

Diagnostic agreement of pediatricians and radiologists on chest x-rays in pediatric emergency department

Diagnostic agreement of pediatricians and radiologists on chest x-rays

Nilüfer Aylanç¹, Fatih Battal², Yusuf Güzel², Koray Öz¹

¹Department of Radiology

²Department of Pediatrics, Çanakkale Onsekiz Mart University, Faculty of Medicine, Çanakkale, Turkey

Abstract

Aim: In this study, we aimed to investigate the diagnostic agreement on chest x-rays between pediatrician and radiologist in anatomical and pathological evaluation.

Material and Methods: Chest radiographs of 700 cases from the pediatric emergency clinic were included in the study. The radiographs were evaluated by the pediatrician and radiologist as double-blind from a technical point of view (posture-position, inspiration adequacy, X-ray dose) and interpretation (hilar, mediastinal, parenchymal pathologies and cardiothoracic ratio). For results, Kappa analysis was used.

Results: The mean age of the patients was 4.5 ± 4.5 (min 0-max 18) years, 335 (47.9%) were females and 365 (52.1%) were males. The diagnostic agreement was found as moderate level with adequate inspiration, optimal position, mediastinal width, and left hilar fullness; low level of disagreement was found in optimal dose, pleural fluid, consolidation, mediastinal fullness, right hilar fullness, infiltration, peribronchial thickening, reticular signs and very good agreement was found in CTR. It was statistically significant in terms of all parameters ($p < 0.001$).

Discussion: There may be differences, especially between interpretations of radiological examinations. To minimize these differences and to establish an accurate and precise diagnosis, the clinician and radiologist should have close contact. Compliance and consensus in the evaluation in the radiological assessment will help to manage the treatment process precisely and correctly.

Keywords

Chest Radiography, Radiologist, Pediatrician, Pediatric Emergency

DOI: 10.4328/ACAM.20690 Received: 2021-05-05 Accepted: 2021-07-01 Published Online: 2021-07-17 Printed: 2021-11-01 Ann Clin Anal Med 2021;12(11):1203-1207

Corresponding Author: Nilüfer Aylanç, Çanakkale 18 Mart University, School of Medicine, Department of Radiology, 17100 Çanakkale, Turkey.

E-mail: niluferaylanc@gmail.com P: +90 286 218 00 18 GSM: +90 5058569606 F: +90 286 218 37 38

Corresponding Author ORCID ID: <https://orcid.org/0000-0002-5889-9763>

Introduction

A chest x-ray is an important and initial diagnostic tool for evaluating the airways, pulmonary parenchyma, mediastinum, heart, pleura and chest wall. It is an imaging technique frequently used in the diagnosis and treatment of pediatric patients, including term and preterm neonates (1). Laboratory tests and radiological evaluations are at least as important as clinical findings (2). Accurate and rapid interpretation of chest radiography can provide early diagnosis, especially in emergency departments. Conditions such as the presence of thymus, some diseases special to children, organs in the developmental stage may make it difficult to interpret the chest radiography in childhood (3, 4). For these reasons, we think that the ability to read chest radiographs is important for radiologists or pediatricians. In the literature, we did not find any study comparing chest radiography evaluations between a radiologist and a pediatrician. In this study, we aimed to investigate the diagnostic agreement between those two observers.

Material and Methods

This study included chest x-rays of children aged 1 month to 18-year who were evaluated at our pediatric emergency clinic, between 01.2017 and 04.2018. After the study approval with the decision of the Clinical Research Ethics Committee dated 18/04/2018 and numbered 08-04, we retrospectively scanned chest radiographs. Children with chronic lung disease (tuberculosis, immune deficiency, ciliary dyskinesia, cystic fibrosis, etc.), congenital heart disease, trauma history and at the age of 0-28-days were not included in the study. The images were evaluated independently by each observer; one radiologist and one pediatrician, both with 15 years of experience. On radiographs, adequate inspiration, patient position, X-ray dose, presence of pleural fluid, consolidation, infiltration, peribronchial thickening, reticular signs, mediastinal fullness, right and left hilar fullness and CTR were evaluated. Mediastinal width was measured in cm with the lateral mediastinal borders at the upper mediastinum by drawing a linear measurement line. CTR was calculated as the ratio of the transverse diameter where the heart is widest to the transverse diameter where the thorax is widest.

Statistical Analysis

The data obtained were analyzed with the statistical package program SPSS 20.0 version (IBM SPSS Statistics 20; Chicago, IL, USA). Spearman’s correlation analysis was used according to the results of the normal distribution compatibility test. Correlation coefficient r: 0-0,24 as weak; r: 0.25-0.49 as moderate; r: 0.50-0.74 as strong; r: 0.75-1.0 as very strong. Kappa analysis (5) was used in categorical data evaluated by two observers. Agreement level in interpretation by Cohen’s kappa coefficient (κ) was <0: worse than chance-related, 0.01-0.20: unimportant, 0.21-0.40: low, 0.41-0.60: moderate, 0, 61-0.80: good, 0.81-1.00: very good (5). A p-value less than 0.05’ was considered statistically significant.

Results

While the mean age of 700 patients was 4.5 ± 4.5 (min 0-max 18) years; 335 (47.9%) were females and 365 (52.1%) were

males.

In terms of adequate inspiration, diagnosis disagreement was experienced in 25.8% of cases, while the agreement was found in 16.4%. A moderate level of diagnostic compliance was found between observers (p <0.001, κ: 0.409). In the evaluation of the optimal position, both observers noted the absence of an optimal position in 34% of cases, while present in 40.6%. However, in 178 cases (25.4%), the diagnoses were discordant with moderate level but were statistically significant (p <0.001, κ: 0.501). In terms of optimal dose, it was present in 56% of cases, while not in 16%. In 196 (28%) cases, the diagnoses were incompatible. Statistically significant agreement was observed among the observers (p <0.001, κ: 0.343). In the evaluation of pleural fluid, both experts commented positively in 668 (98.9%) and negatively in 2 (1.1%) of the cases. The diagnosis was different only in 10 (1.4%) cases, and the compatibility between the observers was statistically significant. (p <0.001, r: 0.28). There was a statistically significant agreement in the consolidation (p <0.001, κ: 0.223). (Table 1)

Both observers stated that there was consolidation in a total of 664 (94.9%) cases, while not in 5 (0.7%) cases. Diagnosis mismatch was present in 31 (4.4%) cases (Figure 1). In the evaluation of infiltration, a very low level of disagreement was found, but it was statistically significant (p <0.001, κ: 0.118). In 414 (59.1%) patients, both observers stated that there was no infiltration, while 38 (5.4%) had it. In 248 patients (35.4%), the diagnoses were incompatible (Figure 2). In terms of peribronchial thickening, a statistically significant but unimportant- very low diagnostic disagreement was found (p <0.001, κ: 0.115). In 39% of cases (n = 273), the diagnoses were discordant (Table 1). Prominence in reticular signs was also evaluated. While both observers agreed in 101 (14.4%) cases that reticular signs were evident, in 351 (50.1%) cases, the findings were negative

Table 1. Qualitative evaluation of chest radiograph

Exam parameters 1	Diagnostic mismatch n (%)	κ	p
Adequate inspiration	181 (25,8)	0,409	
Optimal posture-position	178 (25,4)	0,501	
Optimal shooting dose	196(28)	0,334	
Pleural fluid	10(1,4)	0,280	
Consolidation	31(4,4)	0,234	
Infiltration	248(35,4)	0,118	<0,001
Peribronchial thickening	273(39)	0,115	
Prominence of reticular signs	248(35,4)	0,202	
Mediastinal fullness	185(26,4)	0,267	
Right hilar fullness	204(29,1)	0,360	
Left hilar fullness	153(21,8)	0,525	

κ: Kappa coefficient, p: statistical significance for Kappa analysis

Table 2. Quantitative evaluation of chest radiograph

Exam parameters 2	Rho	p
Cardiothoracic ratio	0,865	<0,001
Mediastinal width	0,375	

Rho: Spearman correlation coefficient, p: statistical significance for Spearman correlation analysis

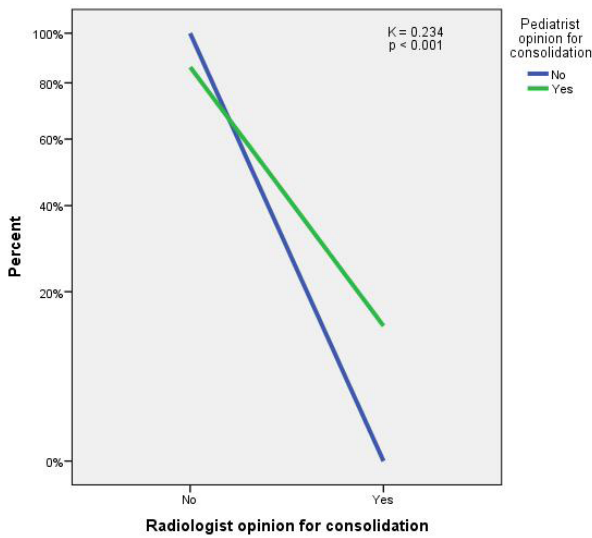


Figure 1. Observers' interpretations on consolidation. Opinions are mostly close to each other.

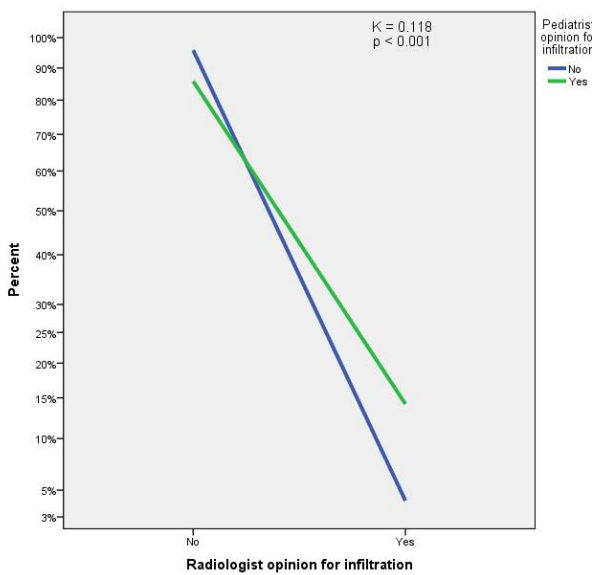


Figure 2. Correlation of interpretations about infiltration. Most of the opinions are similar.

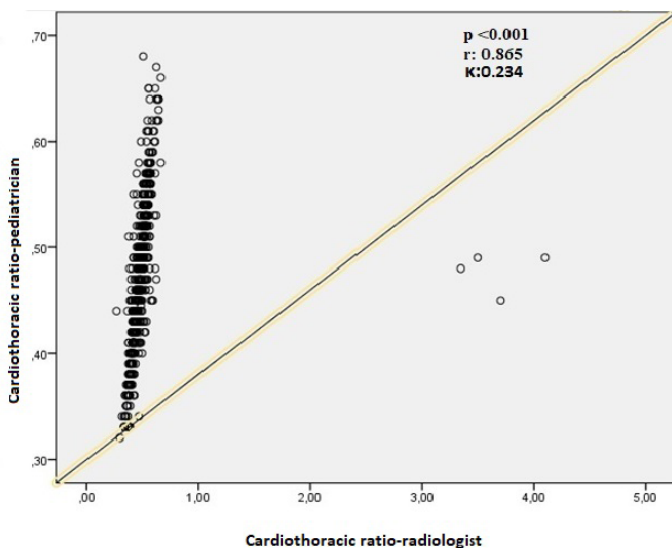


Figure 3. Measurement values of CTR. Nearly all of the values are correlated between observers.

for both physicians. There was a diagnostic mismatch in 248 cases (35.4%), and it was statistically significant but at an unimportant-very low level. ($p < 0.001$, κ : 0.202). In terms of mediastinal fullness, a weak but statistically significant diagnostic disagreement was detected ($p < 0.001$, κ : 0.267). In 185 (26.4%) cases, the diagnoses were incompatible, and both physicians reported that there was no mediastinal fullness in 444 (63.3%) cases, and it was present in 71 (10.1%) cases. Hilar fullness was evaluated separately for both sides. In the evaluation of right hilar fullness, a statistically significant but weak diagnostic disagreement was achieved ($p < 0.001$, κ : 0.36). In 204 cases (29.1%), the diagnoses of the observers differed. In the evaluation of left hilar fullness, a moderate and statistically significant diagnostic agreement was found ($p < 0.001$, κ : 0.525). Only 153 (21.8%) patients had left hilar fullness (Table 1).

Mediastinal width was measured in cm. Statistically significant and moderate agreement was found ($p < 0.001$, r : 0.375). Cardiothoracic ratio values were calculated. There was a statistically significant and very good agreement between observers ($p < 0.001$, r : 0.865) (Figure 3) (Table 2).

Discussion

In our study, chest x-rays of child patients admitted to the pediatric emergency were evaluated and different levels of diagnostic disagreement were obtained among the radiologist and the pediatrician. In radiological examinations, interpretive differences due to observers are inevitable, both between the same or different clinicians [6,7]. Many factors, such as the physician's education, experience and information about the patient's clinic, may play a role in this difference [8]. However, every clinician aims to provide the most appropriate treatment or contribute to this process by making rapid and accurate diagnosis. In order to minimize possible interpretation differences or to standardize the approach to diagnosis, studies on especially artificial intelligence models are conducted [9].

In our study, a moderate agreement was found between observers in terms of adequate inspiration, optimal position, mediastinal width and left hilar fullness. For inspiration adequacy, the result may be related to the clinician's knowledge and experience in chest X-ray. Mediastinal width was determined at the level of the upper mediastinum with both mediastinal lines bordered. Although using numerical data, moderate level agreement may be due to the fact that measurements were made from the upper mediastinum, from different levels, such as a bit upper or a bit lower. In the presence of left hilar fullness, the diagnostic mismatch may be due to asymmetries in posture positions.

The situations in which the disagreement was found to be statistically significant but weak were optimal dose, pleural fluid, consolidation, mediastinal fullness, and right hilar fullness. Disagreement regarding optimal dose may be related to technical knowledge and experience in radiographic evaluation. Although adequate dosage is one of the points to consider at the beginning of the evaluation, in clinical practice, this can often be ignored or forced to be ignored, especially in busy centers. For the presence of pleural fluid, diagnostic disagreement was in only 1.4% of cases, and for the consolidation in only 4.4% of cases. The fact that observers agree in most cases is actually

quite meaningful in clinical practice, and this will contribute patient's treatment management

In mediastinal fullness, diagnostic mismatch was detected in almost $\frac{1}{4}$ of the cases. Mediastinal fullness was interpreted visually and subjectively, rather than by measurement with numerical data. Some cases also have radiographs taken in the supine position, causing the appearance of the mediastinum fuller than normal, which may be a reason for the interpretive difference. Or, the fact that the mediastinum is physiologically seen full due to thymus, especially in the first 1-2 years of age may be a result of the differences. In right hilar fullness, in which the diagnostic agreement was statistically significant but weak, we think that anatomical factors may play a role here as well. The right hilus seems to be fuller compared to the left, because the bronchus reaching the lower lobe is more perpendicular and there are 3 main bronchi for the 3 lobes in right hemithorax. Considering the anatomical changes of the cases with age, this interpretative difference is an expected result.

In infiltration, peribronchial thickening and reticular signs, the disagreement level was very low, but statistically significant. In terms of infiltration, disagreement was in 35.4% of patients. Although the infiltration process can occur in various forms, it is already more difficult to evaluate on radiographs than on computed tomography (CT) until it progresses. However, in daily practice, the first method of imaging is radiography. Also, it may be difficult to differentiate infiltration especially in the presence of insufficient inspiration and dose. Similarly, there was diagnostic disagreement of 35.4% in reticular signs and 39% in peribronchial thickening. Prominence in reticular signs is one of the early findings of interstitial pathologies and can sometimes be confused with opacities of vascular traces. Thus, in this process, CT would be more useful. However, since the cases are pediatric and the first-step imaging is a chest X-ray, the observers' knowledge and experience play an important role here. In peribronchial thickening, the main bronchi may be prominent under 2 years, or when there is suboptimal posture, or anphase appearance of bronchi, then compliance decreases. There was a very good level of agreement in the CTR evaluation and the level was statistically significant. It was thought that, apart from the experience of the observer, the reason is the information obtained by calculating numerical data and therefore, a quantitative evaluation.

There are many studies on the diagnostic agreement in the literature and controversial results can still be obtained. In a review by Usubütün, it is pointed out that this situation is due to the uncertainty areas in medicine [8]. According to Fox [10], the existence of up-to-date scientific information in medicine and situations exceeding the physician's knowledge, constitute the source of these uncertainties. Also, in that review, it was emphasized that physicians' awareness of the uncertainties and learning to overcome problems, will contribute to their professional success. In one study, dentists examined periapical radiographs and investigated the main causes of diagnostic incompatibility. They emphasized that morphological differences in the anatomy and differences in peripheral bone density make evaluation difficult, and uncertainty occurs in radiologic interpretations in difficult cases [11].

In our study, moderate or low compliance in some parameters

also depends on the limitation of chest radiographs. For example, superposition of tissues or postural disturbances make evaluation difficult and decreases diagnostic compliance. In another study, diagnostic compliance between observers was investigated using the fluorescent insitu hybridization method used in the diagnosis of helicobacter pylori, and it was found that the compliance was quite good in cases with high bacterial count, but decreased where the bacterial load was low. This result also shows that in cases where there is a high level of data, the diagnostic compliance is higher, since there are no gaps open to interpretation [12].

In a study by Erdogan et al. [13], humerus proximal fractures were evaluated by orthopaedists. Radiography and tomography findings of patients were evaluated by an experienced upper extremity surgeon and general orthopaedist. The diagnostic agreement was found to be higher in tomography than radiography, since the cross-sectional imaging gave detailed information and did not leave much clarification on interpretation. In another study, mammography examinations were evaluated by radiologists, in terms of diagnostic compatibility based on BIRADS system. Although this system provides diagnostic standardization to a large extent, diagnostic compliance is high for benign and malignant lesions, whereas diagnostic compliance is found to be moderate in the class defined as possible benign, which we call 'category 3'. Therefore, in order to increase diagnostic compliance, it was emphasized that this system should be updated at certain time intervals in accordance with literature studies and scientific opinions [6]. In another study, radiologists evaluated the carotid system with Doppler ultrasound at different time intervals and advocated that the use of the same applicator and same device can increase compliance in cases requiring follow-up [7]; because there may be different interpretations in different times, although by the same observer.

Limitations

This paper attempts to measure diagnostic agreement between radiologists and pediatricians regarding interpretation of chest x rays. We did not aim to determine the accuracy of the diagnosis. The first limitation of our study is that no comparisons were made intra-observers. Another limitation is that only one pediatrician and one radiologist were involved in the study. We think that studies in which researching compliance of interpretations of more than one radiologist and pediatrician would be more valuable.

The strength of our study is that it is the first study in the literature comparing the evaluations of a pediatrician and radiologist over the same chest radiographs.

Radiologists can be expected to have more accurate interpretations, but in daily practice, clinicians have to examine and interpret more chest radiographs than radiologists during the treatment process. Although clinicians do not receive radiology training as much as radiologists, they can be more experienced in terms of radiographs compared to radiologists because they have to examine and interpret graphs in their clinical practice. In addition, because they know patients' clinic, their evaluation on radiographs can be easier and more effective. In conclusion, in radiographic evaluation, there are still areas of disagreements in diagnostic interpretations with

the same or different specialties or physicians with different levels of experience in the same specialty. Clinicians have to make radiological evaluations rapidly during the diagnostic phase and may not always be able to communicate with an experienced radiologist. Ideally, patient's images should be evaluated in detail by different specialists, but this is not possible in practice. However, since diagnostic differences significantly affect the patient's management, they should be minimized as much as possible and eliminated if possible.

Due to the rapid advances in radiologic methods, the variety of imaging modalities is also increasing. As in all diseases, it is necessary to request imaging methods with an appropriate algorithm in the diagnosis, follow-up and evaluation of the treatment response of pediatric respiratory system diseases. It is important to choose the most appropriate exam according to the disease or symptom, with the lowest ionizing radiation, easy accessibility and low cost. In our study, we showed that the clinician and the radiologist can interpret differently chest radiographs in the emergency department. For this reason, we think that clinician-radiologist cooperation is important as well as knowing the general characteristics of imaging examinations. If the chest radiography in children is seen as normal, but cannot not lighten the patient's clinic, then the clinician may act a little more boldly when preferring chest CT. Thus, a good interpretation of radiography findings can prevent many unnecessary CT examinations. Therefore, radiologic examinations should be evaluated systematically using algorithms step by step. During this specialization training, data and parameters should be updated frequently, especially regarding the evaluation of the radiologic images, training seminars should be organized for remembering new data, and the training proficiency should be checked with exams. We strongly recommend that clinicians should contact the radiologist, especially on suspected radiologic findings. Likewise, the radiologist should contact the clinician and try to strengthen his/her diagnosis by obtaining information about patient's clinic. A strong communication between different disciplines will confirm and facilitate patient management. In addition, all these processes will increase the experience of the specialists and keep their level of diagnostic agreement high by enabling them to update their knowledge.

Acknowledgment

We would like to thank Dr. Özgür Özerdoğan for their contribution to the article.

Scientific Responsibility Statement

The authors declare that they are responsible for the article's scientific content including study design, data collection, analysis and interpretation, writing, some of the main line, or all of the preparation and scientific review of the contents and approval of the final version of the article.

Animal and human rights statement

All procedures performed in this study were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. No animal or human studies were carried out by the authors for this article.

Funding: None

Conflict of interest

None of the authors received any type of financial support that could be considered potential conflict of interest regarding the manuscript or its submission.

References

- Walsh BK, Daigle B, Diblasi RM, Restrepo RD. AARC Clinical Practice Guideline. Surfactant Replacement Therapy. *Respir Care*. 2013; 58(2):367–75.
- Lange CD. Radiology in Pediatric Non-Traumatic Thoracic Emergencies. *Insights Imaging*. 2011; 2(5):585–98.
- Aslan A, Yıkılmaz A. Normal and Pathological Pediatric Lung and Thorax Radiography. *TRD Sem*. 2017; 5:98-128.
- Breysem L, Loyer S, Boets A, Proesmans M, Boeck KD, Smet MH. Pediatric Emergencies: Thoracic Emergencies. *Eur Radiol*. 2002; 12(12):2849–65.
- Kılıç S. Kappa test. *Journal of Mood Disorders*. 2015; 5(3):142-4.
- Parlak E, Gürses C, Yıldırım M, Alikanoğlu AS, Özan E, Kargı A, et al. Is Interpretation Differences of Radiologists an Important Issue in Breast Imaging Reporting And Data System? *J Clin Anal Med*. 2013; 4(4):307-9.
- Adapınar B, Akcar N, Kebapçı M, Aslan Kb, Kaya T. Inter-Observer Correlation and Variability in Measurements and Doppler Parameters of Normal Carotis Communis-Interna and Vertebral Arteries. *Turkish Journal of Cerebrovascular Diseases*. 2003; 9(2):51-6.
- Usubütün A. Are the Pathologic Diagnosis Reproducible? *J Curr Pathol*. 2017; 1:8-11.
- Caobelli F. Artificial Intelligence in Medical Imaging: Game Over for Radiologists? *Eur J Radiol*. 2020; 126:108940.
- Anderson RE, Fox RC, Hill RB. Medical Uncertainty And The Autopsy: Occult Benefits For Students. *Hum Pathol*. 1990; 21(2):128–35.
- Keçeci AD, Kaya BÜ, Ünal GÇ, Taç G. Disagreement cases caused interobserver variations during evaluation of periapical status. *SDÜ Medical Journal*. 2006; 13(4):16-21.
- Çiftci İH, Dilek FH, Şafak B, Aşık G, Acartürk G. Is There A Difference Between Observers In The Diagnosis Of Helicobacter Pylori With Fluorescent In Situ Hybridization Method? *Turkish Microbiological Society Association*. 2012; 42(2):61-5.
- Erdoğan Ö, Canbora MK, Uzer G, Demiröz S, Köse Ö. Interobserver agreement in Proximal Humeral Fractures. *Harran University Medical Journal*. 2019; 16(3):417-20

How to cite this article:

Nilüfer Aylanç, Fatih Battal, Yusuf Güzel, Koray Öz. Diagnostic agreement of pediatricians and radiologists on chest x-rays in pediatric emergency department. *Ann Clin Anal Med* 2021;12(11):1203-1207