

**Title:** Clinical hypoxemia score for outpatient child pneumonia care lacking pulse oximetry in Africa and South Asia

\*Holly B. Schuh<sup>1</sup>

\*Shubhada Hooli<sup>2</sup>

Salahuddin Ahmed<sup>3</sup>

Carina King<sup>4</sup>

Arunangshu D. Roy<sup>3</sup>

Norman Lufesi<sup>5</sup>

ASMD Ashraful Islam<sup>3</sup>

Tisungane Mvalo<sup>6,7</sup>

Nabidul H. Chowdhury<sup>3</sup>

Amy Sarah Ginsburg<sup>8</sup>

Tim Colbourn<sup>9</sup>

William Checkley<sup>10,11</sup>

Abdullah H. Baqui<sup>12</sup>

Eric D. McCollum<sup>12,13</sup>

<sup>1</sup>Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

<sup>2</sup>Division of Emergency Medicine, Department of Pediatrics, Baylor College of Medicine, Houston, Texas

<sup>3</sup>Projahnmo Research Foundation, Dhaka, Bangladesh

<sup>4</sup>Department of Global Public Health, Karolinska Institutet, Stockholm, Sweden

<sup>5</sup>Malawi Ministry of Health, Lilongwe, Malawi

<sup>6</sup>University of North Carolina (UNC) Project Malawi, Lilongwe, Malawi

<sup>7</sup>Department of Pediatrics, UNC, Chapel Hill, NC, USA

<sup>8</sup>Clinical Trial Center, University of Washington, Seattle, United States

<sup>9</sup>Institute for Global Health, University College London, London, United Kingdom

<sup>10</sup>Division of Pulmonary and Critical Care, Department of Medicine, School of Medicine, Johns Hopkins University, Baltimore, Maryland, USA

<sup>11</sup>Center for Global Non-Communicable Disease Research and Training, School of Medicine, Johns Hopkins University, Baltimore, Maryland, USA

<sup>12</sup>Health Systems Program, Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

<sup>13</sup>Global Program in Pediatric Respiratory Sciences, Eudowood Division of Pediatric Respiratory Sciences, Department of Pediatrics, School of Medicine, Johns Hopkins University, Baltimore, Maryland, USA

\*Contributed equally

**Corresponding Author:**

Dr. Eric D. McCollum, MD, MPH

Associate Professor of Pediatrics

Eudowood Division of Pediatric Respiratory Sciences

Director, Global Program in Respiratory Sciences

Johns Hopkins University

Baltimore, MD 21287

email: [emccoll3@jhmi.edu](mailto:emccoll3@jhmi.edu)

**Key words:** Respiratory tract infections, Hypoxia, Clinical Decision Rules, Pediatrics,  
Primary Health Care

## ABSTRACT

**Background:** Pulse oximeters are not routinely available in outpatient clinics in low- and middle-income countries. We derived clinical scores to identify hypoxemic child pneumonia.

**Methods:** This was a retrospective pooled analysis of two outpatient datasets of 3-35 month olds with World Health Organization (WHO)-defined pneumonia in Bangladesh and Malawi. We constructed, internally validated, and compared fit & discrimination of four models predicting  $SpO_2 < 93\%$  and  $< 90\%$ : (1) Integrated Management of Childhood Illness guidelines, (2) WHO-composite guidelines, (3) Independent variable least absolute shrinkage and selection operator (LASSO); (4) Composite variable LASSO.

**Results:** 12,712 observations were included. The independent and composite LASSO models discriminated moderately (both C-statistic 0.77) between children with a  $SpO_2 < 93\%$  and  $\geq 94\%$ ; model predictive capacities remained moderate after adjusting for potential overfitting (C-statistic 0.74 and 0.75). The IMCI and WHO-composite models had poorer discrimination (C-statistic 0.56 and 0.68) and identified 20.6% and 56.8% of  $SpO_2 < 93\%$  cases. The highest score stratum of the independent and composite LASSO models identified 46.7% and 49.0% of  $SpO_2 < 93\%$  cases. Both LASSO models had similar performance for a  $SpO_2 < 90\%$ .

**Conclusions:** In the absence of pulse oximeters, both LASSO models better identified outpatient hypoxemic pneumonia cases than the WHO guidelines. Score external validation and implementation are needed.

## *Introduction*

The burden of child pneumonia mortality predominantly occurs in low-income and middle-income countries (LMICs).<sup>1</sup> Hypoxemia – a low blood oxyhemoglobin saturation – conveys increased pneumonia mortality risk, yet most children in LMICs lack pulse oximeter access.<sup>2,3</sup> Pulse oximeters identify hypoxemia by non-invasively measuring the peripheral arterial oxyhemoglobin saturation (SpO<sub>2</sub>).<sup>4</sup> To simplify diagnosis in LMICs the World Health Organization (WHO) Integrated Management of Childhood Illness (IMCI) guidelines consider pneumonia a clinical syndrome.<sup>5</sup> Although IMCI recommends oximeter use when available, it also provides guidance for settings without pulse oximetry.<sup>5</sup> However, when implemented without pulse oximeters, IMCI missed almost 70% of outpatient pneumonia cases with a SpO<sub>2</sub><90% in Malawi and nearly 90% in Bangladesh.<sup>6,7</sup> As most children first access health systems at outpatient clinics, improving outpatient hypoxemia identification may be key to reducing LMIC pneumonia mortality.<sup>4</sup>

Prior research attempted to determine whether clinical signs accurately identify a SpO<sub>2</sub><90% in hospitalized children.<sup>8,9</sup> While research showed high specificity of clinical signs for a SpO<sub>2</sub> <90%, sensitivity was low.<sup>8,9</sup> Additionally, recent data showed elevated mortality among children with a SpO<sub>2</sub> 90-92% than higher SpO<sub>2</sub> levels.<sup>7,10-12</sup> These findings challenge the currently recommended SpO<sub>2</sub><90% threshold for hospitalization and oxygen treatment in LMICs.

We sought to utilize two unique, contemporary outpatient pediatric pneumonia datasets from Bangladesh and Malawi to accomplish three objectives.<sup>13,14</sup> First, evaluate IMCI performance for identifying hypoxemia at a higher SpO<sub>2</sub> threshold (<93%) than recommended. Second, examine whether other combinations of clinical features better identify children with a SpO<sub>2</sub><93%, followed by development and internal validation of hypoxemia clinical scores feasible for LMICs. Third, repeat these analyses using the currently recommended SpO<sub>2</sub><90% threshold.

## **METHODS**

### *Settings*

We used data from Malawi and Bangladesh. Malawi is an African country with an under 5 mortality rate of 49/1,000 births.<sup>15</sup> This study included 18 clinics in Mchinji and Lilongwe districts with a 1.2 million population catchment area, at 1,000-1,100 meters altitude.<sup>13</sup> From October 2011–June 2014, non-physician clinicians and nurses at clinics were IMCI trained and documented care of 0-59 month olds with WHO-defined pneumonia.<sup>13</sup> Providers used Aicare® pulse oximeters with adult clip probes on the big toe if <2 years old or weighing <10kg.<sup>6</sup> Training, data collection, and supervision methodology has been published.<sup>13</sup>

Bangladesh has an under 5 mortality rate of 3/1,000 live births.<sup>16</sup> Since 2001, Projahnmo, a partnership of Johns Hopkins University with the Government of Bangladesh's Ministry of Health and Family Welfare, International Centre for Diarrhoeal Disease Research, Bangladesh, Shimantik, and the Child Health Research Foundation

conducted community-based surveillance in Zakiganj subdistrict of Sylhet district in Bangladesh.<sup>14</sup> From May 2015–September 2017 Projahnmo expanded surveillance into two additional subdistricts (Kanaighat, Beanibazar) and augmented IMCI clinic care of three government Upazila Health Complexes.<sup>14</sup> Altogether these subdistricts have a 770,000 population at 17-23 meters altitude. Upazila Health Complexes provide outpatient and emergency care and limited inpatient pediatric services. IMCI clinic care was provided by Projahnmo physicians per IMCI guidelines.<sup>14</sup> From October 2017 Projahnmo physicians measured the SpO<sub>2</sub> of 3-35 month olds with suspected pneumonia using a Masimo Rad-5® pulse oximeter with a LNCS® Y-I wrap sensor on the big toe. Parent study methodology is published.<sup>14</sup>

### *Inclusion and Exclusion Criteria*

We generated an analytic sample of healthcare visits from Bangladesh and Malawi datasets. Inclusion criteria were: valid SpO<sub>2</sub>, IMCI-defined non-severe or severe pneumonia (2014 guidelines),<sup>5</sup> and age 3-35 months, as the Bangladesh study population with a SpO<sub>2</sub> was limited to this age.<sup>17</sup> See Supplemental Material for study definitions. Implicit to this analysis we assumed SpO<sub>2</sub> was unavailable and excluded it from pneumonia definitions.

### *Variables*

Our primary outcome was a SpO<sub>2</sub><93%; SpO<sub>2</sub><90% was secondary. We explored associations between clinical variables and SpO<sub>2</sub> ranges to evaluate <93% as the primary outcome. We selected variables *a priori*: sex, age, weight-for-age z-score

(WAZ), chest indrawing, wheezing, severe respiratory distress (grunting, head nodding, nasal flaring, and/or age-adjusted very fast breathing), cyanosis, fever (temperature  $\geq 38^{\circ}\text{C}$ ), and WHO-defined general danger signs (stridor, inability to feed/drink, convulsions, and/or lethargy).<sup>8,18</sup>

### *Analysis*

We evaluated missingness using a 5% threshold. We used Chi squared and Fisher's exact tests for proportions, Wilcoxon rank-sum for non-parametric data, and Student's t-test for normally distributed data comparisons. We reviewed individual-level variables and their associated SpO<sub>2</sub><93% and <90% predictive quality. For SpO<sub>2</sub><93% and <90% model development we randomly split the dataset into derivation (70%) and validation (30%) sets balanced by outcome and country. For each model we fit a logistic regression model with hypoxemia as the binary outcome measure. We allowed selection of *country* as a fixed effect to account for significant differences by country. Thereafter, we used a random intercept in each post selection model to control for country after interrogating the suitability of this approach with the Hausman test.<sup>19</sup> All analyses were by Stata 16.1 (StataCorp, College Station, TX).

### *Model development*

#### *IMCI guidelines (IMCI model) and Composite WHO guidelines (WHO-composite model)*

The IMCI model reflects 2014 IMCI referral criteria (Supplemental Material).<sup>5</sup> The WHO-composite model is a composite of four WHO guidelines.<sup>5,18,20,21</sup> We fit both models using the *logistic* command for implementing multivariable, maximum-likelihood logit



models to obtain odds ratios (and 95% confidence intervals) comparing the odds of hypoxemia versus non-hypoxemia.

*Independent variable LASSO model (Independent LASSO model) and Composite variable LASSO model (Composite LASSO model)*

We used the least absolute shrinkage and selection operator (LASSO) reduction method, testing two selection modes ((1) 10-fold cross-validation selection and (2) adaptive selection) using the derivation dataset to develop both models from an expanded variable list of the IMCI and WHO-composite models (Supplemental Material).<sup>22,23</sup> For Independent LASSO we used *singular* variables (i.e., *independent*) whereas for Composite LASSO we used *composite* variables for ‘danger signs’ (WHO-defined general danger signs) and ‘severe respiratory distress.’ For both models we compared the two methods using the C-statistic based on predicted estimates, sensitivity, and specificity. If we found no statistical difference between methods, we used the selection results from the simplest model to implement an unsupervised approach (i.e. selection of the full variable rather than one category of a three-category variable) to refit both models.

*Hypoxemia score development and validation*

We compared discriminatory power and model fit (C-statistic and Bayesian Information Criteria (BIC)) of the four maximum-likelihood logit models using the derivation dataset models.<sup>24,25</sup> Using an unsupervised approach (i.e., if only two age groups were selected, we retained all age categories), each of the LASSO model covariates were

kept on the log scale, rounded to the nearest 0.5, and doubled to form an integer.<sup>10,26,27</sup>

We then split the score into approximately equally sized quintiles to create hypoxemia risk categories.

Using the validation dataset, LASSO model scores were estimated by child, and score discriminatory power to identify children with and without hypoxemia was determined. Scores were not developed using IMCI and WHO-composite models. The C-statistic, sensitivity, specificity, positive and negative predictive value (PPV and NPV), and positive and negative likelihood ratios (LR+ and LR-) were compared across all models. We adjusted C-statistics for optimism by bootstrapping (200 repetitions) to account for any overfitting.<sup>28</sup> A C-statistic 0.71-0.80 was considered moderate and >0.80 as excellent discriminatory power.<sup>29</sup> We applied the same analysis methodology for SpO<sub>2</sub><90%.

### *Ethical Approval*

Ethical approval was provided by National Health Sciences Research Committee of Malawi, the Ethics Committee of University College London, and The Bangladesh Medical Research Counsel, International Centre for Diarrhoeal Diseases Research, Johns Hopkins Bloomberg School of Public Health and School of Medicine Institutional Review Boards. Caregiver written consent was required in Bangladesh but not Malawi.

## **RESULTS**

### *Study Population*

We included 12,712 pneumonia cases; 63.6% (n=8,081) were from Bangladesh (Supplemental Material). Table 1 shows all participant characteristics by SpO<sub>2</sub>. A SpO<sub>2</sub><93% was in 10.4% (1,328/12,712) of cases and 63.6% (845/1,328) of SpO<sub>2</sub><93% cases were in Malawi. Most SpO<sub>2</sub><93% cases had non-severe pneumonia (1,065/1,328; 81.4%) and, without a SpO<sub>2</sub> measurement, were hospitalization ineligible per IMCI guidelines. A larger proportion of severe (263/1,198, 21.9%) than non-severe (1,065/11,514, 9.2%) cases had a SpO<sub>2</sub><93%. A SpO<sub>2</sub><90% was in 4.6% (602/12,712) of cases and in 3.8% (434/11,514) with non-severe disease. 2014 IMCI hospital referral criteria missed 72.0% (434/602) of SpO<sub>2</sub><90% cases. While Bangladesh and Malawi case characteristics differed in frequency, other than WAZ<-3 and danger signs, they had similar crude associations with a SpO<sub>2</sub><93% (Supplemental Material).

We assessed the relationship between referral criteria and hypoxemia at SpO<sub>2</sub><90%, 90%-92%, and <93% for the IMCI and WHO-composite models (Table 2). WAZ<sub>≤</sub>-3, in both models, was associated with a SpO<sub>2</sub> 90-92% and <93% but not <90%. In the WHO-composite model severe respiratory distress was associated with an increased adjusted odds of an abnormal SpO<sub>2</sub> regardless of SpO<sub>2</sub> range. While in the IMCI model danger signs were associated with hypoxemia, including respiratory distress in the WHO-composite model attenuated its effect at each SpO<sub>2</sub> threshold. Respiratory distress was associated with each SpO<sub>2</sub> threshold in the WHO-composite model.

*Multivariable predictive model comparison and score development*

Derivation (n=3,813) and validation (n=8,899) datasets have similar patient characteristic distributions (Supplemental Material, all  $p>0.05$ ). Table 3 presents adjusted odds ratios (aORs) for  $SpO_2<93\%$  using model-specific predictors from the derivation dataset. In the independent LASSO model, individual predictor scores ranged from -1 to 3. The composite LASSO model scores ranged from -1 to 4. Overall predictive performance of the independent (C-statistic=0.774) and composite LASSO models (C-statistic=0.773) was moderate (Table 3) and did not differ in discriminatory power ( $p=0.480$ ), with total score ranges in the Supplemental Material. Model fit improved from the IMCI (BIC 5,536.5) and WHO-composite models (BIC 5,155.9) to the independent (BIC 4,650.6) and composite LASSO models (BIC 4,712.2). Overall, the independent (C-statistic=0.774) and composite LASSO models (C-statistic=0.773) better discriminated between children with and without a  $SpO_2<93\%$  than the IMCI (C-statistic=0.661) and WHO-composite models (C-statistic=0.734).  $SpO_2<90\%$  score development and cross-model comparison are in the Supplemental Material.

#### *Clinical score performance – validation dataset*

In the validation dataset, independent (C-statistic=0.745) and composite LASSO model (C-statistic=0.752) scores moderately discriminated between  $SpO_2<93\%$  and  $\geq 94\%$  cases (Figure 1). When examining the C-statistics adjusted for optimism, there was minimal C-statistic change for both scores (Table 4), and independent and composite LASSO model discriminatory power did not differ ( $p=0.480$ ). For the independent and composite LASSO models about 4% of children with a score in the first stratum had  $SpO_2<93\%$  compared to  $>35\%$  in the last stratum (Table 4).

### *Cross-model comparison: discriminating between hypoxemic and non-hypoxemic children*

Among IMCI and WHO-composite model classified cases, 23% and 24% had a  $SpO_2 < 93\%$ , while among non-cases 9% and 6% had a  $SpO_2 < 93\%$  (Table 4). The ability to discriminate between  $SpO_2 < 93\%$  and  $\geq 94\%$  cases using the IMCI (C-statistic=0.563) and WHO-composite model (C-statistic=0.677) criteria was low. Based on model fit and discrimination, the composite LASSO model was the most predictive of  $SpO_2 < 93\%$ , identifying 49.0% of  $SpO_2 < 93\%$  cases. The independent LASSO model performed similarly, followed by WHO-composite and IMCI models.  $SpO_2 < 90\%$  results on score validation and ability to discriminate hypoxemia from non-hypoxemia cases are in the Supplemental Material.

### *Clinical signs for a $SpO_2 < 93\%$*

The Supplemental Material includes the diagnostic performance of individual clinical signs for a  $SpO_2 < 93\%$  and  $< 90\%$  using the validation datasets.

## **DISCUSSION**

Using pooled data from 12,712 IMCI-defined child pneumonia cases evaluated at 21 clinics in Malawi and Bangladesh we examined WHO IMCI guideline hypoxemia identification performance and developed and internally validated clinical hypoxemia scores for use in LMICs where pulse oximeters suitable for pediatric outpatient care are scarce. Our findings suggest more hypoxemic children could be identified during

outpatient care lacking pulse oximeters if additional signs of respiratory distress are incorporated into IMCI. Notably, >80% of hypoxemic cases with a  $SpO_2 < 93\%$  and >70% with a  $SpO_2 < 90\%$  were misclassified by IMCI as ineligible for hospital referral when pulse oximeters are unavailable. The independent LASSO model added age, severe respiratory distress, chest indrawing, cyanosis, fever, and wheezing parameters into the base IMCI guideline model and better identified children with a  $SpO_2 < 93\%$  than the IMCI and WHO-composite models. The composite LASSO model's simplified composite variables may facilitate implementation. Both LASSO models also achieved excellent discrimination of  $SpO_2 < 90\%$  cases.

Unlike other studies evaluating hypoxemia predictors we focused on ambulatory rather than hospital settings. This distinction is important as our findings should therefore be generalizable to outpatient settings without oximeters and with lower hypoxemia prevalence than hospitals.<sup>30,31</sup> Notably, individual variable sensitivity and specificity for hypoxemia are largely similar to hospital-based studies.<sup>8</sup> However, our WHO-composite model and two LASSO models have good to excellent discriminatory value distinguishing between hypoxemic and non-hypoxemic pneumonia cases, contrasting to other work with smaller samples that limited analyses.<sup>10</sup> An exception is a large, multi-center, hospital-based Nigerian study that found a combination of respiratory distress, inability to feed, cyanosis, lethargy and severe tachypnea had lower discrimination (C-statistic=0.655) for  $SpO_2 < 90\%$  amongst respiratory and non-respiratory cases than in our data.<sup>9</sup> The lower C-statistic may reflect the authors inclusion of older children <15 years.

It is also important our results are interpreted within the broader child pneumonia context of known mortality risk factors. In the IMCI and WHO-composite models  $WAZ \leq -3$  was not associated with a  $SpO_2 < 90\%$ , yet is a known mortality risk factor.<sup>10,32</sup> Conversely wheezing was retained in both LASSO models but, when identified alone without accompanying respiratory distress, it is not associated with radiographic pneumonia or in-hospital mortality.<sup>10,33</sup> Isolated wheeze usually reflects milder, self-limited viral illness (e.g., bronchiolitis) and is susceptible to non-differential misclassification when confused with transmitted upper respiratory sounds.

While our pooling of data from studies in South Asia and Africa aimed to improve generalizability, there are differences between the settings and studies that may be limit. Malawi data had a higher frequency of hypoxemia, danger signs, and respiratory distress than Bangladesh data. Alternatively,  $WAZ \leq -3$  was more frequent in Bangladesh than Malawi. These differences, in part, may reflect higher HIV and malaria prevalence, modestly higher altitude in Malawi, or known challenges in identifying malnutrition in Malawi clinics.<sup>34</sup> These diseases increase susceptibility for hypoxemia or, for malaria, increase the frequency of signs overlapping with IMCI pneumonia.<sup>35</sup> Although the designs differed, all personnel were similarly trained per IMCI. The under 5 mortality rate in Malawi is higher than Bangladesh,<sup>15,16</sup> which aligns with our results suggesting Malawi cases were more severe. Nevertheless, we attempted to account for unmeasured confounders from epidemiological, health system, and methodological differences by fitting models with a country variable.

In sum, these findings may improve hypoxemic pneumonia identification in clinics either without pulse oximetry at all or lacking pulse oximeters suitable for pediatric use. Given a SpO<sub>2</sub> <93% is highly associated with mortality,<sup>2,10,12</sup> earlier hypoxemic case identification with successful referral may reduce fatality. While these models could advance care for hypoxemic children they are an inadequate substitution for pulse oximeters. Pulse oximetry scale up in ambulatory settings must be prioritized, as should pulse oximeter device development targeted for young children in LMICs. Where pulse oximetry is unavailable for children, our models suggest children with chest indrawing and/or with other signs of respiratory distress could be referred. To successfully implement these models ambulatory health workers of varying pediatric experience will need further training to recognize signs of respiratory distress in children. Understanding how these changes may affect referral quality, uptake, and hospitalizations is critical. Next steps include external validation and research evaluating implementation feasibility of the hypoxemia scores.

Figure 1. Comparison of ROC curves for identifying SpO<sub>2</sub> <93% cases (validation dataset)

Table 1. Patient characteristics by SpO<sub>2</sub><93% (full dataset, N=12,712)

Table 2. Association of Models 1 and 2 with SpO<sub>2</sub> ranges (full dataset, N= 12,712)

Table 3. Clinical hypoxemia score for identifying a SpO<sub>2</sub> <93% (derivation dataset)

Table 4. Model performance for identifying a SpO<sub>2</sub><93% (validation dataset)



Supplemental Material:

Supplemental Figure 1: Study Definitions

Supplemental Figure 2: Study flowchart

Supplemental Table 1: Bangladesh: Risk factors for a  $SpO_2 < 93\%$

Supplemental Table 2. Malawi: Risk factors for a  $SpO_2 < 93\%$

Supplemental Table 3. Patient characteristics by  $SpO_2 < 93\%$ , stratified by development versus validation dataset

Supplemental Table 4. Model 3 (independent) scores and associated hypoxemia risk ( $SpO_2 < 93\%$ ), sensitivity, specificity, and positive/negative likelihood ratios (validation dataset)

Supplemental Table 5. Model 4 (composite) scores and associated hypoxemia risk ( $SpO_2 < 93\%$ ), sensitivity, specificity, and positive/negative likelihood ratios (validation dataset)

Supplemental Figure 3. Comparison of ROC curves for identifying  $SpO_2 < 90\%$  cases (validation dataset)

Supplemental Table 6. Association of Models with  $SpO_2 < 90\%$  and performance for identifying  $SpO_2 < 90\%$  (development dataset)

Supplemental Table 7. Model 3 (independent) scores and associated hypoxemia risk ( $SpO_2 < 90\%$ ), sensitivity, specificity, and positive/negative likelihood ratios (validation dataset)

Supplemental Table 8. Model 4 (composite) scores and associated hypoxemia risk ( $SpO_2 < 90\%$ ), sensitivity, specificity, and positive/negative likelihood ratios (validation dataset)

Supplemental Table 9. Model performance and hypoxemia ( $\text{SpO}_2 < 90\%$ ) case rate  
(validation dataset)

Supplemental Table 10. Performance of clinical signs for a  $\text{SpO}_2 < 93\%$  during outpatient  
pediatric care in Bangladesh and Malawi (validation dataset)

Supplemental Table 11. Performance of clinical signs for a  $\text{SpO}_2 < 90\%$  during outpatient  
pediatric care in Bangladesh and Malawi (validation dataset)

### Contributors

EDM had full access to data in the study and takes full responsibility for the integrity of the data and the accuracy of the data analysis. EDM conceived the study idea and HS, SH, and EDM designed the study; HS analyzed the data; SH, HS, and EDM drafted and revised the manuscript. All authors made substantial contributions to the interpretation of results, critical revision of the manuscript, and approved the final version of the manuscript.

### Declaration of Interests

The authors have no competing interests to declare.

### Acknowledgments

We offer our thanks to all of the caregivers and children participating in this research in Bangladesh and Malawi. In Bangladesh we would like to thank the Projahnmo Study Group field and data management staff, the Ministry of Health and Family Welfare, Government of Bangladesh. In Malawi we would like to thank Mr. Charles Makwenda

and Rashid Deula for leading efforts in data collection and curation of the original study, the Republic of Malawi Ministry of Health and the communities and traditional authorities of Lilongwe and Mchinji districts of Malawi, and the dedication and hard work of our field and data collection staff. Lastly, we thank GlaxoSmithKline, Bill & Melinda Gates Foundation, and the National Institute of Health their support of this study.

## References

1. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory infections in 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Infect Dis* 2018;**18**:1191-1210.
2. 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**:e0136166.
3. Ginsburg AS, Van Cleve WC, Thompson MI, English M. Oxygen and pulse oximetry in childhood pneumonia: a survey of healthcare providers in resource-limited settings. *J Trop Pediatr* 2012;**58**:389-93.
4. McCollum ED, King C, Colbourn T, et al. Pulse oximetry in paediatric primary care in low-income and middle-income countries. *Lancet Respir Med* 2019;**7**:1001-1002.
5. World Health Organization, 2014. Integrated Management of Childhood Illness: Chart Booklet. Available at:

[https://apps.who.int/iris/bitstream/handle/10665/104772/9789241506823\\_Chartbook\\_eng.pdf](https://apps.who.int/iris/bitstream/handle/10665/104772/9789241506823_Chartbook_eng.pdf). Accessed August 28, 2022.

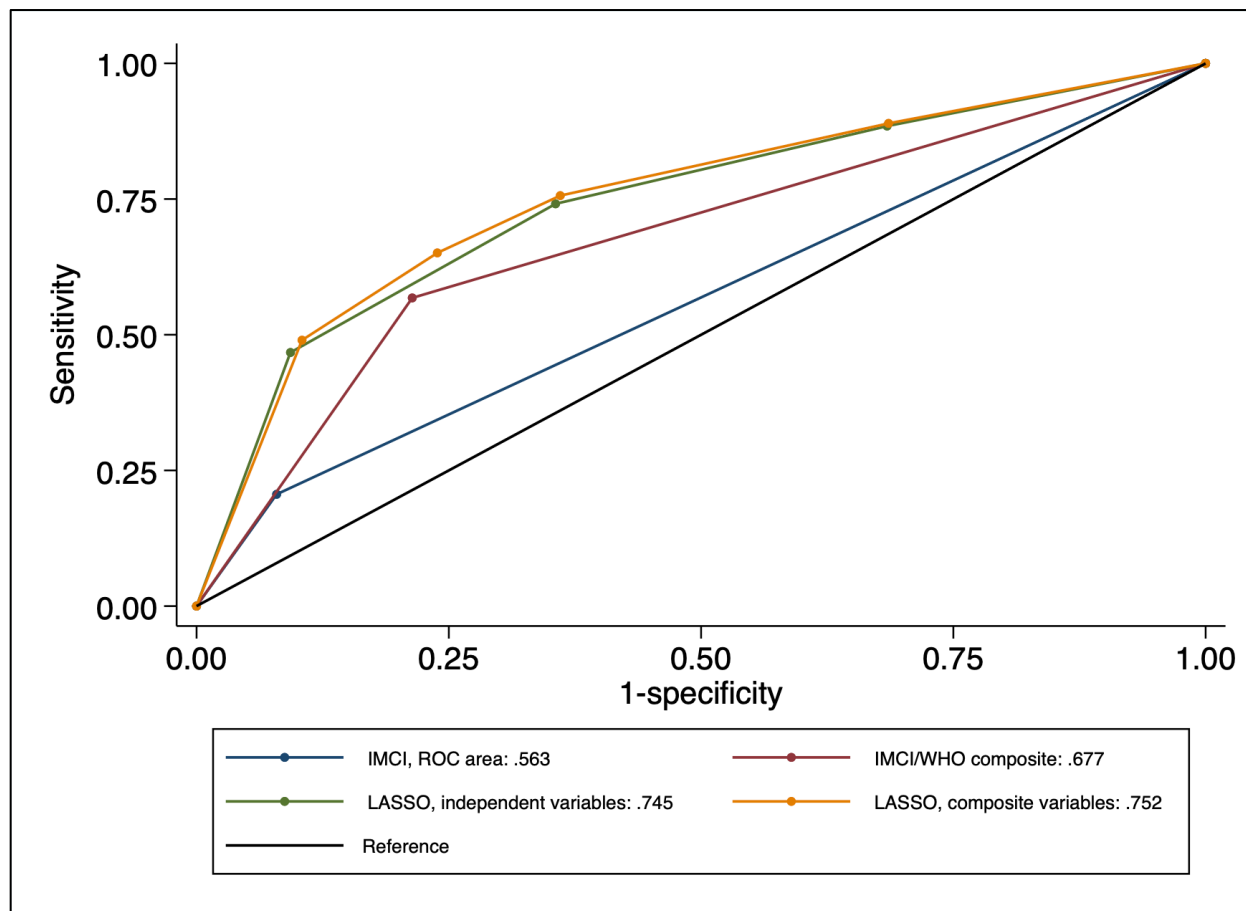
6. McCollum ED, King C, Deula R, et al. Pulse oximetry for children with pneumonia treated as outpatients in rural Malawi. *Bull World Health Organ* 2016;**94**:893-902.
7. McCollum ED, Ahmed S, Roy AD, et al. Risk and accuracy of outpatient-identified hypoxemia for death among suspected child pneumonia cases in rural Bangladesh: a multi-facility cohort study. *Lancet Respir Med* 2023; *In Press*.
8. Zhang L, Mendoza-Sassi R, Santos JC, Lau J. 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**:317-25.
9. Graham H, Bakare AA, Ayede AI, et al. Hypoxaemia in hospitalised children and neonates: A prospective cohort study in Nigerian secondary-level hospitals. *EClinicalMedicine* 2019;**16**:51-63.
10. Hooli S, Colbourn T, Lufesi N, et al. Predicting Hospitalised Paediatric Pneumonia Mortality Risk: An External Validation of RISC and mRISC, and Local Tool Development (RISC-Malawi) from Malawi. *PLoS ONE* 2016;**11**:e0168126.
11. Hooli S, King C, Zadutsa B, et al. The Epidemiology of Hypoxemic Pneumonia among Young Infants in Malawi. *Am J Trop Med Hyg* 2020;**102**:676-683.
12. 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 Med* 2020;**17**:e1003300.

13. 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**:e0168209.
14. Baqui AH, McCollum ED, Saha SK, et al. Pneumococcal Conjugate Vaccine impact assessment in Bangladesh. *Gates Open Res* 2018;**2**:21.
15. UNICEF, 2021. Malawi Country Profile. Available at: <https://data.unicef.org/country/mwi/>. Accessed August 28, 2022.
16. UNICEF, 2021. Bangladesh Country Profile. Available at: <https://data.unicef.org/country/bgd/>. Accessed August 28, 2022.
17. McCollum ED, Ahmed S, Roy AD, et al. Effectiveness of the 10-valent pneumococcal conjugate vaccine against radiographic pneumonia among children in rural Bangladesh: A case-control study. *Vaccine* 2020;**38**:6508-6516.
18. World Health Organization, 2016. Oxygen therapy for children. Available at: [https://apps.who.int/iris/bitstream/handle/10665/204584/9789241549554\\_eng.pdf?sequence=1&isAllowed=y](https://apps.who.int/iris/bitstream/handle/10665/204584/9789241549554_eng.pdf?sequence=1&isAllowed=y). Accessed August 28, 2022.
19. McNeish D, Kelley K. Fixed effects models versus mixed effects models for clustered data: Reviewing the approaches, disentangling the differences, and making recommendations. *Psychol Methods* 2019;**24**:20-35.
20. World Health Organization, 2013. Pocketbook of Hospital care for children: Guidelines for the management of common childhood illnesses; Second Edition. Available at: [https://apps.who.int/iris/bitstream/handle/10665/81170/9789241548373\\_eng.pdf?sequence=1](https://apps.who.int/iris/bitstream/handle/10665/81170/9789241548373_eng.pdf?sequence=1). Accessed August 28, 2022.

21. World Health Organization, 2005. Emergency triage assessment and treatment (ETAT): facilitator guide. Available at:  
[https://apps.who.int/iris/bitstream/handle/10665/43386/9241546883\\_eng.pdf?sequence=2](https://apps.who.int/iris/bitstream/handle/10665/43386/9241546883_eng.pdf?sequence=2). Accessed August 28, 2022.
22. Tibshirani R. Regression shrinkage and selection via the lasso: a retrospective. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 2011;**73**:273-282.
23. Zou H. The Adaptive Lasso and Its Oracle Properties. *Journal of the American Statistical Association* 2006;**101**:1418-1429.
24. Schwarz G. Estimating the Dimension of a Model. *The Annals of Statistics* **1978**;6:461-4.
25. Neath A, Cavanaugh J. The Bayesian information criterion: background, derivation, and applications. *WIREs Computational Statistics* 2012;**4**:199-203.
26. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996;**15**:361-87.
27. Le Gall JR, Lemeshow S, Saulnier F. A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. *JAMA* 1993;**270**:2957-63.
28. Smith GC, Seaman SR, Wood AM, Royston P, White IR. Correcting for optimistic prediction in small data sets. *Am J Epidemiol* 2014;**180**:318-24.

29. Austin PC, Steyerberg EW. Interpreting the concordance statistic of a logistic regression model: relation to the variance and odds ratio of a continuous explanatory variable. *BMC Med Res Methodol* 2012;**12**:82.
30. Subhi R, Adamson M, Campbell H, Weber M, Smith K, Duke T. The prevalence of hypoxaemia among ill children in developing countries: a systematic review. *Lancet Infect Dis* 9(4):219-27.
31. 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. *Lancet Glob Health* 2022;**10**:e348-e359.
32. Deardorff KV, McCollum ED, Ginsburg AS. Pneumonia Risk Stratification Scores for Children in Low-Resource Settings: A Systematic Literature Review. *Pediatr Infect Dis J* 2018;**37**:743-748.
33. McCollum ED, Park DE, Watson NL, et al. Digital auscultation in PERCH: Associations with chest radiography and pneumonia mortality in children. *Pediatr Pulmonol* 2020;**55**:3197-3208.
34. LaCourse SM, Chester FM, Preidis G, et al. Lay-screeners and use of WHO growth standards increase case finding of hospitalized Malawian children with severe acute malnutrition. *J Trop Pediatr* 2015;**61**:44-53.
35. Källander K, Nsungwa-Sabiiti J, Peterson S. Symptom overlap for malaria and pneumonia--policy implications for home management strategies. *Acta Trop* 2004;**90**:211-4.

Figure 1. Comparison of ROC curves for identifying SpO<sub>2</sub><93% cases (validation dataset)



ROC indicates Receiver Operating Characteristic Curve; SpO<sub>2</sub>, peripheral arterial oxyhemoglobin saturation; IMCI, Integrated Management of Childhood Illnesses; WHO, World Health Organization; LASSO, Least Absolute Shrinkage and Selection Operator reduction method.



Table 1. Patient characteristics by SpO<sub>2</sub><93% (full dataset, N=12,712)

	Total	SpO <sub>2</sub> ≥93%	SpO <sub>2</sub> <93%	p-value
Characteristics	N=12,712	N=11,384 (89.6%)	N=1,328 (10.4%)	
Study country, n (%)				<0.001
Bangladesh	8,081 (63.6%)	7,598 (66.7%)	483 (36.4%)	
Malawi	4,631 (36.4%)	3,786 (33.3%)	845 (63.6%)	
Child age (months), median (IQR)	12 (6, 19)	12 (6, 20)	10 (6, 17)	<0.001
Child sex, n (%)				0.36
Male	6,895 (54.2%)	6,207 (54.5%)	688 (51.8%)	
Female	5,530 (43.5%)	4,950 (43.5%)	580 (43.7%)	
Missing	287 (2.3%)	227 (2.0%)	60 (4.5%)	
Weight for age z-score, n (%)				0.056
-2.0	9,441 (74.3%)	8,463 (74.3%)	978 (73.6%)	
-3.0 ≤ z-score < -2.0	2,080 (16.4%)	1,899 (16.7%)	181 (13.6%)	
< -3.0	852 (6.7%)	760 (6.7%)	92 (6.9%)	

Missing	339 (2.7%)	262 (2.3%)	77 (5.8%)	
WHO danger signs, n (%)	364 (2.9%)	183 (1.6%)	181 (13.6%)	<0.001
Stridor at rest	126 (1.0%)	65 (0.6%)	61 (4.7%)	<0.001
Unable to feed	187 (1.5%)	74 (0.7%)	113 (8.5%)	<0.001
Lethargy or unconscious	58 (0.5%)	29 (0.3%)	29 (2.2%)	<0.001
Convulsions	62 (0.5%)	35 (0.3%)	27 (2.0%)	<0.001
Severe respiratory distress, n (%)	2,352 (18.5%)	1,667 (14.6%)	685 (51.6%)	<0.001
Grunting	399 (3.1%)	179 (1.6%)	220 (16.7%)	<0.001
Head nodding	636 (5.0%)	402 (3.5%)	234 (17.8%)	<0.001
Nasal flaring	1,219 (9.6%)	776 (6.8%)	443 (33.5%)	<0.001
Severe fast breathing <sup>1</sup>	993 (7.8%)	739 (6.5%)	254 (19.4%)	<0.001
Central cyanosis, n (%)	104 (0.8%)	26 (0.2%)	78 (6.0%)	<0.001
Temperature $\geq 38^{\circ}$ C, n (%)	3,540 (28.4%)	3,009 (26.8%)	531 (41.9%)	<0.001
Chest indrawing, n (%)	3,944 (31.0%)	3,126 (27.5%)	818 (61.6%)	<0.001
Wheezing, n (%)	462 (3.6%)	248 (2.2%)	214 (16.1%)	<0.001

Non-severe IMCI pneumonia <sup>2</sup> , n (%)	11,514 (90.8%)	10,449 (91.9%)	1,065 (81.4%)	<0.001
Severe IMCI pneumonia <sup>3</sup> , n (%)	1,198 (9.4%)	935 (8.2%)	263 (19.8%)	<0.001

SpO<sub>2</sub> indicates peripheral arterial oxyhemoglobin saturation; IQR, interquartile range; WHO, World Health Organization; IMCI, Integrated Management of Childhood Illnesses.

<sup>1</sup>Respiratory rate  $\geq$  70 breaths per minute for 3-11 month olds,  $\geq$  60 breaths per minute for 12-59 month olds.

<sup>2</sup>Cough and/or difficult breathing plus fast breathing for age (respiratory rate  $\geq$  50 breaths per minute for 3-11 month olds,  $\geq$  40 breaths per minute for 12-59 month olds) or chest wall indrawing without any WHO danger signs (stridor at rest, unable to feed, lethargy or unconscious, convulsions).

<sup>3</sup>Cough and/or difficult breathing plus at least one WHO danger sign (stridor at rest, unable to feed, lethargy or unconscious, convulsions), WAZ <-3, or HIV-infection with chest indrawing.



Table 2. Association of IMCI guideline and WHO-composite models with SpO<sub>2</sub> ranges (full dataset, N= 12,712)

Characteristics	SpO <sub>2</sub> <90% <sup>a</sup>			SpO <sub>2</sub> 90-92% <sup>a</sup>			SpO <sub>2</sub> <93% <sup>a</sup>		
	n = 11,986			n = 12,110			n = 12,712		
	aOR	95% CI	p-value	aOR	95% CI	p-value	aOR	95% CI	p-value
IMCI guideline model									
Intercept	0.04	(0.02, 0.10)	<0.001	0.06	(0.03, 0.12)	<0.001	0.11	(0.05, 0.23)	<0.001
WHO danger signs <sup>1</sup>	9.01	(7.08 – 11.47)	<0.001	2.29	(1.60 - 3.28)	<0.001	5.70	(4.57 - 7.12)	<0.001
Weight for age z-score ≤ -3	1.24	(0.86 - 1.80)	0.206	1.51	(1.14 - 2.01)	0.005	1.43	(1.13 - 1.81)	0.003
WHO-composite model									

Intercept	0.02	(0.01, 0.05)	<0.001	0.05	(0.03, 0.08)	<0.001	0.07	(0.04, 0.13)	<0.001
WHO danger signs <sup>1</sup>	4.82	(3.72 - 6.24)	<0.001	1.41	(0.97 - 2.04)	0.080	3.22	(2.54 - 4.09)	<0.001
Weight for age z-score $\leq$ -3	1.20	(0.82 - 1.74)	0.354	1.47	(1.10 - 1.97)	0.008	1.40	(1.10 - 1.78)	0.007
Severe respiratory distress <sup>2</sup>	5.74	(4.77 - 6.91)	<0.001	3.52	(2.99 - 4.14)	<0.001	4.59	(4.04 - 5.20)	<0.001

SpO<sub>2</sub> indicates peripheral arterial oxyhemoglobin saturation; IMCI, Integrated Management of Childhood Illnesses; WHO, World Health Organization.

<sup>1</sup>stridor at rest, unable to feed or drink, convulsions, lethargy or unconscious

<sup>2</sup>grunting, nasal flaring, head nodding, or severe fast breathing ( $\geq 60$  breaths/min among 3-11 month olds, and  $\geq 70$  breaths/min among 12-35 month olds).

Table 3. Clinical hypoxemia score for identifying a SpO<sub>2</sub> <93% (derivation dataset)

Characteristic	IMCI Guideline Model		WHO-composite Model		Independent LASSO Model			Composite LASSO Model		
	log(odds)	95% CI	log(odds)	95% CI	log(odds)	95% CI	Score	log(odds)	95% CI	Score
WHO danger signs <sup>1</sup>	1.724**	(1.458, 1.990)	1.140**	(0.853, 1.426)				0.521**	(0.173, 0.870)	1
Weight for age z-score ≤ -3	0.321*	(0.039, 0.603)	0.298*	(0.007, 0.588)	0.312*	(0.005, 0.620)	1	0.302+	(0, 0.605)	1
Severe respiratory distress <sup>2</sup>			1.564**	(1.413, 1.715)				0.990**	(0.813, 1.168)	2
Grunt	-	-	-	-	0.740**	(0.407, 1.073)	1			-



Nasal flaring	-	-	-	-	0.531**	(0.314, 0.749)	1			-
Head nodding	-	-	-	-	0.788**	(0.523, 1.053)	2			-
Severe fast breathing <sup>3</sup>	-	-	-	-	0.773**	(0.541, 1.004)	2			-
Age categories in months										
3-5	-	-	-	-	Ref		-	Ref		-
6-11	-	-	-	-	-0.057	(0.269, 0.154)	0	-0.069	(-0.278, 0.140)	0
12-23	-	-	-	-	-0.444**	(-0.663, -0.224)	-1	-0.407**	(-0.621, -0.193)	-1
24-35	-	-	-	-	-0.562**	(-0.844, -0.280)	-1	-0.509**	(-0.785, -0.234)	-1

Unable to feed	-	-	-	-	0.444	(-0.113, 1.001)	1	-	-	-
Lethargy or unconscious	-	-	-	-	-0.332	(1.275, 0.611)	-1	-	-	-
Cyanosis	-	-	-	-	1.701**	(1.052, 2.350)	3	1.763**	(1.171, 2.355)	4
Temperature $\geq 38^\circ$ C	-	-	-	-	0.217*	(0.051, 0.383)	0	0.212*	(0.048, 0.375)	0
Chest indrawing	-	-	-	-	1.060**	(0.880, 1.240)	2	1.090**	(0.915, 1.264)	2
Wheezing	-	-	-	-	1.156**	(0.860, 1.453)	2	1.103**	(0.822, 1.3840)	2
B <sub>0</sub> intercept	-2.246**	(-3.015, -1.476)	-2.692**	(-3.328, -2.057)	-2.855**	(-3.597, -2.113)	-	-2.905**	(-3.635, -2.175)	-
N	8,899		8,899		8,675			8,724		

C-statistic	0.560	(0.546, 0.573)	0.702	(0.684, 0.719)	0.774	0.756, 0.792		0.773	0.755, 0.791	
BIC	5,536.5		5,155.9		4,650.6			4,712.2		

SpO<sub>2</sub> indicates peripheral arterial oxyhemoglobin saturation; IMCI, Integrated Management of Childhood Illnesses; WHO, World Health Organization; LASSO, Least Absolute Shrinkage and Selection Operator reduction method for selection; all models represented here were fit using mixed effects logistic regression with a random effect for country (after selection and refitting for independent and composite LASSO models); BIC, Bayesian Information Criterion.

+ p<0.1

\* p<0.05

\*\* p<0.001

<sup>1</sup>stridor at rest, unable to feed or drink, convulsions, or lethargy or unconsciousness

<sup>2</sup>grunting, nasal flaring, head nodding, or severe fast breathing

<sup>3</sup>respiratory rate of  $\geq 70$  breaths/min if 3-11 months old;  $\geq 60$  breaths/minute if 12-35 months old

Note: all models include a random effect to account for the clustering effect of country



Table 4. Model performance for identifying a SpO<sub>2</sub><93% (validation dataset)

	Hypoxemia risk (n/N) <sup>1</sup>	Crude OR (95% CI)	Mean predicted hypoxemia, %	C-statistic (adjusted for Optimism; <sup>2</sup> 95% CI)	Sensitivity	Specificity	PPV	NPV
IMCI guideline model								
Non- Case	316/3,460	Ref	9.1%	0.563 (0.564; 0.528, 0.600)	20.6%	92.1%	23.2%	90.9%
Case	82/353	3.011 (2.292, 3.955)	23.2%	-				
Total	398 / 3813	-		-				
WHO-composite model								
Non- Case	172/2,857	Ref	6.0%	0.677 (0.677; 0.635, 0.712)	56.8%	78.6%	23.6%	94.0%
Case	226/956	4.833 (3.900, 5.989)	23.6%	-				
Total	398 / 3813	-		-				

Independent LASSO model								
-2 to -1	46/1,124	Ref	4.1%	0.745 (0.742; 0.696, 0.779)	46.7%	90.7%	36.9%	93.6%
0	57/1,179	1.191 (.800, 1.771)	4.8%	-				
1-2	109/1,006	2.848 (1.995, 4.065)	10.8%	-				
3-13	186/504	13.707 (9.697, 19.376)	36.9%	-				
Total	398 / 3813			-				
Composite LASSO model								
-1	44/1,117	Ref	3.9%	0.752 (0.750; 0.701, 0.787)	49.0%	89.5%	35.3%	93.8%
0	53/1,164	1.163 (.773, 1.750)	4.6%	-				

1	42/458	2.462 (1.589, 3.814)	9.2%	-				
2	64/522	3.408 (2.286, 5.079)	12.3%	-				
3-11	195/552	13.320 (9.402, 18.871)	35.3%	-				
Total	398 / 3813							

SpO<sub>2</sub> indicates peripheral arterial oxyhemoglobin saturation; OR, odds ratio; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; IMCI, Integrated Management of Childhood Illnesses; WHO, World Health Organization; LASSO, Least Absolute Shrinkage and Selection Operator reduction method for selection; all models represented here were fit using mixed effects logistic regression with a random effect for country (after selection and refitting for the independent and composite LASSO models).

Hypoxemia risk: numerator (number of hypoxemia cases) and denominator (total number of children with the relevant clinical score)

<sup>1</sup>n is the number of observed hypoxemic participants; N is the total number of participants.

<sup>2</sup>C-statistics are adjusted for optimism.

Note: The independent and composite LASSO models AUC compared using Delong method with Bonferroni correction (p=0.480).