges with regards to automatic lung field segmentation on chest radiography in children compared to adults as the mediastinum and cardiothoracic ratios have different ranges in children younger than 5 years old. Variability in the thymic size in children accounts for variability in the superior mediastinum size on chest radiography in this age group [17–19].

Texture analysis

Chest radiographs demonstrate three main types of abnormalities. Some disease processes may demonstrate a combination of these: texture abnormalities, characterized by diffuse change and structure of the area (e.g., interstitial lesions), focal abnormalities, which appear as isolated changes in density (e.g., pulmonary nodules), and abnormal shape, where the disease process changes the outline of normal anatomy (e.g., cardiomegaly) [5].

In CAD4Kids, texture analysis is used to assign a probability to each location in the lung fields for it to belong to a region of chest radiograph primary-endpoint pneumonia. This is comparable to the approach taken by van Ginneken et al. [16] using the software CAD4TB, where a similar analysis is performed to determine if locations in the lung fields are similar to abnormal regions outlined in chest radiographs of adults investigated for tuberculosis [20]. First, all images were rescaled to a width of 1,024 pixels. The pixels from outlined regions of primary-endpoint pneumonia and other infiltrates were used as examples of abnormal pixels, and pixels in the lung fields from radiographs without outlines (no primary-endpoint pneumonia or other infiltrate) were used as examples of normal pixels. In order to train a classifier to produce a probability that a pixel belongs to the abnormal class, each pixel needs to be described by a feature vector. This vector of numbers should capture the image information at that location. To obtain this vector, pixel data were filtered with Gaussian derivatives through second order (zero order, first order derivatives to x and y, second order derivatives to xx, xy and yy) on scales of 1, 2, 4 and 8 pixels. For every 8^{th} pixel in vertical and horizontal direction, the first four moments (mean, standard deviation, skew, kurtosis) in a local circular neighbourhood with radius 32 pixels were computed for every filtered image plus the original image as a set of texture features. With 25 filtered images and 4 moments, this results in 100 features. Additionally, several spatial context features for each pixel were also calculated: the x- and y-coordinates (normalized to the height of the image), the x- and y-coordinates in the bounding box of the lungs and the distance to the lung wall.

This feature set can be used to classify every 8^{th} pixel and values in between are interpolated. Classification was done with a support-vector machine with radial basis kernel, using

the implementation from Libsvm (https://www.csie.ntu.edu.tw/~cjlin/libsvm/). This classifier has two hyperparameters, the width of the radial basis function and the penalty term C, which are estimated from the data during training. The same system was used as in CAD4TB: filters, every 8th pixel in a 32-pixel radius, 4 moments of filter responses and position features [20].

Training was done in 10-fold cross validation, i.e. chest radiograph data were divided in 10 groups, and every group was classified with a system trained on the other 90% of images. The training set of chest radiographs was not used to test the algorithm's performance. To visualise the working of the texture system, for each chest radiograph, a colour heat map was generated with red representing high likelihood of being abnormal, yellow intermediate, green low, and blue very low.

Image score using a quantile rule

The proposed texture analysis system provides a likelihood of a textural abnormality being present for each pixel in the lung fields. In order to classify the chest radiograph, these pixel scores were summarized into a single score. To compute a single score for each image, all pixel scores were sorted and the 95% percentile of all scores was determined using the support-vector machine classifier system. This simple summary statistic has been shown to provide good performance in previous work [21].

Computer-aided diagnosis on chest radiographs, analysis by HIV exposure status

In order to develop artificial intelligence algorithms involving chest radiographs in predicting clinical outcome, the clinical setting and disease burden profile, including the HIV prevalence, needs to be considered. Hence, analysis using CAD4Kids for primary-endpoint pneumonia versus non-primary-endpoint was also done by three HIV-exposure status groups (HIV-unexposed, HIV-exposed, uninfected [HEU] and HIV-infected children. Chest radiograph primary-endpoint pneumonia was prevalent in 163/428 (38%) of HIV-unexposed children, 94/284 (33%) of HEU children and 65/108 (60%) of HIV-infected children [14] and will be described in detail in a separate paper.

Computer-aided diagnosis versus independent radiologist

A consensus of three readers was used as the reference standard in this study. To put the performance of the computer system in perspective, it is common practice to compare it with an external independent human reading. Blinded to the reference consensus reading, the radiologist with 12 years' clinical experience scored a set of 200 randomly selected chest radiographs from the 858 data set (stratified randomization 100 randomly selected chest radiographs with primary-endpoint pneumonia and 100 randomly selected chest radiographs with no primary-endpoint pneumonia or other infiltrate only). Each radiograph was scored between 0 and 100, where 0 indicates certainty that the image is normal and 100 indicates certainty that the image is abnormal. A receiver operating characteristic (ROC) curve was created for the radiologist.

Data analysis

Data were analysed using STATA software, version 13.0 (StataCorp, College Station, Texas). To compare the results of CAD4Kids to the WHO standardized chest radiograph 3-reader consensus reading as the reference standard, 2x2 tables were generated using the chi-square test or Fisher exact test as appropriate. From these, sensitivity, specificity, positive predictive value and negative predictive values and areas under the ROC curve were generated. Statistical significance was set as a two-tailed P-value <0.05. The area under the ROC curve was utilized as a performance measure of the software CAD4Kids.

Results

During the 2-year study period, 920 cases with WHO-defined clinically severe or very severe pneumonia were enrolled. Frontal chest radiographs were available in 885 (96%) cases. Of the 885 children with available chest radiographs, 27 (3%) were uninterpretable based on the consensus reading.

The median age of the study population with interpretable chest radiographs was 6 months (interquartile range [IQR]: 2 to 12 months) with a mean age of 9.04 months. Of the 858 children with interpretable chest radiographs, 426 (50%) were 1 to 5 months of age, 212 (24%) were 6 to 11 months of age and 220 (26%) were between 12 and 59 months of age. The HIV status of the 858 children included 428 (50%) HIV-unexposed, 284 (33%) HEU, 108 (13%) HIV-infected and 38 (4%) categorised as HIV-unknown exposure, uninfected. Of the 108 HIV-infected children, CD4 (cluster of differentiation 4) results were available in 106 (98%) of children. Thirty-one (29%) HIV-infected children were on antiretroviral therapy before being hospitalized for WHO-defined severe or very severe pneumonia.

Using the majority consensus reading, of 858 interpretable chest radiographs, 333 (39%) had primary-endpoint pneumonia, 208 (24%) had other infiltrate only and 317 (37%) had no primary-endpoint pneumonia or other infiltrate only. Chest radiograph primary-endpoint pneumonia was prevalent in 163/428 (38%) of HIV-unexposed children, 94/284 (33%) of HEU children and 65/108 (60%) of HIV-infected children.

Results of computer-aided diagnosis (CAD4Kids) for WHO chest radiograph primary-endpoint pneumonia

To visualise the working of the texture system, a colour heat map was generated for each test image with red representing a high likelihood of being abnormal texture, yellow intermediate, green low, and blue very low. The colour heat map results on test chest radiograph are demonstrated in Fig. 1.

Compared to the radiologist consensus reading of the presence of primary-endpoint pneumonia (n=333) versus non-primary-endpoint pneumonia (n=525), CAD4Kids had a sensitivity of 76%, specificity of 80% and area under the ROC curve of 0.850 (95% CI 0.823–0.876) (Fig. 2).

For radiologist consensus reading of primary-endpoint pneumonia versus other infiltrate only (n=208), when normal chest radiographs were excluded, CAD4Kids had a sensitivity of 77%, specificity of 73% and area under the ROC curve of 0.810 (95% CI 0.772–0.846) (Fig. 3).

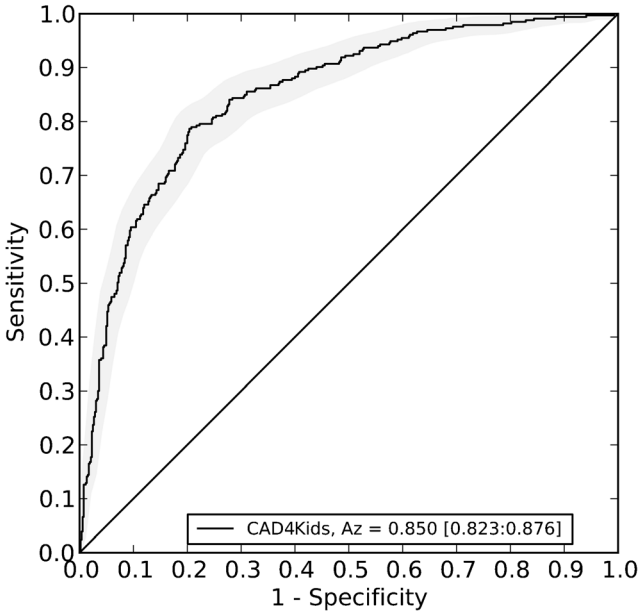

Fig. 2 Receiver operator characteristic (ROC) curve with 95% confidence interval (CI) for chest radiograph primary-endpoint pneumonia versus non-primary-endpoint pneumonia. From the 858 chest radiographs, CAD4Kids had a sensitivity of 76%, specificity of 80% and area under the ROC curve of 0.850 (95% CI 0.823–0.876) using the radiologist consensus reading of WHO standardized chest radiograph interpretation criteria as the reference standard

Analysis by HIV exposure/infection status

Chest radiograph primary-endpoint pneumonia was prevalent in 163/428 (38%) of HIV-unexposed children, 94/284 (33%)

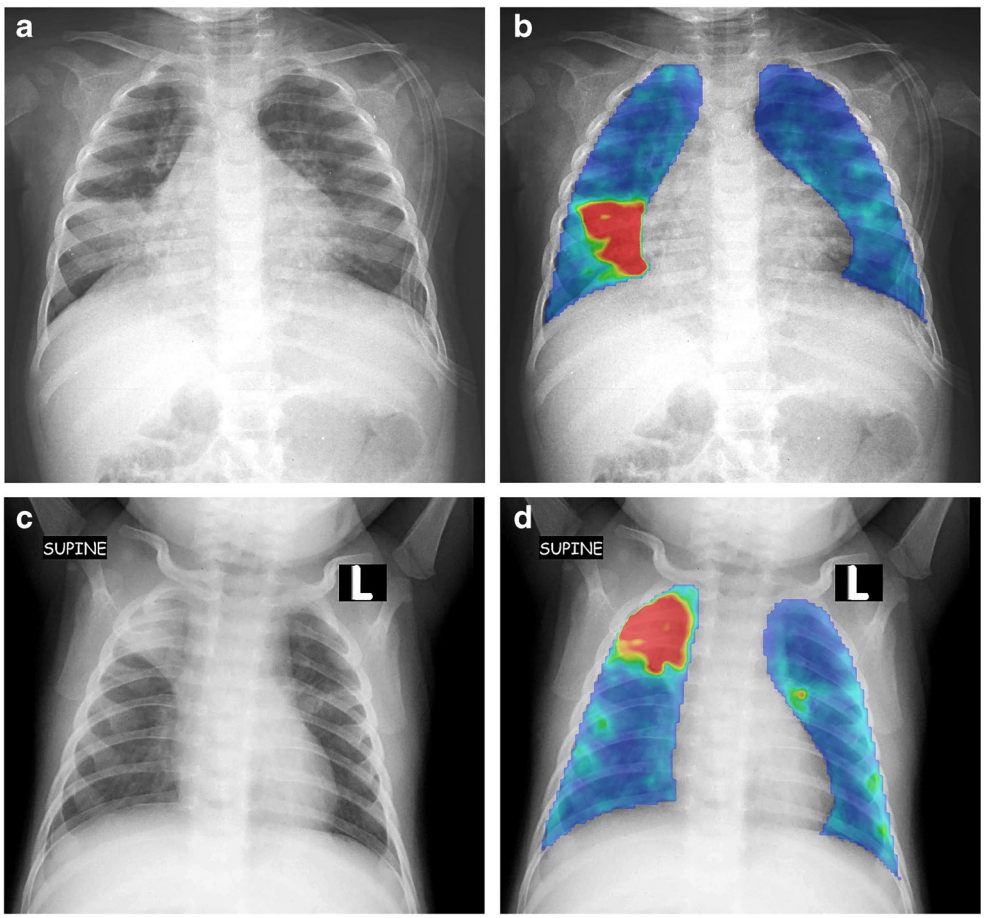

Fig. 1 A colour heat map helps visualise the working of the texture system. For each chest radiograph, a colour heat map was generated with red representing a high likelihood of being abnormal, yellow intermediate, green low and blue very low. **a** A chest radiograph demonstrates right middle lobe primary-endpoint pneumonia. **b** The chest radiograph colour heat map generated by CAD4Kids on the test image demonstrates the right middle lobe primary-endpoint pneumonia as an area of red, correlating with the consensus human interpretation using WHO standardized chest radiograph interpretation criteria. **c** A chest radiograph demonstrates the right upper lobe primary-endpoint pneumonia. **d** The chest radiograph colour heat map generated by CAD4Kids on the test image demonstrates the right upper lobe as an area of red, correlating with the consensus human interpretation using WHO standardized chest radiograph interpretation criteria

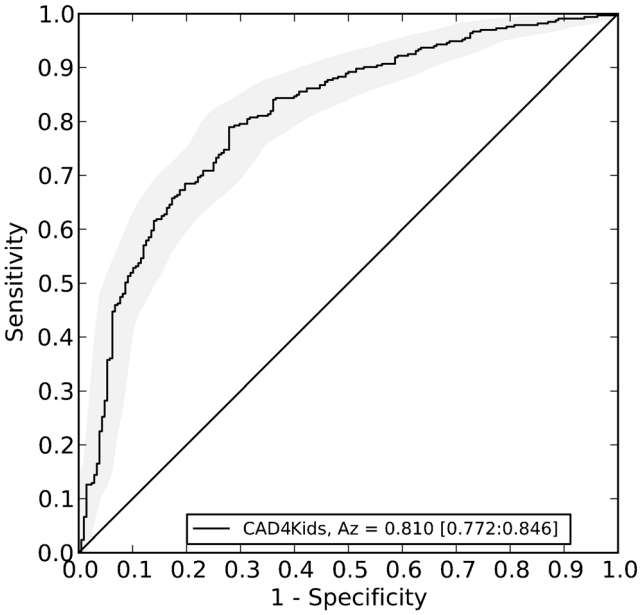

Fig. 3 Receiver operator characteristic (ROC) curve with 95% CI for chest radiograph primary-endpoint pneumonia (n=333) versus other infiltrate only (n=208), where chest radiographs without primary-endpoint pneumonia or other infiltrate were excluded. CAD4Kids had a sensitivity of 77%, specificity of 73%, and area under the ROC curve of 0.810 (95% CI 0.772–0.846) using the radiologist consensus reading with WHO standardized chest radiograph interpretation criteria as the reference standard

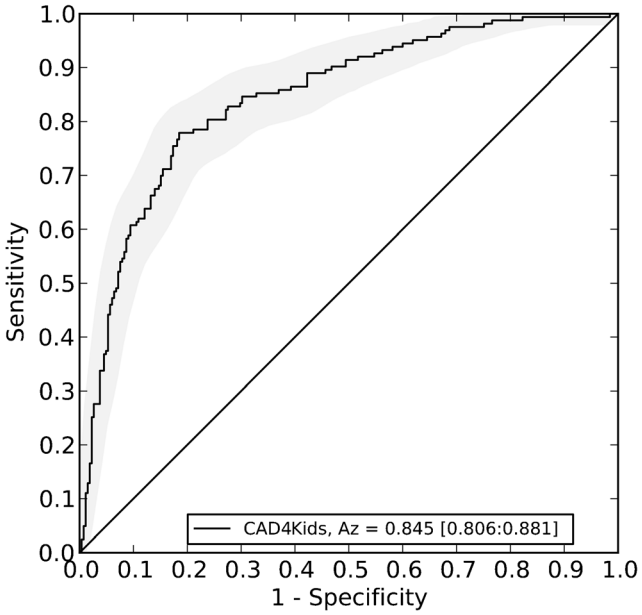

Fig. 4 Receiver operator characteristic (ROC) curve with 95% CI of CAD4Kids for chest radiograph primary-endpoint pneumonia versus non-primary-endpoint pneumonia in 428 children not exposed to HIV. CAD4Kids had a sensitivity of 80%, specificity of 76% and area under the ROC curve of 0.845 (95% CI 0.806–0.881) using the radiologist consensus reading as the reference standard

of HEU children and 65/108 (60%) of HIV-infected children [14]. Analysis for CAD4Kids was further stratified by the three HIV-exposure statuses. The areas under the ROC curve for CAD4Kids was 0.845 (95% CI 0.806–0.881) in HIV-unexposed children (n=428; Fig. 4), 0.850 (95% CI 0.801–0.895) among HEU children (n=284; Fig. 5) and 0.814 (95% CI 0.726–0.892) for HIV-infected children (n=108; Fig. 6).

Computer-aided diagnosis versus independent radiologist

Using the independent scoring system for chest radiograph primary-endpoint pneumonia versus non-chest radiograph primary-endpoint pneumonia, from the 200 randomly selected chest radiographs, the radiologist had a sensitivity of 94%, specificity of 90% and area under the ROC curve of 0.975 (95% CI 0.955–0.991) using the radiologist consensus reading as the reference standard. This was better than the CAD4Kids results (Fig. 7).

Discussion

The chest radiograph remains a major criterion in classifying pneumonia, and despite the acceptance of the recommended WHO criteria for reporting chest radiographs of pneumonia in children, interobserver variability continues to be a problem [22]. Using the WHO standardized chest radiograph interpretation definitions [6], readers interpreting paediatric chest radiographs achieved moderate to substantial agreement (Kappa=0.48–0.62) for WHO-defined primary-endpoint

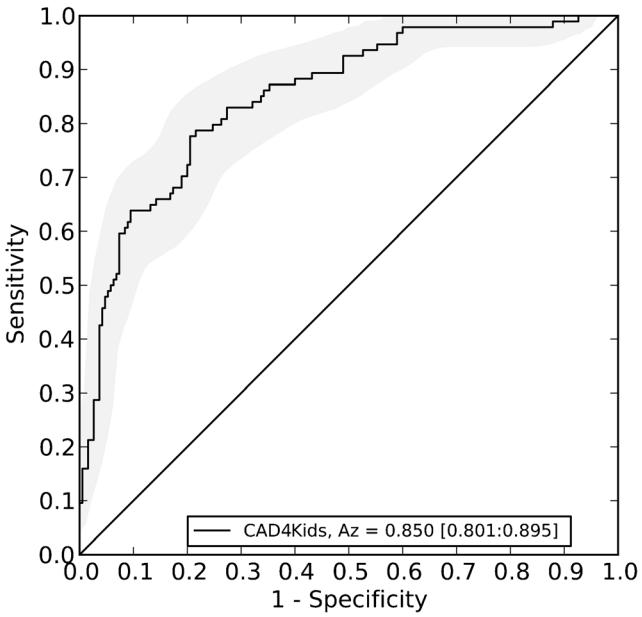

Fig. 5 Receiver operator characteristic (ROC) curve with 95% CI for chest radiograph primary-endpoint pneumonia versus non-primary-endpoint pneumonia in 284 HIV-exposed, uninfected children. CAD4Kids had a sensitivity of 80%, specificity of 75% and area under the ROC curve of 0.850 (95% CI 0.801–0.895) using the radiologist consensus reading as the reference standard

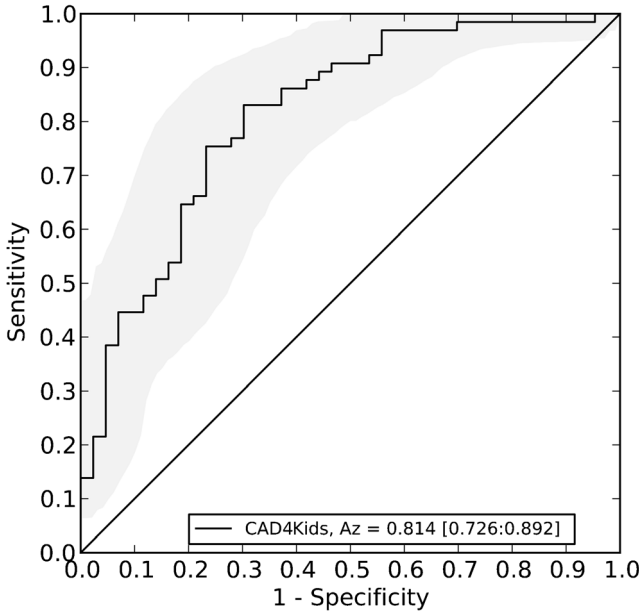

Fig. 6 Receiver operator characteristic (ROC) curve with 95% CI for chest radiograph primary-endpoint pneumonia versus non-primary-endpoint pneumonia in 108 HIV-infected children. CAD4Kids had a sensitivity of 75%, specificity of 77% and area under the ROC curve of 0.814 (95% CI 0.726–0.892) using the radiologist consensus reading as the reference standard

pneumonia and fair to moderate agreement (Kappa=0.27–0.66) for other infiltrate [22–25].

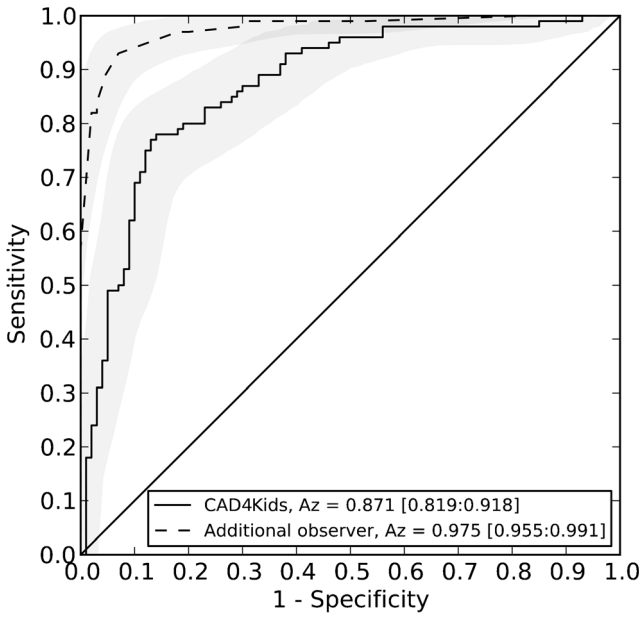

Fig. 7 Receiver operator characteristic (ROC) curve with 95% CI from 200 randomly selected chest radiographs, comparing an independent reader for chest radiograph primary-endpoint pneumonia versus non-primary-endpoint pneumonia using the radiologist consensus reading as the reference standard. The radiologist had a sensitivity of 94%, specificity of 90% and area under the ROC curve of 0.975 (95% CI 0.955–0.991) using the radiologist consensus reading as the reference standard

Elemraid et al. [22] characterized inter-observer variability in the interpretation of chest radiographs for diagnosing pneumonia in children according to the WHO radiologic classification. In this study, 169 patients were identified and treated for pneumonia and/or empyema on chest radiograph and substantial interobserver variability was noted (Kappa=0.7, $P<0.001$), with patchy changes (48.8%) and perihilar changes (28.1%) forming the largest component of this variability in children younger than 5 years of age [22]. False-negative reports between the two interpretations of chest radiographs could lead to an underestimation of likely bacterial pneumonia and also undermine the sensitivity and specificity of study endpoints used to measure the efficacy/effectiveness of vaccines against bacterial pathogens, such as pneumococcal conjugate vaccine [22].

Machine learning and computer-aided diagnosis, a subfield of artificial intelligence, is a "rapidly evolving technology" [26]. There is very limited literature on the use of CAD in paediatric radiology including chest radiographs. In this study, we demonstrated that CAD4Kids software textural analysis generated an area under the ROC curve of 0.850 for identifying chest radiograph primary-endpoint pneumonia in a paediatric population younger than 5 years old, with WHO clinical severe/very severe pneumonia in which the radiologist consensus reading was used as the reference standard. The overall sensitivity and specificity for chest radiograph primary-endpoint pneumonia by CAD4Kids, as evident by the ROC curve of 0.810, were maintained when limiting analysis to those chest radiographs arbitrated by the radiologist panel as having at least chest radiograph primary-endpoint pneumonia or other infiltrate only, i.e. excluding chest radiographs adjudged as being normal.

Oliveira et al. [12] found that the software Pneumo-CAD in paediatric chest radiographs, using the Haar wavelet textural analysis, had sensitivity of 100%, specificity of 80% and area under the ROC curve of 0.97; however, this study was limited by a very small sample size of 60 children and poor chest radiograph quality as chest radiographs were photographed from a light box using a digital camera with low resolution and bit depth [6].

In adult studies using computer-aided diagnosis to classify chest radiographic pneumonia, where a binary radiologist majority consensus was used as a gold standard, the Stanford group developed the software CheXNet, which generated an area under the ROC curve of 0.788 for chest radiograph pneumonia [27]. This is in comparison to You et al. [28] with an area under the ROC curve of 0.713 and Wang et al. [29] with an area under the ROC curve of 0.633 for the classification of chest radiograph pneumonia. There has been further development of CheXNeXt, a convolutional neural network to concurrently detect the 14 chest radiographic patterns. On a test set of 420 frontal chest radiographs with a reference standard of a majority consensus of 3 cardiothoracic specialist

radiologists, CheXNeXt generated an area under the ROC curve for chest radiograph pneumonia of 0.893 (0.859 to 0.924), a sensitivity of 0.594 (95% CI 0.500–0.688) and specificity of 0.927 (95% CI 0.897–0.954).

Using WHO standardized chest radiograph interpretation criteria, chest radiograph primary-endpoint pneumonia is a measure used for evaluating the efficacy and effectiveness of pneumococcal conjugate vaccine against radiologically confirmed pneumonia, and is also being used as a proxy for bacterial pneumonia in epidemiological studies [6]. The CAD4Kids textural analysis software yielded promising concordance in identifying chest radiograph primary-endpoint pneumonia compared to consensus radiologist readings, with further optimization of the software required. It may have an important role especially in settings where there is limited clinical expertise in the interpretation of paediatric chest radiographs using WHO standardized methodology in the future.

This further substantiates the need for CAD-based systems to supplement human reading in interpreting chest radiographs in children in areas with limited resources. CAD4Kids software textural analysis in our study generated an area under the ROC curve of 0.850 for identifying chest radiograph primary-endpoint pneumonia compared to non-chest radiograph primary-endpoint pneumonia, and for an area under the ROC curve of 0.810 for identifying chest radiograph primary-endpoint pneumonia compared to other infiltrate, the software performed better than the described interobserver agreement by human interpretation as described in the literature [22–24]. Automating the interpretation of chest radiographs for the detection of chest radiograph primary-endpoint pneumonia leads to objective reproducible results and a standardized way of chest radiograph reporting globally [5].

Previous studies on CAD in paediatric chest radiographs included that by Mouton et al. [8], who used CAD textural analysis to identify pulmonary tuberculosis, which has a different radiographic pattern than investigated in our study on chest radiograph primary-endpoint pneumonia in the context of childhood community-acquired pneumonia. Nevertheless, the concordance of CAD compared to human readings in our study was comparable to their study on CAD, in which the ROC curve was 0.782 (0.655–0.937) for interstitial lung disease [8]. Notably, the study by Mouton et al. [8] did not use the WHO standardized chest radiograph interpretation and alveolar infiltrate was categorized under interstitial lung disease. Unlike Mouton et al. [8], we did not find significant errors using the CAD4Kids classifier, in the perihilar or basal areas of the lung fields or left lung fields in our paediatric study.

The strength of our study was the large sample size of 858 children, compared to other studies, where the number of features used for classification or feature selection on CAD chest radiographs exceeded the number of chest radiographs with abnormalities. Furthermore, we showed that HIV-exposure and infection status did not materially affect the CAD interpretation of chest radiographs, with the ROC curve ranging from 0.814 to 0.850, even though the percentage of chest radiograph primary-endpoint pneumonia in each of the HIV-infection and -exposure groups varied from 33% to 60%.

Deep learning methods are more accurate in the classification of abnormalities with the added ability to discern various types of suspected diseases simultaneously and directly from the chest radiographs. Despite the developmental potential of deep learning, it is mainly used in limited data sets and is an area for future research and development [5].

Limitations of this study included that JPEG format chest radiographs were used, which is not optimal because there is some loss of image information in the reduction to 8-bit (256) gray values and the compression used in that format. Original Digital Imaging and Communications in Medicine (DICOM) data would have been better, but was unavailable in this study due to the absence of a PACS system.

The CAD4Kids was tested against the radiologist consensus reading as the reference standard, as is currently the practice [8, 12, 30, 31]. However, a more novel approach would be to compare the performance of CAD4Kids software on chest radiographs against a radiologic reference standard like computed tomography or magnetic resonance imaging. The children in our study were only imaged with chest radiographs, hence this was not done.

A further limitation of our study was that CAD4Kids software was tested on chest radiograph primary-endpoint pneumonia from a single study centre. The application of CAD4Kids from chest radiographs from multiple different centres and image sets from different imaging systems would test the robustness of this software on textural analysis.

In the current approach, lung field markings needed manual adjustment in 37% of chest radiographs. It was challenging for the software to automatically segment lung fields due to the variability in mediastinal size on chest radiographs in children younger than 5 years old [17–19]. To improve automatic lung field segmentation, larger training sets of chest radiographs from children younger than 5 may be required. Phase 2 of the CAD4Kids project will involve optimizing automated lung field segmentation. This is being performed using the evolving field of deep learning, including deep convolutional neural networks with machine learning algorithms [26].

The CAD4Kids software was divided into four steps: image pre-processing, extracting regions of interest (ROI), extracting ROI features, and classifying the disease according to the features. With the development of artificial intelligence and accumulation of large volumes of medical images, new opportunities are opening up for CAD systems with deep convoluted neural networks in medical applications [5]. Future research on CAD for paediatric pulmonary tuberculosis would be important, especially in a high prevalence setting.

Conclusion

We showed that HIV-exposure status did not materially affect the CAD4Kids interpretation of chest radiographs, even though the percentage of chest radiograph primary-endpoint pneumonia in each of the HIV-exposure groups varied from 33% to 60%. CAD4Kids software performed better than the interobserver agreement by human interpretation as described in the literature. Further development and validation in multicentre studies are important for future research on computer-aided diagnosis and artificial intelligence in paediatric radiology.

Acknowledgements The study received Carnegie PhD Fellowship funding and a South African Medical Research Council Self-Initiated Research Grant. We thank staff at the Respiratory and Meningeal Pathogens Research Unit involved in enrolling study patients: Azwifarwi Mudau, Debra Katisi, Audrey Kubheka, Magare Lelaka, Tsekabe Makgoba, Matshediso Moilwa, Malebo Motiane, Sibonsile Moya, Thondani Netshishivhe, Minah Nkuna, Mmabatho Selela, Ndulela Titi, Nonhlanhla Tsholetsane, Lerato Mapetla and Gudani Singo.

Compliance with ethical standards

Conflicts of interest None

References

1. Black RE, Cousens S, Johnson HL et al (2010) Global, regional, and national causes of child mortality in 2008: a systematic analysis. Lancet 375:1969–1987
2. O'Brien KL, Wolfson LJ, Watt JP et al (2009) Burden of disease caused by Streptococcus pneumoniae in children younger than 5 years: global estimates. Lancet 374:893–902
3. Watt JP, Wolfson LJ, O'Brien KL et al (2009) Burden of disease caused by Haemophilus influenzae type b in children younger than 5 years: global estimates. Lancet 374:903–911
4. Pitcher RD, Lombard C, Cotton MF et al (2014) Clinical and immunological correlates of chest X-ray abnormalities in HIV-infected South African children with limited access to antiretroviral therapy. Pediatr Pulmonol 49:581–588
5. Qin C, Yao D, Shi Y, Song Z (2018) Computer-aided detection in chest radiography based on artificial intelligence: a survey. Biomed Eng Online 17:113
6. Cherian T, Mulholland EK, Carlin JB et al (2005) Standardized interpretation of paediatric chest radiographs for the diagnosis of pneumonia in epidemiological studies. Bull World Health Organ 83:353–359
7. Mahomed N, Fancourt N, de Campo J et al (2017) Preliminary report from the World Health Organisation Chest Radiography in Epidemiological Studies project. Pediatr Radiol 47:1399–1404
8. Mouton A, Pitcher RD, Douglas TS (2010) Computer-aided detection of pulmonary pathology in pediatric chest radiographs. Med Image Comput Comput Assist Interv 13:619–625
9. Breuninger M, van Ginneken B, Philipsen RH et al (2014) Diagnostic accuracy of computer-aided detection of pulmonary tuberculosis in chest radiographs: a validation study from sub-Saharan Africa. PLoS One 9:e106381
10. Muyoyeta M, Maduskar P, Moyo M et al (2014) The sensitivity and specificity of using a computer aided diagnosis program for automatically scoring chest X-rays of presumptive TB patients compared with Xpert MTB/RIF in Lusaka Zambia. PLoS One 9: e93757
11. van Ginneken B, ter Haar Romeny BM, Viergever MA (2001) Computer-aided diagnosis in chest radiography: a survey. IEEE Trans Med Imaging 20:1228–1241
12. Oliveira LL, Silva SA, Ribeiro LH et al (2008) Computer-aided diagnosis in chest radiography for detection of childhood pneumonia. Int J Med Inform 77:555–564
13. Pneumonia Etiology Research for Child Health (PERCH) Study Group (2019) Causes of severe pneumonia requiring hospital admission in children without HIV infection from Africa and Asia: the PERCH multi-country case-control study. Lancet 394:757–779
14. Mahomed N, Sewchuran T, Moodley H, Madhi SA (2016) Chest x-ray findings in children hospitalized with WHO-defined severe, very severe pneumonia in a high HIV prevalence setting in the era of bacterial conjugate vaccines. Paper presented at the 7th International Pediatric Radiology (IPR) Conjoint Meeting & Exhibition, Chicago, Illinois, 15-20 May 2016
15. Scott JA, Wonodi C, Moisi JC et al (2012) The definition of pneumonia, the assessment of severity, and clinical standardization in the Pneumonia Etiology Research for Child Health study. Clin Infect Dis 54:S109–S116
16. van Ginneken B, Stegmann MB, Loog M (2006) Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database. Med Image Anal 10:19–40
17. Meyers A, Shah A, Cleveland RH et al (2001) Thymic size on chest radiograph and rapid disease progression in human immunodeficiency virus 1-infected children. Pediatr Infect Dis J 20:1112–1118
18. Mendelson DS (2001) Imaging of the thymus. Chest Surg Clin N Am 11:269–293
19. Menashe SJ, Iyer RS, Parisi MT et al (2016) Pediatric chest radiographs: common and less common errors. AJR Am J Roentgenol 207:1–9
20. Hogeweg L, Sanchez CI, Maduskar P et al (2015) Automatic detection of tuberculosis in chest radiographs using a combination of textural, focal, and shape abnormality analysis. IEEE Trans Med Imaging 34:2429–2442
21. Loog M, Van Ginneken B (2004) Static posterior probability fusion for signal detection: applications in the detection of interstitial diseases in chest radiographs. In: Proceedings of the 17th International Conference on Pattern Recognition, 26 August 2004, Cambridge, UK, pp 644-647
22. Elemraid MA, Muller M, Spencer DA et al (2014) Accuracy of the interpretation of chest radiographs for the diagnosis of paediatric pneumonia. PLoS One 9:e106051
23. Patel AB, Amin A, Sortey SZ et al (2007) Impact of training on observer variation in chest radiographs of children with severe pneumonia. Indian Pediatr 44:675–681
24. WHO HiB Initiative Radiology Workshop, Hanoi, Vietnam, 11-12 October 2011
25. Fancourt N, Deloria Knoll M, Baggett HC et al (2017) Chest radiograph findings in childhood pneumonia cases from the multisite PERCH study. Clin Infect Dis 64:S262–S270
26. Moore MM, Slonimsky E, Long AD et al (2019) Machine learning concepts, concerns and opportunities for a pediatric radiologist. Pediatr Radiol 49:509–516
27. Rajpurkar P, Irvin J, Zhu K et al (2018) CheXNet: radiologist-level pneumonia detection on chest X-rays with deep learning. https://arxiv.org/abs/1711.05225. Accessed 4 Jun 2019
28. You D, Hug L, Ejdemyr S et al (2015) Global, regional, and national levels and trends in under-5 mortality between 1990 and 2015, with scenario-based projections to 2030: a systematic analysis by the UN Inter-agency Group for Child Mortality Estimation. Lancet 386:2275–2286

29. Wang X, Peng Y, Lu L et al (2017) ChestX-ray8: hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. https://arxiv.org/abs/1705.02315. Accessed 4 Jun 2019
30. van Ginneken B, Katsuragawa S, ter Haar Romeny BM et al (2002) Automatic detection of abnormalities in chest radiographs using local texture analysis. IEEE Trans Med Imaging 21:139–149
31. Sousa RT, Marques O, Curado GTF et al (2014) Evaluation of classifiers to a childhood pneumonia computer-aided diagnosis system. Proceedings of the IEEE 27th International Symposium on Computer-Based Medical Systems, 477-478. https://doi.org/10.1109/CBMS.2014.98