



## ORIGINAL ARTICLE

# The chest x-ray in acute bronchiolitis: technical quality, findings, and an assessment of its reliability<sup>☆</sup>



Juan Manuel Rius Peris<sup>a,b,\*</sup>, Ana Isabel Maraña Pérez<sup>a,b</sup>, Ana Valiente Armero<sup>a</sup>, Jorge Mateo Sotos<sup>a,b</sup>, Leonor Guardia Nieto<sup>a</sup>, Ana María Torres<sup>a,b</sup>, Elisa María Cueto Calvo<sup>a</sup>

<sup>a</sup> Hospital Virgen de la Luz, Cuenca, Spain

<sup>b</sup> Instituto de Tecnología, Universidad de Castilla-La Mancha, Cuenca, Spain

Received 21 January 2020; accepted 23 March 2020

Available online 5 February 2021

## KEYWORDS

Bronchiolitis;  
Chest x-ray;  
Reproducibility;  
Between-observer  
variability

## Abstract

**Introduction:** Despite the recommendations of the current Clinical Practice Guidelines, the chest x-ray continues to be a widely used diagnostic test in the assessment of infants with acute bronchiolitis (AB). However, there have not been many studies that have assessed its reproducibility in these patients. In the present study, an evaluation is made on the radiographs, describing their quality, their radiological findings, and provides new evidence on the agreement between observers.

**Method:** Out of a total of 281 infants admitted due to acute bronchiolitis, 140 chest x-rays were performed. Twelve doctors from different specialties evaluated the presence or absence of 10 radiological signs previously agreed by consensus. The level of agreement between 2 observers, and in groups of 3 or more, were analysed using the Cohen and Fleiss kappa index, respectively.

**Results:** Only 8.5% of the radiographs showed evidence of a complicated AB. The between-observer agreement in groups of 3 or more was medium, and with little variability (kappa: 0.20–0.40). However, between 2 observers, specialists in radiology, the variability was wider, (kappa: –0.20 to 0.60). This level of agreement was associated with factors including, the sign to evaluate, the medical specialty, and level of professional experience.

<sup>☆</sup> Please cite this article as: Rius Peris JM, Maraña Pérez AI, Valiente Armero A, Mateo Sotos J, Guardia Nieto L, María Torres A, et al. La radiografía de tórax en la bronquiolitis aguda: calidad técnica, hallazgos y evaluación de su fiabilidad. An Pediatr (Barc). 2021;94:129–135.

\* Corresponding author.

E-mail address: [jmrius@sescam.jccm.es](mailto:jmrius@sescam.jccm.es) (J.M. Rius Peris).

**Conclusion:** The low levels of agreement between observers and the wide variability, makes the chest x-ray an unreliable diagnostic tool, and is not recommended for the assessment of infants with AB.

© 2021 Published by Elsevier España, S.L.U. on behalf of Asociación Española de Pediatría. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## PALABRAS CLAVE

Bronquiolitis;  
Radiografía de tórax;  
Reproductibilidad;  
Variabilidad entre  
observadores

## La radiografía de tórax en la bronquiolitis aguda: calidad técnica, hallazgos y evaluación de su fiabilidad

### Resumen

**Introducción:** A pesar de las recomendaciones de las actuales Guías de Práctica Clínica, la radiografía de tórax sigue siendo una prueba diagnóstica ampliamente utilizada en la evaluación de lactantes con bronquiolitis aguda. No obstante, su reproductibilidad en estos pacientes no ha sido muy estudiada. En la presente investigación, se evalúan radiografías describiéndose su calidad técnica, hallazgos radiológicos en las mismas y se aportan nuevas evidencias sobre la concordancia entre observadores.

**Método:** Sobre un total de 281 lactantes ingresados por bronquiolitis aguda, 140 radiografías de tórax fueron realizadas. 12 médicos de diferentes especialidades evaluaron la presencia o ausencia de 10 signos radiográficos previamente consensuados. El nivel de concordancia entre 2 observadores y en grupos de 3 o más, fue estudiado mediante el índice kappa de Cohen y de Fleiss respectivamente.

**Resultados:** Únicamente en el 8,5% de las radiografías se evidenciaron signos de BA complicada. La concordancia entre observadores en grupos de 3 o más fue mediana y con escasa variabilidad (kappa: 0,20–0,40), sin embargo, entre 2 observadores, especialista con radiólogo de referencia, la variabilidad fue más amplia, (kappa: –0,20 to 0,60). Este nivel de concordancia se relacionaba con factores tales como el signo a evaluar, la especialidad médica y el grado de experiencia profesional, entre otros.

**Conclusión:** Los bajos niveles de concordancia entre observadores y su amplia variabilidad, convierten a la radiografía de tórax en una herramienta diagnóstica poco fiable y no recomendable para la evaluación de lactantes con BA.

© 2021 Publicado por Elsevier España, S.L.U. en nombre de Asociación Española de Pediatría. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## Introduction

Acute bronchiolitis (AB) is a disease that entails a substantial utilization of health care services and resources in each epidemic season.<sup>1–3</sup> The plain chest radiograph (CXR) has been and continues to be a widely used diagnostic tool.<sup>4</sup> The most recent clinical practice guidelines and systematic reviews on AB recommend against using CXR in most of these patients<sup>5–7</sup> to avoid an increase in health care costs, reduce waiting times in the emergency department and avoid unnecessary exposure to ionising radiation and irrational use of antibiotherapy.<sup>4</sup> In the past few years, significant efforts have been made to adhere to these recommendations.<sup>8</sup> A diagnostic test must offer an adequate validity, safety, reliability and precision to be recommended for use in clinical practice.<sup>9</sup> The reliability and precision are determined by how reproducible the test is in different clinical contexts. The reproducibility of diagnostic tests is assessed by means of interobserver agreement analyses.<sup>9</sup> Most of the previous studies on the reliability and precision of CXR in children have not focused on infants with AB, but in patients with a broad age range (infants and children) with all sorts of

lower respiratory tract infections. In addition, they mainly compared only 2 observers. As a consequence, the results of different studies are quite heterogeneous.<sup>10</sup> We only found one study in the literature assessing these aspects and specifically focused on infants with AB.<sup>11</sup> The main objective of our study was to determine whether CXR is reliable enough to be used in our setting for diagnosis of AB, exploring whether the medical speciality of the observer and the level of professional experience affected the degree of agreement with the reference radiologist. As secondary objectives, we aimed to describe the radiological findings in a sample of infants admitted to hospital with AB and assess the technical quality of the radiographs.

## Material and methods

### Study design and sample

We conducted a cross-sectional, observational and analytical study in a secondary level hospital (Appendix B) and included physicians that interpreted CXRs of infants with

a diagnosis of AB. The data collection was prospective between 2009 and 2017, including infants aged 12 or fewer months admitted with AB (based on the McConnochie criteria adjusted for age),<sup>12</sup> and retrospective between 2006 and 2008, by reviewing the health records of all infants aged 12 or fewer months, and including those with a diagnosis of AB based on the International Classification of Diseases, ninth revision (ICD-9), specifically with a documented diagnostic code of 466.11 or 466.19.

## Observers

The study involved a total of 12 observers: 5 paediatricians (P1, P2, P3, P4 y P5), 3 emergency doctors (ED1, ED2 and ED3), 2 medical residents in the speciality of Family and Community Medicine (R1, R2 and R3) and 1 paediatric radiologist (PR). With the aim of analysing the potential association between the amount of professional experience of the observing paediatricians and the degree of agreement with the radiologist that served as reference, we classified their professional experience based on an arbitrary breakdown of the years of professional experience following completion of the residency: less than 10 years (low experience), 10–19 years (intermediate experience) and 20 or more years (high experience). Based on this classification, paediatricians P2 and P5 had the least experience, P1 had intermediate experience and P3 and P4 were most experienced. We considered the PR the reference in the interpretation of the CXR. She was specifically trained on paediatric radiology during her residency and has been exclusively devoted to paediatric radiology in our hospital in the past 12 years. All the included physicians that performed as observers gave their verbal consent to participate in the study.

## Radiographs included in the study

Out of a total of 281 infants, 129 (45.9%) underwent at least one CXR. A total of 140 CXRs were performed in this group, either during the initial contact with the emergency department or during the hospital stay. All of them were anteroposterior CXRs.

## Variables under study

### Protocol for observation and interpretation of radiographs

For the purpose of observation and interpretation, each of the 140 CXRs was displayed in the same digital format and in a different order for each of the 12 observers. We set up a computer in each of the departments where the collaborating observers were affiliated to. To prevent the potential identification of the images and potential communication between observers, we changed the sequence of the CXRs to be examined for each observer, although all of them interpreted the same images. During the period in which the X-ray films were available for interpretation, the observers were not allowed to talk to each other or to access any other external source of information.

They were also denied access to the health records of the patients and all they knew about the patients was that they were infants with a diagnosis of AB. The films were viewed on a 24-inch screen with 1920 × 1080 full high-definition resolution. Each observer had to assess each and every radiograph for the presence or absence of the following 10 radiographic signs: hyperinflated lungs, hilar enlargement, perihilar or peribronchial infiltrates, diffuse interstitial infiltrates, subsegmental atelectasis, pulmonary infiltrates or major atelectasis, cardiomegaly, pneumomediastinum, pneumothorax and foreign body. The presence of any of these radiographic signs did not exclude the others. We considered the first 5 signs compatible with “simple or uncomplicated AB”. We considered the presence of the sixth sign compatible with “complicated AB”, and the remaining 4 signs suggestive of a “possible alternative diagnosis other than AB”. When none of these 10 signs were present, the radiograph was considered a “normal CXR”. The categories we used to classify the CXRs taken in infants with AB do not constitute a validated approach but were established by consensus by the research team following the reading and analysis of similar studies in the previous literature.<sup>10–15</sup> In addition, the PR evaluated another 4 features that are indicative of the quality of the image: inclusion of the entire ribcage in the image, proper centring/alignment of the image, adequate hardness/depth of penetration and adequate inspiration. These features were agreed on by the research team after consulting the literature on the subject.<sup>16</sup> All radiographic signs were documented as dichotomous variables with the possible values yes/no.

## Statistical analysis

We have summarised the data as absolute frequencies and percentages with the corresponding 95% confidence intervals (CIs). We assessed the agreement between 2 observers with the Cohen kappa<sup>17,18</sup> and the agreement between 3 or more observers with the Fleiss kappa. The latter is a generalization of the Cohen kappa that can be applied to multinomial data (more than 2 categories), ordinal data, more than 2 observers, incomplete designs and the combination of all of the above, generalizations that involve more complex calculations but resulting in a statistic that is interpreted the same way.<sup>18,19</sup> The statistical analysis was performed with the software MATLAB 2018. To interpret the kappa statistic ( $\kappa$ ) with a widely accepted approach, we took as reference the standards of strength of agreement proposed by Landis and Koch,<sup>20</sup> which have since been endorsed in the scientific literature.<sup>21</sup> According to these standards, kappa values of less than 0.00 indicate poor agreement, 0.00–0.20 slight agreement, 0.21–0.40 fair agreement, 0.41–0.60 moderate agreement, 0.61–0.80 substantial agreement and 0.81–1.00 near perfect agreement.

## Calculation of precision for the sample size

Based on the article published by AB Cantor<sup>22</sup> and using the Epidat 4.2 software, we estimated the precision of the kappa statistics given the obtained sample size (140 radiographs).

**Table 1** Principal characteristics of the sample.

Sex (female)	122 (43.4)
Nationality	
Spanish	242 (86.1)
South/Central American	5 (1.8)
Moroccan	13 (4.6)
Romanian	20 (7.1)
other	1 (0.4)
Age ≤ 28 days	34 (12.1)
Age ≤ 3 months	161 (57.3)
Age ≤ 6 months	230 (81.9)
Age (months)	2.5 (1.5–5)
CXR performed	129 (45.9) <sup>a</sup>
RSV-positive	158 (60) <sup>b</sup>
SaO <sub>2</sub> ≤ 90% at admission	55 (39.3) <sup>c</sup>
Required PICU admission	7 (2.5)
Required ABX	35 (12.5)
LOS	4.48 (2.37)

ABX, antibiotherapy; CXR, chest radiograph; LOS, length of stay; PICU, paediatric intensive care unit; RSV, respiratory syncytial virus; SaO<sub>2</sub>, oxygen saturation.

<sup>a</sup> In some patients, more than one chest radiograph was performed during the hospital stay.

<sup>b</sup> Tests for detection of RSV in nasal secretions performed in 93.6% of patients.

<sup>c</sup> Although oxygen saturation was measured in every patient, it was only documented in 49.8% of cases.

Assuming an expected value of 0.4, with a proportion of positives of 0.8 in observer 1 and 0.5 in observer 2, with a 95% level of confidence and a sample size of 140, the precision of the obtained kappa coefficient would be 0.903.

## Results

At least 1 CXR was ordered in 129 of the 281 hospitalised infants (45.9%). One radiograph was ordered in 120 infants, two in 7 infants and three in 2 infants. A total of 140 CXRs were performed and included in the analysis. Table 1 summarises the main characteristics of the sample under study. When it came to the technical quality of the radiographs, we found adequate inspiration in 96.4% (95% CI, 94.9–97.8), adequate penetration in 90.7% (95% CI, 86.5–94.9) and inclusion of the full ribcage in the image in 86.4% (95% CI, 81.1–91.7). However, only 22.1% (95% CI, 15.0–29.2) were properly centred. Fig. 1 presents the assessment of the presence or absence of the different radiographic signs made by the reference radiologist for the total sample of CXRs. The PR found radiographic signs compatible with complicated AB in 8.5% of the images (95% CI, 3.7–13.3). There were signs compatible with alternative diagnoses other than AB in 2.1% of the radiographs (95% CI, 0.3–4.6), in 2 due to evidence of pneumothorax and in 1 due to indirect signs suggestive of the presence of a foreign body in the airway. Lastly, 92.8% (95% CI, 89.3–96.3) of the radiographs only featured signs compatible with simple or uncomplicated AB and 5.7% (95% CI, 1.7–9.6) were considered normal. For most of the signs under study, the strength of the agreement between the different providers was fair, with kappa values ranging from 0.20 to 0.40, as can be seen in Fig. 2. We only found

slight agreement for the hyperinflation sign, with kappa values ranging from 0.00 to 0.20. Specifically, we found a kappa of 0.17 (0.15–0.18) for paediatricians, a kappa of 0.05 (0.02–0.07) for emergency physicians and a kappa of 0.02 (≤ 0.00–0.04) for medical residents. The strongest agreement corresponded to the foreign body sign, with a kappa of 0.32 (0.32–0.32) for paediatricians, a kappa of 0.34 (0.34–0.34) for emergency physicians and a kappa of 0.32 (0.32–0.33) for medical residents (Fig. 2). We also assessed the strength of agreement between each of the 11 observers (5 paediatricians, 3 emergency physicians and 3 medical residents) and the PR. Fig. 3 shows that all the paediatricians but one (P5) were in moderate agreement with the radiologist for most of the signs, with kappa values between 0.40 and 0.60. Nevertheless, and while one of the most experienced paediatricians (P3) exhibited a strong agreement with the PR, the next most experienced paediatrician (P4) ranked fourth when it came to the strength of agreement with the PR, as P1 and P2 exhibited greater agreement with the radiologist. Emergency physicians achieved kappa values that ranged from 0.20 to 0.40, and 2 of the 3 medical residents exhibited weaker agreement compared to the emergency physicians, with kappa values ranging from –0.20 to 0.20. The kappa values for the remaining medical resident were comparable to the kappa values of the paediatricians. Furthermore, and overall, all professionals exhibited lower agreement with the radiologist in the detection of radiographic signs of uncomplicated AB compared to other signs (complicated AB, other possible diagnoses and normal radiographic findings) (Fig. 3).

## Discussion

In our study, 77.8% of the CXRs were not properly centred in the opinion of the PR. This is one of the basic characteristics used to assess the technical quality of the image, and poor alignment is a frequent cause of false interpretations in paediatric chest radiographs.<sup>23</sup> Our findings regarding the proportion of radiographs with signs compatible with complicated AB, which was of 8.5% in our sample, or suggestive of an alternative diagnosis, which corresponded to 2.1% of the radiographs, were similar to those of other authors. Thus, Schuh et al. reported a proportion of radiographs compatible with complicated AB of 6.9% and a proportion suggestive of alternative diagnoses of 0.7% in a sample of 265 radiographs of patients with AB.<sup>24</sup> A similar study in 140 patients found a percentage compatible with complicated AB of 16% and a percentage compatible with other diagnoses of 0.7%.<sup>25</sup> This already suggests that multiple CXRs and exposure of many infants with a clinical diagnosis of AB to radiation is required to identify signs that would change the diagnosis or the approach to treatment. Studies that analyse interobserver agreement are quite heterogeneous not only in terms of the patients and disorders that they focus on, but also in the radiographic signs chosen to assess the specialists interpreting the radiographs and their level of professional experience, but also the methodology used, for instance the comparison of only 2 observers or 3 or more observers at the same time.<sup>13,14,26,27</sup>

Under these circumstances, studies yield highly heterogeneous results that are difficult to compare, and some

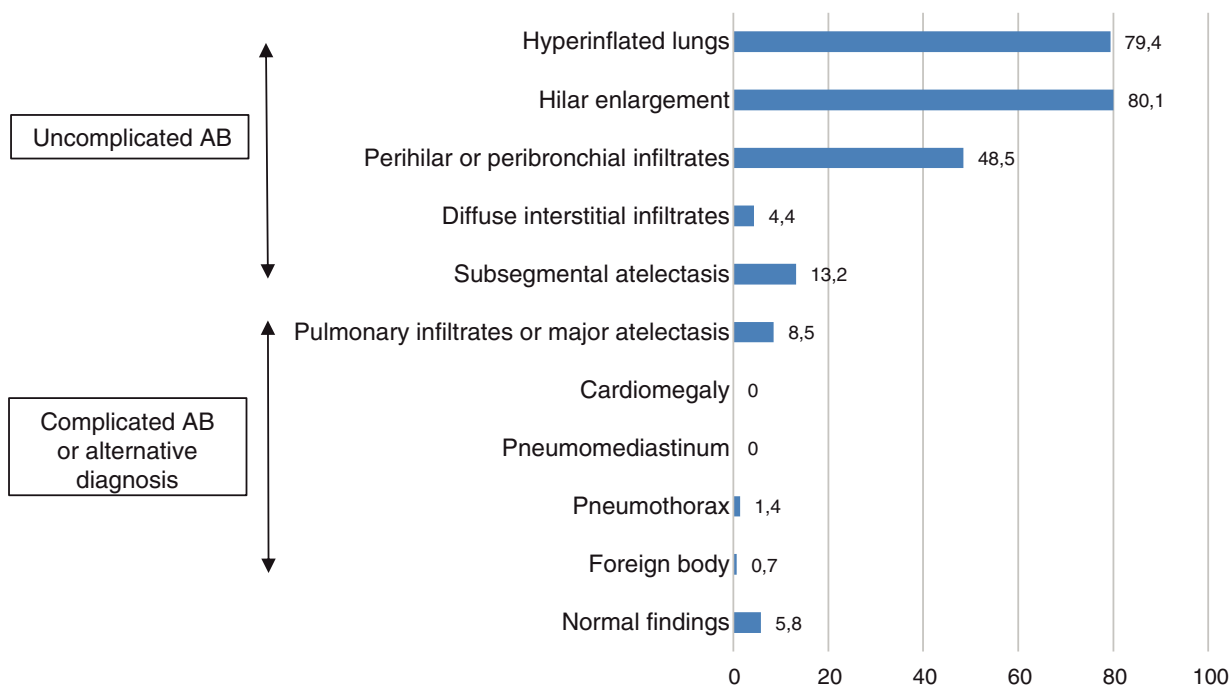


Figure 1 Evaluation by the radiologist of the presence of the different radiographic signs.

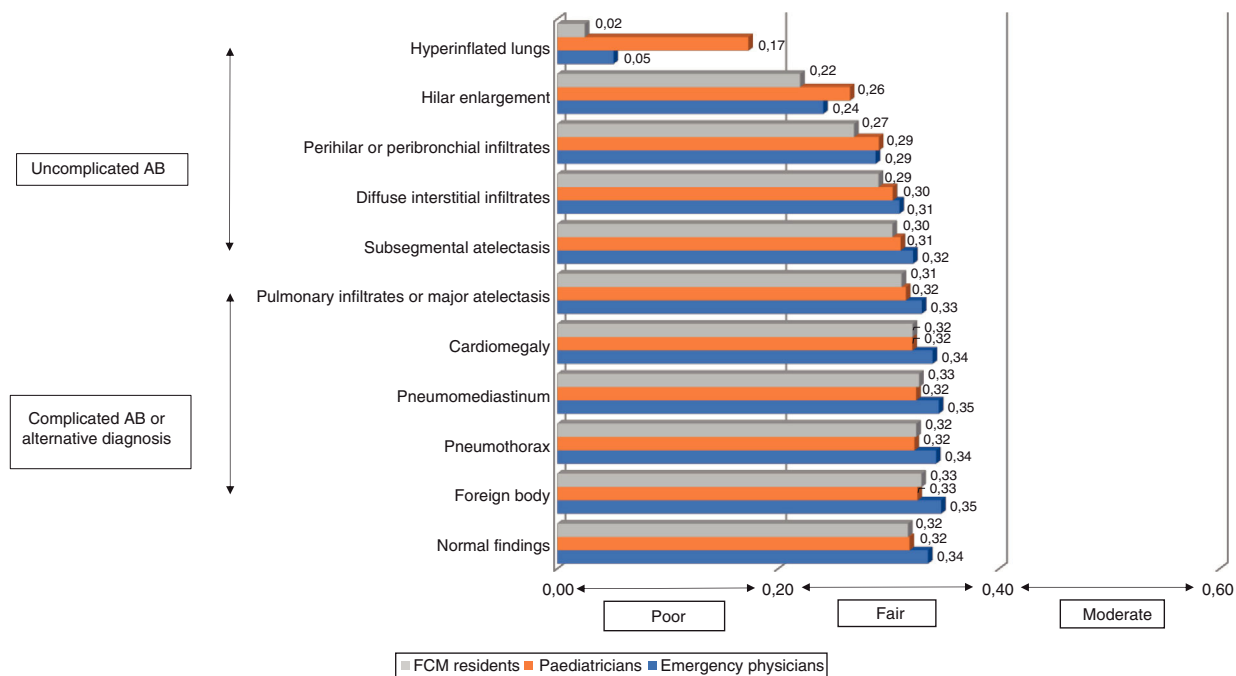
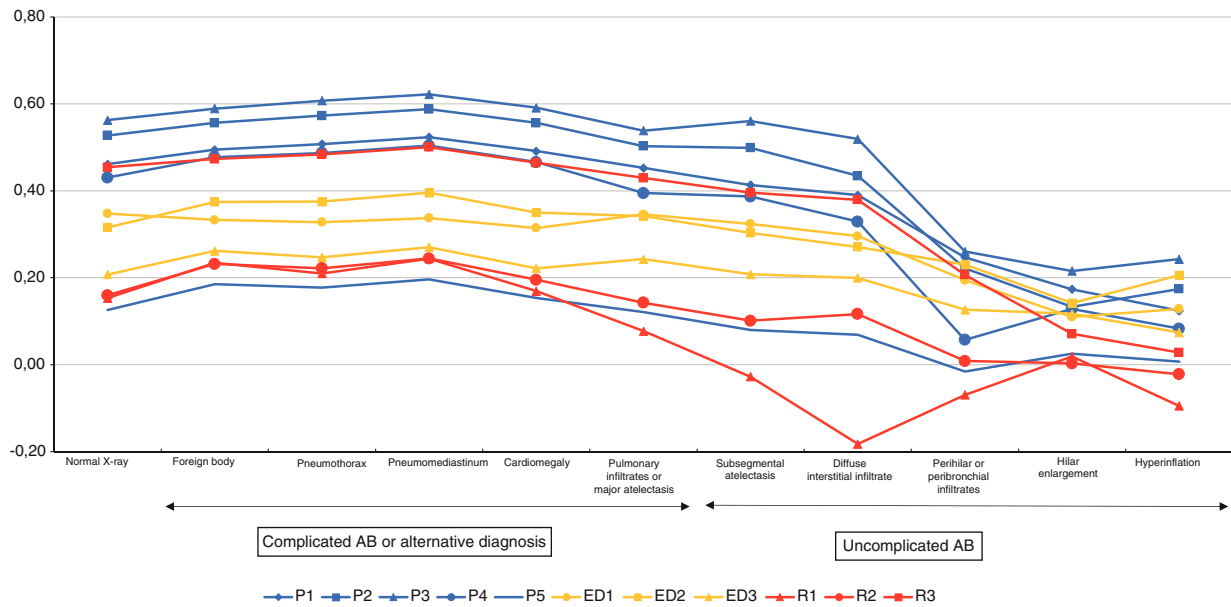


Figure 2 Strength of agreement, measured by means of the Fleiss kappa, between observers grouped by professional category: 5 paediatricians, 3 emergency physicians and 3 medical residents. The strength of agreement categories are shown in the horizontal axis.

AB, acute bronchiolitis.

report considerable agreement between observers<sup>10,14</sup> while others report moderate or fair levels of agreement with kappa values ranging from 0.20 to 0.60,<sup>13,15,27</sup> similar to the levels that we found. In our study, when we analysed the agreement within each group of medical providers, we found

fair agreement with Fleiss kappa values within a narrow range from 0.20 to 0.40 for all signs but one, as can be seen in Fig. 2. However, we found greater variability in the strength of agreement, with Cohen kappa values ranging from -0.20 to 0.60, when we compared the agreement between each



**Figure 3** Interobserver agreement measured with the Cohen kappa for 2 observers. We compared each of the professionals (5 paediatricians, 3 emergency physicians and 3 medical residents) with the paediatric radiologist that served as reference.

professional and the paediatric radiologist, as can be seen in Fig. 3. Our findings were consistent with those of Lewinsky et al.,<sup>15</sup> as we found an association between the strength of the agreement with the reference and the speciality of the observer (Fig. 3), with an overall stronger agreement in paediatricians compared emergency physicians and medical residents. Other authors have suggested that the substantial interrater variability observed could be attributed to the different levels of clinical experience,<sup>15,27</sup> so that more experienced observers would exhibit stronger agreement with the reference.<sup>15</sup> However, our findings do not support this hypothesis, as the strength of agreement with the reference found in paediatric providers was not directly proportional to their level of clinical experience, as one of the paediatricians with substantial experience (P4) exhibited a lower agreement compared to 3 of the 5 paediatricians in the study (Fig. 3). There are limitations to our study, such as the observers knowing from the outset that the radiographs that they would be interpreting corresponded to patients with AB. Although this could have resulted in an increased agreement between observers, this was not the case. Furthermore, our study focused on the reliability of this diagnostic test through the analysis of interobserver agreement, and the information would have been more comprehensive if we had also analysed intraobserver agreement. We also ought to highlight the fact that only one radiologist was used for reference, although she was the only full-time paediatric radiologist in the hospital, and therefore seemed the best possible choice for the purpose of the study. There are also strengths to our study, such as the homogeneous sample of patients, all of them infants aged less than 12 months admitted to hospital with bronchiolitis, and collection of data throughout 11 epidemic seasons. We have found only one other study in the previous literature analysing the reliability of CXR exclusively in infants with AB.<sup>11</sup>

In our study, we found poor to fair interobserver agreement in the interpretation of CXR used for diagnosis of AB in infants. We also described factors that may be related to the low reproducibility observed in our study, such as the quality of the image, the radiographic signs under assessment, the medical speciality of the observer, the clinical experience of the observer and the method used to assess interobserver agreement, comparing 2 observers or 3 or more observers at a time. Based on these findings, CXR offers a low reliability as a diagnostic test in the assessment of infants with AB, and therefore its routine use is not recommended.

## Conflicts of interest

None of the researchers have any conflicts of interest to declare. In addition, none have received a grant or any other form of external funding to carry out this study.

## Acknowledgments

We thank the family medicine residents, Juan Manuel Sánchez, Francisco Javier Cordón and Francisco Alonso, the emergency department physicians, Luís Fernández, Diana Moya and David García, and the physicians in the department of diagnostic radiology, Lourdes Hernández and Daniel Soliva, all of them affiliated to the Hospital Virgen de la Luz in Cuenca, Spain, for their voluntary and unpaid participation in the interpretation of the chest radiographs included in this study.

We also thank our colleague in the department of information technology, María Victoria Carrasco, for managing the blind interpretation of the radiographs by the different health care staff.

Last of all, I thank Rosa Josefina Bertolín Bernades and our children, Juanma and Pedro, for the family time lost to the study, which cannot ever be replaced.

## Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.anpedi.2020.03.011>.

## References

- Bhuiyan MU, Luby SP, Alamgir NI, Homaira N, Sturm-Ramirez K, Gurley ES, et al. Costs of hospitalization with respiratory syncytial virus illness among children aged <5 years and the financial impact on households in Bangladesh, 2010. *J Glob Health*. 2017;7:010412, <http://dx.doi.org/10.7189/jogh.07.010412>.
- Kramer R, Duclos A, Lina B, Casalegno JS, VRS study group in Lyon. Cost and burden of RSV related hospitalisation from 2012 to 2017 in the first year of life in Lyon, France. *Vaccine*. 2018;36:6591–3.
- Reeves RM, Hardeid P, Panagiotopoulos N, Minaji M, Warburton F, Pebody R. Burden of hospital admissions caused by respiratory syncytial virus (RSV) in infants in England: a data linkage modelling study. *J Infect*. 2019, <http://dx.doi.org/10.1016/j.jinf.2019.02.012> [Epub ahead of print].
- Burstein B, Plint AC, Papenburg J. Use of radiography in patients diagnosed as having acute bronchiolitis in US Emergency Departments, 2007-2015. *JAMA*. 2018;320:1598–600.
- Friedman JN, Rieder MJ, Walton JM, Canadian Paediatric Society, Acute Care Committee. Drug Therapy and Hazardous Substances Committee. Bronchiolitis: recommendations for diagnosis, monitoring and management of children one to 24 months of age. *Paediatr Child Health*. 2014;19:485–98.
- Ralston SL, Lieberthal AS, Meissner HC, Alverson BK, Baley JE, Gadomski AM, et al. Clinical Practice Guideline: the diagnosis, management, and prevention of bronchiolitis. *Pediatrics*. 2014;134:e1474–502.
- Florin TA, Plint AC, Zorc JJ. Viral bronchiolitis. *Lancet*. 2017;389:211–24.
- Breakell R, Thorndyke B, Clennett J, Harkensee C. Reducing unnecessary chest x-rays, antibiotics and bronchodilators through implementation of the NICE bronchiolitis guideline. *Eur J Pediatr*. 2018;177:47–51.
- Gómez González C, Pérez Castán JF. Curso de introducción a la investigación clínica. Capítulo 8: Pruebas diagnósticas. *Concordancia. SEMERGEN*. 2007;33:509–19.
- Swingler GH. Observer variation in chest radiography of acute lower respiratory infections in children: a systematic review. *BCM Med Imaging*. 2001;1:1.
- Coblentz CL, Babcock CJ, Alton D, Riley BJ, Norman G. Observer variation in detecting the radiologic features associated with bronchiolitis. *Invest Radiol*. 1991;26:115–8.
- McConnochie KM. Bronchiolitis. What's in the name? *Am J Dis Child*. 1983;137:1–3.
- Correia MA, Mello MJ, Petribú NC, Silva EJ, Bezerra PG, Duarte MC, et al. Agreement on radiological diagnosis of acute lower respiratory tract infection in children. *J Trop Pediatr*. 2011;57:204–7.
- Xavier-Souza G, Vilas-Boas AL, Fontoura MS, Araújo-Neto CA, Andrade SC, Cardoso MR, et al. PNEUMOPAC-Efficacy Study Group. The inter-observer variation of chest radiograph reading in acute lower respiratory tract infection among children. *Pediatr Pulmonol*. 2013;48:464–9.
- Levinsky Y, Mimouni FB, Fisher D, Ehrlichman M. Chest radiography of acute paediatric lower respiratory infections: experience versus interobserver variation. *Acta Paediatr*. 2013;102:e310–4.
- Berman L, De Lacey G, Morley S. Chest radiology: the basic basics. In: *The Chest X-ray*. Philadelphia: Saunders; 2008. p. 2–14.
- Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20:37–46.
- Abraira V. El índice kappa. *Semerger*. 2001;27:247–9.
- Abraira V, Pérez de Vargas A. Generalization of the Kappa coefficient for ordinal categorical data, multiple observers and incomplete designs. *Questiò*. 1999;23:561–71.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–74.
- de Souza TH, Nadal JAH, Peixoto AO, Pereira RM, Giatti MP, Soub ACS, et al. Lung ultrasound in children with pneumonia: inter-operator agreement on specific thoracic regions. *Eur J Pediatr*. 2019;178:1369–77.
- Cantor AB. Sample-size calculations for Cohen's Kappa. *Psychological Methods*. 1996;1:150–3.
- García Asensio D, Fernández Martín M. Pitfalls y variantes del tórax pediátrico. *Radiologia*. 2016;58:58–69.
- Schuh S, Lalani A, Allen U, Manson D, Babyn P, Stephens D, et al. Evaluation of the utility of radiography in acute bronchiolitis. *J Pediatr*. 2007;150:429–33.
- Farah MM, Padgett LB, McLario DJ, Sullivan KM, Simon HK. First-time wheezing in infants during respiratory syncytial virus season: chest radiograph findings. *Pediatr Emerg Care*. 2002;18:333–6.
- Williams GJ, Macaskill P, Kerr M, Fitzgerald DA, Isaacs D, Codarini M, et al. Variability and accuracy in interpretation of consolidation on chest radiography for diagnosing pneumonia in children under 5 years of age. *Pediatr Pulmonol*. 2013;48:1195–200.
- Johnson J, Kline JA. Intraobserver and interobserver agreement of the interpretation of pediatric chest radiographs. *Emerg Radiol*. 2010;17:285–90.