

# Predictive Value of Clinician “Gestalt” in Pediatric Community-Acquired Pneumonia

Hans M. Gao, BS,<sup>a</sup> Lilliam Ambroggio, PhD, MPH,<sup>b</sup> Samir S. Shah, MD, MSCE,<sup>c</sup> Richard M. Ruddy, MD,<sup>d</sup> Todd A. Florin, MD, MSCE<sup>a,e</sup>

abstract

**OBJECTIVES:** Validated prognostic tools for pediatric community-acquired pneumonia (CAP) do not exist. Thus, clinicians rely on “gestalt” in management decisions for children with CAP. We sought to determine the ability of clinician gestalt to predict severe outcomes.

**METHODS:** We performed a prospective cohort study of children 3 months to 18 years old presenting to a pediatric emergency department (ED) with lower respiratory infection and receiving a chest radiograph for suspected CAP from 2013 to 2017. Clinicians reported the probability that the patient would develop severe complications of CAP (defined as respiratory failure, empyema or effusion, lung abscess or necrosis, metastatic infection, sepsis or septic shock, or death). The primary outcome was development of severe complications.

**RESULTS:** Of 634 children, 37 (5.8%) developed severe complications. Of children developing severe complications after the ED visit, 62.1% were predicted as having <10% risk by the ED clinician. Sensitivity was >90% at the <1% predicted risk threshold, whereas specificity was >90% at the 10% risk threshold. Gestalt performance was poor in the low-intermediate predicted risk category (1%–10%). Clinicians had only fair ability to discriminate children developing complications from those who did not (area under the receiver operator characteristic curve 0.747), with worse performance from less experienced clinicians (area under the receiver operator characteristic curve 0.693).

**CONCLUSIONS:** Clinicians have only fair ability to discriminate children with CAP who develop severe complications from those who do not. Clinician gestalt performs best at very low or higher predicted risk thresholds, yet many children fall in the low-moderate predicted risk range in which clinician gestalt is limited. Evidence-based prognostic tools likely can improve on clinician gestalt, particularly when risk is low-moderate.



<sup>a</sup>Feinberg School of Medicine, Northwestern University, Chicago, Illinois; <sup>b</sup>Sections of Emergency Medicine and Hospital Medicine, Department of Pediatrics, Children's Hospital Colorado and University of Colorado, Aurora, Colorado; <sup>c</sup>Divisions of Hospital Medicine and Infectious Diseases and <sup>d</sup>Emergency Medicine, Department of Pediatrics, Cincinnati Children's Hospital Medical Center and College of Medicine, University of Cincinnati, Cincinnati, Ohio; and <sup>e</sup>Division of Emergency Medicine, Department of Pediatrics, Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, Illinois

Mr Gao conceptualized and designed the study, performed statistical analysis and interpreted the data, drafted the initial manuscript, and reviewed and revised the manuscript; Drs Ambroggio, Ruddy, and Shah conceptualized and designed the study, supervised participant enrollment and data acquisition, participated in data interpretation, and reviewed and revised the manuscript; Dr Florin conceptualized and designed the study, supervised participant enrollment and data acquisition, performed statistical analysis and interpreted the data, reviewed and revised the manuscript, and provided overall study supervision; and all authors approved the final manuscript as submitted.

**DOI:** <https://doi.org/10.1542/peds.2020-041582>

**WHAT'S KNOWN ON THIS SUBJECT:** Validated prognostic tools do not exist for children with community-acquired pneumonia, leaving clinicians to rely on gestalt for decision-making. Data on the predictive ability of clinician gestalt for clinical outcomes in children with respiratory infections are limited.

**WHAT THIS STUDY ADDS:** Clinicians have fair ability to discriminate children with community-acquired pneumonia who develop severe complications. Clinician gestalt performs best at very low or higher predicted risk, yet many children fall in the low-intermediate predicted risk range in which clinician gestalt is more limited.

**To cite:** Gao HM, Ambroggio L, Shah SS, et al. Predictive Value of Clinician “Gestalt” in Pediatric Community-Acquired Pneumonia. *Pediatrics*. 2021;147(5):e2020041582

As the second most costly and fifth most prevalent reason for pediatric hospitalization in the United States, community-acquired pneumonia (CAP) remains a significant contributor to health care burden and expenditures.<sup>1-4</sup> Although most children who develop pneumonia fully recover after nonsevere illness, some become severely ill and develop serious complications.<sup>5</sup> No validated prognostic tools exist for children with CAP, leaving clinicians to rely largely on clinical gestalt in making management decisions.

Clinician gestalt informs management decisions on the basis of patient history and physical examination findings.<sup>6,7</sup> Gestalt relies on the ability of the physician to recognize patterns of disease, such as fever, hypoxia, and tachypnea, to distinguish between severe and nonsevere disease.<sup>8</sup> Few studies have been used to examine the predictive value of a clinician's gestalt for subsequent clinical outcomes.<sup>9</sup> One prospective study in the primary care setting revealed that a clinician's gut feeling that a child is seriously ill substantially increased the chance that a severe infection was present, even when the clinical appearance of the child was reassuring.<sup>10</sup> However, authors of another prospective study in a similar setting found no association between gestalt and disease prognosis in children with respiratory tract infections.<sup>11</sup> Given these conflicting results and the lack of evidence-based prognostic tools, it is important to further understand the ability of clinician gestalt to predict severe complications in children with CAP, particularly in the emergency department (ED) setting in which most disposition decisions are made. We sought to determine if clinician gestalt is an accurate predictor of severe outcomes in children with suspected CAP who present to the ED.

## METHODS

### Study Design

This study is a planned analysis from a prospective cohort study of children with suspected CAP, called Catalyzing Ambulatory Research in Pneumonia Etiology and Diagnostic Innovations in Emergency Medicine (CARPE DIEM).<sup>12-14</sup> Patients who presented to the ED at Cincinnati Children's Hospital Medical Center (CCHMC) and enrolled in CARPE DIEM from July 2013 to December 2017 were eligible for inclusion. The study was approved by the CCHMC Institutional Review Board. Informed consent was obtained from all legal guardians, and assent was obtained from children  $\geq 11$  years of age.

### Study Population

Children 3 months to 18 years of age with signs and symptoms of lower respiratory tract infection who received a chest radiograph (CXR) for clinical suspicion of CAP were enrolled.<sup>12,14-16</sup> We excluded children hospitalized  $\leq 14$  days before the study ED visit to exclude possible hospital-acquired pneumonia. Children with immunocompromising or chronic medical conditions known to predispose to severe or recurrent pneumonia (eg, immunodeficiency, chronic corticosteroid use, cystic fibrosis, chronic lung disease, malignancy, sickle cell disease, congenital heart disease, tracheostomy, and neuromuscular disorders) were not included, nor were children with a history of aspiration or aspiration pneumonia. Patients who were enrolled within 30 days before the study ED visit were excluded to ensure a distinct episode of infection.

### Study Procedures

After informed consent, trained research coordinators recorded demographic, historical, and examination findings. After their examination, the treating clinician

was asked about their clinical impressions, including probability of the child developing severe complications of CAP. Responses were recorded on a standardized case report form. At the end of the ED visit, the treating clinicians were asked to complete questions regarding disposition and their final clinical impressions after all data were considered. Severity outcomes occurring after the ED visit were assessed through abstraction from the electronic health record and manual record review.

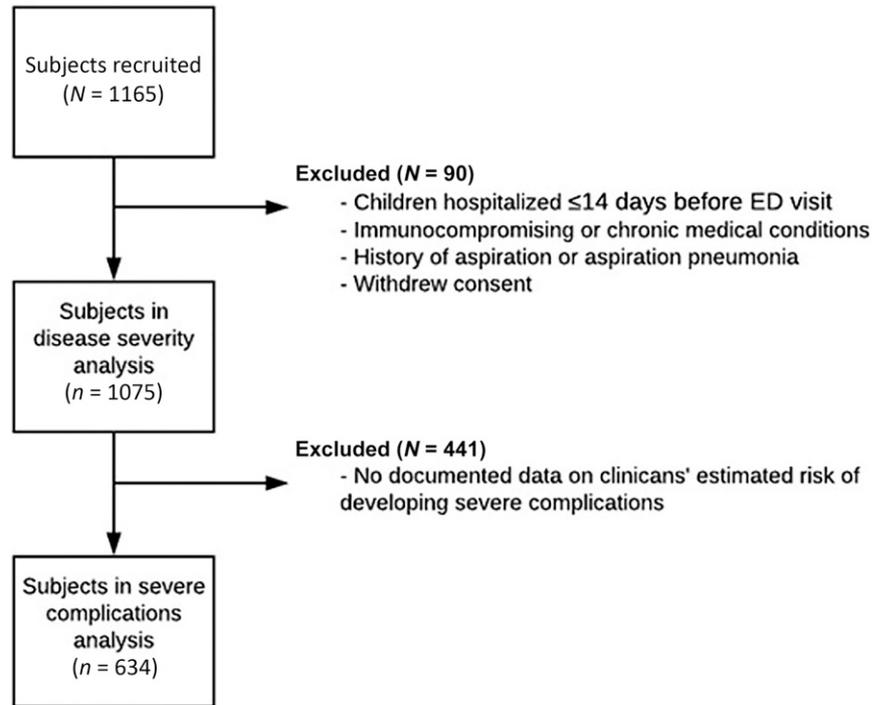
### Measurements and Outcomes

The primary exposure variable for this analysis was the treating clinician's gestalt after reviewing the patient history, laboratory results, and imaging findings, as measured in 2 ways. First, for patients with CXR findings suspicious for radiographic pneumonia, clinicians were asked to "estimate the probability they will develop severe disease or complications of pneumonia." Categorical response options included  $<1\%$ ,  $1\%$  to  $5\%$ ,  $5\%$  to  $10\%$ ,  $10\%$  to  $25\%$ ,  $25\%$  to  $50\%$ ,  $50\%$  to  $75\%$ , or  $75\%$  to  $100\%$ . Clinicians were provided a definition of "severe disease or complications" that included respiratory failure (with or without intubation), empyema or effusion, lung abscess or necrosis, disseminated infection to other sites (eg, osteomyelitis, meningitis), sepsis or septic shock, or death. Second, clinicians for all patients in CARPE DIEM, regardless of CXR findings, were asked, "What is your overall clinical impression of this participant?" with the response options of "mild," "moderate," "severe," and "very severe." These questions were asked at the completion of the ED visit when the clinician had all clinical, laboratory, and imaging data available to them.

The primary outcome for this study was the development of severe

disease or complications, defined identically as presented to the ED clinician (ie, respiratory failure [with or without intubation], empyema or effusion, lung abscess or necrosis, disseminated infection to other sites, sepsis or septic shock, or death). Participants with any of these outcomes at the time of the ED visit were excluded from this analysis.

Our secondary outcome was a composite outcome representing increasing disease severity occurring within 7 days of the study ED visit as a 4-tiered ordinal variable: mild, moderate, severe, and very severe disease. Mild disease was defined as discharge from the ED without return for hospitalization within 7 days. Moderate disease was defined as those hospitalized on initial visit or after revisit within 7 days but not meeting severe or very severe criteria. Severe disease was classified as hospitalization with at least 1 of the following: at least 1 intravenous (IV) fluid bolus, continuous IV fluids for >12 hours, supplemental oxygen, broadening of antibiotics from aminopenicillin to any other antibiotic class, complicated pneumonia (moderate-large pleural effusion, metastatic infection associated with pneumonia, lung abscess, or lung necrosis), or presumed sepsis (systemic inflammatory response syndrome with receipt of antibiotics and  $\geq 40$  mL/kg of IV bolus fluid). Very severe disease required at least 1 of the following: treatment in the ICU, positive-pressure ventilation (including continuous positive-pressure ventilation, bilevel positive-pressure ventilation, or endotracheal intubation with mechanical ventilation), vasoactive infusions, chest drainage, extracorporeal membrane oxygenation, severe sepsis or septic shock (using validated *International Classification of Diseases, Ninth Revision* diagnosis codes), or death.<sup>17,18</sup>



**FIGURE 1**  
Study flow diagram.

### Statistical Analysis

The association of the ED clinician's predicted probability of complications and proportions of patients at each risk threshold who developed the primary outcome of severe complications was compared by using Fisher's exact test with  $P < .05$  accepted for statistical significance. Each reported predicted risk threshold was used as a dichotomous cutoff (above or below the reported threshold), for which sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated. A receiver operator characteristic curve was calculated for the ability of clinical gestalt to discriminate those who go on to develop severe complications from those who do not.<sup>19</sup> The area under the receiver operator characteristic curve (AUC) was calculated. Additional analyses were performed, stratified by clinician years of experience posttraining ( $\leq 5$  or  $>5$  years), patient age ( $\leq 5$  or

$>5$  years old), and days of illness ( $\leq 4$  or  $>4$  days). Statistical analyses were performed in Stata (release 16; Stata Corp, College Station, TX) and R statistical software (version 3.6.1; R Foundation for Statistical Computing, Vienna, Austria).

### RESULTS

#### Cohort Characteristics

Out of the 1142 children enrolled in CARPE DIEM, clinicians documented their overall impressions on 1075 children (Fig 1). Of those, predictions about severe complications of CAP, our primary outcome, were recorded for 634 who had findings on CXR suspicious for radiographic pneumonia. The median age was 3.3 years (interquartile range: 1.4–7.1) (Table 1). Most were boys (54.7%), and most had their race reported by parents or guardians as white (62.8%).

**TABLE 1** Cohort Characteristics

Patient Characteristics	Overall Study Population ( <i>n</i> = 1075), <i>n</i> (%)	Children With Suspected Radiographic CAP ( <i>n</i> = 634), <i>n</i> (%)
Age, y		
Median (IQR)	3.3 (1.4–7.1)	5.2 (1.7–7.9)
Sex		
Male	588 (54.7)	347 (54.7)
Female	487 (45.3)	287 (45.3)
Race		
White	675 (62.8)	410 (64.7)
Black	325 (30.2)	180 (28.4)
Other	75 (7.0)	44 (6.9)
Disease severity		
Mild	517 (48.1)	246 (38.8)
Moderate	243 (22.6)	147 (23.2)
Severe	242 (22.5)	191 (30.1)
Very severe	73 (6.8)	50 (7.9)
Disposition		
Discharge from the hospital	530 (49.3)	255 (40.2)
Admit to floor	498 (46.3)	348 (54.9)
Admit to ICU	47 (4.4)	31 (4.9)

IQR, interquartile range.

### Predicting Severe Complications

Most children with suspected CAP on CXR (*n* = 468; 73.8%) were given a predicted probability of <5% of developing severe complications by their treating ED clinician (Table 2). Of 634 children with suspected radiographic CAP, 37 (5.8%) experienced severe complications. Overall, clinician gestalt was associated with the development of severe complications (*P* < .01); however, in those who developed complications, clinicians tended to underestimate risk. Of the 37 children who developed severe complications, most (*n* = 25; 67.6%) were initially estimated to have a ≤10% chance of developing such complications.

Overall, physicians were more accurate when predicted probability was very low. For example, of those children with predicted risk of 1% to 5%, 5.8% experienced severe complications, but when predicted risk was 50% to 75%, 37.5% experienced severe complications. When examining test characteristics of various thresholds of predicted risk, clinicians predicting <1% probability of severe complications demonstrated high sensitivity and NPV (Table 3). At all remaining thresholds, sensitivity was poor to fair. Clinicians demonstrated high specificity of >90% at risk thresholds of 10% or greater. More experienced clinicians had slightly improved

**TABLE 2** Association of ED Clinician Predicted Probability of Developing Severe Complications and Development of Severe Complications

Predicted Probability of Developing Severe Complications by ED Clinician, %	Total, <i>n</i> (%)	No Severe Complications ( <i>n</i> = 597), <i>n</i> (%)	Severe Complications ( <i>n</i> = 37), <i>n</i> (%)
<1	226 (35.7)	224 (37.5)	2 (5.4)
1–5	242 (38.2)	228 (38.2)	14 (37.8)
5–10	93 (14.7)	86 (14.4)	7 (18.9)
10–25	38 (6.0)	34 (5.7)	4 (10.8)
25–50	12 (1.9)	10 (1.7)	2 (5.4)
50–75	8 (1.3)	5 (0.8)	3 (8.1)
75–100	15 (2.4)	10 (1.7)	5 (13.5)

Fisher's exact test: *P* < .01.

sensitivity for severe complications, with similar specificity. The overall AUC was 0.747, with an AUC of 0.693 for clinicians with ≤5 years of experience and 0.82 for those with >5 years of experience (Fig 2). Results did not change by patient age or illness duration.

### Predicting Disease Severity

When asked their overall clinical impression, no clinician classified children as very severe. Children classified as severe by gestalt (*n* = 25) were likely to develop severe disease (*n* = 6; 24%) or very severe disease (*n* = 16; 64%) (Supplemental Table 4). In children classified as mild (*n* = 617), few children developed severe (*n* = 58; 9.4%) or very severe (*n* = 5; 0.8%) disease. Whereas physicians were more accurate at predicting disease severity in children classified as mild and severe by gestalt, accuracy was lower for those classified as moderate. Of the 433 children classified as moderate by gestalt, 10.6% developed only mild disease, whereas 53.1% developed severe or very severe disease. Of 242 children who developed severe disease, most (*n* = 178; 73.6%) were initially classified as moderate by gestalt. Similarly, out of the 73 children who developed very severe disease, most (*n* = 52; 71%) were initially classified as moderate by gestalt (Supplemental Fig 3).

### DISCUSSION

In this prospective cohort study, although we found a significant association between clinician gestalt and disease severity in children with suspected CAP, ED clinicians had only fair ability to discriminate those who go on to develop complications from those who do not. More experienced clinicians had higher discriminatory capability over those with less experience. In predicting the development of severe complications, physicians were most sensitive when their predicted risk was <1%. As

**TABLE 3** Test Characteristics for Clinician Predicted Probability and Subsequent Development of Severe Complications of Pneumonia

Threshold, %	Sensitivity	Specificity	PPV	NPV
All clinicians (N = 634)				
<1	94.6 (81.8–99.3)	37.5 (33.6–41.5)	8.6 (6.1–11.7)	99.1 (96.8–99.9)
<5	56.8 (39.5–72.9)	75.7 (72.1–79.1)	12.7 (8–18.7)	96.6 (94.5–98)
<10	37.8 (22.5–55.2)	90.1 (87.4–92.4)	19.2 (10.9–30.1)	95.9 (93.9–97.4)
<25	27 (13.8–44.1)	95.8 (93.9–97.3)	28.6 (14.6–46.3)	95.5 (93.5–97)
<50	21.6 (9.8–38.2)	97.5 (95.9–98.6)	34.8 (16.4–57.3)	95.3 (93.3–96.8)
<75	13.5 (4.5–28.8)	98.3 (97–99.2)	33.3 (11.8–61.6)	94.8 (92.8–96.4)
Clinicians with ≤5 y of experience (n = 334; 52.7%)				
<1	90.9 (58.7–99.8)	38.1 (28.5–48.6)	14.3 (7.1–24.7)	97.4 (86.2–99.9)
<5	45.5 (16.7–76.6)	77.3 (67.7–85.2)	18.5 (6.3–38.1)	92.6 (84.6–97.2)
<10	27.3 (6.02–61)	91.8 (84.4–96.4)	27.3 (6.02–61)	91.8 (84.4–96.4)
<25	9.09 (0.23–41.3)	95.9 (89.8–98.9)	20.0 (50.5–71.6)	90.3 (82.9–95.2)
<50	9.09 (0.23–41.3)	96.9 (91.2–99.4)	25.0 (63.1–80.6)	90.4 (83.0–95.3)
<75	0.00 (0.00–28.5)	97.9 (92.7–99.7)	0.00 (0.00–84.2)	89.6 (82.2–94.7)
Clinicians with >5 y of experience (n = 282; 44.5%)				
<1	100 (79.4–100)	35.2 (29.7–41.1)	8 (4.64–12.7)	100 (96.4–100)
<5	75 (47.6–92.7)	74.3 (68.8–79.3)	14.1 (7.51–23.4)	98.1 (95.3–99.5)
<10	50 (24.7–75.3)	88 (83.7–91.6)	19 (8.6–34.1)	96.9 (94–98.7)
<25	37.5 (15.2–64.6)	95.4 (92.3–97.5)	31.6 (12.6–56.6)	96.4 (93.6–98.3)
<50	25 (7.27–52.4)	97.9 (95.5–99.2)	40 (12.2–73.8)	95.9 (92.9–97.8)
<75	25 (7.27–52.4)	98.2 (95.9–99.4)	44.4 (13.7–78.8)	95.9 (92.9–97.9)

All values presented as % (95% CI).

predicted risk increased, sensitivity decreased substantially. Gestalt had higher specificity when predicted risk was higher than 10%. Gestalt was generally most accurate when the clinician classified the patient as either mild or severe. The greatest degree of misclassification occurred in those initially classified as moderate by gestalt.

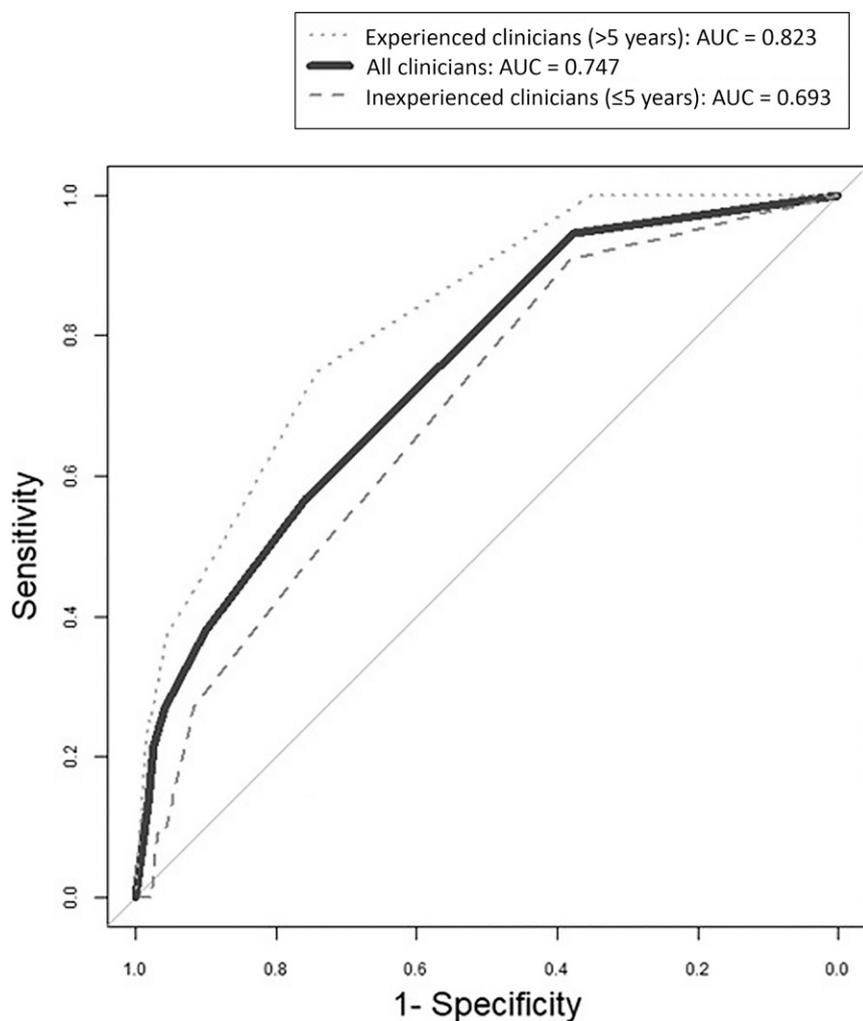
Our results are comparable to the performance of clinician gestalt in several other disease processes, including acute coronary syndrome, sinusitis, and pulmonary embolism. Clinician gestalt has fair diagnostic accuracy in acute coronary syndrome, with an AUC of 0.75 (95% confidence interval [CI]: 0.72–0.79), and sinusitis with an AUC of 0.73 (95% CI: 0.67–0.80).<sup>20,21</sup> Clinician gestalt revealed slightly better accuracy when diagnosing pulmonary embolism, with an AUC of 0.81 (95% CI: 0.78–0.84).<sup>22</sup>

Few studies have been used to examine the role of clinician gestalt in predicting outcomes in CAP. In a study of adults with CAP, researchers found that clinician gestalt is a strong predictor of

hospitalization, although clinicians tended to overestimate short-term mortality.<sup>23</sup> In another study, rates of hospitalization of adults with CAP varied greatly among emergency physicians; the variation was not associated with differing clinical outcomes and could not be explained by objective data.<sup>24</sup> In previous studies, researchers have also examined the accuracy of clinician gestalt in predicting the presence of radiographic pneumonia. In adults presenting to primary care with suspected CAP, CXR was most useful for those in whom the diagnosis was uncertain rather than those at the extremes of definitely having or not having pneumonia.<sup>25</sup> In children, similar results were found, in which CXR was least helpful in altering management in children with high clinical suspicion for CAP; however, if the clinical suspicion was low or uncertain, a CXR did guide clinical management.<sup>26</sup> When ED clinicians were asked to provide their predictions for the presence of radiographic pneumonia, similar to our study, they were most accurate at the extremes (very low suspicion or very high suspicion), whereas they

tended to overestimate the presence of pneumonia when predicted risk was in the middle.<sup>27</sup> Although these previous studies reveal similar results for the prediction of radiographic CAP, none were used to examine the ability of clinicians to prognosticate severe outcomes.

Our study revealed that clinician gestalt has fair ability to predict the development of complications of pneumonia. The NPV of clinician gestalt was high, suggesting that gestalt may be most useful for ruling out the development of complications. Although the NPV was high, the prevalence of severe complications was low. Sensitivity and specificity do not depend on disease prevalence. Sensitivity, although high at a risk threshold of <1%, drops off substantially as risk threshold moves to 5% and greater. Specificity increases when risk exceeds 10% to 25%. The combination of high sensitivity at very low risk thresholds and high specificity at high risk thresholds emphasizes the ability of clinician gestalt to effectively rule out complications at very low levels of predicted risk and to rule in at higher



**FIGURE 2**  
Receiver operator characteristic curves for ability of ED clinicians to predict severe complications of suspected CAP in children.

risk levels but also highlights the challenges of risk prediction when gestalt lies in a low-moderate category. In our study, 373 children, or 58.8% of our study population, had risks ranging from 1% to 25%. Therefore, this low-moderate risk group represents a significant proportion of children with CAP for whom clinician gestalt is not ideal.

Our results suggest that clinicians tend to underestimate CAP severity. Although 29.3% of children in our study developed severe disease or very severe disease, only 2.3% were classified as severe by clinician gestalt. A similar tendency for

clinicians to underestimate the severity and prevalence of pneumonia has also been observed in previous studies.<sup>28</sup> The tendency for clinicians to underestimate pneumonia complications may be related to the low overall prevalence of such complications because most children with pneumonia recover without any complications.<sup>5</sup> Furthermore, because pneumonia complications are rare, clinicians may have more difficulty identifying the patterns of illness associated with these complications.

With our study, we found that experienced physicians may be more

accurate at prognosticating children evaluated for CAP using gestalt than less experienced physicians. Sensitivity, PPV, NPV, and the AUC are all slightly higher in those with more experience, although CIs overlapped. Clinician gestalt depends on clinicians recognizing patterns of disease and making generalizations that allow transfer from one case to another. In previous studies, researchers have shown that clinician experience does improve diagnostic accuracy because experienced clinicians have better pattern recognition skills developed over time than their counterparts with fewer years of clinical experience.<sup>29</sup>

Risk thresholds and tolerance will vary by clinician. Although there are no established risk thresholds for complications in children with CAP, individual clinicians set their own acceptable risks while adjusting for patient characteristics. For instance, if clinicians were to set acceptable risk at 5%, discharging all patients with a <5% chance of developing complications, by gestalt, would result in discharge of 78.1% of patients, of whom 3.4% would develop complications. A more rigorous cutoff, at <1%, would discharge 37.7% of patients, of whom 0.9% would develop complications. Further study is warranted to examine how individual clinicians respond to various risk thresholds.

Overall, clinicians in our study had only fair ability to discriminate those children developing complications from those who did not (AUC 0.747), with worse performance in those clinicians with less experience (AUC 0.693). Consistent with recommendations of key areas for future research in national pediatric CAP guidelines, our results suggest that there is a need for evidence-based clinical prediction rules to assist with risk stratification.<sup>30</sup> Formal clinical prediction rules would be particularly helpful when clinical gestalt is less accurate, as in the case

of moderate disease or for less experienced clinicians, who are refining their clinical reasoning. Our results suggest that all clinicians could benefit from the use of clinical prediction rules to augment their clinical gestalt.<sup>31</sup>

Our study has several limitations. First, few patients in our sample developed very severe disease or complications. This finding, consistent with other studies, emphasizes the need for tools to assist with identifying the rare patient with severe outcomes from a larger pool of patients who generally have a mild disease course. Second, because our study did not track the decision-making process of the clinicians, we cannot provide insight into how gestalt is generated for each patient. Third, our secondary outcome, namely, disease severity, was not explicitly defined to the clinicians before the study, and criteria for each disease category used in our analysis may have differed from the criteria used by individual clinicians during our study. Fourth, sicker patients likely received aggressive treatment, and many potential complications may have been forestalled by emergent medical

care. Finally, our study was conducted in the ED of an urban pediatric tertiary care center. The findings may not be generalizable to other settings.

## CONCLUSIONS

Although an association exists between clinician gestalt and disease severity in pediatric CAP, ED clinicians generally have only fair ability to differentiate those children who go on to develop severe complications from those who do not. Clinician gestalt was highly sensitive when predicted risk was very low and more specific when predicted risk was higher. Clinicians generally did not perform well at predicting outcomes in those with low-moderate predicted risk, which accounts for a substantial proportion of children with CAP. There is thus a need to develop evidence-based clinical decision rules to supplement clinical judgment, particularly for cases in which risk may be unclear or as newer clinicians are developing their clinical acumen.

## ACKNOWLEDGMENTS

We acknowledge Judd Jacobs and Jessi Lipscomb for their role in data

management for the CARPE DIEM study. We are grateful to the entire research team and patient services staff in the divisions of emergency medicine and hospital medicine at CCHMC for their assistance with study procedures. Finally, we are especially grateful to the patients and families who enrolled in CARPE DIEM.

## ABBREVIATIONS

AUC: area under the receiver operator characteristic curve  
CAP: community-acquired pneumonia  
CARPE DIEM: Catalyzing Ambulatory Research in Pneumonia Etiology and Diagnostic Innovations in Emergency Medicine  
CCHMC: Cincinnati Children's Hospital Medical Center  
CI: confidence interval  
CXR: chest radiograph  
ED: emergency department  
IV: intravenous  
NPV: negative predictive value  
PPV: positive predictive value

Accepted for publication Jan 13, 2021

Address correspondence to Todd A. Florin, MD, MSCE, Division of Emergency Medicine, Ann & Robert H. Lurie Children's Hospital of Chicago, 225 E Chicago Ave, Box 62, Chicago, IL 60611. E-mail: tafloirin@luriechildrens.org

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275).

Copyright © 2021 by the American Academy of Pediatrics

**FINANCIAL DISCLOSURE:** The authors have indicated they have no financial relationships relevant to this article to disclose.

**FUNDING:** Supported by the National Institutes of Health National Institute of Allergy and Infectious Diseases (K23AI121325 to Dr Florin and K01AI125413 to Dr Ambroggio), the Gerber Foundation (to Dr Florin), National Institutes of Health National Center for Research Resources, and Cincinnati Center for Clinical and Translational Science and Training (5KL2TR000078 to Dr Florin). The funders did not have any role in the study design, data collection, statistical analysis, or article preparation. Funded by the National Institutes of Health (NIH).

**POTENTIAL CONFLICT OF INTEREST:** The authors have indicated they have no potential conflicts of interest to disclose.

**COMPANION PAPER:** A companion to this article can be found online at [www.pediatrics.org/cgi/doi/10.1542/peds.2020-048637](http://www.pediatrics.org/cgi/doi/10.1542/peds.2020-048637).

## REFERENCES

1. Keren R, Luan X, Localio R, et al.; Pediatric Research in Inpatient Settings (PRIS) Network. Prioritization of comparative effectiveness research topics in hospital pediatrics. *Arch Pediatr Adolesc Med*. 2012;166(12):1155–1164

2. Kronman MP, Hersh AL, Feng R, Huang YS, Lee GE, Shah SS. Ambulatory visit rates and antibiotic prescribing for children with pneumonia, 1994-2007. *Pediatrics*. 2011;127(3):411–418
3. Alpern ER, Stanley RM, Gorelick MH, et al.; Pediatric Emergency Care Applied Research Network. Epidemiology of a pediatric emergency medicine research network: the PECARN Core Data Project. *Pediatr Emerg Care*. 2006; 22(10):689–699
4. Walker CLF, Rudan I, Liu L, et al. Global burden of childhood pneumonia and diarrhoea. *Lancet*. 2013;381(9875): 1405–1416
5. Sandora TJ, Harper MB. Pneumonia in hospitalized children. *Pediatr Clin North Am*. 2005;52(4):1059–1081, viii
6. Sawe HR, Haeffele C, Mfinanga JA, Mwafongo VG, Reynolds TA. Predicting fluid responsiveness using bedside ultrasound measurements of the inferior vena cava and physician gestalt in the emergency department of an urban public hospital in Sub-Saharan Africa. *PLoS One*. 2016;11(9):e0162772
7. Woolley A, Kostopoulou O. Clinical intuition in family medicine: more than first impressions. *Ann Fam Med*. 2013; 11(1):60–66
8. Horwood J, Cabral C, Hay AD, Ingram J. Primary care clinician antibiotic prescribing decisions in consultations for children with RTIs: a qualitative interview study. *Br J Gen Pract*. 2016; 66(644):e207–e213
9. Cook C. Is clinical gestalt good enough? *J Manual Manip Ther*. 2009;17(1):6–7
10. Van den Bruel A, Thompson M, Buntinx F, Mant D. Clinicians' gut feeling about serious infections in children: observational study. *BMJ*. 2012;345: e6144
11. Turnbull S, Lucas PJ, Redmond NM, et al. What gives rise to clinician gut feeling, its influence on management decisions and its prognostic value for children with RTI in primary care: a prospective cohort study. *BMC Fam Pract*. 2018;19(1):25
12. Florin TA, Ambroggio L, Brokamp C, et al. Reliability of examination findings in suspected community-acquired pneumonia. *Pediatrics*. 2017;140(3): e20170310
13. Ambroggio L, Florin TA, Shah SS, et al. Emerging biomarkers of illness severity: urinary metabolites associated with sepsis and necrotizing methicillin-resistant *Staphylococcus aureus* pneumonia. *Pharmacotherapy*. 2017;37(9):1033–1042
14. Florin TA, Ambroggio L, Brokamp C, et al. Biomarkers and disease severity in children with community-acquired pneumonia. *Pediatrics*. 2020;145(6): e20193728
15. Ambroggio L, Brokamp C, Mantyla R, et al. Validation of the British Thoracic Society severity criteria for pediatric community-acquired pneumonia. *Pediatr Infect Dis J*. 2019;38(9):894–899
16. Jain S, Williams DJ, Arnold SR, et al.; CDC EPIC Study Team. Community-acquired pneumonia requiring hospitalization among U.S. children. *N Engl J Med*. 2015;372(9):835–845
17. Williams DJ, Shah SS, Myers A, et al. Identifying pediatric community-acquired pneumonia hospitalizations: accuracy of administrative billing codes. *JAMA Pediatr*. 2013;167(9):851–858
18. Alessandrini EA, Alpern ER, Chamberlain JM, Shea JA, Gorelick MH. A new diagnosis grouping system for child emergency department visits. *Acad Emerg Med*. 2010;17(2):204–213
19. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44(3): 837–845
20. Oliver G, Reynard C, Morris N, Body R. Can emergency physician gestalt “rule in” or “rule out” acute coronary syndrome: validation in a multicenter prospective diagnostic cohort study. *Acad Emerg Med*. 2020;27(1):24–30
21. Williams JW Jr., Simel DL, Roberts L, Samsa GP. Clinical evaluation for sinusitis. Making the diagnosis by history and physical examination. *Ann Intern Med*. 1992;117(9):705–710
22. Penalzo A, Verschuren F, Meyer G, et al. Comparison of the unstructured clinician gestalt, the wells score, and the revised Geneva score to estimate pretest probability for suspected pulmonary embolism. *Ann Emerg Med*. 2013;62(2):117–124.e2
23. Fine MJ, Hough LJ, Medsger AR, et al. The hospital admission decision for patients with community-acquired pneumonia. Results from the pneumonia Patient Outcomes Research Team cohort study. *Arch Intern Med*. 1997;157(1):36–44
24. Dean NC, Jones JP, Aronsky D, et al. Hospital admission decision for patients with community-acquired pneumonia: variability among physicians in an emergency department. *Ann Emerg Med*. 2012; 59(1):35–41
25. Moberg AB, Taléus U, Garvin P, Fransson SG, Falk M. Community-acquired pneumonia in primary care: clinical assessment and the usability of chest radiography. *Scand J Prim Health Care*. 2016;34(1):21–27
26. Nelson KA, Morrow C, Wingerter SL, Bachur RG, Neuman MI. Impact of chest radiography on antibiotic treatment for children with suspected pneumonia. *Pediatr Emerg Care*. 2016;32(8):514–519
27. Neuman MI, Scully KJ, Kim D, Shah S, Bachur RG. Physician assessment of the likelihood of pneumonia in a pediatric emergency department. *Pediatr Emerg Care*. 2010;26(11):817–822
28. Neill AM, Martin IR, Weir R, et al. Community acquired pneumonia: aetiology and usefulness of severity criteria on admission. *Thorax*. 1996; 51(10):1010–1016
29. Kabrhel C, Camargo CA Jr., Goldhaber SZ. Clinical gestalt and the diagnosis of pulmonary embolism: does experience matter? *Chest*. 2005;127(5):1627–1630
30. Bradley JS, Byington CL, Shah SS, et al.; Pediatric Infectious Diseases Society and the Infectious Diseases Society of America. Executive summary: the management of community-acquired pneumonia in infants and children older than 3 months of age: clinical practice guidelines by the Pediatric Infectious Diseases Society and the Infectious Diseases Society of America. *Clin Infect Dis*. 2011;53(7):617–630
31. Mahabee-Gittens EM, Grupp-Phelan J, Brody AS, et al. Identifying children with pneumonia in the emergency department. *Clin Pediatr (Phila)*. 2005; 44(5):427–435