

## Development and validation of a novel clinical risk score to predict hypoxemia in children with pneumonia using the WHO PREPARE dataset

Rainer Tan<sup>1,2</sup>, Arjun Chandna<sup>3,4</sup>, Tim Colbourn<sup>5</sup>, Shubhada Hooli<sup>6</sup>, Carina King<sup>7</sup>, Norman Lufesi<sup>8</sup>, Eric D McCollum<sup>9</sup>, Charles Mwansambo<sup>8</sup>, Joseph L. Mathew<sup>10</sup>, Clare Cutland<sup>11</sup>, Shabir Ahmed Madhi<sup>12</sup>, Marta Nunes<sup>12</sup>, Sudha Basnet<sup>13</sup>, Tor A. Strand<sup>14</sup>, Kerry-Ann O'Grady<sup>15</sup>, Brad Gessner<sup>16</sup>, Emmanuel Addo-Yobo<sup>17</sup>, Noel Chisaka<sup>18</sup>, Patricia L. Hibberd<sup>19</sup>, Prakash Jeena<sup>20</sup>, Juan M. Lozano<sup>21</sup>, William B. MacLeod<sup>19</sup>, Archana Patel<sup>22</sup>, Donald M. Thea<sup>19</sup>, Ngoc Tuong Vy Nguyen<sup>23</sup>, Marilla Lucero<sup>24</sup>, Syed Mohammad Akram uz Zaman<sup>25</sup>, Shinjini Bhatnagar<sup>26</sup>, Nitya Wadhwa<sup>26</sup>, Rakesh Lodha<sup>27</sup>, Satinder Aneja<sup>28</sup>, Mathuram Santosham<sup>29</sup>, Shally Awasthi<sup>30</sup>, Ashish Bavdekar<sup>31</sup>, Monidarin Chou<sup>32</sup>, Pagbajabyn Nymadawa<sup>33</sup>, Jean-William Pape<sup>34</sup>, Glauca Paranhos-Baccala<sup>35</sup>, Valentina S. Picot<sup>35</sup>, Mala Rakoto-Andrianarivelo<sup>36</sup>, Vanessa Rouzier<sup>34</sup>, Graciela Russomando<sup>37</sup>, Mariam Sylla<sup>38</sup>, Philippe Vanhems<sup>39</sup>, Jianwei Wang<sup>40</sup>, Romina Libster<sup>41</sup>, Alexey W. Clara<sup>42</sup>, Fenella Beynon<sup>2</sup>, Gillian Levine<sup>2</sup>, Chris A Rees<sup>43</sup>, Mark I Neuman<sup>44</sup>, Shamim A. Qazi<sup>45</sup>, Yasir Bin Nisar<sup>46\*</sup>, On behalf of the World Health Organization PREPARE study group<sup>47</sup>

<sup>1</sup>Unisanté, Centre for Primary Care and Public Health, University of Lausanne, Switzerland

<sup>2</sup>Swiss Tropical and Public Health Institute, University of Basel, Switzerland

<sup>3</sup>Cambodia Oxford Medical Research Unit, Angkor Hospital for Children, Siem Reap, Cambodia

<sup>4</sup>Centre for Tropical Medicine & Global Health, Nuffield Department of Medicine, University of Oxford, United Kingdom

<sup>5</sup>Institute for Global Health, University College London, London, United Kingdom

<sup>6</sup>Division of Pediatric Emergency Medicine, Texas Children's Hospital, Baylor College of Medicine, Houston, Texas, United States of America

<sup>7</sup>Department of Global Public Health, Karolinska Institutet, Stockholm, Sweden

<sup>8</sup>Ministry of Health, Lilongwe, Malawi

<sup>9</sup>Global Program in Respiratory Sciences, Eudowood Division of Pediatric Respiratory Sciences, Department of Pediatrics, Johns Hopkins School of Medicine, Baltimore, United States of America and Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, United States of America

<sup>10</sup>Advanced Pediatrics Centre, Postgraduate Institute of Medical Education and Research, Chandigarh, India

<sup>11</sup>African Leadership in Vaccinology Expertise, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

<sup>12</sup>South African Medical Research Council: Vaccines and Infectious Diseases Analytics Research Unit, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

<sup>13</sup>Department of Pediatrics, Tribhuvan University Institute of Medicine, Nepal

<sup>14</sup>Research Department, Innlandet Hospital Trust, Lillehammer, Norway

<sup>15</sup>Centre for Healthcare Transformation, Australian Centre for Health Services Innovation, Queensland University of Technology, Kelvin Grove, Queensland, Australia

<sup>16</sup>Pfizer Vaccines, Collegeville, Pennsylvania, United States of America

<sup>17</sup>Kwame Nkrumah University of Science & Technology/Komfo Anokye Teaching Hospital, Kumasi, Ghana

<sup>18</sup>World Bank, Washington DC, United States of America

<sup>19</sup>Department of Global Health, Boston University School of Public Health, Boston, Massachusetts, United States of America

<sup>20</sup>University of KwaZulu-Natal, Durban, South Africa

<sup>21</sup>Florida International University, Miami, United States of America

- <sup>22</sup>Lata Medical Research Foundation, Nagpur and Datta Meghe Institute of Medical Sciences, Sawangi, India
- <sup>23</sup>Children Hospital No 1, Ho Chi Minh City, Vietnam
- <sup>24</sup>Research Institute for Tropical Medicine, Manila, Philippines
- <sup>25</sup>Liverpool School of Tropical Medicine, Liverpool, United Kingdom
- <sup>26</sup>Translational Health Science and Technology Institute, Faridabad, India
- <sup>27</sup>All India Institute of Medical Sciences, New Delhi, India
- <sup>28</sup>School of Medical Sciences & Research, Sharda University, Greater Noida, India
- <sup>29</sup>Vaccine Access Center, Center for Indigenous Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, United States of America
- <sup>30</sup>King George's Medical University, Department of Pediatrics, Lucknow, India
- <sup>31</sup>KEM Hospital Pune, Department of Pediatrics, Pune, India
- <sup>32</sup>University of Health Sciences, Rodolph Mérieux Laboratory & Ministry of Environment, Phnom Penh, Cambodia
- <sup>33</sup>Mongolian Academy of Sciences, Academy of Medical Sciences, Ulaanbaatar, Mongolia
- <sup>34</sup>GHEKIO, Centre GHEKIO, Port au Prince, Haiti
- <sup>35</sup>Fondation Merieux, Lyon, France
- <sup>36</sup>Centre d'Infectiologie Charles Mérieux, Antananarivo, Madagascar
- <sup>37</sup>Universidad Nacional de Asuncion, Departamento de Biología Molecular y Genética, Instituto de Investigaciones en Ciencias de la Salud, Asuncion, Paraguay
- <sup>38</sup>Gabriel Touré Hospital, Department of Pediatrics, Bamako, Mali
- <sup>39</sup>Unité d'Hygiène, Epidémiologie, Infectiovigilance et Prévention, Hospices Civils de Lyon, Lyon, France and Centre International de Recherche en Infectiologie, Institut National de la Santé et de la Recherche Médicale U1111, CNRS Unité Mixte de Recherche 5308, École Nationale Supérieure de Lyon, Université Claude Bernard Lyon 1, Lyon, France
- <sup>40</sup>Chinese Academy of Medical Sciences & Peking Union, Medical College Institute of Pathogen Biology, MOH Key Laboratory of Systems Biology of Pathogens and Dr Christophe Mérieux Laboratory, Beijing, China
- <sup>41</sup>Fundacion INFANT, Argentina
- <sup>42</sup>U.S. Centers for Disease Control, Central American Region, Guatemala City, Guatemala
- <sup>43</sup>Division of Pediatric Emergency Medicine, Emory University School of Medicine, Children's Healthcare of Atlanta, Atlanta, Georgia, United States of America
- <sup>44</sup>Division of Emergency Medicine, Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts, United States of America
- <sup>45</sup>Independent Consultant, Retired Staff, World Health Organization, Geneva, Switzerland
- <sup>46</sup>Department of Maternal, Newborn, Child, and Adolescent Health and Ageing, World Health Organization, Geneva, Switzerland
- <sup>47</sup>Pneumonia Research Partnership to Assess WHO REcommendations

**\*Corresponding Author:**

Yasir Bin Nisar  
Department of Maternal, Newborn, Child and Adolescent Health and Ageing  
World Health Organization  
20 Avenue Appia 1211, 27  
Geneva, Switzerland  
Email: [nisary@who.int](mailto:nisary@who.int)  
Phone: +41-22-791-5595, Fax: +41-22-791-4853

## **ABSTRACT**

### **Background**

Hypoxemia predicts mortality at all levels of care, and appropriate management can reduce preventable deaths. However, pulse oximetry and oxygen therapy remain inaccessible in many primary care health facilities. We aimed to develop and validate a simple risk score comprising commonly evaluated clinical features to predict hypoxemia in 2-59-month-old children with pneumonia.

### **Methods**

Data from 7 studies conducted in 5 countries from the Pneumonia Research Partnership to Assess WHO Recommendations (PREPARE) dataset were included. Readily available clinical features and demographic variables were used to develop a multivariable logistic regression model to predict hypoxemia ( $SpO_2 < 90\%$ ) at presentation to care. The adjusted log coefficients were transformed to derive the PREPARE hypoxemia risk score and its diagnostic value was assessed in a held-out, temporal validation dataset.

### **Results**

We included 14,509 children in the analysis; 9.8% (n=2,515) were hypoxemic at presentation. The multivariable regression model to predict hypoxemia included age, sex, respiratory distress (nasal flaring, grunting and/or head nodding), lower chest indrawing, respiratory rate, body temperature and weight-for-age z-score. The model showed fair discrimination (area under the curve 0.70, 95% CI 0.67 to 0.73) and calibration in the validation dataset. The simplified PREPARE hypoxemia risk score includes 5 variables: age, respiratory distress, lower chest indrawing, respiratory rate and weight-for-age z-score.

### **Conclusion**

The PREPARE hypoxemia risk score, comprising five easily available characteristics, can be used to identify hypoxemia in children with pneumonia with a fair degree of certainty for use in health facilities without pulse oximetry. Its implementation would require careful consideration to limit inappropriate referrals on patients and the health system. Further external validation in community settings in low- and middle-income countries is required.

## **KEY MESSAGES:**

### **What is already known on this topic:**

- Pulse oximetry is unavailable or underutilized in many resource-limited settings in low- and middle-income countries.
- Hypoxemia is a good predictor of mortality and its early identification and further management can reduce mortality.

### **What this study adds:**

- The PREPARE hypoxemia risk score was developed using one of the largest and most geographically diverse datasets on childhood pneumonia to date.
- Using age, lower chest indrawing, respiratory rate, respiratory distress and weight-for-age z-score to calculate the PREPARE hypoxemia risk score could help identify children with hypoxemia in settings without pulse oximeters.

### **How this study might affect research, practice or policy:**

- This study contributes to the important discussion on how best to identify hypoxemic children in the absence of pulse oximetry.
- Further research is warranted to validate the findings in community settings
- Operationalizing and integrating the score within existing clinical management pathways must be tailored to the setting of implementation.

## INTRODUCTION

Hypoxemia – low blood oxygen levels, defined as an oxygen saturation (SpO<sub>2</sub>) of <90%, is common in children with World Health Organization (WHO) classified pneumonia in low- and middle-income countries (LMICs).<sup>1</sup> It is a good predictor of preventable deaths,<sup>2-4</sup> and early identification of hypoxemia and appropriate management can reduce mortality.<sup>5-8</sup> The availability of pulse oximetry to measure SpO<sub>2</sub> impacts clinicians' assessment of illness severity, diagnosis, treatment and need for referral.<sup>6,9</sup> As such, there is increasing rationale to support universal access to pulse oximetry in primary care settings to address the high burden of pneumonia-related mortality.<sup>10-13</sup>

Although the availability of pulse oximetry is increasing, in part due to increased focus during the COVID-19 pandemic,<sup>10</sup> access is limited and inequitable and pulse oximetry remains underutilized in many LMIC settings, particularly in the community and primary care health facilities.<sup>14-17</sup> Barriers to the implementation of pulse oximetry in LMICs include high cost of quality pediatric devices with appropriately sized probes, maintenance, requisite training, battery charging and patient-related factors (e.g. crying and movement), which can be particularly challenging among young children.<sup>18-20</sup> Furthermore, performing pulse oximetry can take time, especially in agitated children; an important consideration for implementation in health facilities with high case-loads or insufficient staff-to-patient ratios.

In primary care level health facilities without pulse oximetry, where the WHO's Integrated Management of Childhood Illnesses (IMCI) chartbook<sup>21</sup> is used to guide the management of sick children, many hypoxemic children are not identified.<sup>6,19</sup> In health facilities where a pulse oximeter is not available, identifying children at the highest risk of being hypoxemic using clinical signs alone could improve the sensitivity of the IMCI chartbook for identification of children in need of oxygen therapy, hospitalization, or referral. Previous clinical scores to predict hypoxemia were limited in sample size, geographical representation, or included patients already identified as requiring referral based on other clinical features.<sup>22-26</sup>

We aimed to develop and validate a clinical risk score to predict hypoxemia (SpO<sub>2</sub> <90%) at presentation to a health facility in children aged 2-59 months with pneumonia according to the 2014 WHO IMCI definition (i.e. cough and/or difficulty breathing with fast breathing and/or lower chest indrawing) using the Pneumonia REsearch Partnership to Assess WHO REcommendations (PREPARE) dataset.<sup>21,27</sup> The goal is not to replace the pulse oximeter, but to guide referral to health facilities where children with suspected hypoxemia can be appropriately managed.

## **METHODS**

### **Study design**

We used the WHO PREPARE dataset, to derive and validate a clinical risk score to identify hypoxemia in children 2-59 months with pneumonia. The PREPARE dataset included primary, patient-level data for children 0-59 months with pneumonia, from 45 separate studies conducted from 1994 to 2018, in over 20 countries in Asia, Africa, Latin America, Oceania and North America.<sup>27</sup> Detailed description of this dataset is found elsewhere.<sup>27</sup> We adhered to the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) guidelines.<sup>28</sup>

### **Patient and public involvement**

The development of the research question was informed by the high rates of pneumonia-related mortality, specifically in children with hypoxemia. Patients were neither advisers in this study nor were they involved in the design, recruitment or conduct of the study. Results of this study will be made publicly available through open-access publication.

### **Study population**

For the present analysis, included studies that reported SpO<sub>2</sub> measured at the same time as assessment of clinical symptoms and signs (within 2 hours) and for which pulse oximetry was performed in  $\geq 80\%$  of cases. Studies in which fewer than 75% of patients had data for each candidate predictor were excluded. Patients aged 2-59 months with IMCI defined pneumonia (irrespective of SpO<sub>2</sub> reading) presenting to health facilities were included. Patients with severe pneumonia (defined as cough and/or difficulty breathing with a general danger sign including stridor in a calm child), or complicated severe acute malnutrition were excluded, since they would already have met IMCI criteria for referral.<sup>21 29</sup>

### **Outcome**

The primary outcome was hypoxemia, defined as a SpO<sub>2</sub> of  $<90\%$ ,<sup>21, 27</sup> assessed at the same time as the clinical predictors (symptoms and signs). A sensitivity analysis assessed the development of a model for the prediction of blood oxygen saturation at a SpO<sub>2</sub> value of  $<92\%$ . This threshold was selected acknowledging the high risk of mortality in children at this threshold,<sup>13</sup> and the use of this threshold by various clinical guidelines.<sup>30 31</sup>

### **Candidate Predictors**

Candidate predictors were selected a priori based on expert knowledge and previous diagnostic value for predicting hypoxemia, identified in the literature,<sup>26</sup> and considering reliability, feasibility and resources available at typical primary care health facilities in LMICs. The seven candidate predictors were: sex, age, chest indrawing, respiratory rate, axillary temperature, weight-for-age z-score and respiratory distress (composite predictor defined as the presence of either nasal flaring, grunting and/or head nodding). A sensitivity analysis assessed the performance of the model without respiratory distress, with and without wheezing. Wheezing was included in a sensitivity analysis because it is difficult to identify without the use of a stethoscope, often absent or infrequently used in primary care health facilities in LMICs, however included within IMCI.<sup>21 32</sup> In the analyses without respiratory distress as a candidate predictor, we excluded all patients with signs of respiratory distress. Given the variety of aims and methods among studies in the PREPARE dataset, some clinical predictors may have been measured with knowledge of the outcome status (peripheral oxygen saturation). Duration of cough or difficulty breathing, history of fever, duration of fever, mid-

upper arm circumference, weight-for-height z-score, tracheal tugging, heart rate, cyanosis and capillary refill time, were considered as candidate predictors based on potential clinical relevance, but were not included due to high levels of missingness in the data.

### Sample size

We used 75% of the dataset for the development of the score and 25% for validation. Each study dataset was split by reserving the last 25% of enrolled patients for the validation dataset. Such a temporal validation strategy is considered intermediate between internal and external validation.<sup>28</sup> Using the approach outlined by Riley et al.,<sup>33</sup> 659 cases of hypoxemia would be required in order to have a sufficient sample size to include the 12 pre-specified parameters (from the 7 candidate predictors) to develop the score (54.9 events per parameter), considering a conservative  $R^2$  Nagelkerke of 0.15, shrinkage factor of 0.9 and outcome prevalence of 9.8% in the development dataset. The development dataset of 10,884 children contained 1,146 cases of hypoxemia (i.e. 95.5 outcome events per parameter). The temporal validation dataset included 275 outcome events, well above the recommended 100 events often recommended for a robust validation.<sup>34 35</sup>

### Analysis

Pulse oximetry measurements were not adjusted for altitude because no sites were greater than 2,500 meters above sea level.<sup>36</sup> Continuous predictors were categorized a priori based on currently accepted thresholds in order to maximize interpretability, encourage uptake amongst healthcare providers and policy makers and align with current clinical practices: age: 2-5 months, 6-11 months, 12-59 months; temperature: <35.5, 35.5-37.4,  $\geq$ 37.5°C; respiratory rate: 0-9, 10-19,  $\geq$ 20 breaths/min above IMCI age-specific cut-offs (age 2-11 months  $\geq$ 50 breaths per minute; age 12-59 months  $\geq$ 40 breaths per minute); and weight-for-age z-score:  $\geq$  -2, < -2 to -3, < -3. Patients with missing data were excluded as multiple imputation was not feasible as it was determined that data were unlikely to be missing at random.<sup>37</sup>

The model was developed using multivariable logistic regression with the least absolute shrinkage and selection operator (LASSO) with L1 regularization. The adjusted model included all parameters and was further adjusted for study. Associations with 95% confidence intervals (CI) for adjusted odds ratios (aOR) that did not cross 1 were considered significant. Discrimination was evaluated by calculating the area under the receiver operating characteristic (AUROC) curve of the model in both the development and validation datasets. Calibration was evaluated using calibration plots (observed risk on y-axis, predicted risk on the x-axis) in the validation dataset.

In order to convert the model into a simple clinical risk score, the regression coefficient of each retained parameter with a statistically significant adjusted odds ratio (CI did not overlap with 1.0) was rounded to the nearest 0.5 and then doubled.<sup>38 39</sup> Sensitivity, specificity, likelihood ratios, miss rate (proportion of patients with hypoxemia not identified), false discovery rate (proportion of patients inappropriately identified as having hypoxemia among all patients identified as having hypoxemia) and proportion of children identified as being hypoxemic at each of the score's cut offs was evaluated in the temporal validation dataset.

Recognizing that the appropriate threshold for referring a child with suspected hypoxemia would be context dependent, a decision curve analysis compared the clinical utility (net benefit) of the model and score across a range of clinically-plausible referral thresholds. The net benefit quantifies the trade-off between true positives (hypoxemic children predicted to have hypoxemia) and false positives (non-hypoxemic children incorrectly predicted to have hypoxemia), weighted according to the relative cost of a false positive (threshold probability), which would vary in different contexts.

The threshold probability reflects the cut-off (predicted probability of hypoxemia) above which a given intervention (in this case, referral) might be considered.

All analyses were performed using STATA version 16.<sup>40</sup>



## RESULTS

### Baseline characteristics

The PREPARE dataset included 294,968 children from 45 studies, of which 14,509 children from 7 studies met inclusion criteria and were included in this analysis (figure 1). Four studies were randomized controlled trials, and three were prospective cohort studies (table 1). All studies included outpatients presenting to hospitals, and one study also included patients presenting to health centers.<sup>41</sup> Among the 7 studies included in this analysis, patients were recruited from 5 countries: Australia, Gambia, India, Malawi and Nepal (table 1). Of the 14,509 patients with pneumonia, 9.8% had a SpO<sub>2</sub> less than 90% at presentation (table 2).

Baseline characteristics between the development (n=10,884, 75%) and validation dataset (n=3,625, 25%) were similar (table 2). Within the total dataset (n=14,509), 54.9% of children were aged less than 1 year, there were more male patients than female patients, 62.0% of children had lower chest indrawing, 34.1% had a temperature of equal to or more than 37.5°C and 6.7% had a very low weight-for-age z-score (< -3).

**Table 1:** Characteristics of studies included in the development and validation of the PREPARE hypoxemia risk score

Study	Study design	Location	Age	Original inclusion criteria from the source studies	Original exclusion criteria from the source studies	Sample size <sup>a</sup> , n/N (%)	Hypoxemia (SpO <sub>2</sub> <90%) n/N (%)
Basnet, 2012 <sup>42</sup>	Randomized controlled trial	Kathmandu, Nepal	2-35 months	- Cough for <14 days - Difficulty breathing ≤72 hours with presence of lower chest indrawing on examination	- Children with recurrent wheezing, heart disease, other severe illness, severe malnutrition, dehydration, hemoglobin <70 g/L, chronic cough, effusion on chest x-ray, or history of documented tuberculosis.	378/641 (59.0%)	251/378 (76.4%)
Cutts, 2005 <sup>41</sup>	Randomized controlled trial	Eastern Gambia	40-364 days	Children with WHO-defined pneumonia of any severity <sup>b</sup>	- Intent to move out of the study area within 4 months - Previous or uncertain receipt of diphtheria-pertussis-tetanus/ <i>Haemophilus influenzae</i> type b (DPT/Hib) or DPT vaccine - Inclusion in a previous vaccine trial - Serious chronic illness	433/1,716 (25.2%)	22/433 (5.1%)
Lazzerini, 2016 <sup>43</sup>	Prospective cohort	Malawi	2-59 months	- Children with WHO-defined pneumonia of any severity <sup>b</sup>	None	7,902/16,475 (48.0%)	737/7,902 (9.3%)
Mathew, 2015 <sup>44</sup>	Prospective cohort	Chandigarh, India	2-59 months	- Children with WHO-defined pneumonia of any severity <sup>b</sup>	- Duration of illness >7 days - Antibiotics for >24 hours - Previous hospitalization within the preceding 30 days - Children with wheeze who received a single dose of bronchodilator and whose symptoms disappeared	1,132/2,400 (47.2%)	134/1,132 (11.8%)
McCollum,	Prospective cohort	Mchinji and Lilongwe	2-59 months	- Children with WHO-defined pneumonia of any severity <sup>b</sup>	None	4,322/6,764 (63.9%)	274/4,322 (6.3%)

2017 <sup>45</sup>		Districts, Malawi					
O'Grady, 2012 <sup>46</sup>	Randomized controlled trial	Central Australia	2-59 months	- Children with WHO-defined pneumonia of any severity <sup>b</sup>	Children with wheezing and chronic conditions	26/147 (17.7%)	0.0% (0/26)
Wadhwa, 2013 <sup>47</sup>	Randomized controlled trial	New Delhi, India	2-24 months	- Children with WHO-defined pneumonia of any severity <sup>b</sup> - Crepitations on auscultation	- Need for mechanical ventilation or inotropic medications - Any other serious underlying medical condition	316/550 (57.5%)	0.9% (3/316)

<sup>a</sup> Sample size included in the present analysis over the sample size from the original study

<sup>b</sup> WHO-defined pneumonia of any severity: presence of age-specific fast breathing, lower chest indrawing, or danger signs, in children with a cough or difficulty breathing

**Table 2.** Baseline characteristics of children with pneumonia (fast breathing and/or chest indrawing) with pulse oximetry measurements and complete data on all candidate predictors (n=14,509)

Characteristic	Total (n=14,509)	Development dataset (n=10,884)	Validation dataset (n=3,625)
<b>Age<sup>a</sup></b>			
2-5 months	4,298 (29.6%)	3,400 (31.2%)	898 (24.8%)
6-11 months	3,663 (25.2%)	2,830 (26.0%)	833 (23.0%)
12-59 months	6,548 (45.1%)	4,654 (42.8%)	1,894 (52.2%)
<b>Sex</b>			
Male	8,193 (56.5%)	6,126 (56.3%)	2,068 (57.1%)
Female	6,316 (43.5%)	4,759 (43.7%)	1,557 (42.9%)
<b>Pneumonia classification</b>			
Fast breathing only	5,519 (38.0%)	4,052 (37.2%)	1,467 (40.5%)
Lower chest indrawing	8,990 (62.0%)	6,832 (62.8%)	2,158 (59.6%)
<b>Any sign of respiratory distress (nasal flaring, grunting or head nodding)</b>			
No	7,942 (54.7%)	5,930 (54.5%)	2,012 (55.5%)
Yes	6,567 (45.3%)	4,954 (45.5%)	1,613 (44.5%)
<b>Respiratory rate</b>			
< age-adjusted tachypnea threshold <sup>b</sup>	163 (1.1%)	124 (1.1%)	39 (1.1%)
0-9 breaths/min above cut-off	6,164 (42.5%)	4,600 (42.3%)	1,563 (43.1%)
10-19 breaths/min above cut-off	5,333 (36.8%)	3,994 (36.7%)	1,340 (37.0%)
>=20 breaths/min above cut-off	2,849 (19.6%)	2,166 (19.9%)	683 (18.8%)
<b>Body temperature</b>			
Normal (>=35.5°C to 37.5°C)	9,402 (64.8%)	6,908 (63.5%)	2,494 (68.8%)
>=37.5°C	4,949 (34.1%)	3,854 (35.4%)	1,095 (30.2%)
<35.5°C	158 (1.1%)	122 (1.1%)	36 (1.0%)
<b>Weight for age z-score</b>			
>=-2	11,767 (81.1%)	8,777 (80.6%)	2,990 (82.5%)
<-2 to -3	1,777 (12.2%)	1,364 (12.5%)	413 (11.4%)
<-3	965 (6.7%)	743 (6.8%)	222 (6.1%)
<b>Oxygen saturation (SpO<sub>2</sub>)</b>			
>=90%	13,088 (90.2%)	9,738 (89.5%)	3,350 (92.4%)
<90%	1,421 (9.8%)	1,146 (10.5%)	275 (7.6%)

<sup>a</sup>Some studies did not include patients in the whole 2-59 month age range; <sup>b</sup>IMCI chartbook threshold for tachypnea: age 2-11 months >=50 breaths per minute; age 12-59 months >=40 breaths per minute

## Model development

The predictors most strongly associated with hypoxemia (SpO<sub>2</sub> <90%) were respiratory rate  $\geq$  20 breaths/minute above the IMCI age-adjusted cut-off (aOR 2.41, 95% CI 1.25 to 4.64), any sign of respiratory distress (aOR 2.21, 95% CI 1.90 to 2.58), weight for age z-score <-3 (aOR 1.96, 95% CI 1.56 to 2.42) and lower chest indrawing (aOR 1.62, 95% CI 1.35 to 1.95) (table 3). None of the 12 candidate parameters were eliminated by the LASSO penalty during model development and hence, all 7 candidate predictors were retained in the final model.

**Table 3.** Multivariable regression model for predicting hypoxemia in the development dataset (n=10,884)

Characteristic	No hypoxemia (n=9,738)	Hypoxemia (n=1,146)	Odds ratio (95% CI)	Adjusted OR (95% CI)
<b>Age</b>				
2-5 months	3,000 (30.8%)	400 (34.9%)	1.34 (1.16, 1.55)	1.29 (1.20, 1.66)
6-11 months	2,506 (25.7%)	324 (28.3%)	1.30 (1.11, 1.52)	1.41 (1.11, 1.50)
12-59 months	4,232 (43.5%)	422 (36.8%)	(reference)	(reference)
<b>Sex</b>				
Male	5,455 (56.0%)	670 (58.5%)	(reference)	(reference)
Female	4,283 (44.0%)	476 (41.5%)	0.90 (0.80, 1.02)	0.91 (0.80, 1.03)
<b>Pneumonia classification</b>				
Fast breathing only	3,836 (39.4%)	216 (18.8%)	(reference)	(reference)
Lower chest indrawing	5,902 (60.6%)	930 (81.2%)	2.80 (2.40, 3.26)	1.62 (1.35, 1.95)
<b>Any sign of respiratory distress (nasal flaring, grunting or head nodding)</b>				
No	5,559 (57.1%)	371 (32.4%)	(reference)	(reference)
Yes	4,179 (42.9%)	775 (67.6%)	2.78 (2.44, 3.16)	2.21 (1.90, 2.58)
<b>Respiratory rate</b>				
< age-adjusted tachypnea threshold*	114 (1.2%)	10 (0.9%)	(reference)	(reference)
0-9 breaths/min above cut-off	4,317 (44.3%)	284 (24.8%)	0.75 (0.39, 1.45)	0.86 (0.44, 1.66)
10-19 breaths/min above cut-off	3,556 (36.5%)	437 (38.1%)	1.40 (0.73, 2.69)	1.36 (0.71, 2.62)
$\geq$ 20 breaths/min above cut-off	1,751 (18.0%)	415 (36.2%)	2.70 (1.40, 5.20)	2.41 (1.25, 4.64)
<b>Body temperature</b>				
Normal ( $\geq$ 35.5°C to 37.5°C)	6,226 (63.9%)	682 (59.5%)	(reference)	(reference)
$\geq$ 37.5°C	3,401 (34.9%)	453 (39.5%)	1.21 (1.07, 1.38)	1.04 (0.91, 1.19)
<35.5°C	111 (1.1%)	11 (1.0%)	0.90 (0.48, 1.69)	0.92 (0.48, 1.75)
<b>Weight for age z-score</b>				
$\geq$ -2	7,939 (81.5%)	838 (73.1%)	(reference)	(reference)
<-2 to -3	1,181 (12.1%)	183 (16.0%)	1.47 (1.24, 1.74)	1.38 (1.15, 1.64)
<-3	618 (6.4%)	125 (10.9%)	1.92 (1.56, 2.35)	1.95 (1.56, 2.42)

\*IMCI chartbook threshold for tachypnea: age 2-11 months  $\geq$ 50 breaths per minute; age 12-59 months  $\geq$ 40 breaths per minute

## Model performance

The PREPARE hypoxemia clinical prediction model had an AUROC of 0.70 (95% CI 0.69 to 0.72) in the development dataset and an AUROC of 0.70 (95% CI 0.68 to 0.73) in the temporal validation dataset (figure 2). Calibration of the PREPARE hypoxemia prediction model in the validation dataset is illustrated in figure 3.

When converted into the PREPARE hypoxemia risk score (table 4), the retained predictors included age 2-11 months (+1 point), lower chest indrawing (+1 points), respiratory rate  $\geq$ 20 breaths/min

above age-adjusted tachypnea threshold (+2 points), any sign of respiratory distress (+2 points) and weight for age z-score <-2 (+1 point), with the score ranging from 0-7 points (table 4 and 5). Sensitivity, specificity, positive and negative likelihood ratios, miss rate, false discovery rate and the proportion of children identified as hypoxemic are included for each score cut-off (table 5).

In the absence of pulse oximetry, the decision curve analysis suggests that in contexts where it might be acceptable to identify (and refer) up to six non-hypoxemic children for each hypoxemic child (up to a threshold probability of ~17%) it would be optimal to use a cut-off of  $\geq 5$  to guide referral decision for further management, whereas in contexts in which greater than 1-in-7 true positives are required, a cut-off of  $\geq 6$  would be preferable (figure S1). The PREPARE hypoxemia model had an equivalent or higher net benefit than the PREPARE hypoxemia risk score (at any cut-off) across all threshold probabilities.

**Table 4:** Components of the PREPARE hypoxemia risk score in the development dataset (n=10,884)

Factor	Adjusted log coefficient	PREPARE hypoxemia risk score <sup>a</sup>
<b>Age</b>		
2-5 months	0.26	+1
6-11 months	0.34	+1
12-59 months	--	---
<b>Pneumonia classification</b>		
Fast breathing only	--	---
Lower chest indrawing	0.49	+1
<b>Any sign of respiratory distress (nasal flaring, grunting or head nodding)</b>		
No	--	--
Yes	0.79	+2
<b>Respiratory rate</b>		
< age-adjusted tachypnea threshold <sup>b</sup>	--	---
0-9 breaths/min above cut-off	-0.15	---
10-19 breaths/min above cut-off	0.31	---
$\geq 20$ breaths/min above cut-off	0.88	+2
<b>Weight for age z-score</b>		
$\geq -2$	--	---
<-2 to -3	0.32	+1
<-3	0.66	+1

<sup>a</sup>Regression coefficients of each retained parameter with a statistically significant adjusted odds ratio (confidence interval (CI) did not overlap with 1.0) was rounded to the nearest 0.5 and then doubled to derive the points contributing to the PREPARE hypoxemia risk score.

<sup>b</sup>IMCI chartbook threshold for tachypnea: age 2-11 months  $\geq 50$  breaths per minute; age 12-59 months  $\geq 40$  breaths per minute

**Table 5.** Performance of the PREPARE hypoxemia risk score at each cut-off in the validation dataset (n=3,625)

Score	Hypoxemia, n/N (%)	Sensitivity (95% CI)*	Specificity (95% CI)*	Positive Likelihood Ratio (95% CI)*	Negative Likelihood Ratio (95%)	Miss rate† % (n/N)	False discovery rate‡ %	Proportion referred^ % (n/N)
-------	--------------------	-----------------------	-----------------------	-------------------------------------	---------------------------------	--------------------	-------------------------	------------------------------

					CI)*		(n/N)	
0	6/450 (1.3%)	100.0 (98.7, 100.0)	0.0 (0.0, 0.1)	1.00	---	0% (0/275)	92.4% (3350/3625)	100% (3625/3625)
1	21/781 (2.7%)	97.8 (95.3, 99.2)	13.3(12.1, 14.5)	1.13 (1.10, 1.15)	0.16 (0.07, 0.36)	2.2% (6/275)	91.5% (2906/3175)	87.6% (3175/3625)
2	44/545 (8.1%)	90.2 (86.0, 93.4)	35.9 (34.3, 37.6)	1.41 (1.34, 1.47)	0.27 (0.19, 0.39)	9.8% (27/275)	89.6% (2146/2394)	66.0% (2394/3625)
3	56/737 (7.6%)	74.2 (68.6, 79.3)	50.9 (49.2, 52.6)	1.51 (1.40, 1.63)	0.51 (0.41, 0.62)	25.8% (71/275)	89.0% (1645/1849)	51.0% (1849/3625)
4	67/689 (9.7%)	53.8 (47.7, 59.8)	71.2 (69.7, 72.8)	1.87 (1.66, 2.11)	0.65 (0.57, 0.74)	46.2% (127/275)	86.7% (964/1112)	30.7% (1112/3625)
5	45/284 (15.8%)	29.5 (24.1, 35.2)	89.8 (88.7, 90.8)	2.89 (2.34, 3.55)	0.79 (0.73, 0.85)	70.5% (194/275)	80.9% (342/423)	11.7% (423/3625)
6	28/119 (23.5%)	13.1 (9.3, 17.7)	96.9 (96.3, 97.5)	4.26 (2.96, 6.10)	0.90 (0.86, 0.94)	86.9% (239/275)	74.1% (103/139)	3.8% (139/3625)
7	8/20 (40.0%)	2.9 (1.3, 5.7)	99.6 (99.4, 99.8)	8.12 (3.35, 19.70)	0.97 (0.95, 0.99)	97.1% (267/275)	60.0% (12/20)	0.6% (20/3625)

\*Calculated at  $\geq$  each respective cut-off

†Miss rate (also known as false negative rate = False Negative/(True Positive + False Negative)

‡False discovery rate = False Positive/(False Positive+True Positive)

^Referral used as a proxy for identification of hypoxemia, assuming that all children identified as at risk of hypoxemia would be referred

### Prediction of hypoxemia at SpO<sub>2</sub> threshold of < 92%

In our sensitivity analysis using the same dataset, 15.7% (2,281/14,509) children had SpO<sub>2</sub> <92%. The proportion with SpO<sub>2</sub> <92% was slightly lower in the validation dataset (14.2%, 514/3,625) compared to the development dataset (16.2%, 1,767/10,884). The predictive value of predictors in predicting SpO<sub>2</sub> <92% compared to SpO<sub>2</sub> <90% were similar (table S1, table 3).

The PREPARE SpO<sub>2</sub> <92% clinical prediction model had a similar AUROC (0.67, 95% CI 0.65 to 0.70) compared to the PREPARE hypoxemia (SpO<sub>2</sub> <90%) prediction model (AUROC 0.70, 95% CI 0.68 to 0.73) (figure S2). The calibration plot for this model is shown in figure S3 and the performance of the risk score at each cut-off is presented in table S3.

The PREPARE SpO<sub>2</sub> <92% risk score integrates 5 clinical predictors for a total score of 6 points: Age 6-11 months (+1 point), lower chest indrawing (+1 point), respiratory rate  $\geq$  20 breaths/min above age-adjusted tachypnea threshold (+2 points), any sign of respiratory distress (+1 point) and weight-for-age z-score <-3 (+1 point) (table S2).

### **PREPARE hypoxemia risk score without respiratory distress**

When excluding respiratory distress as a candidate predictor, the dataset was larger with 15,592 children from 14 studies. Five studies were prospective cohorts, 7 were randomized controlled trials, one was a retrospective cohort and one was a prospective case series (table S4). Among the 14 studies included in this analysis, patients were recruited from 23 countries from North, Central and South America, Africa, Asia and Oceania (table S4). Of the 15,592 patients with pneumonia, 16.1% had a SpO<sub>2</sub> <90% at presentation (table S5). Baseline characteristics were similar in the development and validation datasets (table S5) and did not differ from the primary dataset including respiratory distress. Without respiratory distress, lower chest indrawing (aOR 4.44, 95% CI 3.83 to 5.15) contributed more towards the prediction of hypoxemia (table S6). None of the 11 candidate parameters were eliminated by the LASSO penalty during model development.

The PREPARE hypoxemia clinical prediction model excluding respiratory distress had a higher AUROC than the model that included respiratory distress as a candidate predictor (AUROC 0.78 vs. 0.70) (figure S4). The calibration plot of this model is shown in figure S5.

The final PREPARE hypoxemia risk score when excluding respiratory distress integrated 4 clinical predictors: age 2-11 months (+1 point), lower chest indrawing (+3 points), respiratory rate  $\geq 20$  breaths/min above age-adjusted tachypnea cut-off (+1 point) and weight for age z-score  $< -2$  (+1 point) for a maximum total score of 6 points (table S7).

### **PREPARE hypoxemia risk score with wheezing**

Using the same methodology as the primary analysis (i.e., with respiratory distress), including wheezing as a candidate predictor, resulted in a slightly smaller dataset of 13,792 children from 7 studies (table S9). In this dataset, 9.9% had a SpO<sub>2</sub> of less than 90% at presentation (table S10). Baseline characteristics in these development and validation datasets were similar (table S10) and did not differ from the primary dataset, which excluded wheezing. Wheezing had an adjusted odds ratio of 1.28 (95% CI 1.11 to 1.47), sensitivity of 37.6% and specificity of 73.2% (table S11) for identification of hypoxemia. None of the 12 candidate parameters were eliminated by the LASSO penalty during model development.

The PREPARE hypoxemia clinical prediction model that included wheezing performed similarly to the model without wheezing in the validation datasets (AUROC 0.72 versus 0.70) (figure S6). The calibration plot of the PREPARE hypoxemia clinical prediction model, including wheezing is shown in figure S3. The performance of the PREPARE hypoxemia risk score including wheezing at each cut-off is presented in table S10.

The final PREPARE hypoxemia risk score when including wheezing integrated five clinical parameters: age 2-11 months (+1 points), chest indrawing (+1 point), respiratory distress (+2 point), respiratory rate  $\geq 20$  breaths/min above cut-off (+2 points) and weight for age z-score  $< -2$  (+1 point) for a maximum total score of 7 points (table S12). Wheezing was not retained in the risk score as the adjusted log coefficient was 0.24. A comparison of all risk score weights from the main analysis and sensitivity analyses can be found in table S13.

## DISCUSSION

We developed the PREPARE hypoxemia risk score from one of the largest and most diverse pneumonia datasets, including over 14,500 childhood pneumonia cases (of which 1,421 had hypoxemia) from 7 studies in 5 countries and in 14 studies in 23 countries within a sensitivity analysis. The components of the score include age, respiratory distress, respiratory rate, lower chest-indrawing and weight-for-age z-score. The risk score demonstrated fair discrimination in predicting hypoxemia ( $\text{SpO}_2 < 90\%$ ) in children with pneumonia without danger signs or other indications for referral.

Previous studies predicting hypoxemia included fewer patients from more restricted geographical settings.<sup>8 22-26 48-51</sup> The use of very large and diverse datasets reduces instability of individual predictions, limits overfitting and provides the opportunity for generalizability to a wide range of settings.<sup>52 53</sup> A model developed from a dataset with a similar number of patients, though exclusively evaluated in outpatient clinics from Bangladesh and Malawi, had a higher AUROC (0.79) compared to the PREPARE hypoxemia risk score (0.70), but had a similar AUROC to the PREPARE hypoxemia risk score excluding respiratory distress (0.78).<sup>22</sup> However, this study and many others included patients with IMCI danger signs in whom hospitalization or referral would already be recommended by existing guidelines.<sup>21 22 29</sup> IMCI danger signs were retained in these prior analyses to evaluate their performance in hypoxemia prediction models alongside other candidate parameters and these studies showed IMCI danger signs added little to no predictive value for hypoxemia.<sup>21</sup> While excluding such patients in our study limits our ability to comment on the value of IMCI danger signs in hypoxemia prediction models it aligns more closely with the needs of healthcare providers currently using the IMCI framework, helping them to decide which patients with pneumonia not already identified for hospitalization or referral may benefit from higher-level care. As found in other studies,<sup>21</sup> this approach also underscores the high number of pneumonia patients with hypoxemia (1,421/14,509; 9.8%) that would not have been identified as requiring referral or hospitalization using the current IMCI algorithm in the absence of a pulse oximeter. In line with previous studies, greater tachypnea,<sup>23 24 49 51</sup> chest indrawing,<sup>22 23 51</sup> younger age,<sup>8 22</sup> wheezing,<sup>22</sup> respiratory distress,<sup>8 22 24 49 51</sup> and low and very-low weight-for-age z-score,<sup>8 22</sup> were identified as clinical signs predictive of hypoxemia in children with pneumonia. In the present study and previous studies, individual signs performed poorly in identifying hypoxemia, emphasizing the benefits of a multivariable clinical risk score.

The use of the PREPARE hypoxemia risk score and selection of cut-off to guide referral of children with suspected hypoxemia would need to be tailored to the setting of implementation. Selecting a cut-off with the highest balance between sensitivity and specificity may not be the most appropriate referral threshold in many clinical contexts.<sup>54 55</sup> Care must be taken when deciding to use such a score in order to limit referral of children without hypoxemia or danger signs which could potentially come at great cost to individual patients and/or the health system. Considering that all patients with hypoxemia identified using this score would have otherwise been misidentified as not requiring referral by IMCI without pulse oximetry, using cut-offs with lower sensitivity (and higher specificity) could still make a helpful contribution, limiting unnecessary referrals, whilst still identifying many children with hypoxemia who would have been missed by the IMCI algorithm in settings where pulse oximetry is not available. Nevertheless, even utilising cut-offs of 5 or 6, where between 4-12% of all children would be referred, the majority of referrals (74-81%) would still be for children who were not hypoxemic. It may be that depending on the setting of implementation, a cut-off with relatively low specificity should consider alternative approaches to referral, such as short admission or follow-up within the community, to observe clinical evolution as proposed by Graham et al.<sup>13</sup> Growing

accessibility to “smartphones” might increase feasibility of a digital tool, which would allow more precise predictions and enable referral thresholds to be better tailored to the specific context of implementation. More pragmatically, elements from the PREPARE hypoxemia risk score could be combined with data on their prognostic value to predict death,<sup>56</sup> to update the IMCI criteria for severe pneumonia.

The sensitivity analysis excluding respiratory distress had a higher AUROC value compared to the primary analysis including respiratory distress as a candidate predictor. This however, should not be interpreted to mean that respiratory distress does not add value in identifying hypoxemia. In fact respiratory distress had similar predictive value to identifying hypoxemia compared to lower chest indrawing and very fast breathing for age (i.e., a respiratory rate of 20 breaths or more above age cut-off values). Differences in AUROC values between our analyses and compared to other studies are influenced by differences in the underlying datasets.<sup>52</sup> Indeed, this sensitivity analysis included twice the number of studies and thus the higher discrimination may reflect better generalizability of a score developed from a more diverse range of countries and settings. The value of respiratory distress should thus still be explored to improve the detection of children with hypoxemia and/or severe pneumonia within IMCI consultations, all the while considering the training, mentorship and resources required to reliably assess these signs.

There are several limitations to this study. First, we intended to derive a clinical risk score using data from primary level health facilities since this is where pulse oximetry is often inaccessible and thus the clinical risk score would be most useful. Unfortunately, due to missing data on the outcome of interest and important clinical predictors, almost all studies from primary level health facilities were excluded, leaving a cohort derived predominantly from patients presenting to hospital outpatient departments. While the exclusion of children with danger signs may reduce this spectrum bias, the high proportion of pneumonia patients with chest indrawing (62%) and hypoxemia (9.8%) compared to what is expected at primary level health facilities, demonstrates the more severe presentation of included patients. Nonetheless, a similar study that was limited to ambulatory clinics found that the sensitivity and specificity of individual variables for predicting hypoxemia were similar to hospital-based studies.<sup>22</sup> Secondly, it has been reported that pulse oximetry is less accurate in patients with darker skin,<sup>57,58</sup> and this was not considered in the present clinical risk score. Third, only 9% of children within the target age range with pneumonia were included (14,509/157,433). The vast majority of children and studies were excluded due to missing data (predictors or outcome). While this limited the diversity of the dataset, this decision was taken in order to prioritize better quality data, limiting bias in order to retain studies that systematically assessed the predictors and outcome of interest. Further work is needed to confirm that the score is generalizable to the whole pneumonia population more widely – assuming that patients with pulse oximetry readings are likely to be more severe, the prevalence of hypoxemia in the pneumonia population is likely lower,<sup>59</sup> and the score may therefore benefit from updating.

## CONCLUSION

Hypoxemia predicts mortality at all levels of care,<sup>2-4</sup> and pulse oximetry helps increase sensitivity to identify children with severe pneumonia as defined by IMCI.<sup>6,19</sup> Identification of children with pneumonia who are at high risk of hypoxemia using the PREPARE hypoxemia risk score could improve recognition of children who would benefit from further management and may reduce morbidity and mortality in lower-level health facilities without access to pulse oximetry, even when other well-recognized indications of referral are absent. These benefits must be weighed against the



burden of increased referrals on patients and health systems. Further validation and model updating in community settings is both a research and policy priority.

**ACKNOWLEDGEMENTS:** We would like to thank all study participants, clinicians and researchers involved in the included studies.

**FUNDING:** The study was funded by the Bill & Melinda Gates Foundation (#INV-007927) through a grant to the World Health Organization. The funders had no role in the study design or in the collection, analysis, or interpretation of the data. The funders did not write the report and had no role in the decision to submit the paper for publication.

**AUTHOR CONTRIBUTIONS:** RT, AC, FB, GL and YBN conceptualized and designed the study, interpreted the data and wrote the first draft of the manuscript. YBN verified the underlying data and conducted the statistical analyses. All authors oversaw data collection, verified the underlying data, assisted with the interpretation of the data and reviewed and provided input to the final draft. YBN had final responsibility for the decision to submit for publication.

**COMPETING INTERESTS:** We declare no competing interests. YBN is staff member of the World Health Organization. The authors alone are responsible for the views expressed in this article and they do not necessarily represent the views, decisions or policies of the institutions with which they are affiliated.

## REFERENCES

1. Rahman AE, Hossain AT, Nair H, et al. Prevalence of hypoxaemia in children with pneumonia in low-income and middle-income countries: a systematic review and meta-analysis. *The Lancet Global health* 2022;10(3):e348-e59. doi: 10.1016/s2214-109x(21)00586-6 [published Online First: 2022/02/19]
2. Colbourn T, King C, Beard J, et al. Predictive value of pulse oximetry for mortality in infants and children presenting to primary care with clinical pneumonia in rural Malawi: A data linkage study. *PLoS medicine* 2020;17(10):e1003300. doi: 10.1371/journal.pmed.1003300 [published Online First: 2020/10/24]
3. Chandna A, Tan R, Carter M, et al. Predictors of disease severity in children presenting from the community with febrile illnesses: a systematic review of prognostic studies. *BMJ Glob Health* 2021;6(1) doi: 10.1136/bmjgh-2020-003451 [published Online First: 2021/01/22]
4. Lazzerini M, Sonogo M, Pellegrin MC. Hypoxaemia as a Mortality Risk Factor in Acute Lower Respiratory Infections in Children in Low and Middle-Income Countries: Systematic Review and Meta-Analysis. *PLoS One* 2015;10(9):e0136166. doi: 10.1371/journal.pone.0136166
5. Enoch AJ, English M, Shepperd S. Does pulse oximeter use impact health outcomes? A systematic review. *Archives of disease in childhood* 2016;101(8):694-700. doi: 10.1136/archdischild-2015-309638 [published Online First: 2015/12/25]
6. Hooli S, King C, McCollum ED, et al. In-hospital mortality risk stratification in children aged under 5 years with pneumonia with or without pulse oximetry: A secondary analysis of the Pneumonia REsearch Partnership to Assess WHO REcommendations (PREPARE) dataset. *Int J Infect Dis* 2023;129:240-50. doi: 10.1016/j.ijid.2023.02.005 [published Online First: 2023/02/23]
7. Duke T, Graham S, Cherian M, et al. Oxygen is an essential medicine: a call for international action [Unresolved issues]. *The International journal of tuberculosis and lung disease* 2010;14(11):1362-68.
8. Awasthi S, Rastogi T, Pandey AK, et al. Epidemiology of Hypoxic Community-Acquired Pneumonia in Children Under 5 Years of Age: An Observational Study in Northern India. *Frontiers in pediatrics* 2021;9:790109. doi: 10.3389/fped.2021.790109 [published Online First: 2022/03/01]
9. Sylvies F, Nyirenda L, Blair A, et al. The impact of pulse oximetry and Integrated Management of Childhood Illness (IMCI) training on antibiotic prescribing practices in rural Malawi: A mixed-methods study. *PLoS One* 2020;15(11):e0242440. doi: 10.1371/journal.pone.0242440 [published Online First: 2020/11/20]
10. UNITAID. The ACT-A Oxygen Emergency Taskforce... one year on 2022 [updated 23.02.2022]. Available from: <https://unitaid.org/news-blog/the-act-a-oxygen-emergency-taskforceone-year-on/#en> accessed 09.11.2022 2022.
11. Källander K, Young M, Qazi S. Universal access to pneumonia prevention and care: a call for action. *The Lancet Respiratory Medicine* 2014;2(12):950-52.
12. Ansermino JM. Universal access to essential vital signs monitoring. *Anesthesia & Analgesia* 2013;117(4):883-90.
13. Graham HR, King C, Duke T, et al. Hypoxaemia and risk of death among children: rethinking oxygen saturation, risk-stratification, and the role of pulse oximetry in primary care. *The Lancet Global Health* 2024;12(8):e1359-e64. doi: 10.1016/S2214-109X(24)00209-2
14. Ginsburg AS, Van Cleve WC, Thompson MIW, et al. Oxygen and Pulse Oximetry in Childhood Pneumonia: A Survey of Healthcare Providers in Resource-limited Settings. *Journal of tropical pediatrics* 2012;58(5):389-93. doi: 10.1093/tropej/fmr103
15. McCollum ED, Bjornstad E, Preidis GA, et al. Multicenter study of hypoxemia prevalence and quality of oxygen treatment for hospitalized Malawian children. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2013;107(5):285-92. doi: 10.1093/trstmh/trt017

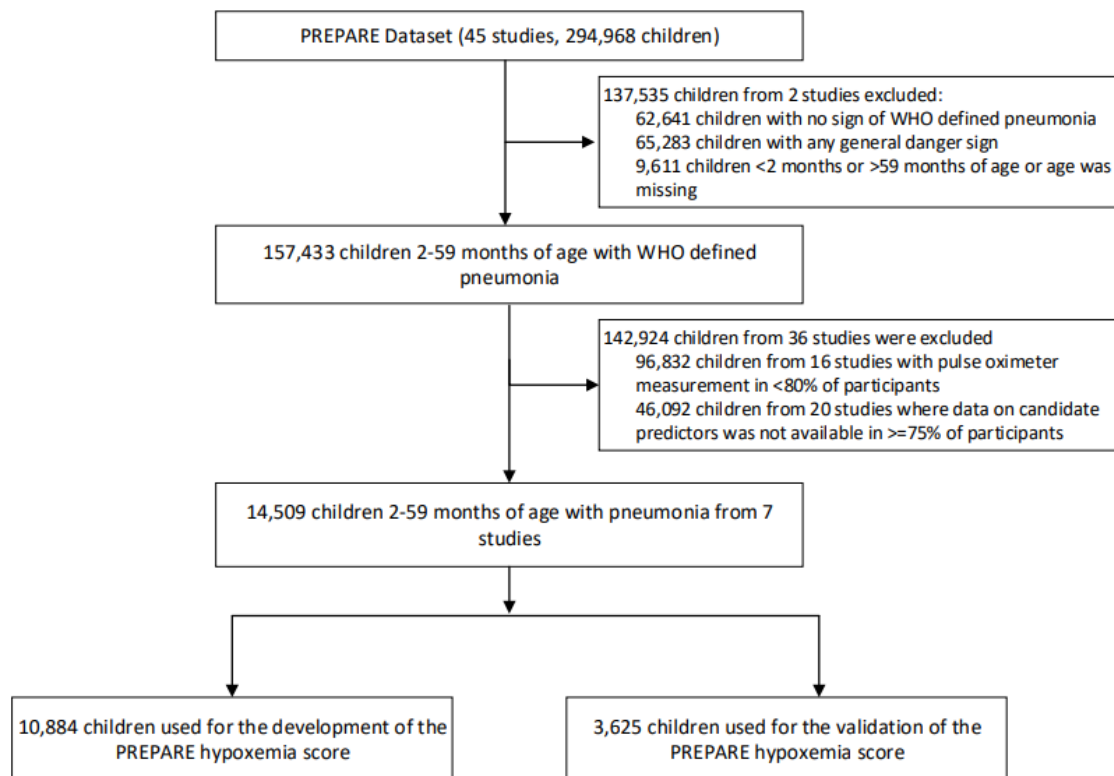
16. English M, Gathara D, Mwinga S, et al. Adoption of recommended practices and basic technologies in a low-income setting. *Archives of disease in childhood* 2014;99(5):452. doi: 10.1136/archdischild-2013-305561
17. Simkovich SM, Underhill LJ, Kirby MA, et al. Resources and Geographic Access to Care for Severe Pediatric Pneumonia in Four Resource-limited Settings. *Am J Respir Crit Care Med* 2022;205(2):183-97. doi: 10.1164/rccm.202104-1013OC [published Online First: 2021/10/19]
18. Enoch AJ, English M, McGivern G, et al. Variability in the use of pulse oximeters with children in Kenyan hospitals: A mixed-methods analysis. *PLoS medicine* 2019;16(12) doi: 10.1371/journal.pmed.1002987; 10.1371/journal.pmed.1002987.r001; 10.1371/journal.pmed.1002987.r002; 10.1371/journal.pmed.1002987.r003; 10.1371/journal.pmed.1002987.r004; 10.1371/journal.pmed.1002987.r005
19. McCollum ED, King C, Deula R, et al. Pulse oximetry for children with pneumonia treated as outpatients in rural Malawi. *Bulletin of the World Health Organization* 2016;94(12):893-902. doi: 10.2471/blt.16.173401 [published Online First: 2016/12/21]
20. King C, Boyd N, Walker I, et al. Opportunities and barriers in paediatric pulse oximetry for pneumonia in low-resource clinical settings: a qualitative evaluation from Malawi and Bangladesh. *BMJ open* 2018;8(1):e019177. doi: 10.1136/bmjopen-2017-019177 [published Online First: 2018/02/01]
21. World Health Organization. *Integrated Management of Childhood Illness Chartbook*. Geneva, 2014.
22. Schuh HB, Hooli S, Ahmed S, et al. Clinical hypoxemia score for outpatient child pneumonia care lacking pulse oximetry in Africa and South Asia. *Frontiers in pediatrics* 2023;11 doi: 10.3389/fped.2023.1233532
23. Lodha R, Bhadauria PS, Kuttikat AV, et al. Can clinical symptoms or signs accurately predict hypoxemia in children with acute lower respiratory tract infections? *Indian pediatrics* 2004;41(2):129-35. [published Online First: 2004/03/09]
24. Usen S, Weber M, Mulholland K, et al. Clinical predictors of hypoxaemia in Gambian children with acute lower respiratory tract infection: prospective cohort study. *Bmj* 1999;318(7176):86-91. doi: 10.1136/bmj.318.7176.86 [published Online First: 1999/01/08]
25. von der Weid L, Gehri M, Camara B, et al. Clinical signs of hypoxaemia in children aged 2 months to 5 years with acute respiratory distress in Switzerland and Senegal. *Paediatrics and International Child Health* 2018;38(2):113-20. doi: 10.1080/20469047.2017.1390828
26. Zhang L, Mendoza-Sassi R, Santos JC, et al. Accuracy of symptoms and signs in predicting hypoxaemia among young children with acute respiratory infection: a meta-analysis. *Int J Tuberc Lung Dis* 2011;15(3):317-25. [published Online First: 2011/02/22]
27. Martin H, Falconer J, Addo-Yobo E, et al. Assembling a global database of child pneumonia studies to inform WHO pneumonia management algorithm: Methodology and applications. *J Glob Health* 2022;12:04075. doi: 10.7189/jogh.12.04075 [published Online First: 2022/12/30]
28. Moons KGM, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration. *Annals of internal medicine* 2015;162(1):W1-W73. doi: 10.7326/M14-0698 %J *Annals of Internal Medicine*
29. World Health Organization. *Pocket Book of Hospital Care for Children: Guidelines for the Management of Common Childhood Illnesses*. 2nd ed. Geneva: World Health Organization 2013.
30. Fields E, Chard J, Murphy MS, et al. Assessment and initial management of feverish illness in children younger than 5 years: summary of updated NICE guidance. *Bmj* 2013;346:f2866. doi: 10.1136/bmj.f2866 [published Online First: 2013/05/24]

31. Harris M, Clark J, Coote N, et al. British Thoracic Society guidelines for the management of community acquired pneumonia in children: update 2011. *Thorax* 2011;66 Suppl 2:ii1-23. doi: 10.1136/thoraxjnl-2011-200598 [published Online First: 2011/10/19]
32. Fung JST, Akech S, Kissoon N, et al. Determining predictors of sepsis at triage among children under 5 years of age in resource-limited settings: A modified Delphi process. *PLoS One* 2019;14(1):e0211274. doi: 10.1371/journal.pone.0211274 [published Online First: 2019/01/29]
33. Riley RD, Snell KI, Ensor J, et al. Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes. *Stat Med* 2019;38(7):1276-96. doi: 10.1002/sim.7992 [published Online First: 2018/10/26]
34. Collins GS, Ogundimu EO, Altman DG. Sample size considerations for the external validation of a multivariable prognostic model: a resampling study. *Statistics in medicine* 2016;35(2):214-26. doi: 10.1002/sim.6787 [published Online First: 2015/11/09]
35. Vergouwe Y, Steyerberg EW, Eijkemans MJ, et al. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. *J Clin Epidemiol* 2005;58(5):475-83. doi: 10.1016/j.jclinepi.2004.06.017 [published Online First: 2005/04/23]
36. World Health Organization. Oxygen therapy for children: a manual for health workers. 2016
37. Jakobsen JC, Gluud C, Wetterslev J, et al. When and how should multiple imputation be used for handling missing data in randomised clinical trials – a practical guide with flowcharts. *BMC Medical Research Methodology* 2017;17(1):162. doi: 10.1186/s12874-017-0442-1
38. Bonnett LJ, Snell KIE, Collins GS, et al. Guide to presenting clinical prediction models for use in clinical settings. *BMJ* 2019;365:l737. doi: 10.1136/bmj.l737
39. Sullivan LM, Massaro JM, D'Agostino Sr RB. Presentation of multivariate data for clinical use: The Framingham Study risk score functions. *Statistics in medicine* 2004;23(10):1631-60.
40. Stata Statistical Software: Release 16. [program]. College Station, TX: StataCorp LLC., 2019.
41. Cutts F, Zaman S, Enwere Gy, et al. Efficacy of nine-valent pneumococcal conjugate vaccine against pneumonia and invasive pneumococcal disease in The Gambia: randomised, double-blind, placebo-controlled trial. *The Lancet* 2005;365(9465):1139-46.
42. Basnet S, Shrestha PS, Sharma A, et al. A randomized controlled trial of zinc as adjuvant therapy for severe pneumonia in young children. *Pediatrics* 2012;129(4):701-08.
43. Lazzerini M, Seward N, Lufesi N, et al. Mortality and its risk factors in Malawian children admitted to hospital with clinical pneumonia, 2001-12: a retrospective observational study. *The Lancet Global health* 2016;4(1):e57-68. doi: 10.1016/s2214-109x(15)00215-6 [published Online First: 2016/01/01]
44. Mathew JL, Singhi S, Ray P, et al. Etiology of community acquired pneumonia among children in India: prospective, cohort study. *Journal of global health* 2015;5(2)
45. McCollum ED, Nambiar B, Deula R, et al. Impact of the 13-valent pneumococcal conjugate vaccine on clinical and hypoxemic childhood pneumonia over three years in central Malawi: an observational study. *PLoS One* 2017;12(1):e0168209.
46. O'Grady KAF, Torzillo PJ, Ruben AR, et al. Identification of radiological alveolar pneumonia in children with high rates of hospitalized respiratory infections: Comparison of WHO-defined and pediatric pulmonologist diagnosis in the clinical context. *Pediatric pulmonology* 2012;47(4):386-92.
47. Wadhwa N, Chandran A, Aneja S, et al. Efficacy of zinc given as an adjunct in the treatment of severe and very severe pneumonia in hospitalized children 2–24 mo of age: a randomized, double-blind, placebo-controlled trial. *The American journal of clinical nutrition* 2013;97(6):1387-94.
48. Rao YK, Midha T, Kumar P, et al. Clinical predictors of hypoxemia in Indian children with acute respiratory tract infection presenting to pediatric emergency department. *World J*

- Pediatr 2012;8(3):247-51. doi: 10.1007/s12519-012-0365-1 [published Online First: 2012/08/14]
49. Alwadhi V, Dewan P, Malhotra RK, et al. Tachypnea and other danger signs vs pulse oximetry for prediction of hypoxia in severe pneumonia/very severe disease. *Indian pediatrics* 2017;54(9):729-34. doi: 10.1007/s13312-017-1163-6
  50. Tesfaye SH, Gebeyehu Y, Loha E, et al. Pulse oximeter with integrated management of childhood illness for diagnosis of severe childhood pneumonia at rural health institutions in Southern Ethiopia: results from a cluster-randomised controlled trial. *BMJ open* 2020;10(6):e036814. doi: 10.1136/bmjopen-2020-036814
  51. Onyango FE, Steinhoff MC, Wafula EM, et al. Hypoxaemia in young Kenyan children with acute lower respiratory infection. *British Medical Journal* 1993;306(6878):612. doi: 10.1136/bmj.306.6878.612
  52. Riley RD, Pate A, Dhiman P, et al. Clinical prediction models and the multiverse of madness. *BMC Medicine* 2023;21(1):502. doi: 10.1186/s12916-023-03212-y
  53. Collins GS, Dhiman P, Ma J, et al. Evaluation of clinical prediction models (part 1): from development to external validation. *BMJ* 2024;384:e074819. doi: 10.1136/bmj-2023-074819
  54. Fackler JC, Rehman M, Winslow RL. Please Welcome the New Team Member: The Algorithm. *Pediatric critical care medicine : a journal of the Society of Critical Care Medicine and the World Federation of Pediatric Intensive and Critical Care Societies* 2019;20(12):1200-01. doi: 10.1097/pcc.0000000000002149 [published Online First: 2019/12/06]
  55. Tustumi F. Choosing the most appropriate cut-point for continuous variables. *Revista do Colégio Brasileiro de Cirurgiões* 2022;49:e20223346.
  56. Rees CA, Colbourn T, Hooli S, et al. Derivation and validation of a novel risk assessment tool to identify children aged 2-59 months at risk of hospitalised pneumonia-related mortality in 20 countries. *BMJ Glob Health* 2022;7(4) doi: 10.1136/bmjgh-2021-008143 [published Online First: 2022/04/17]
  57. Feiner JR, Severinghaus JW, Bickler PE. Dark skin decreases the accuracy of pulse oximeters at low oxygen saturation: the effects of oximeter probe type and gender. *Anesth Analg* 2007;105(6 Suppl):S18-s23. doi: 10.1213/01.ane.0000285988.35174.d9 [published Online First: 2007/12/06]
  58. Sjoding MW, Dickson RP, Iwashyna TJ, et al. Racial Bias in Pulse Oximetry Measurement. *The New England journal of medicine* 2020;383(25):2477-78. doi: 10.1056/NEJMc2029240 [published Online First: 2020/12/17]
  59. EMPIC Study Group. Innovative, enhanced community management of non-hypoxaemic chest-indrawing pneumonia in 2–59-month-old children: a cluster-randomised trial in Africa and Asia. *BMJ Global Health* 2022;7(1):e006405. doi: 10.1136/bmjgh-2021-006405

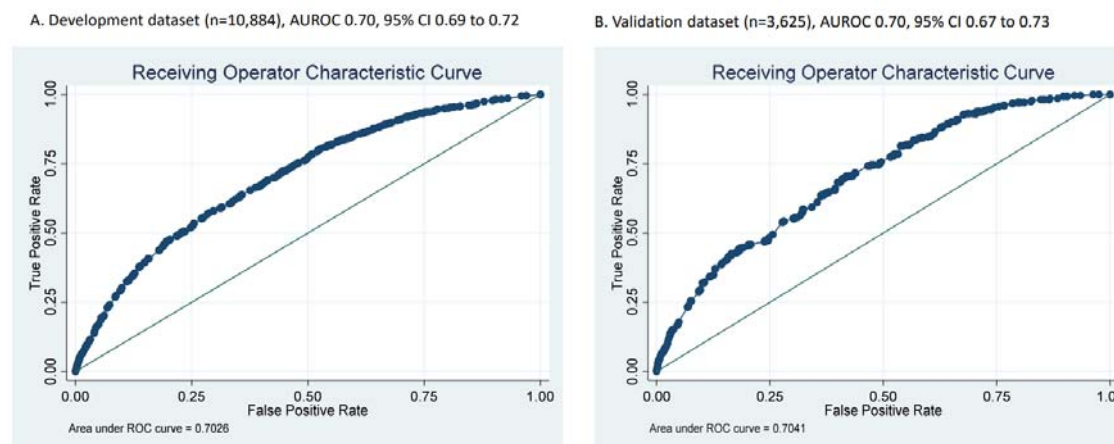
## FIGURES

**Figure 1:** Flow diagram of PREPARE study dataset describing which datasets were excluded/included, the split between development and validation datasets and proportion of patients meeting the primary outcome (SpO<sub>2</sub> <90%)

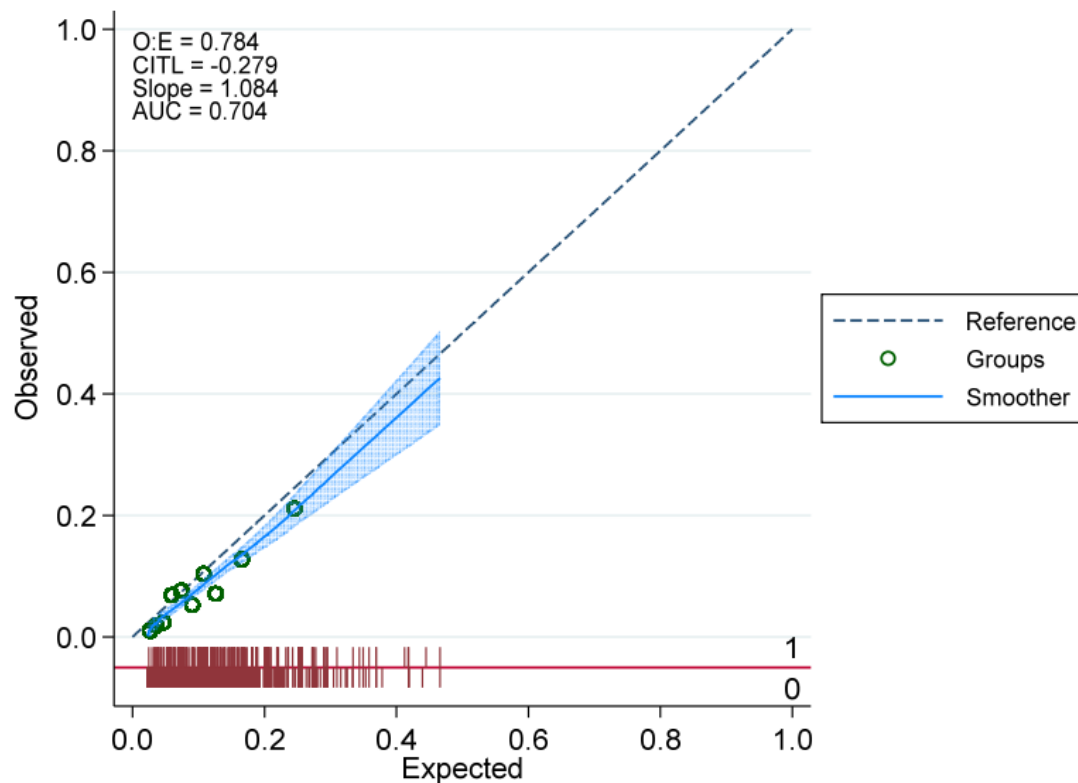


**Figure legend 1:** Danger signs were defined by the IMCI<sup>8</sup> chartbook (i.e., inability to drink, lethargy or unconsciousness, convulsions, vomiting everything, or stridor in a calm child).

**Figure 2.** Receiver operating characteristic curve for the PREPARE hypoxemia clinical prediction model for children 2–59 months of age with pneumonia



**Figure 3:** Calibration plot of the PREPARE hypoxemia clinical prediction model in the validation dataset



**Figure Legend 3:** Reference (dashed line) indicates perfect calibration; Lowess (blue line) indicates calibration slope; Red rug plots indicate the distribution of predicted risk of participants with hypoxemia (above the line) and those without hypoxemia (below the line), E:O = expected to observed events ratio; CITL = calibration-in-the-large; AUC = area under the receiver operating characteristic curve.