PREDICTION OF DISEASE SEVERITY IN PATIENTS WITH FEBRILE ILLNESSES IN RESOURCE-LIMITED SETTINGS

Arjun Chandna

D.Phil. in Biomedical and Clinical Sciences

June 2023

Cambodia Oxford Medical Research Unit

Angkor Hospital for Children

Siem Reap

Cambodia

For R.H.S. Carpenter

Table of Contents

Table of Figures

Table of Tables

Abstract

Febrile illnesses present unique challenges for health systems with scarce resources. Large volumes of mostly self-limiting diseases accompanied by high case-fatality rates for the small proportion of serious infections mean that risk stratification tools must have high sensitivities and/or specificities according to their proposed contexts of use. Unfortunately, accuracy and reliability of existing tools are sub-optimal and they are often impractical for deployment in resource-limited settings.

This thesis explores the development and application of prediction tools for the management of febrile illnesses across a range of resource-constrained settings in South and Southeast Asia. It aims to combine the best prediction model science with a pragmatic field-based reality to address locally-important health issues. Recognising that evaluation of prediction tools should be set in their intended contexts of use, each analysis is framed in a particular clinical use-case, yet draws upon approaches to allow exploration of the generalisability of the findings.

Using a variety of research methodologies, this thesis identifies prognostic factors amongst febrile children presenting to different levels of the health system, and uses these data to externally validate existing severity scores and develop new clinical prediction models suitable for resource-limited settings. Prospective work evaluates the relative contributions of clinical and biomarker-based approaches for the referral of children with respiratory infections from a resource-limited community setting on the Thailand-Myanmar border, whilst retrospective work develops a prognostic model for critically ill children on admission to a paediatric intensive care unit in northern Cambodia. Finally, in response to the arrival of the SARS-CoV-2 pandemic, this thesis reports the development and external validation of prognostic models at two sites in India to support the safe outpatient management of patients presenting with moderate Covid-19.

Important challenges and potential solutions to developing prediction tools in resource-limited settings are discussed, including application of the classical prediction paradigm to the assessment of disease severity, comparative analyses of the clinical utility of different models, and the differential importance of various predictors for identifying both patients who are sick at the time of clinical assessment and those whose illnesses will progress later in their disease course.

1 Preface

1.1 Acknowledgements

There are a number of individuals and organisations without whom this thesis would not have been possible. First and foremost, I would like to thank Angkor Hospital for Children in Siem Reap, Cambodia, for hosting me for the last four years. The organisation's commitment to using clinical research to improve the health of Cambodian children is nothing short of inspiring. In particular, I thank Dr. Pheaktra and the PICU team for their collaboration on the PICU study. I look forward to continuing to work together on the next phases of the project.

I am indebted to the Cambodia Oxford Medical Research Unit team: Sona, Sena, Poda, Pisey, Voleak, Sothea, Roya, Vichet, Thyl, and Keang. You welcomed me into your family and afforded me the space to learn and grow as a clinical researcher. Keang, working alongside you to plan and implement challenging research studies has been an immensely rewarding experience and I am extremely grateful for our friendship. I could not have wished for a better group of research nurses: Sambou, Asean, and Sina, no task was beyond you. I will always admire the flexibility, dedication, and enthusiasm you demonstrated to ensure the disruption to our studies was kept to a minimum during the Covid-19 pandemic. Thank you.

Each of my supervisors contributed greatly to this thesis and my doctoral research training in general. Paul, thank you for your unshakeable and infectious passion for R and for leading by example to show me how research and capacity building go hand-in-hand: global is nothing without local. Yoel, this thesis would not be here without your unwavering support, creativity to find the funds to realise the adapted plans in light of the Covid-19 pandemic, and saintlike patience on Teams calls with He Who Shall Not Be Named. As an adult physician, I could not have hoped for a better mentor to help me deal with my 'paediatric imposter syndrome': Claudia, thank you for giving me the confidence to follow my ideas and for showing me the value of integrating high-quality science with real-world study

designs. Rafael, thank you for agreeing to supervise a D.Phil. student on the other side of the world whom you had never met! You have an extraordinary ability to explain statistical concepts in a way that makes intuitive sense to a clinician. Thank you for hosting me in your department last summer. Finally, Nick, thank you for always having my back: from encouraging me to come out to MORU, to having the faith to send me for my Fellowship interview. You gave me the belief to do things that I thought were beyond me – and most importantly ended my search for Sugar Man, who quickly became the soundtrack to my thesis.

Without question, this journey would have ended very differently without the mentorship of Liz Ashley and Sakib Burza. Liz, you helped me through the most challenging times of my career. Sakib, you were a constant reminder of why it was important to keep going. You are both inspirational people and I can never thank you enough.

I am extremely fortunate to have encountered many brilliant colleagues during my D.Phil. In particular, I am privileged to have had the opportunity to work and interact with Rainer Tan, Melissa Richard-Greenblatt, James Watson, and Nick White. Many of you have become close friends. I am grateful for our conversations – you all made yourselves available whenever I needed to sense-check something or bounce an idea around. My D.Phil. experience was richer because of you.

I would also like to thank the Wellcome Trust for generously funding my Doctoral Training Fellowship and for agreeing to a costed extension to account for the delays caused by the Covid-19 pandemic. I am grateful to Médecins Sans Frontières for funding the PRIORITISE study and to the Tropical Health Education Trust for providing seed funding to kickstart the systematic review. I thank the Medical Sciences Doctoral Training Centre at the University of Oxford for hosting my Fellowship. In particular, I thank Louise Samson for all her support and the interview panel for funding a Fellow based at one of the University's overseas units.

Finally, Kaajal, thank you for putting the vitamins back in the loft and agreeing to join me on this adventure. Obviously, I could not have done it without you.

1.2 Impact of Covid-19 pandemic on research plans

The plans for this doctorate were formulated in late 2018 and early 2019. At the outset, my primary research question was to determine whether measurements of clinical parameters and host biomarkers could be combined to develop simple prognostic clinical prediction models to help risk stratify children presenting with acute febrile illnesses in a resource-limited setting in Cambodia. My Fellowship application was centred around the conduct of a prospective cohort study (with a nested case-control sub-study), into which I planned to recruit 1,300 children. Accordingly, the first six months of my D.Phil. were focussed on preparatory activities (protocol development, ethics submissions, drafting study documents, procuring equipment, and hiring and training three research assistants) before recruitment of the first participant on 5th March 2020.

Unfortunately, the arrival of the Covid-19 pandemic meant that enrolment into the study proceeded slower than anticipated and recruitment eventually had to be temporarily suspended as all non-essential patient-facing activities were paused. As a result, it was not possible to complete the study within a timeframe compatible with including the results in this thesis, and this will now be the focus of my post-doctoral work.

Although it was a major disappointment not to be able to pursue my planned D.Phil., in many ways the Covid-19 pandemic enriched my research training experience. Whilst my original plan had focussed on risk stratification of paediatric febrile illnesses, I broadened the scope of my thesis and was able to apply concepts of prediction modelling more widely. I capitalised on skills and knowledge that I had acquired during the first year of my D.Phil. to rapidly design and initiate the PRIORITISE study, the results of which are reported in Chapter 6 of this thesis. Furthermore, the pandemic forced me to become more flexible in how I plan and implement research studies and enabled me to gain skills in other research methodologies and undertake projects that had not been part of my original plans.

1.3 Funding

Principal funding for this work was provided by the Wellcome Trust via a Doctoral Training Fellowship. In addition, the Wellcome Trust provide core funding to the Cambodia Oxford Medical Research Unit (part of the Thailand Africa Asia Programme). The biomarker assays conducted as part of the work presented in Chapter 4 were funded via a Wellcome Trust project grant (219644/Z/19/Z). Médecins Sans Frontières (MSF) funded the work presented in Chapter 6. The Tropical Health Education Trust provided supplementary funding to support the work presented in Chapter 3.

1.4 Declarations and attributions

The contents of this thesis were prepared by me. All words, figures, and tables are my own, unless otherwise indicated. Clinical research is a multi-disciplinary endeavour and a number of people have made significant contributions to the work presented in this thesis.

1.4.1 Chapter 2

This was conceptualised and written by me. The artwork for Figure 2.3-1 was prepared by John Ross Papa and the original clinical case vignettes drafted by Sakib Burza, Quique Bassat, and Matthew Wiens.

1.4.2 Chapter 3

This study was conceptualised by me. Together with Rainer Tan, I co-led protocol development, article screening, data extraction, data analysis, and manuscript drafting. Michael Carter, Ann Van den Bruel, and Jan Verbakel assisted with quality assessment of the included articles.

I assembled the Technical Advisory Panel of domain experts to provide feedback on the search strategy. The thesis chapter was written by me.

1.4.3 Chapter 4

The original birth cohort study was conducted by a team of researchers led by Paul Turner and Claudia Turner at the Shoklo Malaria Research Unit. I conceptualised this secondary analysis, drafted the protocol, submitted the ethics applications, curated the data, analysed the data, and wrote the manuscripts and thesis chapter. Wanitda Watthanaworawit assisted with specimen management. The biomarker assays were conducted by Phattaranit Tanunchai and Asama Vinitsorn under the supervision of Mohammad Yazid Abdad and Melissa Richard-Greenblatt. Lazaro Mwandigha supported aspects of the statistical analysis and programming to estimate confidence intervals for calibration intercepts and slopes, calculate classification indices, and prepare calibration plots, overlay plots, and locally weighted scatterplot smoothing plots.

1.4.4 Chapter 5

I conceptualised this study, drafted the protocol, submitted the ethics applications, developed the study documents, trained and supervised the research assistants, cleaned and curated the data, analysed the data, and wrote the manuscript and thesis chapter. The study database was developed by Thatsanun Ngernseng. Vichet Pav tested the database and supervised data entry. Arthur Riedel calculated the travel times. Lazaro Mwandigha and Constantinos Koshiaris provided statistical advice and helped troubleshoot programming issues.

1.4.5 Chapter 6

This study was conceptualised by me. I identified the study sites, drafted the protocol, submitted the ethics applications, developed the study documents, provided remote training and supervision for the research teams, led the study management group, cleaned and curated the data, and wrote the manuscript and thesis chapter. Sakib Burza provided mentorship throughout the study. Arthur Cheung conducted the biomarker literature review. Sabine Dittrich and Jennifer Osborn assisted with biomarker shortlisting. Jaruwan Tubprasert oversaw study monitoring. Chonticha Menggred developed the electronic case record forms and managed the study database. Vikash Kumar, Shiril Kumar, and Debasree Kundu performed the laboratory assays, under the supervision of Melissa Richard-Greenblatt. Lazaro Mwandigha and Constantinos Koshiaris performed the statistical analyses to derive and validate the clinical prediction models. I conducted the descriptive analyses and the analysis of the analytical performance of the soluble urokinase plasminogen activator receptor rapid diagnostic test.

1.4.6 Chapter 7

This was conceptualised and written by me.

1.5 Papers arising

The following papers have arisen during the course of this DPhil:

1. **Chandna A**, Keang S, Vorlark M, Sambou B, Chhingsrean C, Sina H, Vichet P, Patel K, Habsreng E, Riedel A, Mwandigha L, Koshiaris C, Perera-Salazar R, Turner P, Chanpheakrta N, and Turner C. *Derivation of a prognostic model for critically ill children in locations with emerging critical care capacity*. Pediatr Crit Care Med (in press).

2. **Chandna A**, Mahajan R, Gautam P, Mwandigha L, Dittrich S, Kumar V, Osborn J, Kumar P, Koshiaris C, Varghese GM, Lubell Y, and Burza S. *Point-of-care prognostication in moderate Covid-19: analytical validation and prognostic accuracy of a soluble urokinase plasminogen activator receptor (suPAR) rapid test.* PLOS Glob Public Health. Aug 21; 3(8): e0001538, 2023

3. **Chandna A**, Lubell Y, Mwandigha L, Tanunchai P, Vinitsorn A, Richard-Greenblatt M, Koshiaris C, Limmathurotsakul D, Nosten F, Abdad MY, Perera-Salazar R, Turner C, and Turner P. *Defining the role of host biomarkers in the diagnosis and prognosis of the severity of childhood pneumonia: a prospective cohort study*. Sci Rep. Jul 25; 13(1): 12024, 2023

4. Adella FJ, Vanna M, Adhikari B, Ol S, Tripura R, Davoeung C, Callery JJ, Sovann Y, **Chandna A**, Bunreth V, Asnong C, von Seidlein L, Dondorp AM, Maude RJ, Lubell Y, Wills B, Lek D, and Peto TJ. *The feasibility of novel point-of-care diagnostics for febrile illnesses at health centres in Southeast Asia: a mixed-methods study*. Trans R Soc Trop Med Hyg. Jun 15: trad036, 2023

5. Akech S, Kwambai T, Wiens MO, **Chandna A**, Berkley JA, and Snow RW. *Tackling post-discharge mortality in children living in LMICs to reduce child deaths*. Lancet Child Adolesc Health. Mar; 7(3): 149-151, 2023

6. **Chandna A**, Mahajan R, Gautam P, Mwandigha L; PRIORITISE Study Investigators; Kumar P, Varghese GM, Koshiaris C, Lubell Y, and Burza S. *Host biomarkers reflect prognosis in patients presenting with moderate Coronavirus disease 2019: a prospective cohort study*. Open Forum Infect Dis. Oct 6; 9(10): ofac526, 2022

7. **Chandna A**, Mahajan R, Gautam P, Mwandigha L, Gunasekaran K, Bhusan D, Cheung ATL, Day N, Dittrich S, Dondorp AM, Geevar T, Ghattamaneni SR, Hussain S, Jimenez C, Karthikeyan R, Kumar S, Kumar S, Kumar V, Kundu D, Lakshmanan A, Manesh A, Menggred C, Moorthy M, Osborn J, Richard-Greenblatt M, Sharma S, Singh VK, Singh VK, Suri J, Suzuki S, Tubprasert J, Turner P, Villanueva AMG, Waithira N, Kumar P, Varghese GM, Koshiaris C, Lubell Y, and Burza S. *Facilitating safe discharge through predicting disease progression in moderate COVID-19: a prospective cohort study to develop and validate a clinical prediction model in resource-limited settings*; Clin Infect Dis. Aug 24; 75(1): e368-e379, 2022

8. Lyvannak S, Sreynich K, Heng S, Thyl M, **Chandna A**, Chanpheaktra N, Pises N, Farrilend P, Jarzembowski J, Leventaki V, Davick J, Neunert C, Keller F, Kean LS, Camitta B, Tarlock K, and Watkins B. *The first case report of visceral leishmaniasis in Cambodia*. Am J Trop Med Hyg. Jun 13; 107(2): 336-8, 2022

9. Roberts T and **Chandna A**, Watthanaworawit W, Thaiprakong A, Soeng S, Simmalavong M, Phoumin P, Saengchun W, Khatta N, Hinfonthong P, Kaewpundoem N, Lee SJ, Perrone C, Amos B, Turner P, Ashley EA and Ling CL. *Impact of delayed processing of positive blood cultures on organism detection: a prospective multi-centre study*. BMC Infect Dis. Jun 4; 22(1): 517, 2022

10. **Chandna A** and Richard-Greenblatt M, Tustin R, Lee SJ, Kain KC, Burza S, Lubell Y, and Turner P. *Practical methods to permit the analysis of host biomarkers in resource-limited settings*. Am J Trop Med Hyg. Apr 18; 106(6): 1765-69, 2022

11. **Chandna A**, Chew R, Shwe Nwe Htun N, Peto TJ, Zhang M, Liverani M, Brummaier T, Phommasone K, Perrone C, Pyae Phyo A, Sattabongkot J, Roobsoong W, Nguitragool W, Sen A, Ibna Zaman S, Sandar Zaw A, Batty E, Waithira N, Abdad MY, Blacksell SD, Bodhidatta L, Callery JJ, Fagnark W, Huangsuranun W, Islam S, Lertcharoenchoke S, Lohavittayavikant S, Mukaka M, Moul V, Kumer Neogi A, Nedsuwan S, Pongvongsa T, Ponsap P, Richard-Greenblatt M, Schilling WHK, Thaipadungpanit J, Tripura R, Dondorp AM, Mayxay M, White NJ, Nosten F, Smithuis F, Ashley EA, Maude RJ, Day NPJ, and Lubell Y. *Defining the burden of febrile illness in rural South and Southeast Asia: an open letter to announce the launch of the Rural Febrile Illness project*. Wellcome Open Res. Mar 10, 6:64, 2022

12. Chew R, Zhang M, **Chandna A,** and Lubell Y. *The impact of pulse oximetry on diagnosis, management and outcomes of acute febrile illness in low-income and middle-income countries: a systematic review*. BMJ Glob Health. Nov; 6(11): e007282, 2021

13. Schilling WH, Callery JJ, **Chandna A**, Hamers RL, Watson JA, and White NJ. *The WHO guideline on drugs to prevent COVID-19: small numbers – big conclusions*. Wellcome Open Res. Sep 21; 6: 71, 2021

14. **Chandna A**, Osborn J, Bassat Q, Bell D, Burza S, D'Acremont V, Fernandez-Carballo BL, Kain KC, Mayxay M, Wiens M, and Dittrich S. *Anticipating the future: prognostic tools as a complementary strategy to improve care for patients with febrile illnesses in resource-limited settings*. BMJ Glob Health. Jul; 6(7): e006057, 2021

15. Mawji A, Li E, **Chandna A**, Kortz T, Akech S, Wiens MO, Kissoon N, and Ansermino M. *Common data elements for predictors of pediatric sepsis: A framework to standardize data collection*. PLOS One. Jun 10; 16(6): e0253051, 2021

16. **Chandna A** and Bonhoeffer M, Miliya T, Suy K, Sao S, and Turner P. *Improving Treatment and Outcomes for Melioidosis in Children, Northern Cambodia, 2009–2018*. Emerg Infect Dis. Apr; 27(4):1169-1172, 2021

17. **Chandna A**, Aderie EM, Ahmad R, Arguni E, Ashley EA, Cope T, Dat VQ, Day NPJ, Dondorp AM, Illanes V, De Jesus J, Jimenez C, Kain K, Suy K, Koshiaris C, Lasry E, Mayxay M, Mondal D, Perera R, Pongvongsa T, Rattanavong S, Rekart M, Richard-Greenblatt M, Shomik M, Souvannasing P, Tallo V, Turner C, Turner P, Waithira N, Watson JA, Yosia M, Burza S, and Lubell Y. *Prediction of disease severity in young children presenting with acute febrile illness in resource-limited settings: a protocol for a prospective observational study*. BMJ Open. Jan 25; 11(1): e045826, 2021

18. **Chandna A** and Tan R, Carter M, Van Den Bruel A, Verbakel J, Koshiaris C, Salim N, Lubell Y, and Turner P and Keitel K. *Predictors of disease severity in children presenting from the community with febrile illnesses: a systematic review of prognostic studies*. BMJ Glob Health. Jan; 6(1): e003451, 2021

1.6 Abbreviations

Abbreviations are defined at the time of first use in each chapter to assist those who may choose to read this thesis out of order. For ease of interpretation, abbreviations used in figures or tables are defined underneath, irrespective of whether they have previously been defined in the text. A full list of abbreviations in alphabetical order is provided here (Table 1.6-1).

Table 1.6-1: Abbreviations used in thesis.

Abbreviation	Definition
AHC	Angkor Hospital for Children
AIIMS	All India Institute of Medical Sciences
AIOS	Acute Infantile Observation Score

Abbreviation	Definition
Ang-1	Angiopoietin-1
Ang-2	Angiopoietin-2
APTT	Activated partial thromboplastin clotting time
AQUAMAT	African Quinine Artesunate Malaria Trial
ARI	Acute respiratory infection
ARDS	Acute respiratory distress syndrome
AUC	Area under the receiver operating characteristic curve
AVPU	Alert Voice Pain Unresponsive scale
BCS	Blantyre Coma Scale
BMI	Body mass index
BP	Blood pressure
BPM	Beats or breaths per minute
BUN	Blood urea nitrogen
CHARMS	Checklist for Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies
CHARMS-PF	CHARMS-Prognostic Factors
CHI3L1	Chitinase-3-like protein-1
cHIS	Covid-19 associated hyperinflammatory syndrome
CI	Confidence interval
CLSI	Clinical and Laboratory Standards Institute
CMC	Christian Medical College
CPM	Clinical prediction model
CRF	Case report form
CRP	C-reactive protein
CRS	Clinical Recognition Signs
CRT	Capillary refill time
Ct	Cycle threshold
CXCL-10	C-X-C motif chemokine-10
DIC	Disseminated intravascular coagulation
DOI	Digital object identifier
ED	Emergency department
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
EPP	Events per parameter

Abbreviation	Definition
ER	Emergency room
FEAST-PET	Fluid Expansion As Supportive Therapy-Paediatric Emergency Triage
FEAST-PETaL	Fluid Expansion As Supportive Therapy-Paediatric Emergency Triage and Laboratory
FiO_2	Fraction of inspired oxygen
FN	False negative
FP	False positive
GCS	Glasgow Coma Scale
GE	Gastroenteritis
HIS	Hospital Information System
HIV	Human immunodeficiency virus
HR	Heart rate
iCCM	integrated Community Case Management
ICU	Intensive care unit
IL-10	Interleukin-10
IL-1ra	Interleukin-1 receptor antagonist
IL-6	Interleukin-6
IL-8	Interleukin-8
IMCI	Integrated Management of Childhood Illnesses
IP-10	Interferon-y induced protein-10
IPSCC	International pediatric sepsis consensus conference
IQR	Interquartile range
ISARIC	International Severe Acute Respiratory and Emerging Infection Consortium
ISRCTN	International Standard Randomised Controlled Trial Number
ITAT	Inpatient Triage and Treatment
ITN	Insecticide-treated bednet
LDH	Lactate dehydrogenase
LLTG	Lower limb temperature gradient
LMIC	Low- and middle-income country
LODS	Lambaréné Organ Dysfunction Score
LOQ	Limit of quantification
LOWESS	Locally weighted scatterplot smoothing
LqSOFA	Liverpool quick Sequential Organ Failure Assessment
LRT	Likelihood ratio test
MHRA	Medicines and Healthcare products Regulatory Agency

Abbreviation	Definition
MICE	Multiple imputation with chained equations
mRISC	modified Respiratory Index of Severity in Children
MR-proADM	Mid-regional pro-adrenomedullin
MSF	Médecins Sans Frontières
MSF ERB	MSF Ethical Review Board
mSIRS	modified Systemic Inflammatory Response Syndrome
MUAC	Mid-upper arm circumference
MV	Mechanical ventilation
NECHR	National Ethics Committee for Health Research
NICU	Neonatal intensive care unit
NIV	Non-invasive ventilation
NLR	Negative likelihood ratio
NLR	Neutrophil-to-lymphocyte ratio
NNR	Number needed to refer
NPS	Nasopharyngeal swab
ODK	Open Data Kit
OPD	Outpatient department
OPS	Oropharyngeal swab
OR	Odds ratio
OxTREC	Oxford Tropical Research Ethics Committee
PAWS	Paediatric Advanced Warning Score
PCT	Procalcitonin
PEDIA-e	Paediatric Early Death Index for Africa-Early
PEDIA-i	Paediatric Early Death Index for Africa-Immediate
PEDIA-l	Paediatric Early Death Index for Africa-Late
PERISKOPE-TB	Personalised risk predictor for incident tuberculosis
PEWS	Pediatric Early Warning Score
PEWS-RL	PEWS-Resource Limited
PF	Prognostic factor
PHC	Primary health centre
PICOTS	Population Intervention Comparator Outcome Timing Setting
PICU	Paediatric intensive care unit
PLR	Positive likelihood ratio

Abbreviation	Definition
PRIORITISE	Prognostication of Oxygen Requirement in Patients with Non-severe SARS-CoV-2 Infection
PRISM III	Pediatric Risk of Mortality III
PRISMA	Preferred Reporting Items for Systematic Review and Meta-Analysis
PROBAST	Prediction Model Risk of Bias Assessment Tool
PT	Prothrombin time
qPELOD-2	quick Pediatric Logistic Organ Dysfunction-2
qSOFA	quick Sequential Organ Failure Assessment
QUIPS	Quality In Prognosis Studies
RA	Room air
RDT	Rapid diagnostic test
RECOVERY	Randomised Evaluation of Covid-19 Therapy
RISC	Respiratory Index of Severity in Children
RMRI	Rajendra Memorial Research Institute of Medical Sciences
RR	Respiratory rate
SAM	Severe acute malnutrition
SBI	Serious bacterial infection
SBP	Systolic blood pressure
SD	Standard deviation
SEAIDCRN	Southeast Asia Infectious Disease Clinical Research Network
sFlt-1	Soluble fms-like tyrosine kinase-1
SICK	Signs of Inflammation in Children that Kill
SIRS	Systemic Inflammatory Response Syndrome
SNOMED-CT	Systematized Nomenclature of Medicine – Clinical Terms
SpO_2	Peripheral oxygen saturation
SSA	sub-Saharan Africa
STARD	Standards for Reporting Diagnostic Accuracy Studies
sTNFR-1	Soluble tumour necrosis factor receptor-1
sTREM-1	Soluble triggering receptor expressed on myeloid cells-1
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
suPAR	Soluble urokinase plasminogen activator receptor
TMEC	Mahidol University Tropical Medicine Ethics Committee
TN	True negative
TP	True positive

Abbreviation	Definition
TRIPOD	Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis
USA	United States of America
US-CDC	United States Centers for Disease Control
VOC	Variant of concern
WAZ	Weight-for-age z-score
WHO	World Health Organization
WHO-CPS	WHO Clinical Progression Scale
YOS	Yale Observation Score

2 Introduction

This chapter draws upon work published in: **Chandna A**, Osborn J, Bassat Q, et al. *Anticipating the future: prognostic tools as a complementary strategy to improve care for patients with febrile illnesses in resource-limited settings*. BMJ Glob Health. Jul; 6(7): e006057, 2021.

2.1 Challenges and burden of febrile illnesses in resource-limited settings

Globally, fever is amongst the most common reasons to seek healthcare.[1,2] Although estimates vary depending on methodology, epidemiology, and season, febrile syndromes are thought to account for over a billion episodes of illness each year.[3] Costs in terms of lives and healthcare expenditure are substantial.[4,5] The burden is especially large in children: in high-income settings, one in three unscheduled paediatric primary care contacts is fever-related,[6] whilst in resource-limited settings this proportion increases to 80%,[1] with a child living in rural sub-Saharan Africa experiencing up to 50 febrile illnesses before the age of five.[7]

In the rural tropics, fever is principally caused by infection and management of febrile patients presents specific challenges for many low- and middle-income country (LMIC) health systems. In part, this is due to a wider range of infectious agents and greater burden of antimicrobial resistance,[8] which combined with sparse epidemiological data concentrated from specialist urban centres, complicate formulation of empirical treatment strategies that are relevant for the rural locations in which the majority of populations live. Inadequate access to therapeutics constrains treatment options further.[9] Seasonal surges in vector-borne and communicable diseases can overwhelm health systems.[10] In many LMICs maldistribution of healthcare resources exacerbates these problems and results in considerable rural-urban inequity in availability of quality care.[11,12]

Notably, these issues do not occur in isolation. Limited investment in health workers and systems, together with absence of effective diagnostic and treatment options, contribute to both under- and overuse of antimicrobials. This, in turn, drives emergence and spread of antimicrobial resistance, further compromising care for febrile patients. For these and many other reasons, sepsis and severe febrile illnesses continue to disproportionately affect individuals living in LMICs and remain leading causes of morbidity and mortality.[4,13,14]

Understanding the underlying causes of febrile illnesses is critical to the development of successful treatment and control strategies. Accordingly, several recent multi-country initiatives have addressed this topic.[15-21] Nevertheless, approaches that focus solely on diagnosis can struggle to reconcile colonisation from infection,[22] the potential role of co-infections,[14] and the fact that patients with the same infection or syndrome can have markedly different illness trajectories, potentially due to variations in inoculum but perhaps also reflecting differing host nutritional and other susceptibility states.[23,24] Furthermore, at a pragmatic level, syndromic diagnoses are challenging for the many patients that present with undifferentiated illnesses, while even well-resourced fever studies fail to identify an aetiology in a substantial proportion of participants.[25]

In rural regions of LMICs, these problems are magnified. Healthcare providers often have inadequate access to the necessary training, supervision, and diagnostic testing to support their clinical decision-making. In such contexts where access to health facilities and/or specialist care can be particularly challenging and costly, in addition to determining the *cause* of a patient's illness, an equally pertinent question is: what is the probability that my patient's condition will progress and require a higher level of care?

Used appropriately, prognosis can complement diagnosis to improve precision and efficiency of management algorithms for febrile illnesses. This could be particularly impactful in resource-constrained settings where diagnosis remains most challenging, triaging practices predominantly rely on clinical evaluation, and decisions to refer or escalate care must be made early due to complex

context-specific reasons including difficulties in access (due to geography, climate, or inadequate infrastructure), sociocultural beliefs, and political or financial insecurity.

2.2 The prediction paradigm and the assessment of disease severity

Healthcare providers regularly integrate multiple sources of data, including patient demographics, comorbidities, clinical signs and symptoms, and results of radiological and laboratory investigations to determine the 'true' underlying disease or health state of their patient. Depending on the temporal relationship between these baseline data (predictors) and the disease or health state in question (the outcome), these predictions are either diagnostic or prognostic.[26]

In contrast to a *diagnostic* test which determines whether a specific disease or health state is present at the moment the test is performed, a *prognostic* test provides information on the likelihood of a particular outcome occurring in the future.[26] Often the distinction between diagnosis and prognosis is clear: integration of clinical, laboratory, and radiological information to predict whether a patient may have infective endocarditis (Duke criteria)[27] is a *diagnostic* process, whereas predicting the probability that an individual may develop active tuberculosis over the next two years based on their demographics, medical history, and tuberculin skin test result (personalised risk predictor for incident tuberculosis; PERISKOPE-TB)[28] is easily recognisable as *prognosis*.

For predictors of severity these concepts can become blurred. A patient's severity reflects their likelihood of a poor outcome and hence predictors of severity are inherently prognostic. However, certain predictors are also used to indicate a patient's 'level of severity' at the time of measurement. For example, a peripheral oxygen saturation < 90% in a patient with pneumonia is indicative of severe pneumonia.[29] In this sense, as well as providing prognostic information, such predictors can also be considered diagnostic of a patient's severity *at that moment in time*. This is in contrast to other predictors that are primarily harbingers of *future deterioration* in patients who appear otherwise well

(for example, various clinical and laboratory parameters measured during the febrile phase of a dengue or Covid-19 infection).[24,30,31] Application of the classical prediction paradigm when considering different predictors of severity can provide a framework to evaluate their potential roles in patient assessment and better understand how they might fit within existing clinical workflows (Figure 2.2-1).

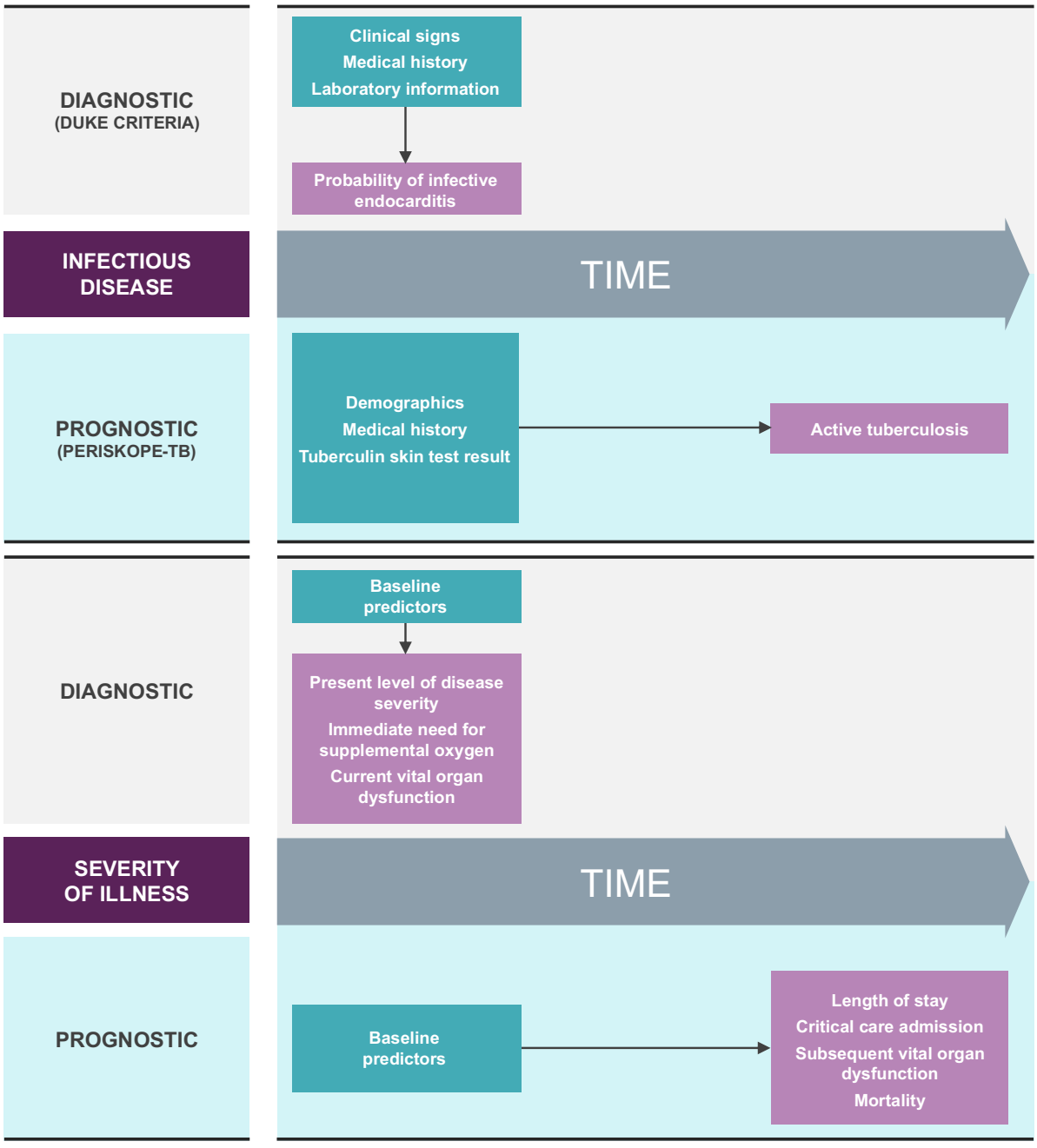


Figure 2.2-1: The classical prediction paradigm applied to infectious diseases and the assessment of disease severity. Green boxes contain examples of baseline data (predictors) and pink boxes contain examples of diseases or health states (outcomes). Thin arrows indicate temporal relationship between predictors and outcomes. Note that the PERISKOPE-TB tool is validated for use in low-endemicity settings and is included here as an illustrative example of a prognostic tool for an infectious disease unrelated to disease severity. PERISKOPE-TB = personalised risk predictor for incident tuberculosis.

Most predictors used for the assessment of severity serve both diagnostic and prognostic functions. For example, many guidelines and tools devised to inform the management of febrile illnesses in resource-limited settings use 'Danger Signs' to identify patients who are severely ill at the time of assessment *and* at high risk of mortality.[32,33] Hence, these 'Danger Signs' can be considered both diagnostic (of the severity of illness at the time of assessment) and prognostic (for future risk of death). However, their lack of sensitivity and specificity, and high interobserver variability, make their performance poor for these particular purposes.[34,35]

In addition to developing prognostic tools that are reliable and practical in resource-limited settings, improving recognition of impending serious illness in patients that lack clinical signs of severity as determined by existing management algorithms is a global public health priority.[36] A growing body of evidence suggests that host response biomarkers of final common pathways to severe febrile illness, that exist across a range of microbial aetiologies and reflect endothelial injury, immune activation, and coagulation, appear to add value to clinical risk scores.[37-41] It is possible that rather than being crude laboratory surrogates for bedside assessment, certain biomarkers might reflect subclinical pathophysiological changes and facilitate identification of patients with a poor prognosis whom may have otherwise been incorrectly categorised as low risk. Reliable and practicable tests for these markers could help risk stratify febrile patients and inform management decisions at critical junctures in patient care pathways.

While a standalone test for a host response biomarker could provide useful prognostic information, these tests would be more effective as part of an algorithm, combining measurement of a biomarker(s) with other clinical parameters (clinical signs, anthropometrics, demographic information, etc.) to more accurately assess risk and guide rational management. However, laboratory tests always carry an opportunity cost, especially in settings where resources are particularly scarce. It is therefore essential that the potential benefits of candidate biomarker tests are assessed in the context of clinical assessments commensurate with the level of the health system at which they are

being proposed for use. In this regard, specific use-cases for prognostic tools in the management of febrile illnesses at different levels of the health system can help frame evaluations of prediction tools: contextualising interpretation of results; encouraging consistency of reporting; and maximising comparability of findings from disparate studies.

2.3 Use-cases for prognostic tools in the management of febrile illnesses

Unlike diagnosis, prognosis is inherently context-dependent: a patient's eventual outcome is inextricably influenced by the available resources and quality of care.[42] Hence, an informed evaluation of prognostic testing in febrile illnesses requires specific use-cases to be defined. Each use-case should detail the clinical problem and consider the resources available to treat febrile illnesses in that setting (for example, health worker and laboratory capabilities, referral capacity, and availability of essential resources such as oxygen, fluids, antimicrobials, and provision of vital organ support), in order to contextualise the outcomes against which a candidate prognostic test or algorithm is to be assessed (Table 2.3-1; Figure 2.3-1). This is especially important given the heterogeneous circumstances in which care is often provided for patients with febrile illnesses in many LMIC health systems: widely-used labels applicable to more advanced health systems (for example, primary care, secondary care, etc.) may not adequately describe the reality of care provision in less well-regulated and fragile healthcare contexts.

Table 2.3-1: Use-cases for prognostic tools in the management of febrile illnesses in resource-limited health systems. The three use-cases are described in sections 2.3-1, 2.3-2, and 2.3-3, and examples provided in Figure 2.3-1. *Human and technical capacity varies greatly within and across countries: the examples given are for illustrative purposes and do not reflect all settings. MUAC = mid-upper arm circumference; PHC = primary health centre; RDT = rapid diagnostic test.

Use-case	Healthcare context	Typical human and technical capacity for the management of febrile illnesses*	Relevant outcomes to assess candidate prognostic factors
Identifying which patient to refer, admit to hospital, or provide maximal pre-referral care if referral is not feasible	Community health worker or village health volunteer in rural village	Health workers are often lay people with a few days to months training, supervised by staff from the PHC or other actors implementing community-based healthcare programmes. A very limited range of equipment (for example, MUAC tapes, thermometers, respiratory rate counters), diagnostics (qualitative RDTs for malaria), and treatments (antipyretics, oral antibiotics or antimalarials, oral rehydration solution and nutritional supplements) may be available.	Persistence or worsening of symptoms; referral to health facility; admission to hospital
	Healthcare provider at primary health centre	Primary healthcare providers typically include clinical officers, nurses, or midwives with a few months to years training. A greater range of clinical equipment (for example, pulse oximeters, weighing scales, stethoscopes) and diagnostic tests (for example, RDTs for other diseases and basic haematology) may be available. Some facilities may have capacity for overnight observation and the delivery of intravenous fluids, antibiotics, or nebulisers.	Persistence or worsening of symptoms; referral to hospital; admission to hospital
	Healthcare provider in district hospital outpatient department	Healthcare staff can range from clinical officers with a few years training to experienced physicians. Similar clinical equipment available as at a PHC. Laboratory tests can also include instrumented platforms (which may be batched, depending on patient throughput). Proximity to inpatient care areas means threshold for admission may be lower.	Admission to hospital; length of hospital stay; admission to high-dependency area; measures of vital organ dysfunction
Resource prioritisation for hospitalised patients	Healthcare provider in district hospital inpatient department	Healthcare staff can range from clinical officers with a few years training to experienced physicians. A range of clinical equipment is available, as well as variable access to radiological (for example, point-of-care ultrasound), microbiological, and laboratory testing. Frequent vital sign observations and delivery of supplemental oxygen therapy, intravenous medications, and surgical interventions may be possible. Admission of patients also permits evaluation of trends in clinical or laboratory parameters over time and response to therapeutic interventions to be observed.	Length of hospital stay; admission to high-dependency area; measures of vital organ dysfunction; mortality

Use-case	Healthcare context	Typical human and technical capacity for the management of febrile illnesses*	Relevant outcomes to assess candidate prognostic factors
	Physician at admission to high dependency area or critical care unit	Experienced physicians with access to clinical equipment and radiological, microbiological, and laboratory testing. Near-patient tests such as blood gas machines and point-of-care ultrasound may be available in some settings, as may continuous vital sign monitoring and vital organ support (for example, inotropic therapy and non-invasive or mechanical ventilation).	Length of stay in high-dependency area; length of hospital stay; measures of vital organ dysfunction; mortality
Allocation of peri- and post-discharge care interventions	Healthcare provider at hospital peri-discharge	Range of healthcare staff, clinical equipment, and radiological, microbiological, and laboratory testing depending on the level of the health facility. Feasible to compare discharge measurements to those taken during the hospital stay. Some facilities may have community outreach teams or links with nearby community health facilities to assist with patient follow-up after discharge.	Readmission to a health facility; return to baseline health status; acute sequelae resolution; neurocognitive outcomes; mortality

Community referral	In-hospital prioritisation	Post-discharge follow-up
A chronic conflict involving a separatist border province has flared up, with multiple airstrikes damaging the already crumbling health infrastructure. The nearest functional hospital is across the border, a journey of 20km involving crossing the conflict front-line and evading an unfriendly border army who are actively trying to prevent a refugee influx. A male patient presents to one of the functional health posts with a shrapnel wound to his lower leg and a low-grade fever. He is otherwise feeling well. He is seen by a community health worker who is unsure if oral antibiotics would be sufficient yet is concerned about the risk of the patient making the journey across the border. A referral mechanism exists but is risky from a security perspective and will utilise the only available vehicle. She uses a simple prognostic algorithm (including temperature, pulse, and measurement of a hypothetical biochemical biomarker using a lateral flow assay similar to a malaria rapid diagnostic test), which suggests that the patient's risk is above the threshold for safe discharge back to the community. She therefore mobilises the vehicle and advises the patient, with informed consent, that the cross-border journey is likely to be necessary. Fortunately, the patient arrives safely. His condition deteriorates over the next 48 hours but timely intervention with intravenous antibiotics and fluid resuscitation result in a positive outcome and discharge after five days.	A 9 year old boy presents to a district hospital with high parasitaemia malaria during peak malaria season. His mother reports a history which is suggestive of a fit but this cannot be confirmed. The single clinical officer on duty assesses the child with the assistance of a prognostic algorithm (including the patient's age, comorbidities, illness history, presenting vital signs, and measurement of two hypothetical biochemical biomarkers). The child is deemed to be at high risk of an adverse outcome. For this reason, rather than being sent to the normal paediatric ward, the child is admitted to the four-bedded high-dependency unit. This proves the right decision when four hours after admission the child enters into status epilepticus and requires urgent resuscitation. Due to the higher nurse-to-patient ratio the child is attended promptly and the seizure terminated. Clinical evolution is slow but there is a response to treatment, and the child is weaned off parenteral artesunate and discharge planning can begin.	A 4 year old girl, referred from a community health centre, was admitted to a district hospital with signs and symptoms of pneumonia. On admission, she was moderately malnourished with a mid-upper arm circumference of 121mm, but has no other comorbidities. Her mother had travelled for two and a half hours with her daughter using motorcycle taxi and walking as the primary modes of transportation to reach the hospital. Her hospital course was uncomplicated and by the time of her discharge, four days later, her vital signs have normalised. A prognostic score (including the degree of malnutrition, severity of her respiratory illness at admission [oxygen saturation and presence of chest indrawing], the geographic location of her residence, and measurement of a hypothetical biochemical marker) indicates that she is at high risk of post-discharge mortality. Based on her risk score she is provided with a robust plan for follow-up, which includes three follow-up visits within the next fortnight at a health centre near her home. Extra care is taken to provide information and counselling to the mother on monitoring her daughter's recovery, how and when to seek further care, hygiene, and nutrition. Seven days after discharge, during a routine follow-up at the nearby health centre, the nurse notes worsening respiratory symptoms and refers the child back to the hospital, where the child is readmitted for recurrent pneumonia.

Figure 2.3-1: Clinical vignettes illustrating use-cases for hypothetical prognostic tools in the management of febrile illnesses in resource-limited settings.

2.3.1 Referral for higher-level medical care by community healthcare providers

Most patients with febrile illnesses present to peripheral levels of the health system.[1] Distinguishing those that require referral or admission can be difficult, and once identified the decision to transfer may not be straightforward. Particularly in rural areas and conflict settings, poorly functioning infrastructure, as well as geographic, climatic, social, and political challenges mean that referral decisions often involve complex mechanisms and incur costs and risks for both patient and provider.[43]

Even with optimal deployment of existing algorithms such as the World Health Organization's Integrated Management of Childhood Illnesses or integrated Community Case Management guidelines, cases of serious illness can be missed and patients are inappropriately referred.[34] In many settings these algorithms are not regularly used or are improperly applied due to various constraints common in peripheral healthcare locations in LMICs, such as health workers who have received limited training, high patient volumes, and poor diagnostic capacity.[35,44] A prognostic test that could give community healthcare providers increased confidence in their decision to refer (or not) would have great potential to improve appropriateness of referrals and reduce resource misallocation.[45] Increased confidence may also lead to better communication between providers and patients, which is important in contexts where strong traditional beliefs about causes and treatments of febrile illnesses exist.[46]

In settings where referral is not immediately feasible, accurate prognosis could guide provision of pre-referral care, such as the first dose of parenteral antibiotics for suspected serious bacterial infections.[47] During epidemics, identifying patients suitable for home-based management could prevent overburdening of health facilities.[10] Such a prognostic test or algorithm would need to function within the limited human and material resources available at the peripheral levels of most LMIC health systems, and the threshold for referral adjusted according to the risks and benefits present in particular contexts, reflecting the relative importance of ruling-in or ruling-out serious disease.

Rather than being mutually exclusive, prognostic and diagnostic tests should be considered complementary: an algorithm integrating prognostic and diagnostic components can be envisioned as being highly useful in this context. First, the algorithm could identify patients likely to benefit from referral for higher-level care, as described above, and second, it could guide the further management of individuals who were identified as suitable for care at the community level (for example, informing appropriate antimicrobial prescription).

2.3.2 Resource allocation for patients admitted to hospital

In many LMICs, febrile illnesses remain the leading cause of hospitalisation.[48] Particularly during seasonal outbreaks (for example, due to malaria, acute bacterial meningitis, respiratory syncytial virus, or dengue), health facilities are vulnerable to overcrowding and limited resources stretched further.[10]

Being able to predict the likely course of a patient's illness given the resources available at a typical district-level hospital could enable better resource prioritisation – from simple measures such as facilitating early discharge or increased frequency of vital observations, to admission to high dependency or intensive care areas with limited bed capacities, or informing timely transfer for more specialist services.

At provincial- and tertiary-level hospitals, accurate prognostication might help direct scarce resources towards patients more likely to benefit, for example early institution of high-cost therapies and adjunctive procedures. Effective resource prioritisation is essential to ensure sustainability in many settings where demand for and provision of critical care services are growing.[49] Furthermore, this may reduce the likelihood of prolonged admission and subsequent long-term morbidity, and the financial burden of this on patients, their families, and the health system. Data driven risk stratification can help reduce pressure on individual doctors and provide a framework for discussions with patients and relatives regarding appropriate individualised ceilings of care.[50]

2.3.3 Identification of patients requiring closer follow-up after discharge from hospital

Survivors of severe infections are at increased risk of morbidity and mortality but this risk is modifiable with post-discharge care.[51,52] However, outpatient follow-up and safety-netting is typically very difficult in rural and remote healthcare contexts. A systematic review found that in many LMICs paediatric post-discharge mortality rates are often as high as those occurring in-hospital.[53]

Hospitalisation represents a rare opportunity for healthcare workers to engage with some of society's most vulnerable individuals.[54] Risk stratification of patients using data collected in the lead up to discharge would enable resources to be focussed on more comprehensive follow-up of those at highest risk of post-discharge morbidity and mortality. Appropriate risk thresholds could be determined based on resources available for such a program. Prognostic factors and algorithms that predict poor outcome following hospitalisation have been identified.[41,55] Operationalising these for routine use might enable better targeting of peri- and post-discharge interventions.[56,57]

2.3.4 Prognostic tools to support clinical research and quality improvement initiatives

Prognostic tools could also improve management of febrile illnesses indirectly. Stratifying participant recruitment into trials of novel therapeutics by expected prognosis would ensure comparability between different sites, as well as selection of a study population in whom the outcome frequency is sufficiently high to adequately power the trial. Outside of clinical trials, accurate prognostication could help assess the impact of quality improvement initiatives, training programmes, and organisational changes, as well as facilitating inter-unit comparisons and benchmarking.[58]

2.4 Challenges to developing prediction tools for disease severity

There are many aspects that challenge the development of accurate, reliable, and practical tools to support healthcare providers working in resource-constrained contexts assess disease severity

in their patients. Whilst many studies evaluate clinical and laboratory prognostic factors in febrile patients, few examine the comparative utility of *combining* laboratory tests with clinical features. Methodological advances now allow comparison of different prediction algorithms in ways that acknowledge that the relative benefits of ruling-in and ruling-out serious disease are often dynamic and context-dependent.[59]

Reporting of clinical prediction research is sub-optimal.[60] Study design is often unclear with poor reporting of the proportion of participants who had met the pre-defined severity endpoint at the time baseline predictors were measured, making it difficult to understand whether the proposed prediction tool is serving predominantly diagnostic or prognostic functions.[61-63] Further, many studies only provide summary measures of model performance (for example, area under the receiver operating characteristic curves), which are poor correlates of clinical utility and of limited use to a health worker confronted with an individual patient.[64,65] Recent guidance on the design, reporting, and assessment of prediction studies aims to address this.[66,67]

2.5 Overview of this thesis

This thesis will explore the development and application of diagnostic and prognostic prediction tools to the problem of febrile illnesses in resource-limited settings. Specific objectives include: (i) identifying existing prognostic factors, severity scores, and clinical prediction models that can risk stratify febrile children presenting from the community, with a particular focus on acute respiratory infections [Chapters 3 and 4]; (ii) determining the differential contribution of clinical, biomarker-based, and combinatorial approaches to the diagnosis and prognosis of childhood pneumonia at the community-level [Chapter 4]; (iii) development of a prognostic model for critically ill children in locations with emerging critical care capacity [Chapter 5]; and (iv) development and external validation of a readily implementable prognostic model, combining simple clinical parameters

and point-of-care biomarkers, to facilitate the safe outpatient management of patients presenting with moderate Covid-19 [Chapter 6].

The overarching ethos is to combine the best clinical prediction model science with a pragmatic field-based reality to address locally-important healthcare issues in a manner that permits exploration of the generalisability of the findings.

3 Predictors of severity in paediatric community febrile illness

This chapter is based upon work published in: **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. Jan; 6(1): e003451, 2021.

3.1 Introduction

Acute febrile illnesses are amongst the most common reasons that parents seek medical care for their children.[1,2] Whilst most episodes are mild, an important minority of children progress to severe disease.[4] Early recognition of low-incidence serious disease is challenging,[68] especially in the peripheral healthcare settings of many countries where health workers receive limited training, patient volumes are high, diagnostic capacity is poor, and different acute febrile illnesses are often clinically indistinguishable.[14,42]

Clinical and laboratory prognostic factors that enable early and accurate identification of children at risk of developing severe disease could improve patient outcomes and reduce resource misallocation.[3,69] An increasing number of clinical decision-support algorithms and risk stratification tools integrate clinical and laboratory predictors to guide referral, admission, and treatment decisions.[44] Whilst no unified strategy exists to guide selection of candidate predictors, those already reported as prognostic should normally be considered, especially those demonstrating generalisability across different geographies and endemicities.[70,71]

Previous reviews have evaluated predictors of 'serious bacterial infections'.[72,73] However, these studies are diagnostic rather than prognostic.[70] Furthermore, 'serious bacterial infection' is an imperfect measure of disease severity: microbiological tests for bacterial infections lack sensitivity, especially in the context of prior antibiotic consumption;[74] 'serious bacterial infections' are not always

severe (for example, the majority of children with enteric fever are successfully managed as outpatients);[75] and severe febrile illnesses are frequently caused by non-bacterial pathogens, especially in low- and middle-income countries (LMICs),[14,34,76] in part secondary to the introduction of widespread vaccination against prevalent bacterial pathogens of childhood.[77]

This systematic review identifies which clinical and laboratory factors – alone or as part of clinical prediction models – predict progression to severe disease in febrile children presenting from the community to a community health worker, primary health centre, or hospital outpatient or emergency department. The aim was to understand which prognostic factors might support health workers faced with this difficult and common clinical question and to inform variable selection for future prospective studies aiming to develop data driven triage tools.

3.2 Methods

3.2.1 Protocol and registration

The methods for this systematic review were specified in advance (PROSPERO protocol: CRD42019140542). Data extraction processes adhered to the CHARMS (CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies)[78] and CHARMS-PF (a modification of CHARMS for Prognostic Factor studies)[79] checklists. Assessments of study quality (risk of bias and applicability) followed the QUIPS (QUality In Prognosis Studies)[80] and PROBAST (Prediction model Risk Of Bias ASsessment Tool) guidelines.[81] Reporting of the study is in accordance with PRISMA (Preferred Reporting Items for Systematic Review and Meta-Analysis) guidelines (Appendix 9.1).[82]

3.2.2 Eligibility criteria

All prognostic studies (prognostic factor and clinical prediction model) including ≥ 20 patients were eligible. The target population was children aged > 28 days and < 19 years, presenting from the

community with an acute febrile illness, defined as a documented abnormal temperature (fever or hypothermia) or history of fever, or suspected sepsis. Temperature is a dynamic parameter and a proportion of children with acute infection will not have an abnormal temperature at the time of presentation. For example, 29.1% (554/1,902) of children aged > 28 days and < 5 years hospitalised with suspected sepsis at Angkor Hospital for Children, Siem Reap, Cambodia presented with a normal temperature (unpublished data, 2017-19). Therefore, 'history of fever' was included to mitigate this. Furthermore, while sepsis is not always well-defined in children,[83] 'suspected sepsis' was included to ensure that the majority of studies targeting children with suspected infections were captured.

Studies were excluded if disaggregated paediatric data were not presented or if ≥ 50% of paediatric data pertained to neonates. Studies only reporting on specific clinical syndromes (for example, neurological presentations), particular pathogens (for example, *Plasmodium* spp., influenza, etc.), or children with comorbidities uncommon in the settings of interest (for example, studies focussing on febrile neutropenia were excluded but studies focussing on human immunodeficiency virus [HIV], sickle cell trait/disease, thalassaemia, or malnutrition were considered) were excluded.

Demographic, anthropometric, socioeconomic, clinical, and historical variables were considered, as well as laboratory parameters measured at presentation to care. Studies only reporting variables that would not be available at the time of presentation (for example, blood culture results) were excluded. Studies where participants were recruited partway through receipt of inpatient treatment or where authors identified that a substantial proportion were recruited following transfer from another health facility were excluded, as the aim of the review was to identify prognostic variables measured at presentation.

3.2.3 Primary outcome

The primary outcome was any objective measure of disease severity (for example, mortality, admission to critical care, length of stay in hospital and/or critical care unit, a pre-defined severity

score, or duration of symptoms), occurring within 30 days of measurement of the predictors or during hospitalisation. Studies assessing outcome at the same timepoint as baseline predictor measurements (diagnostic studies) were excluded.

3.2.4 Search strategy and selection criteria

MEDLINE, Embase, and Web of Science databases were searched, without language restriction, for publications between 31st May 1999 and 30th April 2020. Narrative reviews, letters, editorials, opinion pieces, comments, conference abstracts, and case series of less than 20 patients were not eligible for inclusion as primary articles but were considered for 'snowballing' (forward and reverse crosschecking of reference lists). Cochrane Prognosis Methods Group recommendations were followed to build the search strategy (Table 3.2-1). The search strategy was structured according to the PICOTS (Population, Intervention, Comparator, Outcome, Timing, Setting) framework and previously validated search strings optimised for systematic reviews of clinical prediction research were adapted to retrieve a list of studies tailored to the topic of interest.[84-86] Prior to implementing the search, the strategy was peer-reviewed by an independent Technical Advisory Panel of domain experts (Appendix 9.2).

Table 3.2-1: Search strategies used to retrieve articles for screening.

	MEDLINE	Embase	Science Citation Index via Web of Science
1	Fever[MeSH Terms] OR Fever[Title/Abstract] OR Febrile[Title/Abstract] OR “suspected sepsis”[Title/Abstract]	fever/ or (fever* or febrile or suspected sepsis).ti,ab,kw.	TS=(fever* or febrile or "suspected sepsis")
2	pediatrics[MeSH Terms] OR pediatric*[Title/Abstract] OR paediatric*[Title/Abstract] OR child[MeSH Terms] OR child*[Title/Abstract] OR Infant[Mesh:NoExp] OR infant[Title/Abstract]	pediatrics/ or child/ or infant/ or preschool child/ or school child/ or toddler/ or boy/ or girl/ or (pediatric* or paediatric* or child* or infant*).mp.	TS=(pediatric* or paediatric* or child* or infant*)
3	(((((((Validat*[tw] OR Predict*[ti] OR Rule*[tw]) OR (Predict*[tw] AND (Outcome*[tw] OR Risk*[tw] OR Model*[tw])) OR ((History OR Variable*[tw] OR Criteria OR Scor*[tw] OR Characteristic*[tw] OR Finding*[tw] OR Factor*[tw]) AND (Predict*[tw] OR Model*[tw] OR Decision*[tw] OR Identif*[tw] OR Prognos*[tw])) OR (Decision*[tw] AND (Model*[tw] OR Clinical*[tw] OR “Logistic Models”[MeSH Terms])) OR (Prognostic AND (History OR Variable*[tw] OR Criteria OR Scor*[tw] OR Characteristic*[tw] OR Finding*[tw] OR Factor*[tw] OR Model*[tw]))))) OR ("Stratification" OR "ROC Curve"[MeSH Terms] OR "Discrimination" OR "Discriminate" OR "c-statistic" OR "c statistic" OR "Area under the curve" OR "AUC" OR "Calibration" OR "Indices" OR "Algorithm" OR "Multivariable")))))	predict*.ti. or (validat* or rule* or (predict and (outcome* or risk* or model*)) or ((history or variable or criteria or scor* or characteristic* or finding* or factor*) and (predict* or model* or decision* or identif* or prognos*)) or (decision* and (model* or clinical*)) or (prognostic and (history or variable* or criteria or scor* or characteristic* or finding* or factor* or model*)) or stratification or discrimination or discriminate or c-statistic or "c statistic" or auc or calibration or indices or algorithm or multivariable).mp. or statistical model/ or "receiver operating characteristic"/ or "area under the curve"/	TI=(predict*) OR TS=(validat* or rule*) OR TS=((predict and (outcome* or risk* or model*))) OR TS=(((history or variable or criteria or scor* or characteristic* or finding* or factor*) and (predict* or model* or decision* or identif* or prognos*))) OR TS=((decision* and (model* or clinical*))) OR TS=((prognostic and (history or variable* or criteria or scor* or characteristic* or finding* or factor* or model*))) OR TS=(stratification or discrimination or discriminate or c-statistic or "c statistic" or auc or calibration or indices or algorithm or multivariable)

	MEDLINE	Embase	Science Citation Index via Web of Science
4	death[MeSH Terms] OR death[Title/Abstract] OR mortality[MeSH Terms] OR mortality[Title/Abstract] OR systemic inflammatory response syndrome[MeSH Terms] OR "systemic inflammatory response syndrome"[Title/Abstract] OR SIRS[Title/Abstract] OR sepsis[Title/Abstract] OR septic*[Title/Abstract] OR "severe disease*"[Title/Abstract] OR "severe infection*"[Title/Abstract] OR "severe bacterial infection*"[Title/Abstract] OR "severe illness"[Title/Abstract] OR "severe febrile illness"[Title/Abstract] OR "serious disease*"[Title/Abstract] OR "serious infection*"[Title/Abstract] OR "serious bacterial infection*"[Title/Abstract] OR "serious illness"[Title/Abstract] OR "serious febrile illness"[Title/Abstract]	mortality/ or childhood mortality/ or infant mortality/ or exp mortality rate/ or death/ or child death/ or fatality/ sepsis/ or systemic inflammatory response syndrome/ or exp septic shock/ or septicemia/ or (death or mortality or systemic inflammatory response or sirs or sepsis or septic* or ((severe or serious) adj2 (disease or illness* or infection*))).mp.	TS=(death or mortality or "systemic inflammatory response" or sirs or sepsis or septic*) OR TS=(((severe or serious) NEAR/2 (disease or illness* or infection*)))
5	1 AND 2 AND 3 AND 4	1 and 2 and 3 and 4	#4 AND #3 AND #2 AND #1
6	("1999/05/31"[Date - Publication] : "2020/04/30"[Date – Publication])	conference*.pt.	#4 AND #3 AND #2 AND #1 Refined by: PUBLICATION YEARS: (2020 OR 2019 OR 2010 OR 2002 OR 2018 OR 2009 OR 2001 OR 2017 OR 2008 OR 2000 OR 2016 OR 2007 OR 1999 OR 2015 OR 2006 OR 2014 OR 2005 OR 2013 OR 2004 OR 2012 OR 2003 OR 2011)
7	5 AND 6	5 not 6	#4 AND #3 AND #2 AND #1 Refined by: PUBLICATION YEARS: (2020 OR 2019 OR 2010 OR 2002 OR 2018 OR 2009 OR 2001 OR 2017 OR 2008 OR 2000 OR 2016 OR 2007 OR 1999 OR 2015 OR 2006 OR 2014 OR 2005 OR 2013 OR 2004 OR 2012 OR 2003 OR 2011) AND [excluding] DOCUMENT TYPES: (MEETING ABSTRACT OR PROCEEDINGS PAPER)

3.2.5 Study selection

Title, abstract, and full-text screening were performed independently by two reviewers. Agreement was checked after the first 20 and 250 articles. Discrepancies were resolved by discussion or independent assessment by a third reviewer.

Eligible studies and relevant review articles were 'snowballed' to identify additional studies. The list of eligible studies was presented to the Technical Advisory Panel who were asked to identify obvious omissions and suggest key authors whose publication lists were subsequently reviewed for additional eligible studies (Appendix 9.2).

3.2.6 Data extraction

A data extraction sheet was developed based on the CHARMS (prediction model studies) and CHARMS-PF (prognostic factor studies) checklists (Appendix 9.3).[78,79] Data were extracted independently by one reviewer and checked by a second. Discrepancies were discussed and resolved between the two reviewers. Authors of studies not reporting likelihood ratios (prognostic factors) or areas under the receiver operating characteristic curves (AUCs; clinical prediction models), or the data to allow their calculation, were contacted. In total, nine authors were contacted: seven responded to requests for clarifications and six provided additional data not available in the published manuscript. All predictors were harmonised using the Systematized Nomenclature of Medicine Clinical Terms (SNOMED-CT) to facilitate summaries of the predictive performance of the different prognostic factors.

3.2.7 Data analysis: prognostic factors

Contingency tables were constructed and positive and negative likelihood ratios (PLRs and NLRs) calculated for each prognostic factor. In the case of an empty cell, 0.5 was added to each cell (Haldane-Anscombe correction). Confidence intervals were calculated on the basis of the standard error of a proportion. Likelihood ratios were selected as the principal effect estimate as they allow

estimation of post-test probabilities, are intuitive for clinicians, and are frequently used to compare performance of predictors in diagnostic and prognostic studies.[72,73,87,88] Prognostic factors are presented in the main analysis if at least one study reported a PLR ≥ 5.0 (i.e., a rule-in test), or a NLR ≤ 0.2 (i.e., a rule-out test).[87] To contextualise the results, the outcome prevalence of individual studies was used to calculate the pre-test probability, and display positive and negative post-test probabilities on dumbbell plots (R v3.6.1).[89] Post-test probabilities were calculated using post-test odds: post-test probability = post-test odds / 1 + post-test odds. Post-test odds were calculated by multiplying the pre-test odds by the likelihood ratio.

3.2.8 Data analysis: clinical prediction models

For clinical prediction models, AUCs are presented on forest plots. When available, likelihood ratios associated with specific cut-offs of the models are presented.

3.2.9 Synthesis of results

Due to expected heterogeneity between studies (as a result of variations in case-mix and baseline risk), few common predictors for comparison, and absence of well-defined subgroups, no formal meta-analysis nor comparison of variability and bias between studies was planned, as these comparisons are recognised as being prone to bias.[90] Qualitative comparisons are described considering major differences between populations and study design. Prevalence of severe disease was used to group studies into low (<2.5%), moderate (2.5-7.5%), and high (>7.5%) prevalence settings, as a proxy for the case-mix of the study population.

3.2.10 Quality assessment

Risk of bias and applicability of studies were assessed using the QUIPS tool for prognostic factor studies,[80] and PROBAST for studies developing, validating, and/or updating prediction models.[81] Each

study was independently assessed using QUIPS or PROBAST by three reviewers. All discrepancies were resolved by discussion. For prognostic factor studies, risk of bias was categorised as low, medium, or high, whilst in clinical prediction model studies risk was categorised as low, high, or unclear. For all studies, applicability was assessed as being of low, high, or unclear concern.

3.3 Results

3.3.1 Summary of included studies

The electronic search retrieved 5,930 articles, and 19 additional articles were identified through snowballing and consultation with the Technical Advisory Panel (Figure 3.3-1). Eighteen studies were included in the review: 16 studies evaluated 200 prognostic factors, from 75 SNOMED-CT categories,[91-106] and eight evaluated 33 clinical prediction model/outcome pairs, using 25 distinct models.[92,93,95,97,98,106-108] Reasons for exclusion of studies qualifying for full-text screening could be grouped into: the wrong outcome (for example, bacterial infections); the wrong population (for example, adults or neonates); the wrong target condition (for example, febrile neutropenia); the wrong setting (for example, children recruited from the intensive care unit); the wrong design (for example, diagnostic studies measuring predictors and outcome at the same moment in time); or the wrong predictors (for example, microbiological tests, the results of which would not be available at the time the predictor or model was intended for use).

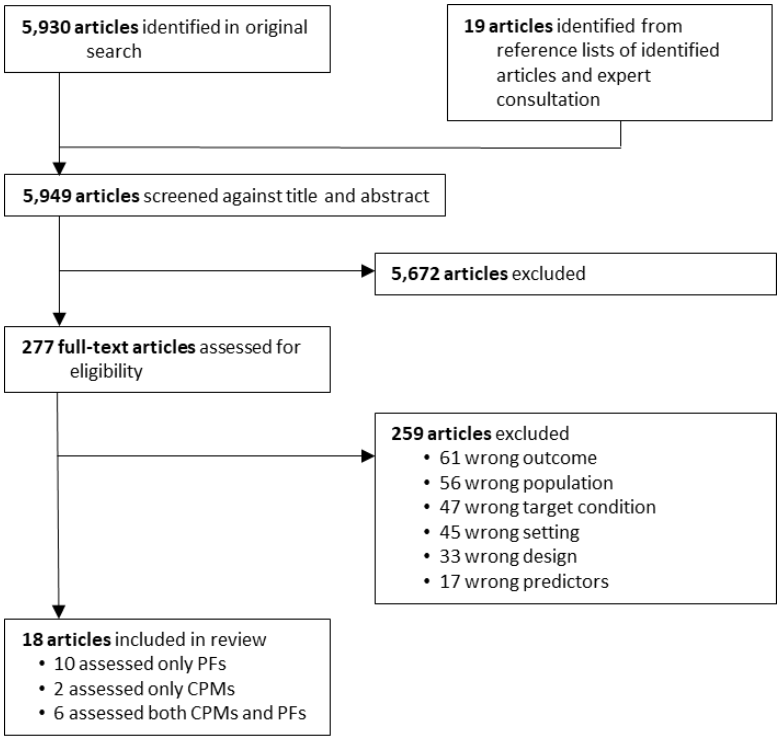


Figure 3.3-1: Selection of studies for systematic review. Only one reason for exclusion per study is listed. CPM = clinical prediction model; PF = prognostic factor.

In total 24,530 children were included, with overlap across eight studies.[91,98,99,101-104,108] The majority (11/18) included only hospitalised patients. Two studies recruited children from primary care,[95,100] and five recruited both children admitted and those sent home directly from hospital outpatient or emergency departments.[94,102-104,108] Seven studies included children aged five years and under,[92,93,96,99-101,107] with the remainder including patients up to 19 years of age. Definition of fever varied between studies, ranging from an axillary temperature (or equivalent) of ≥ 37.5°C to > 38.1°C. Five studies did not include a temperature measurement in their eligibility criteria and enrolled children on the basis of suspected infection or sepsis.[93,97,102,106,108] Eight studies were conducted in sub-Saharan Africa,[91-93,97-101] four in North America,[102-104,108] three in Europe,[95,96,106] two in Asia,[105,107] and one in South America.[94] Six were multi-centre studies.[91,93,98,100,105,108] Most used 'hard' outcomes to define severe disease, such as mortality, organ dysfunction, or need for organ support, while four used 'softer' outcomes, such as prolonged length of stay or persistence of symptoms.[95,96,100,106] Characteristics of the 18 studies are summarised in Table 3.3-1.

Table 3.3-1: Characteristics of studies included in systematic review. Studies are grouped according to the type of outcome they used: 'hard' (death, organ dysfunction, organ support, PICU admission) or 'soft' (length of stay, persistence of symptoms). [*]Studies evaluating both PFs and CPMs were categorised on the basis of their primary analysis to facilitate review using the appropriate quality assessment tool; [#]only children the treating physician decided to admit were eligible but recruitment occurred at the time of admission to the health facility; [†]hypotensive systolic blood pressure on arrival with receipt of a fluid bolus or vasoactive agent within 30 minutes; [‡]hypotension plus receipt of ≥ 30ml/kg isotonic crystalloids or vasoactive medication; [§]respiratory distress (increased work of breathing or deep breathing) and/or impaired consciousness (coma or prostration) AND evidence of poor peripheral perfusion (CRT > 2s or LLTG or weak radial pulse or severe tachycardia); [¶]initiation of antibiotics within 24h of arrival in the ED; [||]decreased mental status or perfusion in the setting of suspected infection; [**]rectal temperature > 38.5°C or < 35°C (or equivalent) AND heart rate > 2 SD above normal for age (unless hypothermic) AND respiratory rate > 2 SD above normal for age AND altered mental status OR systolic blood pressure < 2 SD below normal for age OR pulse pressure < 20mmHg OR CRT > 2s OR SpO_2 < 95% OR leucocyte count > $12x10^3$ cells/µl or < $5x10^3$ cells/µl; [***]convulsions, repeated vomiting, lethargy, severe anaemia, or loss of consciousness. CPM = clinical prediction model; CRT = capillary refill time; ED = emergency department; HIV = human immunodeficiency virus; LLTG = lower limb temperature gradient; OPD = outpatient department; PF = prognostic factor; PICU = paediatric intensive care unit; SAM = severe acute malnutrition; SD = standard deviation; SEAIDCRN = Southeast Asia Infectious Disease Clinical Research Network; SIRS = Systemic Inflammatory Response Syndrome; USA = United States of America.

Study (year); Setting, country	Cohort	Design	Quality assessment		CPM or PF[*]	Sample size	Population		Outcome	Outcome prevalence (n/N)
			Risk of bias	Applicability concern			Inclusion criteria	Exclusion criteria		
Outcomes including death, organ dysfunction, organ support, and PICU admission										
Scott (2020);[108] Secondary and tertiary care hospitals, USA	Hospital OPD/ED	Retrospective cohort	High	Low	CPM	2,464	Age 60d-18y; Clinician-suspected sepsis	Hypotensive septic shock on arrival;[†] transfer to another centre; leaving ED before formal evaluation; incorrect registration	Hypotensive septic shock[‡] ≤ 24h	11.4% (282/2,464)
Walia (2016);[107] Tertiary care hospital, India	Hospitalised[#]	Prospective cohort	High	High	CPM	100	Age 3-36m; Axillary temperature > 36.9°C (early morning) or > 37.4°C	Non-infectious cause of fever; immunisation ≤ 2d; immunodeficiency, autoimmune disorder	In-hospital mortality; Mechanical ventilation	11.0% (11/100); 17.0% (17/100)

Study (year); Setting, country	Cohort	Design	Quality assessment		CPM or PF*	Sample size	Population		Outcome	Outcome prevalence (n/N)
			Risk of bias	Applicability concern			Inclusion criteria	Exclusion criteria		
Aramburo (2018);[91] Secondary and tertiary care hospitals, Kenya, Tanzania and Uganda	Hospitalised#	Randomised controlled trial	Moderate	High	PF	3,008	Age 60d-12y; history of fever or axillary temperature ≥ 37.5°C or < 36°C; severe febrile illness§	Non-infectious cause of illness; SAM, gastroenteritis, burns, chronic kidney disease, pulmonary oedema, intoxication, surgical conditions, receipt of isotonic fluids during the same illness	In-hospital mortality (72h)	10.3% (309/3,008)
George (2015);[98] Secondary and tertiary care hospitals, Kenya, Tanzania and Uganda	Hospitalised#	Randomised controlled trial	High	High	CPM	3,121	Age 60d-12y; history of fever or axillary temperature ≥ 37.5°C or < 36°C; severe febrile illness§	Non-infectious cause of illness; SAM, gastroenteritis, burns, chronic kidney disease, pulmonary oedema, intoxication, surgical conditions, receipt of isotonic fluids during the same illness	In-hospital mortality (48h)	9.8% (306/3,121)
Scott (2012);[104] Tertiary care hospital, USA	Hospital OPD/ED	Prospective cohort	High	High	PF	239	Age < 19y; temperature > 38.5°C or < 36°C and heart rate > 2 SD above normal for age; underwent phlebotomy as part of usual care	Transfer from another health facility; known inborn errors of metabolism; receipt of > 15 min of intravenous therapy	24h organ dysfunction	5.4% (13/239)

Study (year); Setting, country	Cohort	Design	Quality assessment		CPM or PF*	Sample size	Population		Outcome	Outcome prevalence (n/N)
			Risk of bias	Applicability concern			Inclusion criteria	Exclusion criteria		
Scott (2014);[103] Tertiary care hospital, USA	Hospital OPD/ED	Prospective cohort	High	High	PF	239	Age < 19y; temperature > 38.5°C or < 36°C and heart rate > 2 SD above normal for age; undergoing phlebotomy as part of routine care	Transfer from another health facility; known inborn errors of metabolism; receipt of > 15 min of intravenous therapy	24h organ dysfunction	5.4% (13/239)
Nadjm (2013);[101] Secondary care hospital, Tanzania	Hospitalised#	Prospective cohort	Moderate	High	PF	3,319	Age 2m-5y; history of fever in last 48h or axillary temperature ≥ 37.5°C	Chronic illness (excluding HIV and malnutrition); trauma; surgical conditions	In-hospital mortality	5.1% (170/3,319)
Mtove (2011);[99] Secondary care hospital, Tanzania	Hospitalised#	Prospective cohort	Moderate	High	PF	3,248	Age 2m-13y; history of fever in last 48h or axillary temperature ≥ 37.5°C	Chronic illness (excluding HIV and malnutrition); trauma; surgical conditions	In-hospital mortality	5.0% (164/3,248)
Lowlaavar (2016);[93] Secondary and tertiary care hospitals, Uganda	Hospitalised#	Prospective cohort	High	High	CPM	1,307	Age 6-60m; admitted during study working hours or within 8h of study shift with a proven or suspected infection	Previous enrolment; residence outside study catchment area	In-hospital mortality	5.0% (65/1,307)
Conroy (2015);[92] Tertiary care hospital, Uganda	Hospitalised#	Prospective cohort	High	High	CPM	2,502	Age 2m-5y; history of fever in last 48h or axillary temperature > 37.5°C	None reported	In-hospital mortality	4.7% (99/2,089)

Study (year); Setting, country	Cohort	Design	Quality assessment		CPM or PF*	Sample size	Population		Outcome	Outcome prevalence (n/N)
			Risk of bias	Applicability concern			Inclusion criteria	Exclusion criteria		
van Nassau (2018);[106] Secondary care hospital, The Netherlands	Hospitalised#	Retrospective cohort	High	High	CPM	864	Age < 18y; suspected bacterial infection¶	Surgical conditions	PICU transfer and/or in-hospital mortality	2.7% (24/864)
Scott (2017);[102] Tertiary care hospital, USA	Hospital OPD/ED	Retrospective cohort	Low	High	PF	1,299	Age 60d-18y; suspected sepsis;‖ measurement of venous lactate as part of routine care within 8h of ED arrival	Transfer from another health facility	30d mortality	1.9% (25/1,299)
SEAIDCRN (2017);[105] Tertiary care hospitals, Indonesia, Thailand and Vietnam	Hospitalised#	Prospective cohort	High	High	PF	763	Age 30d-18y; modified SIRS criteria**	Suspicion of hospital-acquired infection; admission to hospital within previous 30d; transfer from another health facility after > 72h admission; weight < 3kg; enrolment in another clinical study	28d mortality	1.9% (14/731)
Costa de Santana (2017);[94] Tertiary care hospital, Brazil	Hospital OPD/ED	Retrospective cohort	High	High	PF	254	Age < 13y; axillary temperature > 38.5°C; measurement of respiratory rate and heart rate on three occasions in absence of fever; measurement of leucocyte count as part of routine care	Congenital malformations, bronchopulmonary dysplasia, medullary aplasia and cardiac, renal or hepatic insufficiency	In-hospital mortality	1.6% (4/254)

Study (year); Setting, country	Cohort	Design	Quality assessment		CPM or PF*	Sample size	Population		Outcome	Outcome prevalence (n/N)
			Risk of bias	Applicability concern			Inclusion criteria	Exclusion criteria		
Kwizera (2019);[97] Secondary care hospital, Rwanda	Hospitalised#	Prospective cohort	High	High	CPM	949	Age 28d-18y; confirmed acute infectious disease; symptom onset < 14d prior to hospital admission	Allergy to antimicrobials to treat sepsis (antibiotics, artesunate, artemether-lumefantrine); terminal disease	In-hospital mortality	1.5% (14/949)
Outcomes including length of stay and persistence of symptoms										
Freyne (2013);[96] Secondary care hospital, Ireland	Hospitalised#	Prospective cohort	High	High	PF	46	Age 6-36m; axillary temperature > 38.1°C	Chronic illness; immunisation ≤ 2d, antipyretic use ≤ 2h	Length of stay > 96h	26.1% (12/46)
van Nassau (2018);[106] Secondary care hospital, The Netherlands	Hospitalised#	Retrospective cohort	High	High	CPM	864	Age < 18y; suspected bacterial infection¶	Surgical conditions	Length of stay ≥ 7d	22.2% (179/806)
Elshout (2015);[95] General Practice (out of hours), The Netherlands	Primary care	Prospective cohort	High	High	PF	480	Age 3m-6y; history of fever	Communication in Dutch not possible; enrolment in last 2 weeks; direct referral to hospital required	Persistent fever at D3	13.1% (63/480)
Mwandama (2016);[100] Community Health Workers, Malawi	Primary care	Prospective cohort	High	High	PF	285	Age 2-59m; history of fever in last 48h or temperature ≥ 37.5°C; negative malaria rapid diagnostic test	Receipt of antimalarial in last 2 weeks; presence of danger signs***	Persistent symptoms at D7	10.4% (19/182)

3.3.2 Prognostic factors

Figures 3.3-2, 3.3-3, 3.3-4, and 3.3-5 present prognostic factors identified as having rule-in (PLR ≥ 5.0) or rule-out (NLR ≤ 0.2) value in at least one study. Prognostic factors that met neither of these pre-specified cut-offs are presented in Appendix 9.4. In settings with moderate prevalence of severe disease, both high lactate (PLR range 4.97 to 5.13) and hypoglycaemia (PLR range 12.63 to 13.36) were useful for ruling in severe disease (Figure 3.3-2),[99,101,104] whereas a lactate ≤ 5mM was more useful as a rule-out test (NLR range 0.13 to 0.39) amongst a population in whom prevalence of severe disease was high (febrile children with signs of poor organ perfusion).[91,98]

Hypoxia was most useful to rule-in severe disease in moderate prevalence settings (PLR range 8.11 to 9.49).[92,101] Some studies found hypotension and bedside markers of poor peripheral perfusion (capillary refill time [CRT], limb-core temperature gradient, and pulse character) to have useful rule-in value, but this was inconsistent (PLR range 1.89 to 9.57 and 1.78 to 17.38 respectively).[91,92,98,99,101-103,106] Bradycardia was evaluated in a multi-centre study conducted across three East African countries and found to have useful rule-in value (PLR range 5.95 to 14.59) for severe disease in those high prevalence settings (Figure 3.3-3).[91,98]

Impaired consciousness, assessed using bedside coma scales, was a useful predictor of severe disease, particularly in moderate prevalence settings (PLR range 3.38 to 14.02), with the post-test probability of poor outcome increasing with the degree of neurological impairment (Figure 3.3-4).[92,93,97,99,101,106]

In sub-Saharan African settings, severe malnutrition (weight-for-age z-score [WAZ] ≤ -3 or mid-upper arm circumference [MUAC] ≤ 11.5cm; PLR range 1.56 to 11.23),[92,97,99,101] HIV positive status (PLR range 4.09 to 12.48),[92,93,97] and bedside correlates of metabolic derangement such as deep breathing and jaundice (PLR range 3.57 to 7.71) were useful rule-in predictors, across a range of low and moderate prevalence settings (Figure 3.3-5).[92,99,101]

Study (setting)	Outcome	Prevalence	Cutoff	PLR (95% CI)	NLR (95% CI)
OUTCOME: Mortality, organ dysfunction, or PICU admission					
Lactate					
Scott 2017 (USA)	30d mortality	1.9 %	≥ 4mM	2.60 (1.16 to 5.83)	0.87 (0.71 to 1.05)
Mtove 2011 (Tanzania)	In-hospital mortality	5.0 %	> 5mM	5.13 (4.34 to 6.08)	0.49 (0.41 to 0.59)
Nadjm 2013 (Tanzania)	In-hospital mortality	5.1 %	> 5mM	5.00 (4.21 to 5.93)	0.48 (0.40 to 0.58)
Scott 2012 (USA)	Organ dysfunction (24h)	5.4 %	≥ 4mM	4.97 (1.90 to 12.98)	0.74 (0.51 to 1.06)
George 2015 (East Africa)	In-hospital mortality (48h)	9.9 %	> 5mM	2.28 (2.10 to 2.47)	0.34 (0.27 to 0.42)
Aramburo 2018 (East Africa)	In-hospital mortality (72h)	10.3 %	≥ 5mM	2.63 (2.47 to 2.79)	0.13 (0.09 to 0.19)
Glucose					
Mtove 2011 (Tanzania)	In-hospital mortality	5.0 %	< 2.5mM	12.63 (8.88 to 17.99)	0.75 (0.69 to 0.83)
Nadjm 2013 (Tanzania)	In-hospital mortality	5.1 %	< 2.5mM	13.36 (9.36 to 19.07)	0.76 (0.69 to 0.83)
Aramburo 2018 (East Africa)	In-hospital mortality (72h)	10.3 %	< 2.5mM	5.10 (3.65 to 7.14)	0.87 (0.83 to 0.91)
Potassium					
Aramburo 2018 (East Africa)	In-hospital mortality (72h)	10.3 %	≥ 6mM	6.64 (4.46 to 9.89)	0.84 (0.78 to 0.89)

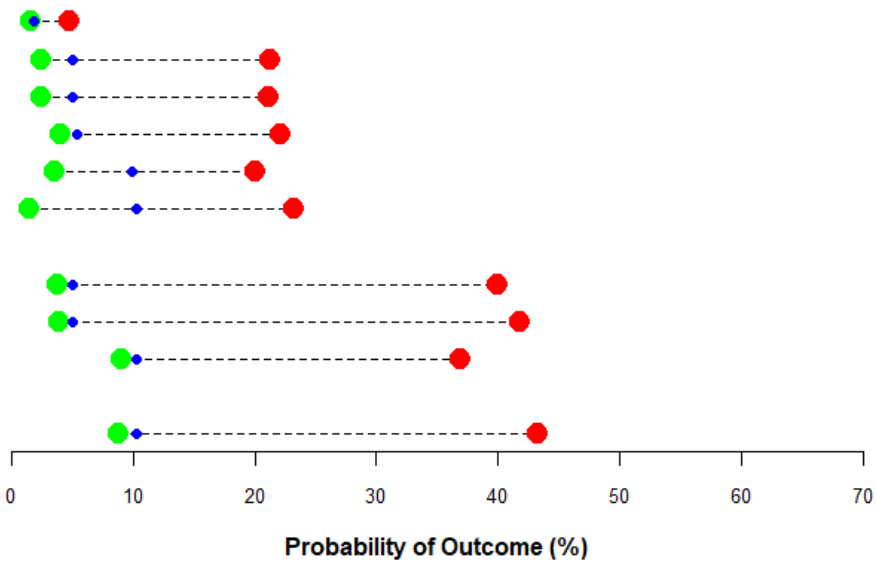


Figure 3.3-2: Prognostic factors with rule-in or rule-out value for severe disease – laboratory tests. Blue dot = prevalence or pre-test probability, green dot = negative post-test probability, and red dot = positive post-test probability. CI = confidence interval; NLR = negative likelihood ratio; PICU = paediatric intensive care unit; PLR = positive likelihood ratio.

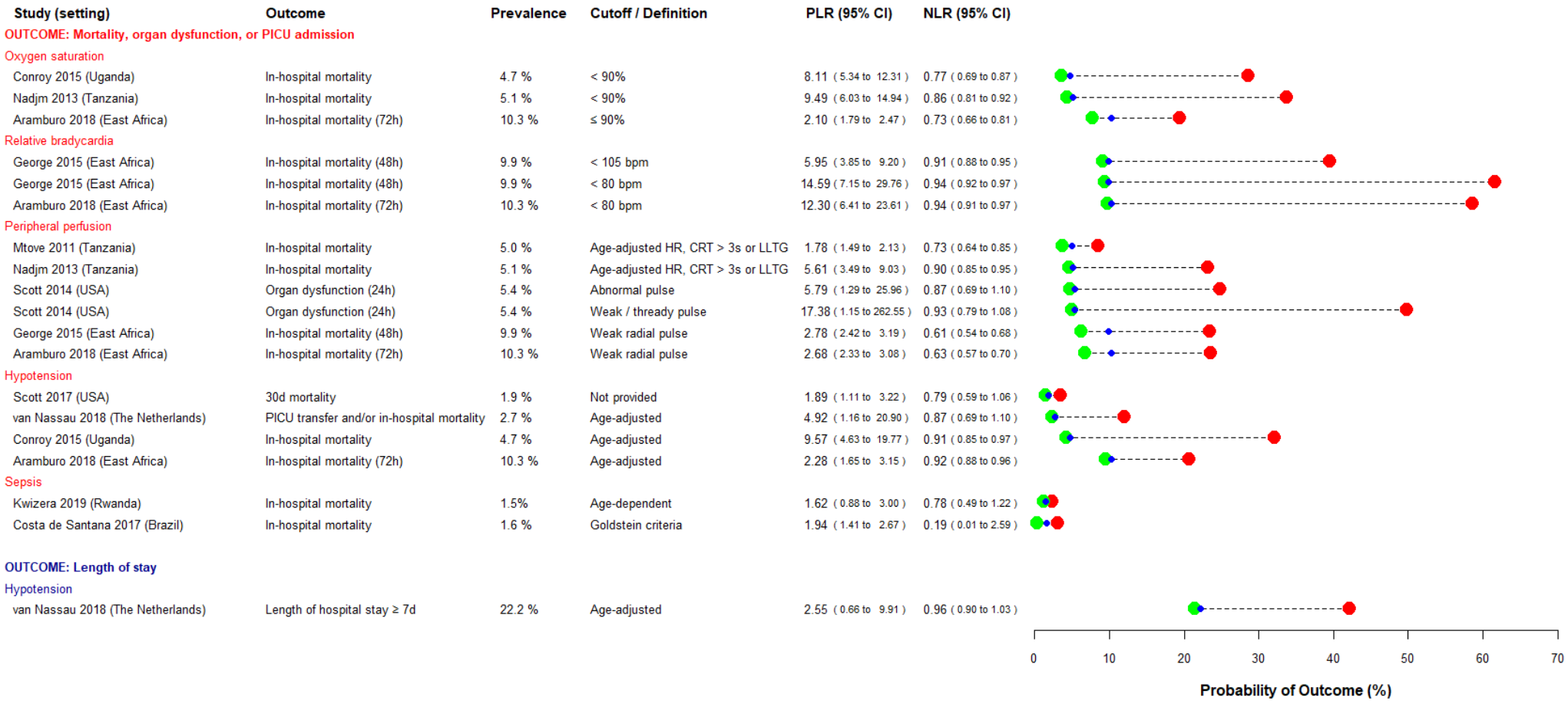


Figure 3.3-3: Prognostic factors with rule-in or rule-out value for severe disease – cardiovascular or respiratory signs. 'Sepsis' was defined according to Goldstein criteria (Costa de Santana et al.)[109] and the quick Sequential Organ Failure Assessment score in children aged ≥ 15 years, and using a combination of temperature, mental status, respiratory distress, prostration, and seizures in children aged < 15 years (Kwizera et al.). Blue dot = prevalence or pre-test probability, green dot = negative post-test probability, and red dot = positive post-test probability. Bpm = beats per minute; CI = confidence interval; CRT = capillary refill time; HR = heart rate; LLTG = lower-limb temperature gradient; NLR = negative likelihood ratio; PICU = paediatric intensive care unit; PLR = positive likelihood ratio.

Study (setting)	Outcome	Prevalence	Cutoff / Definition	PLR (95% CI)	NLR (95% CI)
OUTCOME: Mortality, organ dysfunction, or PICU admission					
Coma scales					
Kwizera 2019 (Rwanda)	In-hospital mortality	1.5%	Not provided	0.95 (0.69 to 1.29)	1.04 (0.82 to 1.33)
van Nassau 2018 (The Netherlands)	PICU transfer and/or in-hospital mortality	2.7 %	AVPU < A or GCS < 15	6.04 (3.51 to 10.39)	0.47 (0.26 to 0.87)
Conroy 2015 (Uganda)	In-hospital mortality	4.7 %	AVPU < A	5.42 (4.65 to 6.31)	0.27 (0.19 to 0.39)
Scott 2014 (USA)	Organ dysfunction (24h)	5.4 %	AVPU < A	3.38 (1.88 to 6.07)	0.55 (0.30 to 0.99)
George 2015 (East Africa)	In-hospital mortality (48h)	9.9 %	AVPU < A	1.24 (1.19 to 1.29)	0.30 (0.20 to 0.44)
Lowlaavar 2016 (Uganda)	In-hospital mortality	5.0%	BCS < 5	2.80 (1.63 to 4.81)	0.61 (0.36 to 1.03)
Mtove 2011 (Tanzania)	In-hospital mortality	5.0 %	BCS < 5	8.50 (6.99 to 10.32)	0.49 (0.41 to 0.58)
Nadjm 2013 (Tanzania)	In-hospital mortality	5.1 %	BCS < 5	8.07 (6.66 to 9.77)	0.50 (0.42 to 0.59)
Conroy 2015 (Uganda)	In-hospital mortality	4.7 %	BCS < 3	14.02 (10.52 to 18.68)	0.48 (0.39 to 0.60)
George 2015 (East Africa)	In-hospital mortality (48h)	9.9 %	Inability to localise pain	3.26 (2.75 to 3.86)	0.69 (0.63 to 0.75)
Aramburo 2018 (East Africa)	In-hospital mortality (72h)	10.3 %	Inability to localise pain	3.11 (2.61 to 3.71)	0.71 (0.65 to 0.78)
Prostration					
Conroy 2015 (Uganda)	In-hospital mortality	4.7 %	Inability to sit or breastfeed	3.88 (3.46 to 4.36)	0.18 (0.11 to 0.30)
Aramburo 2018 (East Africa)	In-hospital mortality (72h)	10.3 %	Inability to sit or breastfeed	0.87 (0.78 to 0.97)	1.23 (1.07 to 1.40)
OUTCOME: Length of stay, duration of symptoms					
Coma scales					
van Nassau 2018 (The Netherlands)	Length of hospital stay ≥ 7d	22.2 %	AVPU < A or GCS < 15	2.13 (1.16 to 3.92)	0.91 (0.82 to 1.00)

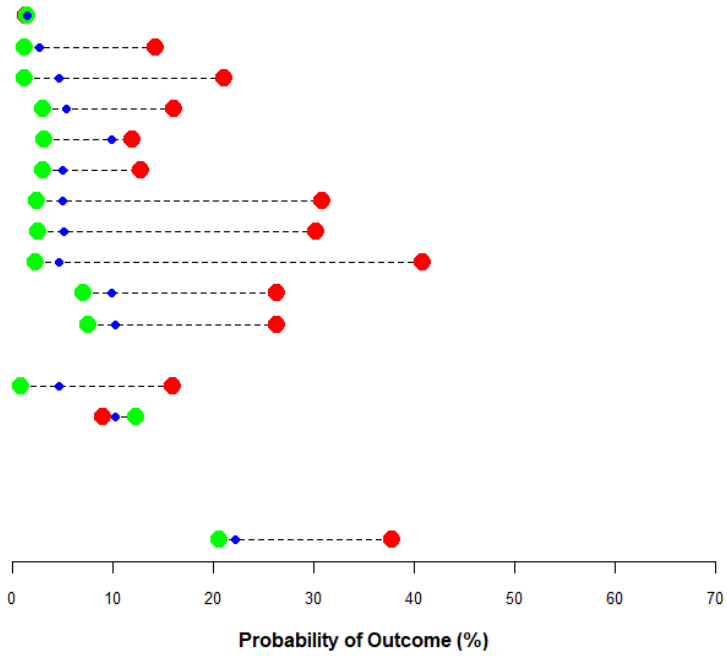


Figure 3.3-4: Prognostic factors with rule-in or rule-out value for severe disease – neurological signs. Blue dot = prevalence or pre-test probability, green dot = negative post-test probability, and red dot = positive post-test probability. AVPU = Alert Voice Pain Unresponsive scale; BCS = Blantyre Coma Scale; CI = confidence interval; GCS = Glasgow Coma Scale; NLR = negative likelihood ratio; PICU = paediatric intensive care unit; PLR = positive likelihood ratio.

Study (setting)	Outcome	Prevalence	Cutoff / Definition	PLR (95% CI)	NLR (95% CI)
OUTCOME: Mortality, organ dysfunction, or PICU admission					
Malnutrition					
Kwizera 2019 (Rwanda)	In-hospital mortality	1.5%	WAZ < 3*	11.13 (1.43 to 86.48)	0.93 (0.81 to 1.08)
Conroy 2015 (Uganda)	In-hospital mortality	4.7%	WAZ < 3*	1.56 (0.94 to 2.59)	0.94 (0.87 to 1.02)
Mtove 2011 (Tanzania)	In-hospital mortality	5.0%	WAZ ≤ 3*	3.90 (2.37 to 6.42)	0.92 (0.87 to 0.97)
Nadjm 2013 (Tanzania)	In-hospital mortality	5.1%	WAZ ≤ 3 or MUAC < 11.5cm*	4.95 (3.33 to 7.35)	0.87 (0.81 to 0.93)
Aramburo 2018 (East Africa)	In-hospital mortality (72h)	10.3%	MUAC ≤ 11.5cm	3.08 (1.72 to 5.54)	0.95 (0.92 to 0.99)
Deep breathing					
Conroy 2015 (Uganda)	In-hospital mortality	4.7 %	Bedside assessment	3.57 (3.12 to 4.08)	0.28 (0.20 to 0.41)
Mtove 2011 (Tanzania)	In-hospital mortality	5.0 %	Bedside assessment	7.71 (6.03 to 9.85)	0.64 (0.57 to 0.73)
Nadjm 2013 (Tanzania)	In-hospital mortality	5.1 %	Bedside assessment	7.63 (6.02 to 9.66)	0.63 (0.56 to 0.72)
George 2015 (East Africa)	In-hospital mortality (48h)	9.9 %	Bedside assessment	1.36 (1.29 to 1.44)	0.40 (0.31 to 0.52)
Aramburo 2018 (East Africa)	In-hospital mortality (72h)	10.3 %	Bedside assessment	1.37 (1.29 to 1.44)	0.40 (0.31 to 0.53)
Jaundice					
Conroy 2015 (Uganda)	In-hospital mortality	4.7 %	Bedside assessment	5.42 (3.65 to 8.06)	0.78 (0.70 to 0.88)
Focus of infection					
Kwizera 2019 (Rwanda)	In-hospital mortality	1.5%	Abdominal	8.91 (0.48 to 165.11)	0.97 (0.88 to 1.07)
Comorbidity					
Kwizera 2019 (Rwanda)	In-hospital mortality	1.5%	HIV positive	12.48 (0.63 to 248.97)	0.97 (0.88 to 1.06)
Lowlaavar 2016 (Uganda)	In-hospital mortality	5.0%	HIV positive	4.22 (2.16 to 8.23)	0.84 (0.72 to 0.98)
Conroy 2015 (Uganda)	In-hospital mortality	5.7%	HIV positive	4.09 (1.86 to 9.02)	0.86 (0.74 to 1.01)
Aramburo 2018 (East Africa)	In-hospital mortality (72h)	11.3%	HIV positive	2.32 (1.46 to 3.70)	0.95 (0.91 to 0.99)
Scott 2017 (USA)	30d mortality	1.9%	Any comorbidity	1.35 (1.25 to 1.46)	0.12 (0.02 to 0.81)
Scott 2014 (USA)	Organ dysfunction (24h)	5.4%	Any comorbidity	2.03 (1.17 to 3.51)	0.63 (0.35 to 1.14)
OUTCOME: Length of stay, duration of symptoms					
Comorbidity					
Elshout 2015 (The Netherlands)	Fever > 3d after recruitment	13.1%	Under routine follow-up	0.86 (0.48 to 1.51)	1.04 (0.92 to 1.17)

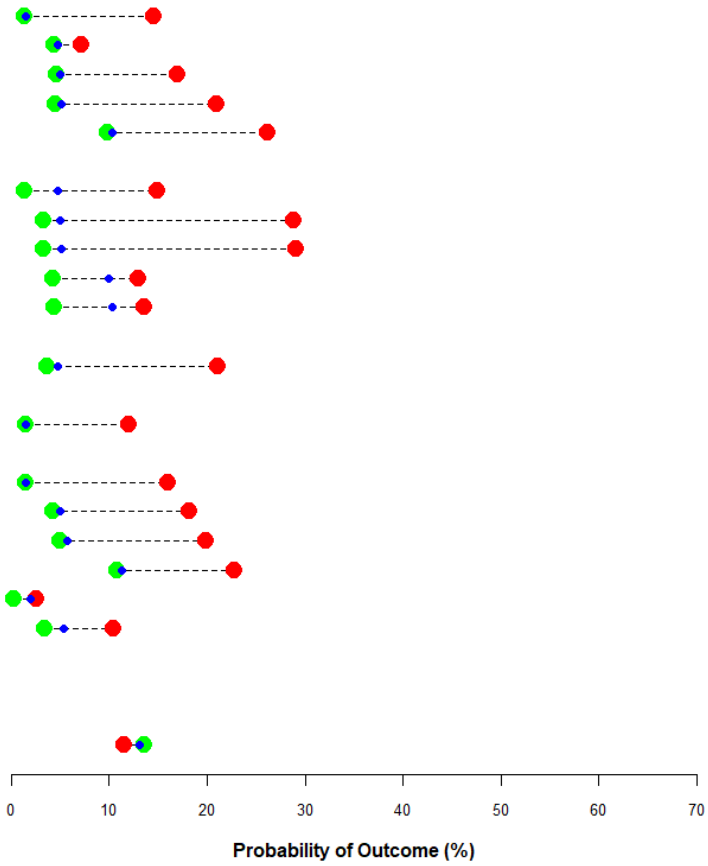


Figure 3.3-5: Prognostic factors with rule-in or rule-out value for severe disease – historical, anthropometric, and metabolic variables. *Visible wasting or nutritional oedema included in malnutrition definition. 'Comorbidity' defined as under routine care of paediatrician or ear nose and throat specialist (Elshout et al.). Blue dot = prevalence or pre-test probability, green dot = negative post-test probability, and red dot = positive post-test probability. CI = confidence interval; HIV = human immunodeficiency virus; MUAC = mid-upper arm circumference; NLR = negative likelihood ratio; PICU = paediatric intensive care unit; PLR = positive likelihood ratio; WAZ = weight-for-age z-score.

Very few individual prognostic factors were satisfactorily able to rule-out progression to severe disease: presence of comorbidities (NLR range 0.12 to 1.04), sepsis at admission (NLR range 0.19 to 0.78), and prostration (NLR range 0.18 to 1.23) were each identified in only one study.[92,94,102]

3.3.3 Clinical prediction models

Figure 3.3-6 illustrates the discrimination (AUCs) of 25 clinical prediction models for 33 different outcomes assessed in eight studies: most (18/33) were external validations of existing models;[92,98,106,107] 13 were newly derived models;[93,95,97,98,108] and two were updates and external validations of an existing model.[106] Components of the clinical prediction models are summarised in Table 3.3-2.

Three models (Lambaréné Organ Dysfunction Score [LODS], Paediatric Early Death Index for Africa-Early [PEDIA-e], and Signs of Inflammation in Children that Kill [SICK]) demonstrated AUCs ≥ 0.80 in a Ugandan setting where in-hospital mortality occurred at a prevalence of 4.7% (AUC range 0.85 to 0.90).[92] Two of these (LODS and PEDIA-e) were also assessed in a multi-country study in East Africa where discrimination was lower (AUCs of 0.77 and 0.70).[98] This study also derived two models (Fluid Expansion As Supportive Therapy-Paediatric Emergency Triage [FEAST-PET] and FEAST-PET and Laboratory [FEAST-PETaL]), which achieved AUCs of 0.82 and 0.86 respectively.[98] Two other East African studies used combinations of simple clinico-demographic variables to derive a number of prediction models, four of which had AUCs ≥ 0.80.[93,97]

One North American study, including both inpatients and outpatients, derived a model to predict hypotensive shock in children presenting with suspected sepsis, which showed an AUC of 0.87 in an external geographic validation.[108] The Yale Observation Score also demonstrated AUCs of 0.97 for mortality and 0.89 for mechanical ventilation in India, however the small sample size (n = 100) rendered the results difficult to interpret.[107] In general, models assessed against 'softer' outcomes (for example, persistence of symptoms or length of stay) had lower discrimination.

Study (setting)	Model	Outcome	Outcome prevalence (N)	AUC (95% CI)
OUTCOME: Mortality, organ dysfunction, or PICU admission				
George 2015 (East Africa)	PEDIA Late death	In-hospital mortality (>48h)	1.0% (959)	0.55 (0.40 to 0.69)
Kwizera 2019 (Rwanda)	Model 1	In-hospital mortality	1.5% (949)	0.79 (0.75 to 0.81)
Kwizera 2019 (Rwanda)	Model 2	In-hospital mortality	1.5% (949)	0.70 (0.60 to 0.79)
Kwizera 2019 (Rwanda)	Model 3	In-hospital mortality	1.5% (949)	0.72 (0.64 to 0.91)
Kwizera 2019 (Rwanda)	Model 4	In-hospital mortality	1.5% (949)	0.80 (0.78 to 0.80)
Kwizera 2019 (Rwanda)	Model 5	In-hospital mortality	1.5% (949)	0.72 (0.57 to 0.80)
van Nassau 2018 (The Netherlands)	SIRS	PICU transfer and/or in-hospital mortality	2.9% (755)	0.64 (0.53 to 0.74)
George 2015 (East Africa)	PEDIA Immediate death	In-hospital mortality (<4h)	3.0% (1,007)	0.75 (0.68 to 0.83)
van Nassau 2018 (The Netherlands)	qPELOD-2	PICU transfer and/or in-hospital mortality	3.3% (545)	0.60 (0.45 to 0.76)
van Nassau 2018 (The Netherlands)	qSOFA	PICU transfer and/or in-hospital mortality	3.7% (484)	0.72 (0.57 to 0.86)
George 2015 (East Africa)	PEDIA Early death	In-hospital mortality (4-48h)	4.0% (1,003)	0.70 (0.63 to 0.77)
Conroy 2015 (Uganda)	SICK	In-hospital mortality	4.7% (2,089)	0.85 (0.83 to 0.86)
Conroy 2015 (Uganda)	PEDIA Early death	In-hospital mortality	4.7% (2,089)	0.90 (0.88 to 0.91)
Conroy 2015 (Uganda)	LODS	In-hospital mortality	4.7% (2,089)	0.90 (0.88 to 0.91)
George 2015 (East Africa)	AQUAMAT	In-hospital mortality (48h)	5.0% (648)	0.74 (0.65 to 0.83)
Lowlaavar 2016 (Uganda)	Model 1	In-hospital mortality	5.0% (1,307)	0.85 (0.80 to 0.89)
Lowlaavar 2016 (Uganda)	Model 2	In-hospital mortality	5.0% (1,307)	0.84 (0.79 to 0.89)
Lowlaavar 2016 (Uganda)	Model 3	In-hospital mortality	5.0% (1,307)	0.82 (0.72 to 0.91)
George 2015 (East Africa)	PRISM III	In-hospital mortality (48h)	6.0% (627)	0.71 (0.61 to 0.81)
George 2015 (East Africa)	LODS	In-hospital mortality (48h)	7.0% (1,038)	0.77 (0.72 to 0.82)
George 2015 (East Africa)	FEAST-PET	In-hospital mortality (48h)	7.0% (1,024)	0.82 (0.77 to 0.87)
George 2015 (East Africa)	FEAST-PETaL	In-hospital mortality (48h)	7.0% (1,024)	0.86 (0.82 to 0.90)
van Nassau 2018 (The Netherlands)	qSOFA-lactate	PICU transfer and/or in-hospital mortality	7.9% (151)	0.67 (0.50 to 0.84)
George 2015 (East Africa)	Bedside PEWS	In-hospital mortality (48h)	9.0% (1,000)	0.64 (0.56 to 0.71)
Walia 2016 (India)	YOS	Mortality	11.0% (100)	0.89 (0.89 to 0.89)
Scott 2020 (USA)	Scott model (Temporal set)	Hypotensive septic shock (24h)	11.0% (698)	0.75 (0.69 to 0.81)
Scott 2020 (USA)	Scott model (Geographic set)	Hypotensive septic shock (24h)	14.0% (69)	0.87 (0.73 to 1.00)
Walia 2016 (India)	YOS	Mechanical ventilation	17.0% (100)	0.97 (0.97 to 0.97)
OUTCOME: Length of stay or duration of symptoms				
Elshout 2015 (The Netherlands)	Elshout model	Fever >3d after recruitment	13.1% (480)	0.64 (0.58 to 0.70)
van Nassau 2018 (The Netherlands)	qPELOD-2	Length of hospital stay >/=7d	22.2% (545*)	0.51 (0.45 to 0.57)
van Nassau 2018 (The Netherlands)	qSOFA	Length of hospital stay >/=7d	22.2% (484*)	0.53 (0.46 to 0.59)
van Nassau 2018 (The Netherlands)	qSOFA-lactate	Length of hospital stay >/=7d	22.2% (151*)	0.56 (0.46 to 0.67)
van Nassau 2018 (The Netherlands)	SIRS	Length of hospital stay >/=7d	22.2% (755*)	0.49 (0.44 to 0.54)

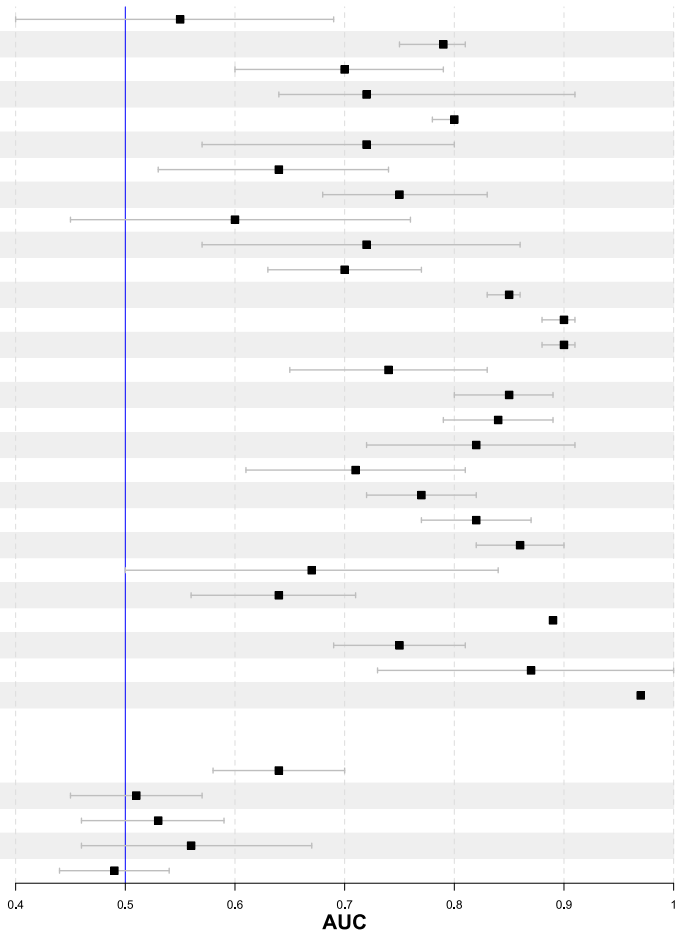


Figure 3.3-6: Discrimination of clinical prediction models to identify children at risk of severe disease. Individual studies evaluated different clinical prediction models using datasets with varying numbers of children with severe disease, depending on the data available. The outcome prevalence reflects the proportion of children with severe disease in the dataset used to evaluate that particular prediction model/outcome pair. This may be different from the overall prevalence of

children with severe disease in the study, which is listed in Table 3.3-1 and used to classify studies into low, moderate, or high prevalence settings. No confidence intervals were provided for the AUC estimates in the study by Walia et al. *Number of participants included in analyses for 'length of stay' outcome in van Nassau et al. study is an estimate: 93% of participants eligible for the primary analysis had data on length of stay available. AQUAMAT = African Quinine Artesunate Malaria Trial; AUC = area under the receiver operating characteristic; CI = confidence interval; FEAST-PET = FEAST-Paediatric Emergency Triage; FEAST-PETaL = FEAST-Paediatric Emergency Triage and Laboratory; LODS = Lambaréné Organ Dysfunction Score; PEDIA = Pediatric Early Death Index for Africa; PEWS = Pediatric Early Warning Score; PICU = paediatric intensive care unit; PRISM III = Pediatric Risk of Mortality; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2; qSOFA = quick Sequential Organ Failure Assessment; SICK = Signs of Inflammation in Children that Kill; SIRS = Systemic Inflammatory Response Syndrome; YOS = Yale Observation Score.

Table 3.3-2: Constituent variables of clinical prediction models. *Kwashiorkor was not included in the PEDIA-e score in the Conroy et al. study; †Receipt of oxygen therapy and respiratory effort included in the original PEWS but not measured in the George et al. study; ‡pupillary reflexes, pH, total CO_2, arterial PaO_2, creatinine, urea, white blood cells, prothrombin time, and platelets included in the original PRISM III score but not measured in the George et al. study; +serious illness was defined as any abnormal test result (for example, lumbar puncture, chest radiograph, serum electrolytes, or urine, stool, or blood cultures) in a series of 262 febrile children seen in primary care in New Haven, USA. AQUAMAT = African Quinine Artesunate Malaria Trial; CI = confidence interval; FEAST-PET = FEAST-Paediatric Emergency Triage; FEAST-PETaL = FEAST-Paediatric Emergency Triage and Laboratory; LODS = Lambaréné Organ Dysfunction Score; PEDIA = Pediatric Early Death Index for Africa; PEWS = Pediatric Early Warning Score; PICU = paediatric intensive care unit; PRISM III = Pediatric Risk of Mortality; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2; qSOFA = quick Sequential Organ Failure Assessment; SICK = Signs of Inflammation in Children that Kill; SIRS = Systemic Inflammatory Response Syndrome; YOS = Yale Observation Score.

Clinical Prediction Model	Predictor variables used in the original studies	Original outcome	Included study evaluating the model	Original study developing the model
AQUAMAT	Base deficit, impaired consciousness, convulsions, elevated blood urea, underlying chronic illness	In-hospital mortality	George 2015[98]	von Seidlein 2012[110]
ELSHOUT model	Sore throat, palpable lymph nodes, duration of fever before consultation, C-reactive protein	Duration of fever	Elshout 2015[95]	Elshout 2015[95]
FEAST-PET	Axillary temperature, heart rate, capillary refill time, conscious level, respiratory distress, lung crepitations, severe pallor, weak pulse	48h mortality	George 2015[98]	George 2015[98]
FEAST-PETaL	FEAST-PET with the addition of lactate, pH, blood urea nitrogen	48h mortality	George 2015[98]	George 2015[98]
KWIZERA model 1	Age, respiratory rate, heart rate, temperature, capillary refill time, altered mental state	In-hospital mortality	Kwizera 2019[97]	Kwizera 2019[97]
KWIZERA model 2	Age, respiratory rate, heart rate, capillary refill time, altered mental state	In-hospital mortality	Kwizera 2019[97]	Kwizera 2019[97]
KWIZERA model 3	Age, respiratory rate, temperature, capillary refill time, altered mental state	In-hospital mortality	Kwizera 2019[97]	Kwizera 2019[97]
KWIZERA model 4	Age, respiratory rate, capillary refill time, altered mental state	In-hospital mortality	Kwizera 2019[97]	Kwizera 2019[97]
KWIZERA model 5	Age, respiratory rate, altered mental state	In-hospital mortality	Kwizera 2019[97]	Kwizera 2019[97]
LODS	Deep breathing, coma, and prostration	In-hospital mortality	George 2015;[98] Conroy 2015[92]	Helbok 2009[111]
LOWLAAVAR model 1	Conscious level, HIV, weight-for-age z-score	In-hospital mortality	Lowlaavar 2016[93]	Lowlaavar 2016[93]
LOWLAAVAR model 2	Conscious level, HIV, mid-upper arm circumference	In-hospital mortality	Lowlaavar 2016[93]	Lowlaavar 2016[93]

Clinical Prediction Model	Predictor variables used in the original studies	Original outcome	Included study evaluating the model	Original study developing the model
LOWLAAVAR model 3	Conscious level, mid-upper arm circumference	In-hospital mortality	Lowlaavar 2016[93]	Lowlaavar 2016[93]
PEDIA – immediate	Anaemia, jaundice, indrawing, deep breathing, conscious level, prostration, convulsions/seizures, temperature	In-hospital mortality (< 4h)	George 2015[98]	Berkley 2003[112]
PEDIA – early	Jaundice, indrawing, conscious level, prostration, convulsions/seizures, wasting, kwashiorkor*	In-hospital mortality (4-48h)	George 2015;[98] Conroy 2015[92]	Berkley 2003[112]
PEDIA – late	History > 7d, conscious level, prostration, convulsions/seizures, temperature, wasting, kwashiorkor	In-hospital mortality (> 48h)	George 2015[98]	Berkley 2003[112]
PEWS†	Heart rate, capillary refill time, respiratory rate, oxygen saturation, systolic blood pressure	PICU admission (immediate)	George 2015[98]	Parshuram 2009[113]
PRISM III‡	Heart rate, temperature, conscious level, systolic blood pressure, glucose, potassium, PCO_2, pH, acidosis, pupillary reflexes	In-PICU mortality	George 2015[98]	Pollack, 1996[114]
qPELOD-2	Systolic or mean arterial pressure, heart rate, altered mentation	In-PICU mortality	van Nassau 2018[106]	Leclerc 2017[115]
qSOFA	Respiratory rate, altered mentation, systolic blood pressure	In-hospital mortality	van Nassau 2018[106]	Seymour 2016[116]
qSOFA-lactate	qSOFA with the addition of lactate	In-hospital mortality and/or PICU transfer	van Nassau 2018[106]	van Nassau 2018[106]
SCOTT model	Systolic blood pressure, diastolic blood pressure, temperature, age, respiratory rate, heart rate, arrival via emergency medical services, oncological comorbidity, indwelling central line on arrival, hospitalised in the last year	Hypotensive septic shock (24h)	Scott 2020[108]	Scott 2020[108]
SICK	Level of consciousness, temperature, heart rate, respiratory rate, systolic blood pressure, SpO_2, capillary refill time, age	In-hospital mortality	Conroy 2015[92]	Kumar 2003[117]
SIRS	Heart rate, respiratory rate, leukocyte count, temperature	NA (expert consensus)	van Nassau 2018[106]	Goldstein 2005[109]
YOS	Quality of cry, reaction to parent stimulation, state variation, colour, hydration, response to social overtures	Serious illness+	Walia 2016[107]	McCarthy 1982[118]

3.3.4 Quality assessment

Only one prognostic factor study was at low risk of bias,[102] while another was judged to be at low risk of bias in all but one domain.[91] The domains at highest risk of bias were: study confounding, related to omission of important covariates; study participants, often due to requirement for the measurement of specific laboratory parameters (for example, leukocyte count); and statistical analysis, as a result of inadequate reporting or inappropriate exclusion of participants from the analysis (Figure 3.3-7).

Each clinical prediction model/outcome pair was assessed independently and all judged to be at high risk of bias (Figure 3.3-7). Most often this was due to inadequate reporting of model performance (studies reporting discrimination but not calibration), circularity between predictors and outcomes, or having fewer than 100 participants with severe outcomes for model validation. It is noteworthy that one study which externally validated three models included 99 children who died.[92] Another study which derived and/or validated nine models undertook an additional external validation in a population of acutely unwell but non-febrile children (and thus not eligible for consideration in this review), which included more than 100 children who died.[98]

In all but one study there was high concern regarding applicability to the review question.[108] This was largely due to the majority of studies including only children requiring hospitalisation, with recruitment occurring after the decision to admit had been made by the treating physician. Full details on risk of bias and applicability assessments are provided in Table 3.3-3 and Table 3.3-4.

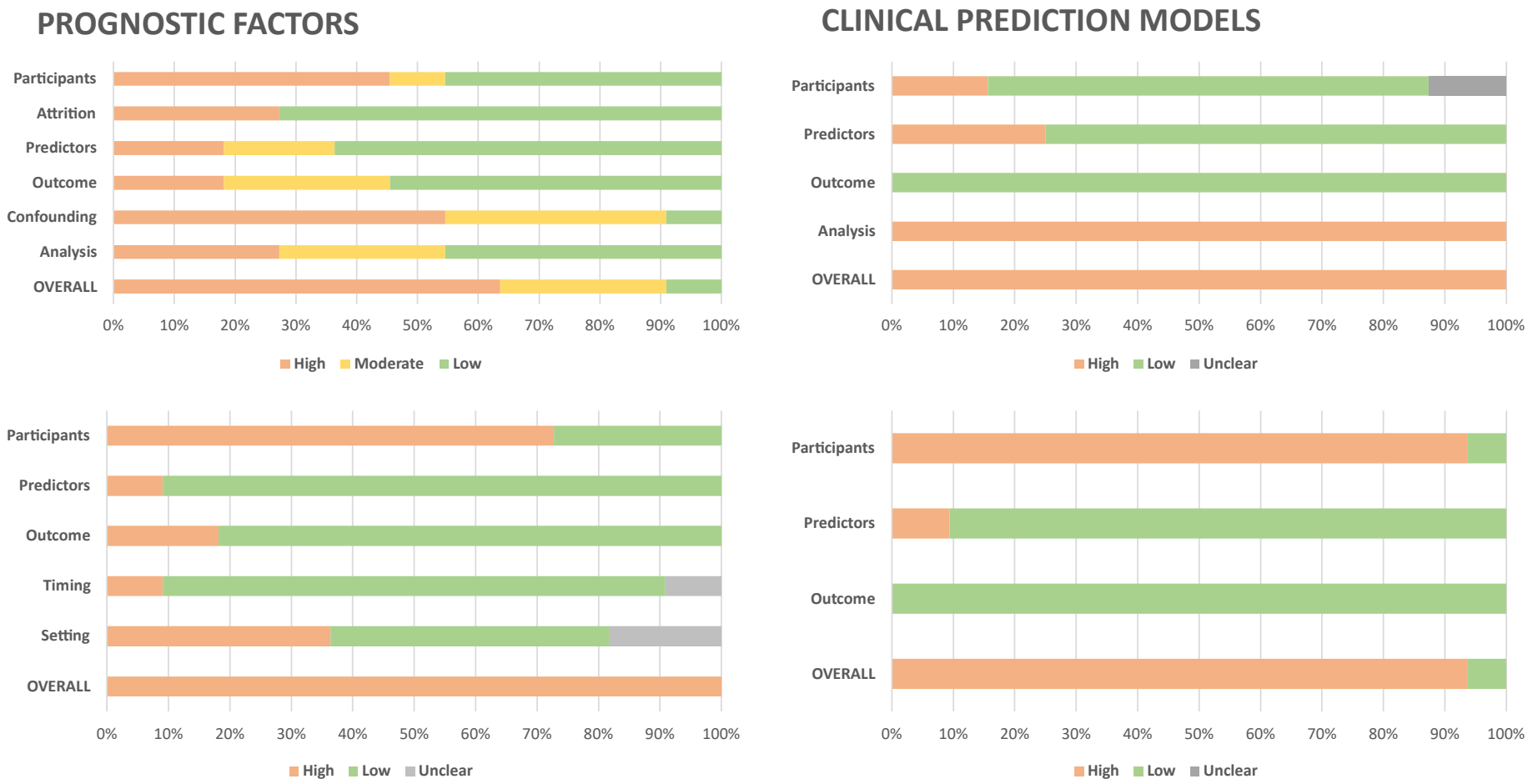


Figure 3.3-7: Overall risk of bias and applicability of included studies. Left panel: prognostic factor studies (n = 11) assessed using the QUIPS tool: risk of bias (top) and applicability concern (bottom). Right panel: clinical prediction model studies (n = 32 model/outcome pairs) assessed using PROBAST: risk of bias (top) and applicability concern (bottom). The study by Elshout et al. was primarily a prognostic factor study and was assessed using the QUIPS tool. Red = high risk of bias / applicability concern, orange = moderate risk of bias, green = low risk of bias / applicability concern, grey = unclear applicability concern. PROBAST = Prediction model Risk Of Bias ASsessment Tool; QUIPS = Quality In Prognosis Studies.

Table 3.3-3: Risk of bias and applicability of prognostic factor studies. Eleven prognostic factor studies assessed using the QUIPS tool. The study by Elshout et al. was primarily a prognostic factor study and was assessed using the QUIPS tool. H = high risk/concern; L = low risk/concern; M = moderate risk; QUIPS = Quality in Prognosis Studies; U = unclear concern.

Study	Risk of Bias							Applicability Concern					
	Overall	Analysis	Confounding	Outcome	Predictors	Attrition	Participants	Overall	Setting	Timing	Outcome	Predictors	Participants
Elshout 2015	H	M	M	H	M	H	H	H	L	L	H	L	H
Scott 2012	H	L	H	M	L	L	H	H	L	H	L	L	H
Scott 2014	H	L	H	H	L	L	L	H	L	L	L	L	H
Scott 2017	L	L	L	L	L	L	L	H	L	L	L	L	H
Freyne 2013	H	H	H	M	L	L	H	H	U	L	L	L	H
Mtove 2011	M	L	M	L	L	L	L	H	H	L	L	L	L
Nadjm 2013	M	M	M	L	L	L	L	H	H	L	L	L	L
Aramburo 2018	M	L	M	L	L	L	L	H	H	L	L	L	H
Costa 2017	H	H	H	L	H	L	H	H	U	U	L	L	H
Mwandama 2016	H	M	H	M	M	H	H	H	L	L	H	L	H
SEAIDCRN 2017	H	H	H	L	H	H	M	H	H	L	L	H	L

Table 3.3-4: Risk of bias and applicability of clinical prediction model studies. Seven clinical prediction model studies assessed using PROBAST. Each prediction model/outcome pair (n = 32) is assessed independently. AQUAMAT = African Quinine Artesunate Malaria Trial; D = derivation; FEAST-PET = FEAST-Paediatric Emergency Triage; FEAST-PETaL = FEAST-Paediatric Emergency Triage and Laboratory; H = high risk/concern; L = low risk/concern; LODS = Lambaréné Organ Dysfunction Score; PEDIA = Pediatric Early Death Index for Africa; PEWS = Pediatric Early Warning Score; PICU = paediatric intensive care unit; PRISM III = Pediatric Risk of Mortality; PROBAST = Prediction model Risk Of Bias ASsessment Tool; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2; qSOFA = quick Sequential Organ Failure Assessment; SICK = Signs of Inflammation in Children that Kill; SIRS = Systemic Inflammatory Response Syndrome; U = unclear concern; V = validation; YOS = Yale Observation Score.

Study	Clinical prediction model	Outcome	Risk of Bias					Applicability Concern			
			Overall	Analysis	Outcome	Predictors	Participants	Overall	Outcome	Predictors	Participants
George 2015	FEAST-PET (D)	In-hospital mortality (48h)	H	H	L	L	L	H	L	L	H
George 2015	FEAST-PETaL (D)	In-hospital mortality (48h)	H	H	L	L	L	H	L	L	H
George 2015	LODS (D)	In-hospital mortality (48h)	H	H	L	L	L	H	L	L	H
George 2015	PEDIA-i (V)	In-hospital mortality (<4h)	H	H	L	L	L	H	L	L	H
George 2015	PEDIA-e (V)	In-hospital mortality (4-48h)	H	H	L	L	L	H	L	L	H
George 2015	PEDIA-l (V)	In-hospital mortality (>48h)	H	H	L	L	L	H	L	L	H
George 2015	PRISM (V)	In-hospital mortality (48h)	H	H	L	L	L	H	L	L	H
George 2015	PEWS (V)	In-hospital mortality (48h)	H	H	L	L	L	H	L	L	H
George 2015	AQUAMAT (V)	In-hospital mortality (48h)	H	H	L	L	L	H	L	L	H
Conroy 2015	LODS (V)	In-hospital mortality	H	H	L	L	L	H	L	L	H
Conroy 2015	SICK (V)	In-hospital mortality	H	H	L	L	L	H	L	L	H
Conroy 2015	PEDIA-i (V)	In-hospital mortality	H	H	L	L	L	H	L	L	H
Lowlaavar 2016	Model 1 (D)	In-hospital mortality	H	H	L	H	L	H	L	H	H
Lowlaavar 2016	Model 2 (D)	In-hospital mortality	H	H	L	H	L	H	L	H	H
Lowlaavar 2016	Model 3 (D)	In-hospital mortality	H	H	L	H	L	H	L	H	H

Study	Clinical prediction model	Outcome	Risk of Bias					Applicability Concern			
			Overall	Analysis	Outcome	Predictors	Participants	Overall	Outcome	Predictors	Participants
Walia 2016	YOS (V)	Mortality	H	H	L	L	U	H	L	L	H
Walia 2016	YOS (V)	Mechanical ventilation	H	H	L	L	U	H	L	L	H
van Nassau 2018	qSOFA (V)	PICU transfer and/or in-hospital mortality	H	H	L	L	L	H	L	L	H
van Nassau 2018	qPELOD-2 (V)	PICU transfer and/or in-hospital mortality	H	H	L	L	L	H	L	L	H
van Nassau 2018	SIRS (V)	PICU transfer and/or in-hospital mortality	H	H	L	L	L	H	L	L	H
van Nassau 2018	qSOFA-lactate (V)	PICU transfer and/or in-hospital mortality	H	H	L	L	L	H	L	L	H
van Nassau 2018	qSOFA (V)	Length of stay ≥ 7d	H	H	L	L	L	H	L	L	H
van Nassau 2018	qPELOD-2 (V)	Length of stay ≥ 7d	H	H	L	L	L	H	L	L	H
van Nassau 2018	SIRS (V)	Length of stay ≥ 7d	H	H	L	L	L	H	L	L	H
van Nassau 2018	qSOFA-lactate (V)	Length of stay ≥ 7d	H	H	L	L	L	H	L	L	H
Kwizera 2019	Model 1 (D)	In-hospital mortality	H	H	L	H	H	H	L	L	H
Kwizera 2019	Model 2 (D)	In-hospital mortality	H	H	L	H	H	H	L	L	H
Kwizera 2019	Model 3 (D)	In-hospital mortality	H	H	L	H	H	H	L	L	H
Kwizera 2019	Model 4 (D)	In-hospital mortality	H	H	L	H	H	H	L	L	H
Kwizera 2019	Model 5 (D)	In-hospital mortality	H	H	L	H	H	H	L	L	H
Scott 2020	Temporal (V)	Hypotensive shock ≤ 24h	H	H	H	L	H	L	L	L	L
Scott 2020	Geographic (V)	Hypotensive shock ≤ 24h	H	H	H	L	H	L	L	L	L

3.4 Discussion

This systematic review of prognostic factors and clinical prediction models assessing severity of disease in febrile children highlights that few well-conducted studies address this important public health question, particularly in unselected children presenting from the community. One of its main strengths is the inclusion of studies from a wide geographic context, aiding understanding of how predictive performance can vary across settings. By focusing on prognostic studies, features that predict the likelihood that a child's illness might progress were identified, rather than features associated with illness severity at the moment of assessment.

Most prognostic factors identified as valuable for predicting severe childhood febrile illness (PLRs ≥ 5.0) overlapped with individual components of the most promising clinical prediction models (AUCs ≥ 0.80): nutritional and HIV status, hypoxia, altered consciousness, and markers of acidosis (raised venous lactate or deep breathing) and poor peripheral perfusion (weak pulse, limb-core temperature gradient, or prolonged CRT).[92,93,98,99,101,103,106] Hypoglycaemia was a useful prognostic factor identified in this review, but omitted in most clinical prediction models. Many of these features, however, indicate a child that is already very unwell, reflecting the fact that most studies included only hospitalised children and focused on predicting mortality.

3.4.1 Prognostic factors

The majority of prognostic factors identified as having diagnostic utility were 'red flags' (PLRs ≥ 5.0) with few able to adequately rule-out the possibility of progression to severe disease (NLRs ≤ 0.2). This finding is consistent with a previous systematic review evaluating the diagnostic utility of clinical features for serious bacterial infections,[72] and is not surprising as it is highly unlikely that absence of a single feature can accurately rule-out serious disease. Ruling out serious disease requires a combinatorial approach and is a strong rationale for the development of multivariate clinical prediction models. Rule-out tests or algorithms may be particularly valuable for presentations as common as

acute febrile illnesses where specificity of risk stratification tools is crucial, especially in under-resourced health systems which are vulnerable to overburdening.

Performance of prognostic factors varied across settings. The importance of considering context (both the resources available and population studied) when interpreting performance of different prognostic factors cannot be overstated. The results for venous lactate (Figure 3.3-2) typify this. It is striking that the PLR of a raised lactate for predicting in-hospital mortality in Tanzania was similar to that for predicting 24-hour organ dysfunction in the USA;[99,104] presumably if the rule-in ability of an elevated lactate had been judged against mortality in the USA performance would have appeared far inferior. Similarly, even in settings with comparable resources, differences in the studied populations influence predictor performance: amongst febrile Tanzanian children with moderate prevalence of severe disease lactate appears to be a useful rule-in test (PLR = 5.13; 95% CI = 4.34 to 6.08),[99] whereas amongst critically ill children from the same region (febrile illnesses accompanied by poor organ perfusion) lactate appears to have most value as a rule-out test (NLR = 0.13; 95% CI = 0.09 to 0.19).[91]

Only 30/200 (15%) prognostic factors met the pre-specified threshold for clinical relevance (PLR ≥ 5.0 or NLR ≤ 0.2). This may reflect the difficulty of identifying parsimonious predictors for all febrile children. While common pathophysiological pathways for severe disease have been identified across a spectrum of microbial aetiologies,[37,119] certain predictors may perform better for specific syndromes or pathogens, compared to all-cause febrile illness. Five studies in this review reported a high proportion of children as being either slide- or rapid diagnostic test-positive for malaria. Notwithstanding the issues of co-infection and/or concomitant incidental parasitaemia in settings of high malaria endemicity, it is possible that the findings of these studies are more pertinent to children with malaria. However, four of these studies compared the prognostic performances of hyperlactataemia, hypoglycaemia, and the prediction models SICK, LODS, and PEDIA, and found them to be broadly consistent between children with malaria, non-malarial fever, and invasive bacterial

disease.[91,92,99,101] Furthermore, as can be seen in Figure 3.3-2, Figure 3.3-3, and Figure 3.3-4, a number of predictors identified in malaria endemic regions, also demonstrated prognostic utility in contexts where malaria is not endemic (for example, venous lactate, impaired peripheral perfusion, hypotension, and altered consciousness). This, in conjunction with the subgroup analyses performed in the original studies, provides reassurance that the prognostic factors identified are generalisable across different infecting pathogens. Nonetheless, future reviews using search strategies developed to retrieve syndrome- or pathogen-specific studies are needed.

Another potential explanation for the relatively few valuable prognostic factors identified is work-up bias. In most studies predictors were available to the treating clinicians: abnormal values are likely to have been acted upon and, at least for studies aiming to predict mortality, predictive performance underestimated. For most predictors, particularly clinical signs, this is unavoidable as blinding is often neither possible nor ethical. It is thus important to qualify performance of a particular predictor with the context within which it was assessed. When feasible, randomisation is required to definitively assess their potential impact.[120] This is particularly important for new tests proposed in resource-limited settings. For example, both lactate and hypoxia were identified as potentially of value in this review but introducing tests for these parameters at all first-line health facilities would incur substantial cost, and as their predictive value may vary in different settings, could result in unnecessary or missed referrals. Randomisation can help determine whether new tests such as these add value to simple clinical assessment.[121]

3.4.2 Clinical prediction models

Clinical prediction models performed better when derived and validated in similar populations:[92] in East Africa LODS and PEDIA-e (both derived in sub-Saharan Africa)[111,112] were superior to SICK (originally derived in India).[117] Model performance also improved when predicting the same outcome as the derivation study: the quick Sequential Organ Failure Assessment (qSOFA) and quick

Pediatric Logistic Organ Dysfunction-2 (qPELOD-2) scores, derived to predict mortality, performed poorly when predicting prolonged length of stay.[106,115,116] These findings highlight the importance of deriving prediction models using populations and outcomes appropriate to the clinical question. Whilst mortality is a 'hard' outcome, it seldom occurs in primary care. Furthermore, its reflection of disease severity is influenced (mediated) by the level and quality of available care, as exemplified by the comparable rule-in ability of elevated lactate for two very different outcomes in Tanzania and the USA as described above.[99,104] Rather than relying on models derived in secondary care to generalise to outpatient settings across different epidemiological landscapes, alternative ways to quantify disease severity (for example, receipt of vital organ support or other interventions indicative of disease severity), which consider local context yet avoid circularity between predictor variables and outcome definitions, will be important to develop risk prediction tools for community-based use-cases. Finally, the fact that most studies summarised model performance using only the AUC means that it is difficult to appreciate what the impact might be on clinical decision making.[64,65]

3.4.3 Limitations

The major limitation of this review arises from the heterogeneity of studies, which precludes a meta-analysis or quantitative comparisons of effect estimates. Secondly, it is difficult to determine if studies included children presenting to first-line health workers. Studies were not excluded solely based on the designated 'level' of a health facility: concerned caregivers in all settings use primary, secondary, and tertiary care facilities as their first point-of-access and children's illness journeys in many LMICs are often unpredictable and convoluted.[122] Thirdly, most studies included only hospitalised children. This is a major barrier to understanding the potential for prognostic factors and prediction models to guide referral or admission decisions for children presenting from the community. Follow-up of children assessed as 'low risk' (i.e., those managed in the community) must be a priority for future studies seeking to determine the validity of prognostic factors and prediction

models in outpatient settings.[123] Fourthly, in line with other reviews most studies were found to be of low quality.[60] Recent guidelines on the reporting of prediction research may help address this.[70] Finally, the review was framed around 'febrile illness', rather than, for example, 'clinically-suspected infection'. The rationale was to ensure the findings were as relevant as possible for lesser-trained community healthcare providers working in resource-constrained settings, for whom a presumptive diagnosis of suspected infection can be challenging. Febrile illness is an accepted 'pragmatic point-of-entry' in these settings,[32] however, despite the deliberately broad definition of febrile illness used in this review (documented abnormal temperature and history of fever), and the inclusion of studies of children with 'suspected sepsis', relevant studies may still have been missed.

Acknowledging that most deaths amongst hospitalised children in LMICs occur within the first 48 hours of admission,[124] few studies aimed to predict outcomes more distant in time after the point of baseline predictor measurements.[95,96,98,100,102,106] Amongst studies using 'hard' outcomes (for which comparison with studies using more proximal outcomes is most valid),[98,102,106] predictive performance generally deteriorated as the time horizon increased, likely reflecting growing complexity of the prediction problem as the clinical question moves from diagnosis to prognosis.

3.4.4 Specific considerations for resource-limited settings

In many LMIC community care contexts certain variables are not feasible to collect,[125] and as noted above, some may incur substantial cost. Interestingly, HIV and nutritional status were both identified in this review and represent the only prognostic factors meeting the pre-specified threshold for clinical relevance that may not necessarily reflect a child that is overtly very unwell. Whilst biological plausibility for the prognostic utility of these two variables is high, it should be noted that the study which identified them was small and correspondingly the confidence interval for the PLR is wide.[97] More studies that approach risk assessment in a holistic manner, including variables reflecting

the background of a child (for example, demographics, nutritional indices, comorbidities) as well as important contextual determinants of outcome (for example, a child's journey to care) are needed.

The World Health Organization's Integrated Management of Childhood Illnesses and integrated Community Case Management guidelines recommend certain 'Danger Signs' to guide referrals from community healthcare providers in resource-constrained settings.[32,126] Of these, only altered consciousness was widely represented amongst included studies, and most found it to be a good predictor of severe disease.[91-93,97-99,101,103,106] History of convulsions and prostration were each examined in two studies whilst "vomiting everything" was not evaluated.[91,92]

3.4.5 Conclusions

Overall, the findings of this systematic review emphasise the limitations of individual prognostic factors. Performance varies widely across settings and clinicians must be cognisant not to over interpret individual predictors. Whilst multivariate prediction models can support clinical decision making, they must be derived and validated using appropriate methodology, as well as populations and outcomes relevant to the clinical problem. For the identification of children at risk of severe febrile illness, this will require multiple, large, collaborative research initiatives across different settings, which collect harmonised data on predictors and outcomes,[127,128] and include unselected children presenting from the community.

3.5 Additional evidence arising since completion of the systematic review

An updated MEDLINE search (8th April 2023) using the identical search strategy as the original systematic review retrieved an additional 952 articles. After screening these articles against the same eligibility criteria, 841 articles were excluded upon review of their titles and/or abstracts, and six would

have been eligible for inclusion had they been available at the time the original systematic review was conducted (Figure 3.5-1).[129-134]

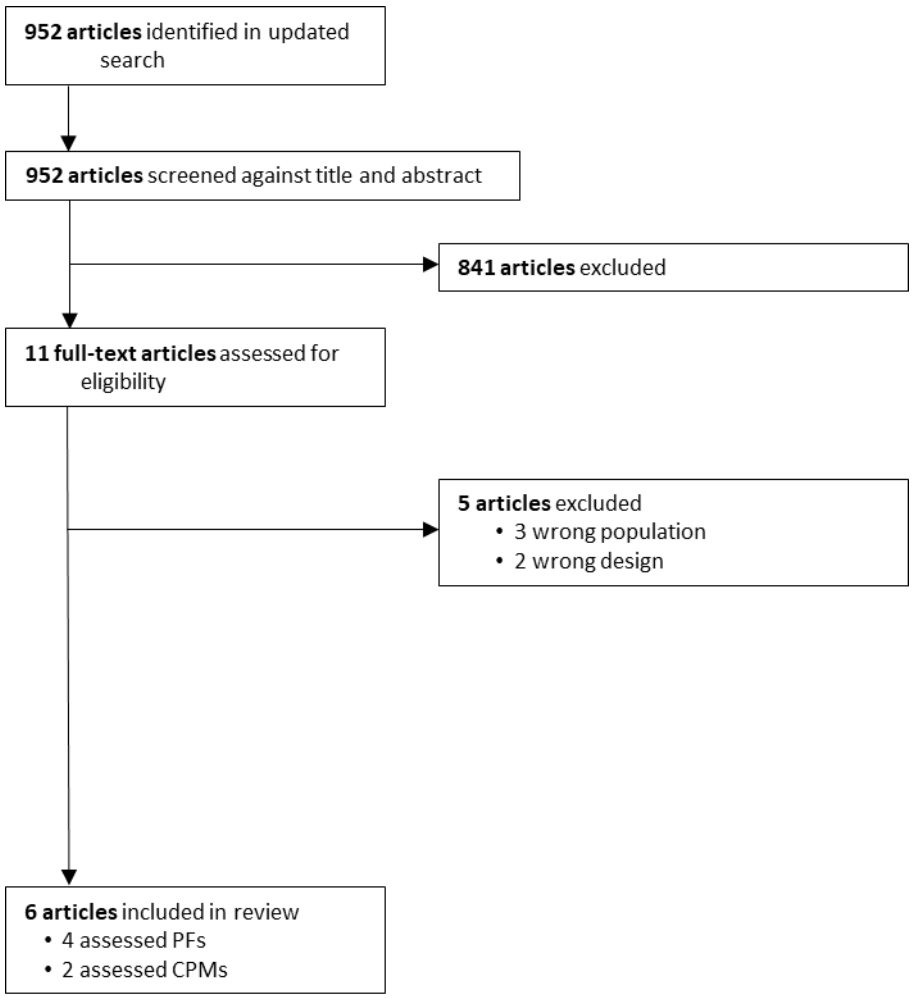


Figure 3.5-1: Additional articles identified by updated MEDLINE search. Only one reason for exclusion per study is listed. CPM = clinical prediction model; PF = prognostic factor**.**

Encouragingly the majority of additional studies (4/6; Table 3.5-1) included children managed in the community,[130-133] although three were from the same centre in the United Kingdom.[130,131,133] Nine clinical severity scores were evaluated, including the external validation of a variety of Paediatric Early Warning Scores (PEWS) and the qSOFA score, as well as the derivation and validation of the modified Liverpool-qSOFA (LqSOFA) score. Whilst PEWS were found to have superior discrimination (AUCs 0.91 to 0.95), work-up bias likely inflated performance, as a PEWS was in place to guide clinical decisions during the studies.[130,133] Previous studies have found that PEWS are impractical for use in

resource-limited settings, partly due to the large number of constituent variables, including systolic blood pressure (SBP).[130] It may be that the simpler LqSOFA score, which contains only four parameters and uses heart rate and CRT in place of SBP, may be a more practical alternative. LqSOFA demonstrated reasonable rule-in ability (PLR = 4.8; 95% CI = 4.3 to 5.4) and discrimination (AUC = 0.81; 95% CI = 0.76 to 0.86), although low sensitivity (0.72; 95% CI = 0.63 to 0.79) may limit potential as a screening tool for unselected febrile children presenting from the community.[133]

Three additional studies evaluated laboratory factors,[129,131,134] including host response biomarkers of immune activation, endothelial injury, and coagulation, which have previously been implicated in final common pathways to severe febrile illness and sepsis.[37,38,135] One of these studies demonstrated encouraging ability of a single marker (soluble triggering receptor expressed on myeloid cells-1 [sTREM-1]) to risk stratify children hospitalised with febrile illnesses in Uganda.[134] However, it is unclear if measurements of sTREM-1 add value to clinical assessment and if these findings apply to unselected febrile children presenting from the community. In this study population, discrimination of sTREM-1 (AUC = 0.90; 95% CI = 0.86 to 0.95) was comparable to LODS (AUC = 0.91 95% CI = 0.88 to 0.95), suggesting that children at risk of deterioration were readily identified using a simple bedside clinical severity score. However, it is noteworthy that the same marker (sTREM-1) has been found to predict mortality amongst febrile adult outpatients in Tanzania,[39] and that other host response biomarkers reflecting endothelial function (for example, mid-regional pro-adrenomedullin [MR-proADM]) appear to be useful for risk stratification of unselected febrile children who were not identified by clinical severity scores.[131]

The findings from these additional studies, along with those from the original systematic review, have been used to inform design and conduct of the prospective and retrospective work that is reported in the following chapters of this thesis.

Table 3.5-1: Characteristics of eligible studies identified by updated MEDLINE search. CPM = clinical prediction model; DIC = disseminated intravascular coagulation; ED = emergency department; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; MR-proADM = mid-regional pro-adrenomedullin; NICU = neonatal intensive care unit; OPD = outpatient department; PEWS = Pediatric Early Warning Score; PF = prognostic factor; PICU = paediatric intensive care unit; sTREM-1 = soluble triggering receptor expressed on myeloid cells-1; UK = United Kingdom; USA = United States of America.

Study (year); Setting, country	Cohort	Design	PF or CPM	Sample size	Population		Outcome	Outcome prevalence (n/N)
					Inclusion criteria	Exclusion criteria		
Slatnick (2021);[129] Tertiary care hospital, USA	Hospitalised	Prospective cohort	Disseminated Intravascular Coagulation	284	Age 60d-18y; clinician-suspected sepsis; undergoing DIC order set as part of routine care	Length of stay < 24h; transfer from another centre; anticoagulation; NICU admission; inotropes prior to enrolment	Vasopressor initiation; Mechanical ventilation; PICU admission; Mortality (30d, 90d, 1y)	32.4% (92/284); 23.2% (66/284); 70.4% (200/284); 2.1% (6/284), 3.5% (10/284), 8.1% (23/284)
Hagedoorn (2022);[132] Tertiary care hospitals, Austria, Germany, Greece, Latvia, Netherlands, Spain, Slovenia, UK	Hospital OPD/ED	Prospective cohort	Shock Index	5,622	Age 0-18y; temperature ≥ 38.0°C or history of fever < 72h; undergoing blood pressure measurement as part of routine care		PICU admission	1.2% (69/5,622)
Lenihan (2022);[131] Tertiary care hospital, UK	Hospital OPD/ED	Prospective cohort	MR-proADM	1,183	Age 0-16y; temperature ≥ 38.0°C or history of fever < 24h; undergoing blood tests as part of routine care	Primary immunodeficiency	PICU admission (48h); fluid resuscitation	4.1% (48/1,183); 12.3% (146/1,183)
Romaine (2021);[130] Tertiary care hospital, UK	Hospital OPD/ED	Retrospective cohort	PEWS	11,449	Age 0-16y; temperature ≥ 38.0°C or history of fever < 72h	Transfer from another hospital; ≥ 2 missing baseline variables	PICU admission (48h); in-hospital sepsis-related mortality	1.2% (134/11,449); 0.04% (5/11,449)
Romaine (2020);[133] Tertiary care hospital, UK	Hospital OPD/ED	Retrospective cohort	LqSOFA	12,241	Age 0-16y; temperature ≥ 38.0°C or history of fever < 72h	Transfer from another hospital; ≥ 2 missing baseline variables	PICU admission (48h); in-hospital sepsis-related mortality	1.1% (135/12,241); 0.04% (5/12,241)
Leligdowicz (2021);[134] Tertiary care hospital, Uganda	Hospitalised	Prospective cohort	sTREM-1	2,502	Age 2m-5y; axillary temperature > 37.5°C or history of fever < 48h	Isolated diarrhoeal illness	In-hospital mortality (7d)	3.8% (95/2,502)

4 Risk stratification of childhood respiratory infections at a medical clinic in a refugee camp on the Thailand-Myanmar border

This chapter is based upon work published in: **Chandna A**, Lubell Y, Mwandigha L, et al. *Defining the role of host biomarkers in the diagnosis and prognosis of the severity of childhood pneumonia – a prospective cohort study*. Sci Rep. Jul 25; 13(1): 12024, 2023; and work under review: **Chandna A**, Mwandigha L, Koshiaris C, et al. *External validation and updating of clinical severity scores to guide referral of young children with acute respiratory infections in resource-limited primary care settings*. medRxiv. 2022 (doi: https://doi.org/10.1101/2022.12.06.22283016).

4.1 Introduction

Acute respiratory infections (ARIs) are the leading reason for unscheduled childhood medical consultations worldwide.[136,137] Primary care workers function as gatekeepers to the formal health system, aiming to distinguish the minority of ARIs requiring onward referral from the vast majority suitable for community-based care.[68]

Existing tools to support community healthcare providers in their assessment of unwell children, such as the World Health Organization's (WHO) Integrated Management of Childhood Illnesses (IMCI) and integrated Community Case Management (iCCM) guidelines, recommend certain 'Danger Signs' to guide referrals.[32,126] However, these lack sensitivity and specificity, and suffer from considerable interobserver variability.[34,138] A systematic review of paediatric triage tools in resource-constrained settings (including IMCI) highlighted concerns with regards reliability and that lack of follow-up data on children managed in the community rendered the validity of existing tools uncertain.[123]

Pneumonia is a common manifestation of lower ARIs and most children with pneumonia can be successfully managed at home.[139,140] In rural locations of many low- and middle-income countries

(LMICs), where accessing hospital-level care may incur substantial cost, community-based care is often preferred by families.[141] However, pneumonia remains the leading cause of death and disability for young children living in LMICs,[142,143] and clear criteria for safe outpatient management are lacking.[140]

A growing body of evidence indicates that final common pathophysiological pathways reflecting endothelial injury and immune activation are shared across a range of infectious diseases,[37,40] including in young children hospitalised with pneumonia.[144-147] Measurements of markers of these pathways improve performance of clinical severity scores,[39,40,148] and consequently they have been proposed as adjuncts to paediatric triage tools.[149] However, although altered microvascular function has been demonstrated in ambulatory children with mild ARIs,[150] it is unknown whether markers of endothelial injury are elevated sufficiently early in the natural history of childhood pneumonia for them to be useful for risk stratification in community settings.

This study identifies paediatric severity scores suitable for use in resource-limited primary care settings and externally validates their ability to guide referral of unselected young children presenting with ARIs.[151] Improvements in performance that might be expected if the scores were deployed as simple clinical prediction models and updated to include variables pertinent to children presenting in rural LMIC settings are characterised. Next, in the subset of children meeting WHO-pneumonia criteria, serum concentrations of biomarkers of endothelial injury, immune activation, and inflammation are quantified, and the value added to the best-performing clinical severity score explored. The hypothesis was that clinical assessment and host response biomarkers would make differential contributions to the diagnosis and prognosis of the severity of childhood pneumonia.

4.2 Methods

4.2.1 Study population

Data were collected during a prospective birth cohort study at a medical clinic for refugees and internally displaced people on the Thailand-Myanmar border.[151] Between September 2007 and September 2008 consecutive pregnant women receiving antenatal care at the clinic were invited to participate in the study. Children of consenting women were reviewed at birth and followed-up each month (routine visit) and during any intercurrent illness (illness visit) until 24 months of age. The local circumstances (inability of the population to move freely out of the camp and lack of other medical providers) contributed to low attrition rates and capture of the majority of acute illnesses for which care was sought.

All ARI illness visits were included in the evaluation of the clinical severity scores. An ARI was defined as (A) a presentation with rhinorrhoea, nasal congestion, cough, respiratory distress (chest indrawing, nasal flaring, grunting, tracheal tug, and/or head bobbing), stridor, and/or abnormal lung auscultation (crepitations and/or wheeze), and (B) a compatible contemporaneous syndromic diagnosis (rhinitis, croup, bronchiolitis, influenza-like illness, pneumonia, viral infection, and/or wheeze) for children sent home directly from the clinic. All illness visits meeting WHO-pneumonia criteria (cough or difficulty breathing associated with age-adjusted tachypnoea) were included in the pneumonia subgroup analyses.[152]

4.2.2 Data collection

All data were measured by study staff and entered onto structured case report forms (Appendix 9.5). With the exception of anthropometric data, all clinical data were collected at the time of presentation. Core (rectal) temperature was measured for neonates and infants and adjusted to axillary temperature by subtracting 0.5°C.[32] Mental status was assessed using the Alert Voice Pain Unresponsive (AVPU) scale. Capillary refill time (CRT) was measured following the release of gentle

pressure on the child's sternum. For children admitted to the clinic, weight was measured at the time of presentation (seca scale; precision ± 5g for neonates or ± 50g after birth). In addition, all children had their mid-upper arm circumference (MUAC), weight, and height measured at each monthly routine visit. For the purposes of these analyses, age-adjusted z-scores (R package: *z scorer*)[153] were calculated using the closest anthropometric data to the index illness visit within the following window periods: height ≤ 28 days; MUAC ≤ 28 days without intervening admission; weight ≤ 14 days without intervening admission. Median time between the illness visit and each anthropometric measurement is reported. Participants were followed-up each day during admission to the clinic and at monthly routine visits conducted as part of the longitudinal birth cohort study.

Serum samples were collected in plain tubes at presentation, centrifuged within two hours (ambient temperature, at 3,000 rpm, for 10 minutes), and stored at 2-8°C. Each day, samples were transported using a cold-chain to the off-site laboratory, aliquoted, and stored at -80°C within 12 hours of collection. Samples collected on Saturday evening or Sunday were transported at the end of the working day on Monday (≤ 48 hours after collection).

4.2.3 Primary and secondary outcomes

The primary outcome was receipt of supplemental oxygen at any time during the illness visit. Clinic treatment protocols specified that peripheral oxygen saturation (SpO_2) must be checked prior to initiation of supplemental oxygen, with therapy only indicated if SpO_2 was < 90%, in line with the WHO definition of severe pneumonia requiring hospital referral.[32] All staff were trained on the treatment protocols prior to study commencement. Study staff were unaware which baseline variables were to be used as candidate predictors at the time of ascertaining outcome status.

To explore the diagnostic vs. prognostic value of the biomarkers, the secondary outcomes were: SpO_2 < 90% at presentation (diagnostic outcome); and then considering only participants who did not meet the diagnostic outcome, receipt of supplemental oxygen during the illness visit

(prognostic outcome 1), and receipt of supplemental oxygen within 28 days of presentation (prognostic outcome 2).

4.2.4 Identification and shortlisting of clinical severity scores

Building on the results of two recent systematic reviews, 16 severity scores that might risk stratify young children presenting from the community with ARIs were longlisted (Table 4.2-1).[154,155] After considering reliability, validity, and feasibility for implementation eight scores that required specialist equipment and/or laboratory tests (for example, a pulse oximeter, leucocyte count, or venous lactate) unlikely to be practical for the assessment of young children in busy LMIC community care settings were excluded.[98,106,109,113,156-159] Four others were excluded as ≥ 25% of the constituent variables were unavailable in the primary dataset.[98,111,112,160]

Two of the remaining scores (quick Sequential Organ Failure Assessment [qSOFA] and quick Pediatric Logistic Organ Dysfunction-2 [qPELOD-2]) contained blood pressure.[115,161] Hypotension is a late sign in paediatric sepsis and not suitable for early recognition of impending serious illness at the community level.[162] Furthermore, accurate use and maintenance of sphygmomanometers and stethoscopes may not be feasible in resource-limited settings.[125] Recently, Romaine et al. replaced systolic blood pressure (SBP) with alternate signs of circulatory compromise (heart rate and CRT) to develop the Liverpool-qSOFA (LqSOFA) score, and demonstrated superior performance compared to qSOFA in febrile children presenting from the community.[133] Hence, the LqSOFA score (in preference to qSOFA) and an adapted qPELOD-2 score (replacing SBP with CRT and assessing mental status using the simpler AVPU scale rather than the Glasgow Coma Scale [GCS]) were selected. Ultimately, the three scores shortlisted for external validation were LqSOFA, qPELOD-2, and the modified Systemic Inflammatory Response Syndrome (mSIRS). The scores are presented and comparisons between the original derivation settings and study population detailed in Table 4.2-2.[115,133,163]

Table 4.2-1: Longlisted severity scores with reasons for exclusion for those not selected. AVPU = Alert Voice Pain Unresponsive scale; FEAST-PET = Fluid Expansion as Supportive Therapy-Paediatric Emergency Triage; FEAST-PETaL = FEAST-PET and Laboratory; GCS = Glasgow Coma Scale; ITAT = Inpatient Triage and Treatment; LMIC = low- and middle-income country; LODS = Lambaréné Organ Dysfunction Score; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; mRISC = modified Respiratory Index of Severity in Children; mSIRS = modified Systemic Inflammatory Response Syndrome; PAWS = Paediatric Advanced Warning Score; PEDIA = Paediatric Early Death Index for Africa; PEWS = Paediatric Early Warning System; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2; qSOFA = quick Sequential Organ Failure Assessment; RISC = Respiratory Index of Severity in Children; SIRS = Systemic Inflammatory Response Syndrome.

Score	Constituent variables	Included / Excluded
FEAST-PET[98]	Heart rate, temperature, capillary refill time, respiratory distress, lung crepitations, pulse character, prostration, pallor	Excluded – pulse character and prostration not available in primary dataset
FEAST-PETaL[98]	Heart rate, temperature, capillary refill time, respiratory distress, lung crepitations, pulse character, prostration, pallor, lactate, pH, blood urea nitrogen	Excluded – large number of laboratory parameters not appropriate for high-throughput LMIC primary care settings and not available in primary dataset; pulse character and prostration not available in primary dataset
ITAT[156]	Heart rate, temperature, respiratory rate, oxygen saturation	Excluded – pulse oximetry not feasible for young infants in high-throughput LMIC primary care settings
LODS[111]	Deep breathing, coma, prostration	Excluded – deep breathing and prostration not available in primary dataset
LqSOFA[133]	Heart rate, respiratory rate, capillary refill time, mental status	**Included**
mRISC[160]	Chest indrawing, prostration, weight-for-age z-score, dehydration, history of unconsciousness, night sweats, mental status, lab-confirmed malaria	Excluded – prostration, history of unconsciousness, and night sweats not available in primary dataset
mSIRS[163]	Heart rate, respiratory rate, temperature	**Included**
PAWS[159]	Heart rate, respiratory rate, temperature, capillary refill time, mental status, oxygen saturation, respiratory distress	Excluded – pulse oximetry not feasible for young infants in high-throughput LMIC primary care settings

Score	Constituent variables	Included / Excluded
PEDIA(s)[112]	Anaemia, jaundice, respiratory distress, deep breathing, mental status, prostration, seizures, temperature, wasting, kwashiorkor, symptom duration	Excluded – deep breathing, prostration, and seizures not available in primary dataset
PEWS[113]	Heart rate, respiratory rate, systolic blood pressure, capillary refill time, oxygen saturation, supplemental oxygen, respiratory distress	Excluded – supplemental oxygen therapy not relevant for primary care-based score; pulse oximetry not feasible for young infants in high-throughput LMIC primary care settings
qPELOD-2[115]	Heart rate, blood pressure, mental status	**Included** – blood pressure replaced with capillary refill time; mental status assessed using AVPU instead of GCS
qSOFA[161]	Respiratory rate, systolic blood pressure, mental status	Excluded – LqSOFA selected instead as capillary refill time and heart rate more appropriate for circulatory assessment in primary care
qSOFA-lactate[106]	Respiratory rate, systolic blood pressure, mental status, lactate	Excluded – lactate not feasible for high-throughput LMIC primary care settings and not available in primary database
RISC[158]	Oxygen saturation, chest indrawing, wheezing, prostration, weight-for-age z-score	Excluded – pulse oximetry not feasible for young infants in high-throughput LMIC primary care settings; prostration not available in primary dataset
RISC-Malawi[157]	Oxygen saturation, mid-upper arm circumference, sex, wheeze, mental status	Excluded – pulse oximetry not feasible for young infants in high-throughput LMIC primary care settings
SIRS[109]	Heart rate, respiratory rate, temperature, leukocyte count	Excluded – leukocyte count not feasible for high-throughput LMIC primary care settings; mSIRS selected instead

Table 4.2-2: Severity scores shortlisted for external validation. *Age-adjusted thresholds for heart rate and respiratory rate from each of the original derivation studies were used to calculate the scores in the study dataset. AVPU = Alert Voice Pain Unresponsive scale; bpm = beats / breaths per minute; ED = emergency department; GCS = Glasgow Coma Scale; ICU = intensive care unit; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; mSIRS = modified Systemic Inflammatory Response Syndrome; PICU = paediatric intensive care unit; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2.

Score	Constituent variables	Population	Outcome
LqSOFA[133]	1. Capillary refill time > 2 seconds 2. Mental status < alert on AVPU 3. Heart rate > age-adjusted threshold 4. Respiratory rate > age-adjusted threshold *Each variable allocated one point to give score of 0-4*	Derivation: 1,121 febrile children < 16y attending the ED and requiring a blood test at a specialist paediatric hospital in the United Kingdom Validation: 12,241 febrile children < 16y attending the ED at a specialist paediatric hospital in the United Kingdom	Critical care admission within 48h of ED attendance Prevalence: 4.2% (derivation) and 1.1% (validation)
mSIRS[163]	1. Core temperature > 38.5°C or < 36°C 2. Heart rate > or < age-adjusted threshold 3. Respiratory rate > age-adjusted threshold *Each variable allocated one point to give score of 0-3*	Derivation: expert consensus (original SIRS score)[109] Validation: 1,184 adults > 18y admitted to a hospital in Sri Lanka with suspected infection	In-hospital mortality, cardiac arrest or ICU admission (validation) Prevalence: 3.6% (validation)
qPELOD-2[115]	1. Mental status < 11 on GCS 2. Heart rate > age-adjusted threshold 3. Blood pressure < age-adjusted threshold *Each variable allocated one point to give score of 0-3*	Derivation: 862 children < 18y admitted to nine European PICUs with suspected infection Validation: 545 children < 18y admitted to a hospital in the Netherlands with suspected bacterial infection[106]	In-PICU mortality (derivation) or PICU admission and/or mortality (validation) Prevalence: 7.0% (derivation) and 3.3% (validation)
This study	1. Capillary refill time > 2 seconds 2. Mental status < alert on AVPU 3. Heart rate > age-adjusted threshold* 4. Respiratory rate > age-adjusted threshold* 5. Axillary temperature > 38°C or < 35.5°C	3,010 ARI presentations from 756 children < 2y presenting to a primary care clinic on the Thailand-Myanmar border	Supplemental oxygen therapy Prevalence: 3.5%

4.2.5 Selection of variables for score updating

The existing points-based scores were converted to clinical prediction models and additional variables relevant for children presenting with ARIs in LMIC primary care settings were considered for inclusion. Nutritional status (weight-for-age z-score [WAZ]) and presence of respiratory distress were selected *a priori* for model updating, after considering resource constraints, reliability, validity, biological plausibility, availability of data in the primary dataset, and sample size (Table 4.2-3).[125] Birthweight, gestational age, and chest auscultation were excluded due to concerns regarding reliability, validity, and/or feasibility.

Table 4.2-3: Variables available in the primary dataset considered for model updating. Nutritional status and respiratory distress were selected for model updating. LMIC = low- and middle-income country; MUAC = mid-upper arm circumference; WAZ = weight-for-age z-score.

Variable	Reliability, validity, and feasibility considerations
Birthweight	Validity concerns with regards measurement at time of birth. Reliability concerns with regards recall by caregivers at time of presentation.
Chest auscultation	Reliability concerns with regards capacity of limited-skill primary care providers for auscultation. Feasibility concerns with regards maintenance of stethoscopes.
Gestational age	Validity concerns with regards measurement at time of birth. Reliability concerns with regards recall by caregivers at time of presentation.
Nutritional status	Feasibility and reliability concerns with regards accurate measurement of length (length-for-age and weight-for-length). Validity concerns with regards MUAC, particularly in children < 3 months old. WAZ chosen as compromise between feasibility, reliability, and validity to reflect nutritional status.
Respiratory distress	Feasibility and validity concerns with regards inclusion of multiple different measures of respiratory distress. Presence of respiratory distress chosen as binary variable for updating of score.

4.2.6 Identification and selection of biomarkers

Host biomarkers were selected for the pneumonia subgroup analyses following review of the literature and expert consultation (Table 4.2-4). A range of viral and bacterial pathogens commonly cause pneumonia in children and it is not possible to obtain a microbiological diagnosis in the vast majority of cases presenting to primary care. Acknowledging this and recognising therefore that

clinically-useful biomarkers would need to be predictive across a spectrum of infecting organisms, biomarkers implicated in 'pathogen agnostic' final common pathways to severe febrile illness and sepsis, including those reflecting endothelial injury (angiopoietin-1 [Ang-1], angiopoietin-2 [Ang-2], and soluble fms-like tyrosine kinase-1 [sFlt-1; sVEGFR-1]) and immune activation (chitinase-3-like protein-1 [CHI3L1], interferon-gamma-inducible protein-10 [IP-10; CXCL-10], interleukin-1 receptor antagonist [IL-1ra], interleukin-6 [IL-6], interleukin-8 [IL-8], interleukin-10 [IL-10], soluble tumour necrosis factor receptor-1 [sTNFR-1], and soluble triggering receptor expressed on myeloid cells-1 [sTREM-1]) were prioritised.[37,39-41,134,144-146,148,164-166] Two acute phase proteins (C-reactive protein [CRP] and procalcitonin [PCT]) were also included. Although previous studies had found CRP and PCT to be of modest utility for predicting the severity of childhood pneumonia,[167] they are measurable using inexpensive commercially-available rapid tests and are familiar to many clinicians.

Table 4.2-4: Host biomarkers of endothelial injury, immune activation, and inflammation selected for evaluation. Ang-1 = angiopoietin-1; Ang-2 = angiopoietin-2; CHI3L1 = chitinase-3-like protein-1; CRP = C-reactive protein; CXCL-10 = C-X-C motif chemokine-10; IL-1ra = interleukin-1 receptor antagonist; IL-6 interleukin-6; IL-8 = interleukin-8; IL-10 – interleukin-10; IP-10 = interferon-y induced protein-10; PCT = procalcitonin; SBI = serious bacterial infection; SSA = sub-Saharan Africa; sFlt-1 = soluble fms-like tyrosine kinase-1; sTREM-1 = soluble triggering receptor expressed on myeloid cells-1; sTNFR-1 = soluble tumour necrosis factor receptor-1; sVEGFR-1 = soluble vascular endothelial growth factor receptor-1.

Biomarker	Overview of supportive data
Endothelial injury	
Ang-1 and -2	Supportive data from Asia/SSA/Europe in children/adults, that increases in Ang-2, decreases in Ang-1, and/or the Ang-2:1 ratio predicts mortality in pneumonia, malaria, SBI, and all-cause febrile illnesses.[37,39,41,134,144-146,148,164,165]
sFlt-1 / sVEGFR-1	Supportive data from SSA that increases in sFlt-1 predict mortality in children hospitalised with pneumonia, severe malaria, and all-cause febrile illnesses, and adults with all-cause febrile illnesses.[39,41,134,146]
Immune activation	
CHI3L1	Supportive data from SSA that increases in CHI3L1 predict mortality in children hospitalised with pneumonia and all-cause febrile illnesses, and adults with all-cause febrile illnesses.[39,134,146]
IL-1ra	Supportive data that increases in IL-1ra are associated with severity in children with meningococcal disease, adults with SARS-CoV-2 infection, and predict need for longer antibiotic duration in children with febrile lower respiratory tract infections.[166,168,169]

Biomarker	Overview of supportive data
Biomarker	**Overview of supportive data**
IL-6	Supportive data from India that increases in IL-6 are predictive of mortality in children with dengue;[170] in Switzerland, supportive data that increases in IL-6 predict need for longer antibiotic duration in children with febrile lower respiratory tract infection, and disease severity in adults with SARS-CoV-2 infection.[166,171]
IL-8	Supportive data from India that increases in IL-8 predict mortality in children with dengue;[170] in the UK, supportive data that increases in IL-8 predict disease severity in children with meningococcal disease.[168]
IL-10	Supportive data from India that increases in IL-10 predict of mortality in children with dengue.[170]
IP-10 / CXCL-10	Supportive data from Uganda that increases in IP-10 predict mortality in children hospitalised with severe malaria.[41]
sTNFR-1	Supportive data from SSA that increases in sTNFR-1 predict mortality in children hospitalised with pneumonia and all-cause febrile illnesses, and adults with all-cause febrile illnesses.[39,134,146]
sTREM-1	Supportive data from SSA that increases in sTREM-1 predict mortality in children hospitalised with pneumonia, severe malaria, and all-cause febrile illnesses, and adults with all-cause febrile illnesses;[39,41,134,146] in Asia, increased sTREM-1 predicted length of stay in infant febrile illness and in-hospital mortality in adults hospitalised with infection and children hospitalised with pneumonia.[40,145,172]
Inflammation	
PCT	Supportive evidence that increases in PCT predict severe illness in hospitalised children with suspected bacterial infections or meningococcal disease.[173,174]
CRP	Although there is limited supportive evidence for the use of CRP as a prognostic marker for disease severity,[167] as it is the most widely studied biomarker in the region, and numerous point-of-care tests already exist, further evaluation is warranted.

4.2.7 Laboratory procedures

Frozen serum aliquots were transported on dry ice to the research laboratories of the Mahidol-Oxford Tropical Medicine Research Unit in Bangkok, Thailand. Aliquots were thawed overnight and concentrations of host biomarkers were quantified using custom built Simple Plex multi-analyte assays on the Ella microfluidic platform (ProteinSimple, San Jose, California, USA; Appendix 9.6).[175] Analytes below the limit of quantification (LOQ) were assigned a value one-third of the lower limit of the standard curve (Table 4.2-5).

Table 4.2-5: Proportion of host biomarkers below the lower limit of quantification. Analytes below the limit of quantification were assigned a value one-third of the lower limit of the standard curve. Ang-1 = angiopoietin-1; CHI3L1 = chitinase-3-like protein-1; CRP = C-reactive protein; IL-6 = interleukin-6; IL-8 = interleukin-8; IL-10 = interleukin-10; sFlt-1 = soluble fms-like tyrosine kinase-1; sTNFR-1 = soluble tumour necrosis factor receptor-1.

Biomarker	Proportion below limit of quantification % (n/N)
Ang-1	0.2% (2/900)
CHI3L1	0.4% (4/900)
CRP	1.4% (13/900)
IL-6	0.8% (7/900)
IL-8	0.1% (1/900)
IL-10	1.2% (11/900)
sFlt-1	0.6% (5/900)
sTNFR-1	0.2% (2/900)

4.2.8 Missing data

Missingness in the ARI and pneumonia cohorts were considered and addressed separately. Among the ARI cohort, 616 presentations were missing data on one or more candidate predictors (616/3,010; 20.5%) with CRT containing the highest proportion of missingness (442/3,010; 14.7%; Table 4.2-6). For external validation of the existing clinical severity scores, under a missing-at-random assumption (Figure 4.2-1), multiple imputation with chained equations (MICE) was used to deal with missing data (R package: *mice*).[176] Analyses were done in each of 100 imputed datasets and results pooled. Variables were selected for inclusion in the imputation model if they were included in any of the final analysis models (explanatory or response variables) or might plausibly be associated with CRT, the variable with the highest proportion of missing data (Table 4.2-7).

Table 4.2-6: Proportion of missing data amongst candidate predictors.

Variable	Missing data % (n/N)
Age	0%
Heart rate	0.3% (9/3,010)
Respiratory rate	0.3% (8/3,010)
Axillary temperature	0.1% (3/3,010)
Capillary refill time	14.7% (442/3,010)
Mental status	0.1% (37/3,010)
Weight-for-age z-score	4.9% (147/3,010)
Respiratory distress	< 0.01% (1/3,010)

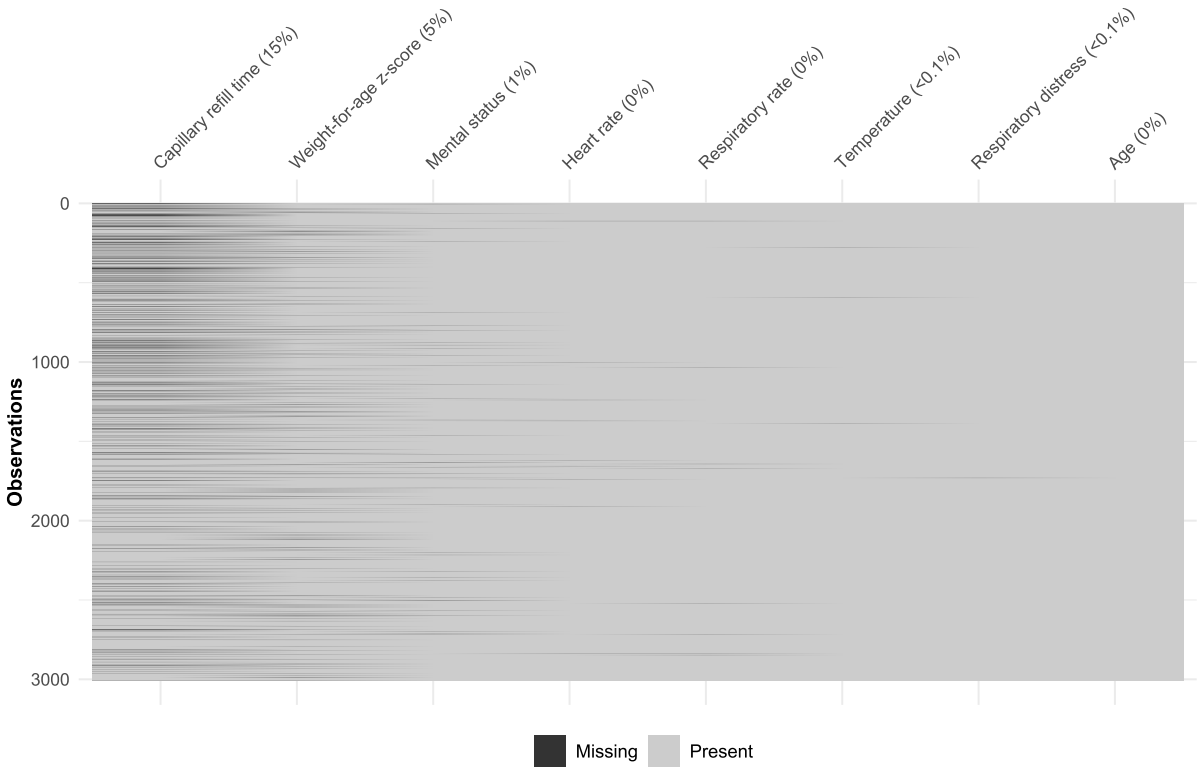


Figure 4.2-1: Number and pattern of missing data amongst candidate predictors. Predictors are ordered by proportion of missing data from left to right.

Table 4.2-7: Variables selected for inclusion in the imputation model. All explanatory and response variables in the final analysis models, along with variables plausibly associated with CRT (the variable with the highest proportion of missingness) were included in the imputation model. CRT = capillary refill time.

Variable	Reason for inclusion in imputation model
Age	Final analysis models (explanatory variable)
Heart rate	Final analysis models (explanatory variable)
Respiratory rate	Final analysis models (explanatory variable)
Axillary temperature	Final analysis models (explanatory variable)
Capillary refill time	Final analysis models (explanatory variable)
Mental status	Final analysis models (explanatory variable)
Birthweight	Plausibly associated with CRT
Gestational age	Plausibly associated with CRT
Known comorbidity	Plausibly associated with CRT
Weight-for-age z-score	Final analysis models (explanatory variable)
Weight-for-length z-score	Plausibly associated with CRT
Length-for-age z-score	Plausibly associated with CRT
Head bobbing	Final analysis models (explanatory variable)
Nasal flaring	Final analysis models (explanatory variable)
Chest indrawing	Final analysis models (explanatory variable)
Tracheal tug	Final analysis models (explanatory variable)
Grunting	Final analysis models (explanatory variable)
Abnormal lung auscultation	Plausibly associated with CRT
Dehydration	Plausibly associated with CRT
Clinic disposition	Plausibly associated with CRT
Receipt of intravenous fluids	Plausibly associated with CRT
Receipt of supplemental oxygen	Final analysis models (response variable)

Prior to model updating, sensitivity analyses were performed to explore the potential impact of different methods for handling missing data. These confirmed that both full-case analyses and single (median) imputation conditional on outcome status produced similar results to MICE (Table 4.2-8). In

order to build the simplest and most practical models, variable selection was included during model updating. To avoid conflicts across multiply imputed datasets (different variables being selected in different imputed datasets), single (median) imputation conditional on outcome status was used to address missing data for model updating.

Table 4.2-8: Discrimination of models when missing data were addressed using different methods. *Pooled discrimination (AUC) across multiply imputed datasets reported. †Optimism-adjusted discrimination reported. $Dataset for full-case analyses created using pairwise deletion, omitting presentations missing data for any explanatory variable or the response variable: LqSOFA = 2,525 presentations with 81 outcome events; mSIRS = 2,992 presentations with 99 outcome events; qPELOD-2 = 2,531 presentations with 83 outcome events. AUC = area under the receiver operating characteristic curve; CI = confidence interval; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; MICE = multiple imputation with chained equations; mSIRS = modified Systemic Inflammatory Response Syndrome; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2.

Method to address missing data	LqSOFA	mSIRS	qPELOD-2
MICE (95% CI)*†	0.90 (0.86 to 0.94)	0.81 (0.76 to 0.86)	0.84 (0.79 to 0.89)
Median imputation (95% CI)†	0.89 (0.84 to 0.94)	0.80 (0.75 to 0.85)	0.83 (0.78 to 0.88)
Full-case analysis (95% CI)†$	0.89 (0.85 to 0.94)	0.81 (0.75 to 0.86)	0.83 (0.78 to 0.88)

Of the 900 pneumonia presentations, 827 (91.9%; 827/900) had complete data for all baseline predictor variables (Table 4.2-9). Hence, due to fewer missing data, single (median) imputation conditional on outcome status was also used to address missingness for these analyses.

Table 4.2-9: Proportion of candidate predictors with missing data for pneumonia subgroup analyses. LqSOFA = Liverpool quick Sequential Organ Failure Assessment.

Variables	Missing data % (n/N)
Heart rate	0.2% (2/900)
Respiratory rate	0.1% (1/900)
Capillary refill time	7.0% (63/900)
Mental status	1.3% (12/900)
LqSOFA score	8.1% (73/900)

4.2.9 Statistical methods

Discrimination and calibration of each clinical severity score was assessed by quantifying the area under the receiver operating characteristic curve (AUC) and plotting the observed proportion of participants that met the primary outcome at each level of a score. Predicted classifications (sensitivity, specificity, negative likelihood ratio [NLR], and positive likelihood ratio [PLR]) were examined at each of the scores' cut-offs.

Prior to model building the relationship between continuous predictors and the primary outcome was explored using locally weighted scatterplot smoothing (LOWESS) to identify non-linear patterns. Accordingly, temperature was modelled using restricted cubic splines (R package: *rms*)[177] with three knots placed at prespecified locations based on percentiles (5^{th} and 95^{th}) and recognised physiological thresholds (36°C).[29,178] Logistic regression was used to derive the models and interactions between age and each of respiratory rate and heart rate were tested for using likelihood ratio tests (LRTs). Random-effects were not modelled as 22% (169/756) of children presented only once. All predictors were prespecified and no predictor selection was performed during model development. Internal validation was performed using 100 bootstrap samples with replacement and optimism-adjusted discrimination and calibration reported (R package: *rms*).[177]

Models were updated by including respiratory distress and WAZ as additional candidate predictors. Penalised (lasso) logistic regression was used for model updating, variable selection, and shrinkage to minimise overfitting (R package: *glmnet*).[179] Discrimination and calibration of the updated models were assessed, predicted classifications examined at clinically-relevant referral thresholds, and the clinical utility (net benefit) of the models compared to the best-performing points-based severity score using decision curve analyses (R package: *dcurves*).[180] Sensitivity analyses were performed: (a) excluding children who were hypoxaemic (SpO_2 < 90%) at the time of presentation; and (b) classifying children sent away from the clinic without admission but who received supplemental oxygen within the next 28 days as meeting the primary outcome.

For the pneumonia subgroup analyses, LOWESS was used to explore the relationship between each biomarker and the primary outcome. Univariable logistic regression was used to quantify the ability of the best-performing points-based severity score from the first phase of the analysis and of each individual biomarker to discriminate (quantified using the AUC) children presenting with pneumonia who required supplemental oxygen during their illness visit.

Discrimination of the clinical score was compared to the biomarkers and to a combinatorial approach, including the clinical score plus one biomarker (R package: *pROC*; DeLong method).[181,182] To reduce the risk of multiple testing, comparisons with the combinatorial clinical-biomarker approach were limited to the five top-performing biomarkers, selected on the basis of their univariate discrimination, after confirming that none of the biomarker concentrations were strongly correlated with baseline clinical severity scores (R package: *polycor*).[183] Decision curve analyses were used to determine the net benefit of including each of the five biomarkers alongside the clinical score across a range of clinically-relevant referral thresholds. Sensitivity analyses were performed excluding serum samples collected at the weekend to assess any potential impact of the delay in sample transfer from 2-8°C to definitive -80°C storage.

To explore the differential contribution of clinical assessment and laboratory tests to the diagnosis and prognosis of pneumonia severity, univariable logistic regression was used to assess the ability of the clinical score and the five top-performing biomarkers to discriminate children who were hypoxaemic at presentation (diagnostic outcome), and to discriminate children who were not initially hypoxaemic but whose disease progressed to require supplemental oxygen in the 28 days following presentation (prognostic outcomes 1 and 2). For the diagnostic outcome, sensitivity analyses were performed assuming that the 139 attendances missing baseline SpO_2 data were not hypoxaemic (SpO_2 < 90%) at presentation.

Finally, recognising that a management strategy requiring measurement of a biomarker in every child presenting with pneumonia may not be practical, recursive partitioning was used to

construct a proof-of-concept algorithm sequentially combining the clinical score and the top-performing biomarker to identify children who might be safe for community-based management (R package: *rpart*).[184] Acknowledging that safety and simplicity were paramount for community triage, a loss-matrix of 10:1 and maximum tree depth of two were prespecified.

All analyses were done in R, version 4.0.2.[89]

4.2.10 Sample size

Of the 3,010 eligible ARI presentations, 104 met the primary outcome, ensuring sufficient outcome events for external validation of the existing clinical severity scores.[185]

For derivation and updating of the clinical prediction models, as recommended by Riley et al., a conservative R^2 Nagelkerke of 0.15 and shrinkage factor of 0.9 were assumed:[186] at an outcome prevalence of 3.5% (104/3,010) up to 13 candidate predictors (events per parameter [EPP] = 8) could safely be used to build the prediction models whilst minimising the risk of overfitting (R package: *pmsampsize*).[187]

No formal sample size calculation was performed for the pneumonia subgroup analyses. All available data were used to maximise power and generalisability.

4.2.11 Ethics and reporting

Ethical approvals were provided by the Mahidol University Ethics Committee (TMEC 21-023) and Oxford Tropical Research Ethics Committee (OxTREC 511-21). Informed consent was obtained from the caretakers of all participants. The study is reported in accordance with the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) and STrengthening the Reporting of OBservational Studies in Epidemiology (STROBE) guidelines (Appendix 9.7 and Appendix 9.8).[66,188]

4.3 Results

4.3.1 Study population

From September 2007 to September 2008, 999 pregnant women were enrolled, with 965 children born into the cohort. Amongst 4,061 acute illness presentations, 3,064 (75.4%; 3,064/4,061) were for ARIs. Fifty-four ARI presentations were excluded (1.8%; 54/3,010) as information on oxygen therapy was not available in the study database, leaving 3,010 ARI presentations from 756 individual children for the external validation and updating of the clinical severity scores.

For the pneumonia subgroup analyses, 1,164 acute illness presentations met WHO-pneumonia criteria (28.7%; 1,164/4,061), 905 of which (77.8%; 905/1,164) had a stored serum sample available for quantification of biomarker concentrations. Data on supplemental oxygen therapy were missing in five of these attendances, leaving 900 presentations from 444 individual children for the primary analysis (Figure 4.3-1).

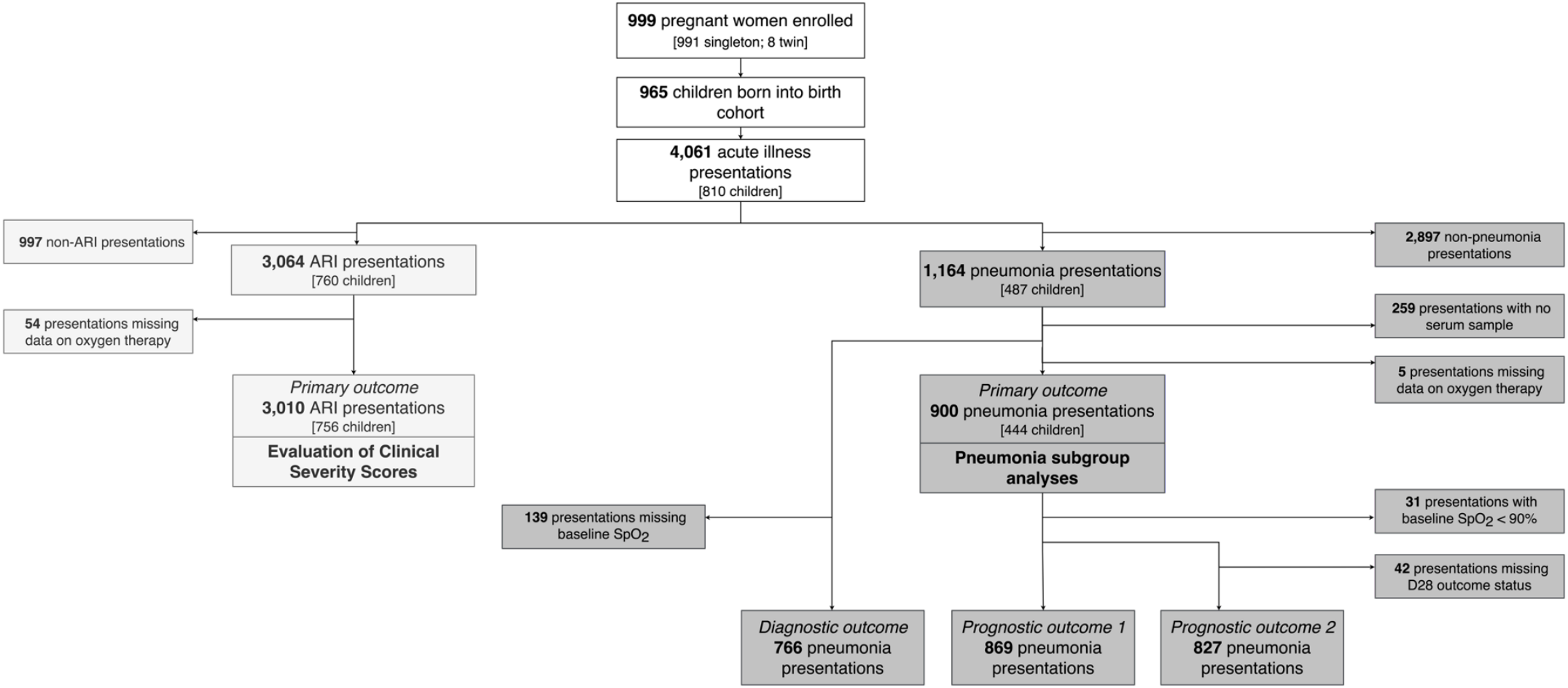


Figure 4.3-1: Eligibility of acute illness visits for inclusion in the study analyses. Primary outcome for ARI and pneumonia cohorts = receipt of supplemental oxygen during illness visit. Amongst pneumonia cohort: diagnostic outcome = SpO_2 < 90% at presentation; prognostic outcome 1 = receipt of supplemental oxygen during illness visit; prognostic outcome 2 = receipt of supplemental oxygen during 28d following presentation. Prognostic analyses exclude presentations meeting diagnostic outcome. ARI = acute respiratory infection.

4.3.2 Clinical characteristics of the ARI cohort

Median age at presentation was 8.1 months (interquartile range [IQR] = 3.7 to 13.7 months) and 52.9% (1,592/3,010) of attendances were for male children (Table 4.3-1). The majority of children were born at term with normal birthweight. Median number of previous illness visits was three (IQR = 2 to 6 visits). Children had been ill for a median of three days (IQR = 2 to 5 days) prior to presentation and presentations that met the primary outcome had been preceded by another illness visit more recently (11 vs. 31 days; $p < 0.001$). Most presentations included fever (65.2%; 1,958/3,005) and cough (92.1%; 2,767/3,005), with three-quarters (72.3%; 2,175/3,010) of ARIs managed in the community. Median length of stay at the clinic for the 835 admissions was three days (IQR = 2 to 4 days).

One hundred and four (3.5%; 104/3,010) presentations met the primary outcome, with those with signs of respiratory distress (92/104 [88%] vs. 416/2,905 [14%]), age-adjusted tachycardia (150 bpm vs. 140 bpm; 147 bpm vs. 136 bpm; 140 bpm vs. 128 bpm) and/or tachypnoea (64.5 bpm vs. 48.0 bpm; 58.0 bpm vs. 48.0 bpm; 57.0 bpm vs. 44.0 bpm), abnormal lung auscultation (83/99 [84%] vs. 1,372/2,852 [48%]), lower baseline SpO_2 (88.0% vs. 95.0%), prolonged CRTs (9/92 [9.8%] vs. 27/2,476 [1.1%]), altered mental status (66/98 [67%] vs. 306/2,875 [11%]), lower nutritional indices, known comorbidities (14/102 [14%] vs. 39/2,898 [1.3%]), and reported prior antibiotic use (20/104 [19%] vs. 125/2,906 [4.3%]) more likely to require supplemental oxygen (Table 4.3-1). The median time from the index illness presentation to the weight measurement used to calculate the WAZ was four days (IQR = 0 to 10 days), without intervening admission.

Table 4.3-1: Baseline characteristics of the acute respiratory infection cohort, stratified by primary outcome status. [#]Respiratory distress defined as head bobbing, tracheal tug, grunting, and/or chest indrawing; [†]abnormal chest auscultation defined as crepitations and/or wheeze; [‡]rectal temperature converted to axillary temperature for neonates and infants;[32] [§]median interval between anthropometric measurement and index illness presentation: length = 8 days (IQR = 4 to 12 days); MUAC = 9 days (IQR = 4 to 13 days); weight = 4 days (IQR = 0 to 10 days). [*]Missing data: gestation = 5; birthweight = 14; comorbidity = 10; symptom duration = 21; unwell family member = 10; fever = 5; runny nose = 2; noisy breathing = 6; stridor = 1; respiratory distress = 1; head bobbing = 1; tracheal tug = 1; grunting = 1; chest indrawing = 1; abnormal lung auscultation = 59; lung crepitations = 69; wheeze = 79; dehydration = 7; colour = 50; heart rate = 9; respiratory rate = 8; temperature = 3; oxygen saturation = 1,645; capillary refill time = 442; mental status = 37; weight-for-length z-score = 158; weight-for-age z-score = 147; MUAC-for-age-z-score = 682; length-for-age z-score = 14. ARI = acute respiratory infection; bpm = beats / breaths per minute; IQR = interquartile range; MUAC = mid-upper arm circumference.

Baseline characteristic	Overall N = 3,010 Median (IQR); n/N (%)	Supplemental oxygen		p-value[1]
		No N = 2,906 Median (IQR); n/N (%)	Yes N = 104 Median (IQR); n/N (%)	
Demographics				
Age (months)	8.1 (3.7, 13.7)	8.2 (3.8, 13.8)	7.3 (3.4, 12.7)	0.40
Male sex	1,592 / 3,010 (53%)	1,541 / 2,906 (53%)	51 / 104 (49%)	0.40
Birth history				
Gestation (weeks)[*]	39.1 (38.1, 40.0)	39.2 (38.2, 40.0)	38.4 (37.3, 39.7)	0.001
Birthweight (kg)[*]	2.9 (2.6, 3.2)	2.9 (2.6, 3.2)	2.6 (2.0, 3.0)	< 0.001
Previous medical history				
Number of previous illness visits	3.0 (2.0, 6.0)	3.0 (2.0, 6.0)	4.0 (2.0, 9.0)	0.043
Time since last illness visit (days)	29.0 (3.0, 81.0)	31.0 (3.0, 82.0)	11.0 (2.0, 36.5)	< 0.001
Number of previous ARI visits	3.0 (2.0, 5.0)	3.0 (2.0, 5.0)	3.5 (2.0, 8.0)	0.006
Known comorbidity[*]	53 / 3,000 (1.8%)	39 / 2,898 (1.3%)	14 / 102 (14%)	< 0.001
History of current illness				
Duration of symptoms (days)[*]	3.0 (2.0, 5.0)	3.0 (2.0, 5.0)	3.0 (2.0, 5.0)	0.30

Baseline characteristic	Overall N = 3,010 Median (IQR); n/N (%)	Supplemental oxygen		p-value[1]
		No N = 2,906 Median (IQR); n/N (%)	Yes N = 104 Median (IQR); n/N (%)	
Antibiotics prior to presentation	145 / 3,010 (4.8%)	125 / 2,906 (4.3%)	20 / 104 (19%)	< 0.001
Family member unwell[*]	287 / 3,000 (9.6%)	276 / 2,898 (9.5%)	11 / 102 (11%)	0.70
Presenting symptoms and signs				
Fever[*]	1,958 / 3,005 (65%)	1,885 / 2,901 (65%)	73 / 104 (70%)	0.30
Cough	2,767 / 3,010 (92%)	2,667 / 2,906 (92%)	100 / 104 (96%)	0.11
Runny nose[*]	2,565 / 3,008 (85%)	2,491 / 2,904 (86%)	74 / 104 (71%)	< 0.001
Noisy breathing[*]	447 / 3,004 (15%)	430 / 2,901 (15%)	17 / 103 (17%)	0.60
Stridor[*]	6 / 3,009 (0.2%)	6 / 2,905 (0.2%)	0 / 104 (0%)	> 0.90
Respiratory distress[#*]	508 / 3,009 (17%)	416 / 2,905 (14%)	92 / 104 (88%)	< 0.001
Head bobbing[*]	52 / 3,009 (1.7%)	27 / 2,905 (0.9%)	25 / 104 (24%)	< 0.001
Tracheal tug[*]	134 / 3,009 (4.5%)	96 / 2,905 (3.3%)	38 / 104 (37%)	< 0.001
Grunting[*]	26 / 3,009 (0.9%)	11 / 2,905 (0.4%)	15 / 104 (14%)	< 0.001
Chest indrawing[*]	493 / 3,009 (16%)	402 / 2,905 (14%)	91 / 104 (88%)	< 0.001
Abnormal lung auscultation[†*]	1,455 / 2,951 (49%)	1,372 / 2,852 (48%)	83 / 99 (84%)	< 0.001
Crepitations[*]	1,158 / 2,941 (39%)	1,085 / 2,844 (38%)	73 / 97 (75%)	< 0.001
Wheeze[*]	794 / 2,931 (27%)	751 / 2,833 (27%)	43 / 98 (44%)	< 0.001
Dehydration[*]	127 / 3,003 (4.2%)	121 / 2,899 (4.2%)	6 / 104 (5.8%)	0.40
Pale, mottled or cyanosed[*]	107 / 2,960 (3.6%)	91 / 2,862 (3.2%)	16 / 98 (16%)	< 0.001
Vital signs				
Heart rate (bpm)[*]				
Neonate	140.0 (132.0, 150.0)	140.0 (132.0, 148.0)	150.0 (140.0, 165.0)	0.014

Baseline characteristic	Overall N = 3,010 Median (IQR); n/N (%)	Supplemental oxygen		p-value[1]
		No N = 2,906 Median (IQR); n/N (%)	Yes N = 104 Median (IQR); n/N (%)	
Infant	138.0 (128.0, 144.0)	136.0 (128.0, 144.0)	147.0 (136.5, 154.0)	< 0.001
Child	128.0 (120.0, 140.0)	128.0 (120.0, 140.0)	140.0 (127.5, 149.0)	0.002
Respiratory rate (bpm)*				
Neonate	48.0 (45.0, 56.0)	48.0 (44.2, 54.0)	64.5 (54.0, 77.0)	0.008
Infant	48.0 (42.0, 56.0)	48.0 (42.0, 56.0)	58.0 (54.0, 66.0)	< 0.001
Child	45.0 (38.0, 52.0)	44.0 (38.0, 52.0)	57.0 (46.5, 62.0)	< 0.001
Axillary temperature (°C)‡*	36.6 (36.0, 37.5)	36.6 (36.0, 37.4)	36.8 (36.2, 37.8)	0.040
Oxygen saturation (%)*	95.0 (93.0, 96.0)	95.0 (93.0, 96.0)	88.0 (85.0, 93.0)	< 0.001
Capillary refill time > 2 secs*	36 / 2,568 (1.4%)	27 / 2,476 (1.1%)	9 / 92 (9.8%)	< 0.001
Not alert*	372 / 2,973 (13%)	306 / 2,875 (11%)	66 / 98 (67%)	< 0.001
Anthropometrics				
Weight-for-length z-score§*	0.0 (-0.8, 0.8)	0.0 (-0.8, 0.8)	-0.5 (-1.8, 0.7)	< 0.001
Weight-for-age z-score§*	-0.9 (-1.6, -0.2)	-0.9 (-1.6, -0.2)	-1.9 (-3.4, -0.8)	< 0.001
MUAC-for-age z-score§*	0.2 (-0.4, 0.8)	0.2 (-0.4, 0.8)	-0.7 (-1.9, 0.6)	< 0.001
Length-for-age z-score§*	-1.5 (-2.3, -0.7)	-1.4 (-2.2, -0.7)	-2.4 (-3.4, -1.4)	< 0.001

[1]Wilcoxon rank sum test; Pearson's Chi-squared test; Fisher's exact test

4.3.3 LqSOFA and qPELOD-2 scores outperform the mSIRS score for risk stratification of ARIs

Discrimination and calibration of the LqSOFA (AUC = 0.84; 95% CI = 0.79 to 0.89) and qPELOD-2 (AUC = 0.79; 95% CI = 0.74 to 0.84) scores were considerably better than the mSIRS score (AUC = 0.57; 95% CI = 0.51 to 0.63; Figure 4.3-2; Table 4.3-2). Little variability was observed between each of the imputed datasets (Figure 4.3-3). In contrast to the LqSOFA and qPELOD-2 scores, it was notable that the majority of participants had the same mSIRS score (one), irrespective of their outcome status. At a cut-off of ≥ 1 the LqSOFA score demonstrated a sensitivity of 0.80 (95% CI = 0.72 to 0.89) and specificity of 0.86 (95% CI = 0.85 to 0.88), and modest rule-in (PLR = 5.89; 95% CI = 5.08 to 6.82) and rule-out (NLR = 0.23; 95% CI = 0.15 to 0.36) potential; neither the mSIRS nor qPELOD-2 scores achieved a sensitivity and specificity > 0.70 at any cut-off (Table 4.3-3).

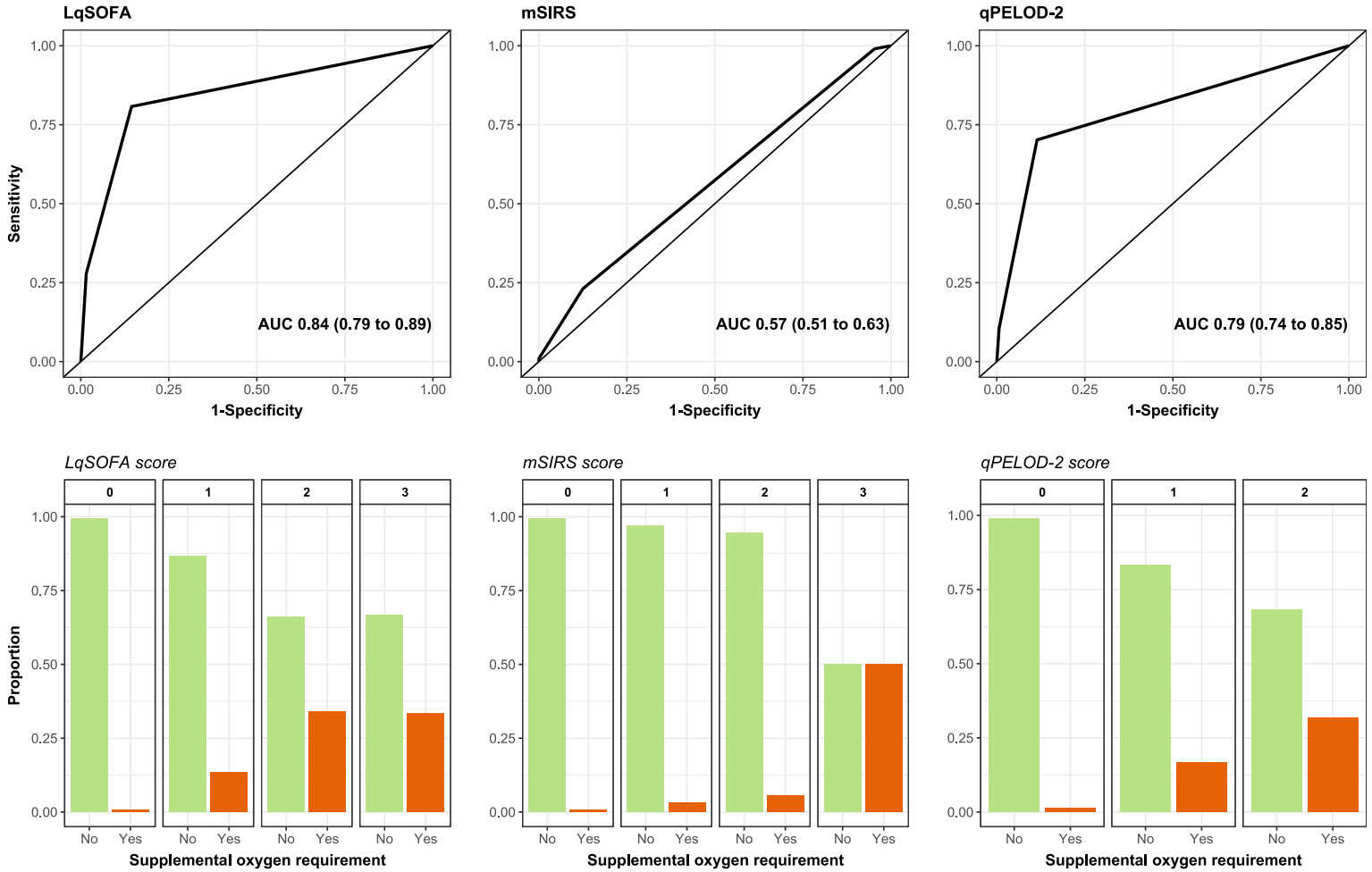


Figure 4.3-2: Discrimination and calibration of existing severity scores. Receiver operating characteristic curve for one imputed dataset shown. Variability across multiply imputed datasets shown in Figure 4.3-3. Pooled AUC reported. Bar plots present data using full case analysis: LqSOFA = 2,525 presentations (81 met primary outcome); mSIRS = 2,992 presentations (99 met primary outcome); qPELOD-2 = 2,531 presentations (83 met primary outcome). AUC = area under receiver operating characteristic curve; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; mSIRS = modified Systematic Inflammatory Response Syndrome; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2.

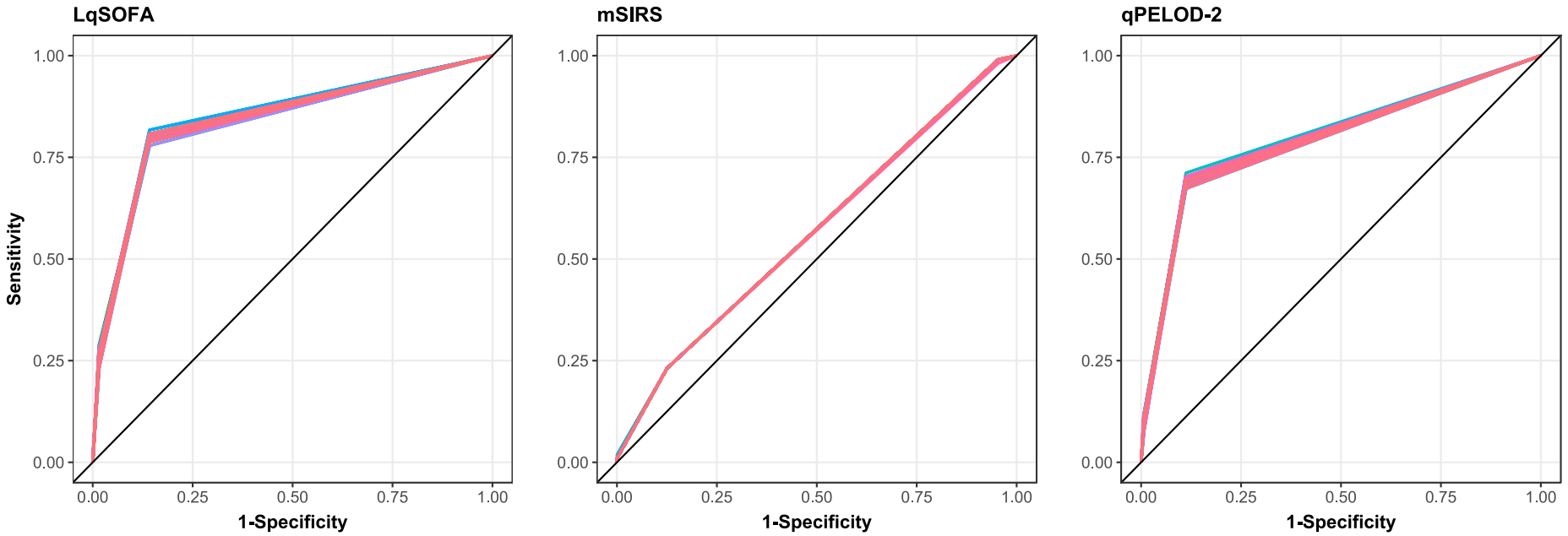


Figure 4.3-3: Variability in discrimination of the existing severity scores across multiply imputed datasets. Overlay plot to illustrate variability in receiver operating characteristic curves across multiply imputed datasets. LqSOFA = Liverpool quick Sequential Organ Failure Assessment; mSIRS = modified Systemic Inflammatory Response Syndrome; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2.

Table 4.3-2: Proportion of participants that met the primary outcome for constituent variables and levels of each score. Scores calculated using full-case analysis: LqSOFA = 2,525 presentations (81 met primary outcome); mSIRS = 2,992 presentations (99 met primary outcome); qPELOD-2 = 2,531 presentations (83 met primary outcome). Bpm = beats per minute; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; mSIRS = modified Systemic Inflammatory Response Syndrome; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2.

Baseline characteristic	Overall N = 3,010 n/N (%)	Supplemental oxygen		p-value[1]
		No N = 2,906 n/N (%)	Yes N = 104 n/N (%)	
LqSOFA				
Respiratory rate > 99th centile[189]	136 / 3,002 (4.5%)	111 / 2,901 (3.8%)	25 / 101 (25%)	< 0.001
Heart rate > 99th centile[189]	7 / 3,001 (0.2%)	4 / 2,900 (0.1%)	3 / 101 (3.0%)	0.001
Not alert	372 / 2,973 (13%)	306 / 2,875 (11%)	66 / 98 (67%)	< 0.001
Capillary refill time > 2 secs	36 / 2,568 (1.4%)	27 / 2,476 (1.1%)	9 / 92 (9.8%)	< 0.001
LqSOFA = 0	2,127 / 2,525 (84%)	2,111 / 2,444 (86%)	16 / 81 (20%)	< 0.001
LqSOFA = 1	342 / 2,525 (14%)	296 / 2,444 (12%)	46 / 81 (57%)	
LqSOFA = 2	53 / 2,525 (2.1%)	35 / 2,444 (1.4%)	18 / 81 (22%)	
LqSOFA = 3	3 / 2,525 (0.1%)	2 / 2,444 (<0.1%)	1 / 81 (1.2%)	
mSIRS				
Axillary temperature > 38°C	391 / 3,007 (13%)	370 / 2,903 (13%)	21 / 104 (20%)	0.027
Axillary temperature < 35.5°C	7 / 3,007 (0.2%)	5 / 2,903 (0.2%)	2 / 104 (1.9%)	0.022
Heart rate > 95th centile	6 / 3,001 (0.2%)	3 / 2,900 (0.1%)	3 / 101 (3.0%)	< 0.001
Heart rate < 5th centile	0 / 3,001 (0%)	0 / 2,900 (0%)	0 / 101 (0%)	NA
Respiratory rate > 95th centile	2,852 / 3,002 (95%)	2,753 / 2,901 (95%)	99 / 101 (98%)	0.20

Baseline characteristic	Overall N = 3,010 n/N (%)	Supplemental oxygen		p-value[1]
		No N = 2,906 n/N (%)	Yes N = 104 n/N (%)	
mSIRS = 0	135 / 2,992 (4.5%)	134 / 2,893 (4.6%)	1 / 99 (1.0%)	0.002
mSIRS = 1	2,475 / 2,992 (83%)	2,399 / 2,893 (83%)	76 / 99 (77%)	
mSIRS = 2	380 / 2,992 (13%)	359 / 2,893 (12%)	21 / 99 (21%)	
mSIRS = 3	2 / 2,992 (<0.1%)	1 / 2,893 (<0.1%)	1 / 99 (1.0%)	
qPELOD-2				
Heart rate > 195 bpm	0 / 3,001 (0%)	0 / 2,900 (0%)	0 / 101 (0%)	NA
Not alert	372 / 2,973 (13%)	306 / 2,875 (11%)	66 / 98 (67%)	< 0.001
Capillary refill time > 2 secs	36 / 2,568 (1.4%)	27 / 2,476 (1.1%)	9 / 92 (9.8%)	< 0.001
qPELOD-2 = 0	2,217 / 2,531 (88%)	2,190 / 2,448 (89%)	27 / 83 (33%)	< 0.001
qPELOD-2 = 1	292 / 2,531 (12%)	243 / 2,448 (9.9%)	49 / 83 (59%)	
qPELOD-2 = 2	22 / 2,531 (0.9%)	15 / 2,448 (0.6%)	7 / 83 (8.4%)	

[1]Fisher's exact test; Pearson's Chi-squared test

Table 4.3-3: Classifications at each level of a score. Classifications calculated using full-case analysis: LqSOFA = 2,525 presentations (81 met primary outcome); mSIRS = 2,992 presentations (99 met primary outcome); qPELOD-2 = 2,531 presentations (83 met primary outcome). CI = confidence interval; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; mSIRS = modified Systemic Inflammatory Response Syndrome; NLR = negative likelihood ratio; PLR = positive likelihood ratio; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2.

Cut-off	Sensitivity (95% CI)	Specificity (95% CI)	NLR (95% CI)	PLR (95% CI)	Cases referred (%)	Cases managed in community (%)	Ratio of Incorrect to Correct referrals	Ratio of Correct to Incorrect cases managed in community
LqSOFA score								
≥ 1	0.80 (0.72 to 0.89)	0.86 (0.85 to 0.88)	0.23 (0.15 to 0.36)	5.89 (5.08 to 6.82)	407 (16.1%)	2118 (83.9%)	5 to 1	131 to 1
≥ 2	0.23 (0.14 to 0.33)	0.98 (0.98 to 0.99)	0.78 (0.69 to 0.88)	15.49 (9.33 to 25.72)	68 (2.7%)	2457 (97.3%)	3 to 1	39 to 1
≥ 3	0.01 (-0.01 to 0.04)	1.00 (1.00 to 1.00)	0.99 (0.96 to 1.01)	15.09 (1.38 to 164.69)	1 (< 0.01%)	2524 (> 99.9%)	0 to 1	31 to 1
mSIRS score								
≥ 1	0.99 (0.97 to 1.00)	0.05 (0.04 to 0.05)	0.22 (0.03 to 1.54)	1.04 (1.02 to 1.06)	2846 (95.1%)	146 (4.9%)	28 to 1	145 to 1
≥ 2	0.22 (0.14 to 0.30)	0.88 (0.86 to 0.89)	0.89 (0.80 to 0.99)	1.79 (1.22 to 2.61)	369 (12.3%)	2623 (87.7%)	16 to 1	33 to 1
≥ 3	0.01 (-0.01 to 0.03)	1.00 (1.00 to 1.00)	0.99 (0.97 to 1.01)	29.22 (1.84 to 463.84)	1 (< 0.1%)	2991 (> 99.9%)	0 to 1	30 to 1
qPELOD-2 score								
≥ 1	0.68 (0.57 to 0.78)	0.90 (0.88 to 0.91)	0.36 (0.27 to 0.50)	6.40 (5.30 to 7.73)	301 (11.9%)	2230 (88.1%)	4 to 1	82 to 1
≥ 2	0.08 (0.03 to 0.14)	0.99 (0.99 to 1.00)	0.92 (0.86 to 0.98)	13.76 (5.77 to 32.86)	31 (1.2%)	2500 (98.8%)	3 to 1	32 to 1

4.3.4 Improved performance of severity scores when deployed as clinical prediction models

Relationships between continuous predictors and the primary outcome are illustrated in Figure 4.3-4. Approximate linear (heart rate and respiratory rate) or uniform (age) relationships with probability of supplemental oxygen therapy were observed. As expected, the relationship between temperature and the primary outcome was non-linear, with both hypothermia and fever associated with increased probability of receiving supplemental oxygen. Given known age-dependent variations in heart rate and respiratory rate associated with physiological maturation in young children, age-specific relationships were explored. Visualisations did not suggest interaction and this was supported by the results of LRTs between age and each of heart rate (LRT = 2.09; p = 0.35) and respiratory rate (LRT = 0.77; p = 0.68).

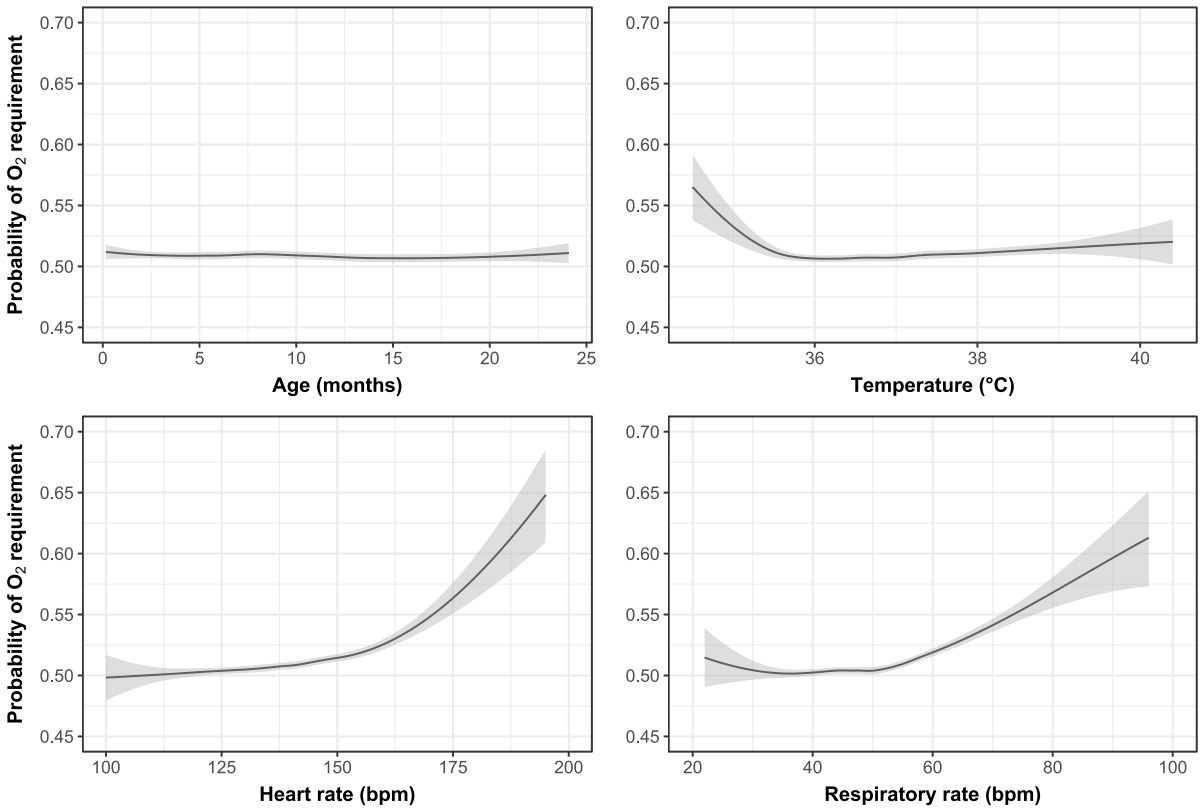


Figure 4.3-4: Association between continuous predictors and the primary outcome. Black line = probability of oxygen requirement. Grey shaded areas = 95% confidence interval. Bpm = beats / breaths per minute.

Optimism-adjusted discrimination of the three models ranged from an AUC of 0.81 to 0.90, with the LqSOFA model appearing most promising (AUC = 0.90; 95% CI = 0.86 to 0.94; Figure 4.3-5). Calibration of the qPELOD-2 model was good. The LqSOFA and mSIRS models overestimated risk in children with predicted probabilities of supplemental oxygen therapy $\geq$ 30%. Again, little variability was observed between each of the imputed datasets (Figure 4.3-6).

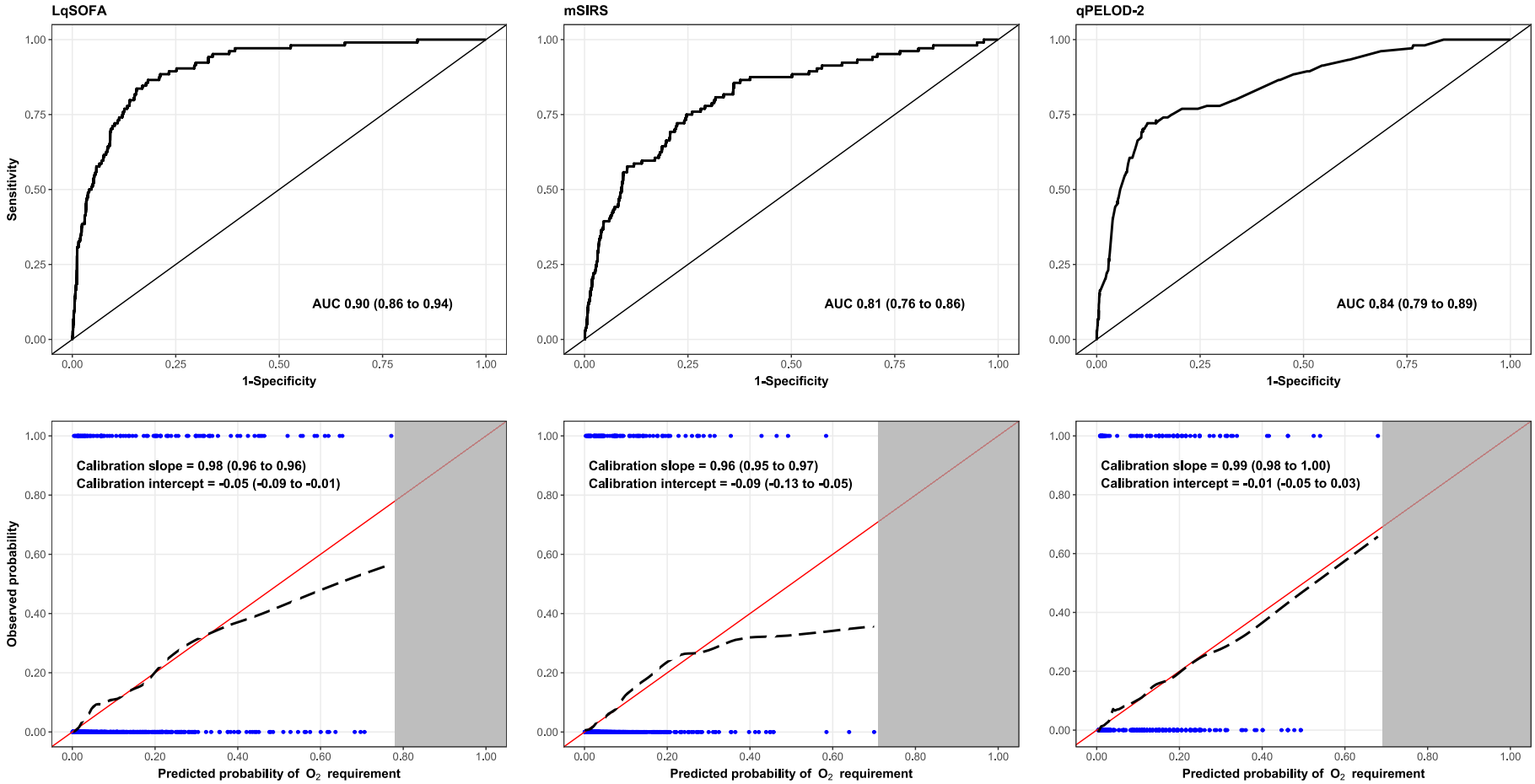


Figure 4.3-5: Discrimination and calibration of the clinical prediction models. Receiver operating characteristic curve and calibration plot for one imputed dataset shown. Pooled optimism-adjusted AUCs and calibration slopes reported. On calibration plots, red line indicates perfect calibration; black dashed line indicates calibration slope for that particular model; blue rug plots indicate distribution of predicted risks for participants who did (top) and did not (bottom) meet the primary outcome. AUC = area under receiver operating characteristic curve; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; mSIRS = modified Systemic Inflammatory Response Syndrome; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2.

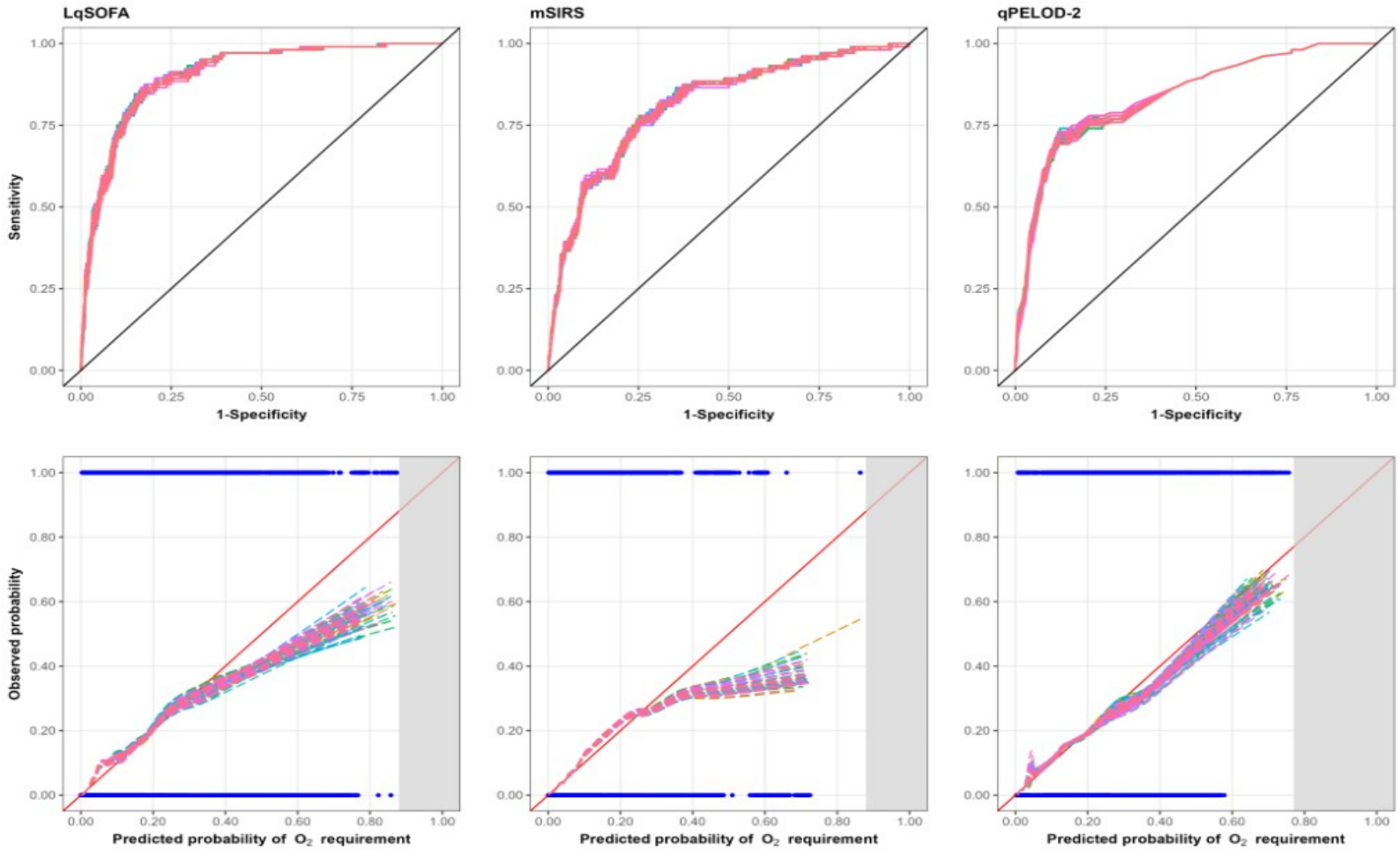


Figure 4.3-6: Variability in discrimination and calibration of the clinical prediction models across multiply imputed datasets. Top panel: receiver operating characteristic curves. Bottom panel: calibration plots. On calibration plots, solid red line indicates perfect calibration; coloured dashed line indicates calibration slope for each imputed dataset; blue rug plots indicate distribution of predicted risks for participants who did (top) and did not (bottom) meet the primary outcome. LqSOFA = Liverpool quick Sequential Organ Failure Assessment; mSIRS = modified Systemic Inflammatory Response Syndrome; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2.

Discrimination of all three updated models containing respiratory distress and WAZ improved (AUCs = 0.93 to 0.95). Notably, improvements were more substantial for the qPELOD-2 and mSIRS models, compared to the LqSOFA model, which already had comparably high discrimination prior to inclusion of the additional variables. Calibration of the updated LqSOFA and qPELOD-2 models was good, whereas the updated mSIRS model underestimated risk in children with predicted probabilities of supplemental oxygen therapy $\geq$ 50% (Figure 4.3-7). The full model equations are reported in Equations 4.3-1 to allow estimation of the predicted probability that an individual child presenting with an ARI will require supplemental oxygen therapy during their illness visit.

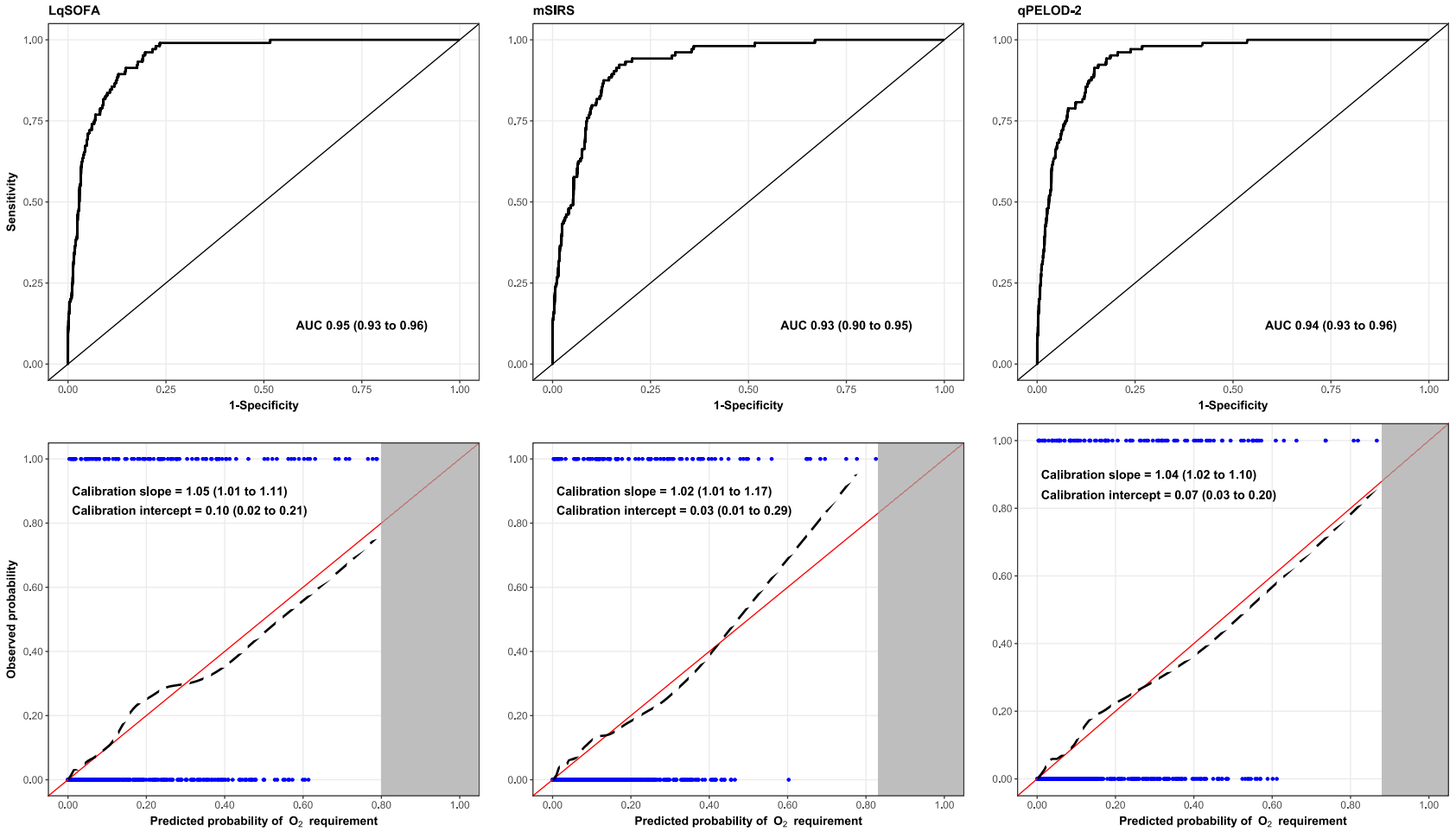


Figure 4.3-7: Discrimination and calibration of the updated clinical prediction models. On calibration plots, red line indicates perfect calibration; black dashed line indicates calibration slope for that particular model; blue rug plots indicate distribution of predicted risks for participants who did (top) and did not (bottom) meet the primary outcome. AUC = area under receiver operating characteristic curve; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; mSIRS = modified Systemic Inflammatory Response Syndrome; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2.

$$\Pr(\boldsymbol{Oxygen\ requirement}) = \frac{\boldsymbol{e^{LP}}}{\boldsymbol{1 + e^{LP}}}$$

where LP is the linear predictor

The LP should be estimated for each model using the following equations:

Original models

LqSOFA model

$$LP(\text{LqSOFA model}) = -11.097 + \begin{cases} 0\ if\ CRT \leq 2 \\ 0.999\ if\ CRT > 2 \end{cases} + \begin{cases} 0\ if\ AVPU = A \\ 2.366\ if\ AVPU < A \end{cases} + 0.025 * age + 0.016 * hr + 0.084 * rr$$

mSIRS model

$$LP(\text{mSIRS model}) = 23.819 + \begin{cases} -1.0.47 * temp\ if\ temp < 36 \\ 2.930 * temp\ if\ temp \geq 36 \end{cases} + 0.036 * age + 0.038 * hr + 0.094 * rr$$

qPELOD-2 model

$$LP(\text{qPELOD2 model}) = -8.783 + \begin{cases} 0\ if\ CRT \leq 2 \\ 0.966\ if\ CRT > 2 \end{cases} + \begin{cases} 0\ if\ AVPU = A \\ 2.475\ if\ AVPU < A \end{cases} + 0.033 * hr$$

Updated models

LqSOFA model

$$LP(\text{LqSOFA model}) = -8.727 + \begin{cases} 0\ if\ CRT \leq 2 \\ 0.526\ if\ CRT > 2 \end{cases} + \begin{cases} 0\ if\ AVPU = A \\ 1.685\ if\ AVPU < A \end{cases} - 0.013 * age + 0.008 * hr + 0.035 * rr +$$

$$\begin{cases} 0\ if\ no\ respiratory\ distress \\ 2.722\ if\ respiratory\ distress \end{cases} - 0.364 * WAZ$$

mSIRS model

$$LP(\text{mSIRS model}) = 25.623 + \begin{cases} -0.993 * temp\ if\ temp < 36 \\ 3.428 * temp\ if\ temp \geq 36 \end{cases} - 0.024 * age + 0.019 * hr + 0.036 * rr +$$

$$\begin{cases} 0\ if\ no\ respiratory\ distress \\ 3.208\ if\ respiratory\ distress \end{cases} - 0.439 * WAZ$$

qPELOD-2 model

$$LP(\text{qPELOD2 model}) = -8.112 + \begin{cases} 0\ if\ CRT \leq 2 \\ 0.558\ if\ CRT > 2 \end{cases} + \begin{cases} 0\ if\ AVPU = A \\ 1.646\ if\ AVPU < A \end{cases} + 0.016 * hr +$$

$$\begin{cases} 0\ if\ no\ respiratory\ distress \\ 3.070\ if\ respiratory\ distress \end{cases} - 0.325 * WAZ$$

Equation 4.3-1: Model equations to estimate the probability of oxygen requirement. AVPU = Alert Voice Pain Unresponsive scale; CRT = capillary refill time; hr = heart rate; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; mSIRS = modified Systemic Inflammatory Response Syndrome; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2; rr = respiratory rate; temp = temperature; WAZ = weight-for-age z-score.

4.3.5 Gains in performance associated with updating the points-based severity scores

The ability of each updated model to guide referrals at thresholds ranging from 1% to 40% is shown (Table 4.3-4). A referral threshold of 5% reflects a strategy whereby any child with a predicted probability of requiring oxygen therapy ≥ 5% is referred. At this cut off, the models would suggest referral in ~15% of all presentations, correctly identifying ~86-87% of children requiring referral, at a cost of also recommending referral in ~12-13% of children not requiring referral; i.e., a number needed to refer (NNR; the number of children referred to identify one child who would require oxygen) of five.

Using the most promising cut-off amongst the points-based severity scores (the LqSOFA score ≥ 1) would recommend referral in a similar proportion of presentations (16.1%). However, compared to using the clinical prediction models to guide referrals, the points-based score would result in a ~25% increase in incorrect referrals (a NNR of six vs. five) and a ~25-30% increase in the number of children incorrectly identified as safe for community-based management (a ratio of correct to incorrect cases managed in the community of 171 to 193:1 vs. 131:1; Table 4.3-4).

Table 4.3-4: Classifications at different referral thresholds using the updated clinical prediction models. A referral threshold of 5% reflects a management strategy whereby any child with a predicted probability of requiring oxygen ≥ 5% is referred. CI = confidence interval; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; mSIRS = modified Systemic Inflammatory Response Syndrome; NLR = negative likelihood ratio; PLR = positive likelihood ratio; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2.

Model	Sensitivity (95% CI)	Specificity (95% CI)	NLR (95% CI)	PLR (95% CI)	Cases referred (%)	Cases managed in community (%)	Ratio of Incorrect to Correct referrals	Ratio of Correct to Incorrect cases managed in community
Referral threshold = 1%								
LqSOFA	0.97 (0.93 to 1.00)	0.78 (0.73 to 0.82)	0.03 (0.00 to 0.09)	4.48 (3.65 to 5.57)	723 (24.0%)	2287 (76.0%)	6 to 1	571 to 1
qPELOD-2	0.96 (0.93 to 0.99)	0.79 (0.75 to 0.83)	0.05 (0.01 to 0.10)	4.55 (3.97 to 5.87)	715 (23.8%)	2295 (76.2%)	6 to 1	573 to 1
mSIRS	0.94 (0.90 to 0.98)	0.78 (0.65 to 0.84)	0.08 (0.04 to 0.14)	4.36 (3.00 to 6.21)	714 (23.7%)	2296 (76.3%)	6 to 1	382 to 1
Referral threshold = 5%								
LqSOFA	0.87 (0.78 to 0.93)	0.88 (0.86 to 0.91)	0.15 (0.09 to 0.25)	7.40 (6.22 to 9.50)	431 (14.3%)	2579 (85.7%)	4 to 1	171 to 1
qPELOD-2	0.87 (0.78 to 0.93)	0.87 (0.85 to 0.91)	0.15 (0.08 to 0.25)	6.79 (5.96 to 8.98)	471 (15.6%)	2539 (84.4%)	4 to 1	180 to 1
mSIRS	0.86 (0.77 to 0.93)	0.87 (0.85 to 0.89)	0.16 (0.08 to 0.26)	6.55 (5.80 to 7.74)	482 (16.0%)	2528 (84.0%)	4 to 1	193 to 1
Referral threshold = 10%								
LqSOFA	0.75 (0.66 to 0.83)	0.93 (0.91 to 0.95)	0.26 (0.18 to 0.37)	11.76 (9.04 to 16.80)	262 (8.7%)	2748 (91.3%)	2 to 1	101 to 1
qPELOD-2	0.73 (0.61 to 0.82)	0.93 (0.90 to 0.95)	0.29 (0.20 to 0.41)	11.57 (8.21 to 17.02)	259 (8.6%)	2751 (91.4%)	2 to 1	94 to 1
mSIRS	0.76 (0.63 to 0.86)	0.91 (0.88 to 0.93)	0.27 (0.16 to 0.41)	8.22 (6.83 to 10.18)	354 (11.8%)	2632 (88.3%)	3 to 1	110 to 1

Referral threshold = 20%								
LqSOFA	0.59 (0.45 to 0.69)	0.97 (0.96 to 0.97)	0.42 (0.32 to 0.56)	17.82 (13.83 to 23.17)	157 (5.2%)	2853 (94.8%)	2 to 1	65 to 1
qPELOD-2	0.56 (0.41 to 0.65)	0.97 (0.96 to 0.97)	0.46 (0.36 to 0.60)	16.81 (12.98 to 22.87)	152 (5.0%)	2858 (95.0%)	2 to 1	57 to 1
mSIRS	0.49 (0.37 to 0.61)	0.96 (0.95 to 0.97)	0.53 (0.40 to 0.65)	12.97 (9.85 to 19.44)	161 (5.3%)	2849 (94.7%)	2 to 1	50 to 1
Referral threshold = 40%								
LqSOFA	0.28 (0.16 to 0.41)	0.99 (0.98 to 1.00)	0.73 (0.60 to 0.85)	27.50 (17.38 to 56.63)	56 (1.9%)	2954 (98.1%)	1 to 1	36 to 1
qPELOD-2	0.28 (0.13 to 0.41)	0.99 (0.98 to 0.99)	0.73 (0.59 to 0.87)	27.90 (16.46 to 47.56)	61 (2.0%)	2949 (98.0%)	1 to 1	39 to 1
mSIRS	0.21 (0.09 to 0.35)	1.00 (0.99 to 1.00)	0.80 (0.66 to 0.91)	Inf	32 (1.1%)	2978 (98.9%)	1 to 1	34 to 1

4.3.6 Promising utility of LqSOFA and qPELOD-2 models to guide referrals from primary care

The relative value of correct and incorrect referrals is highly context-dependent, reflecting resource availability, practicalities of referral, and capacity for follow-up. Decision curve analyses accounting for differing circumstances suggest that the updated models could provide greater utility (net benefit) compared to the points-based LqSOFA score, with the LqSOFA and qPELOD-2 models appearing most promising over a wide range of plausible referral thresholds (Figure 4.3-8).

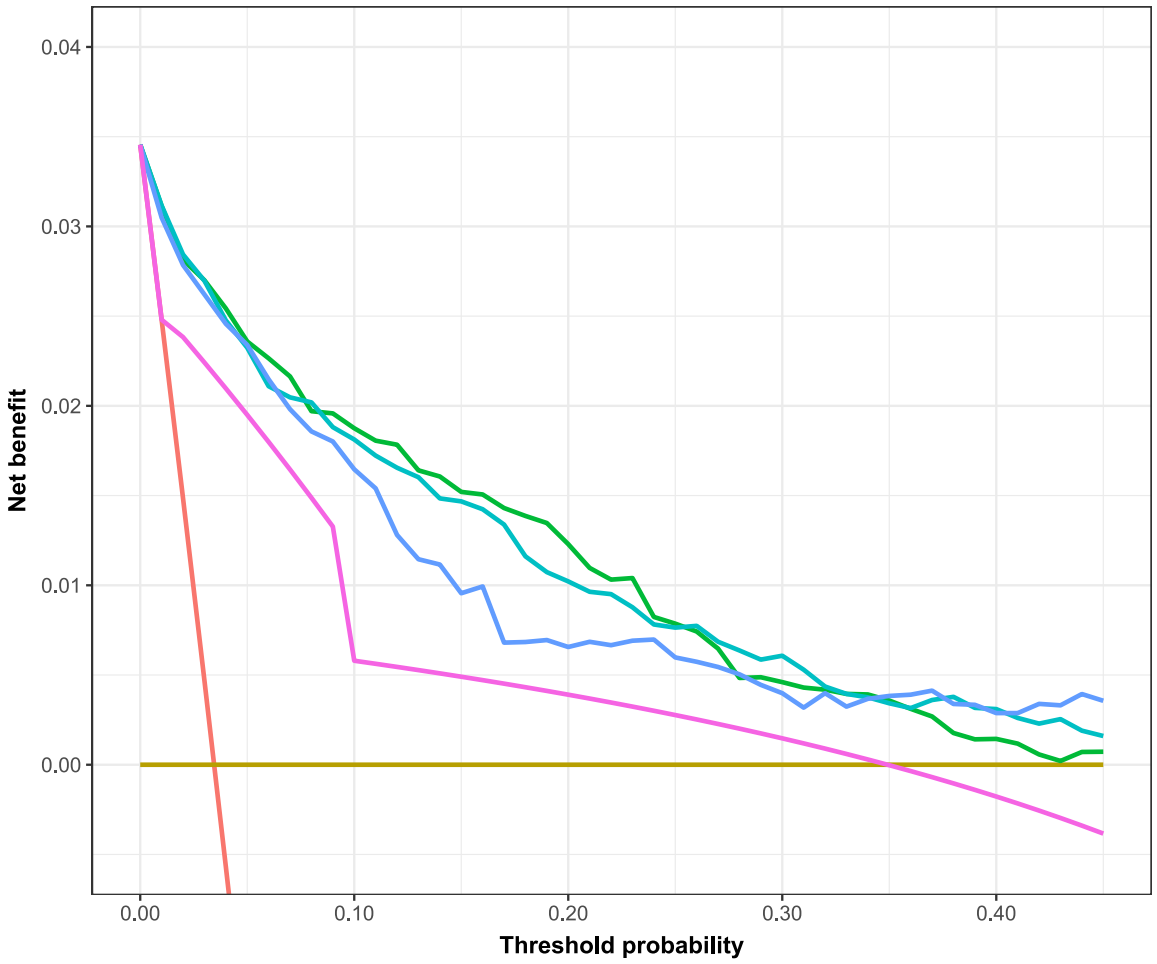


Figure 4.3-8: Utility of the updated clinical prediction models for the identification of children suitable for community-based management of pneumonia. The net benefit of the updated models (green [LqSOFA], turquoise [qPELOD-2], and blue [mSIRS] lines) and original LqSOFA score (pink line), are compared to a 'refer-all' (red line) and 'refer-none' (brown line) approach. A threshold probability of 5% indicates a management strategy whereby any child with a ≥ 5% probability of requiring oxygen is referred (i.e., a scenario where the value of one correct referral is equivalent to 19 incorrect referrals or a NNR of 20). Moving from left to right along the x-axis (increasing referral threshold) reflects increasing 'penalisation' of an incorrect referral (false positive), indicative of contexts in which referrals may be more challenging or costly. LqSOFA = Liverpool quick Sequential Organ Failure Assessment; mSIRS = modified Systemic Inflammatory Response Syndrome; NNR = number needed to refer; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2.

4.3.7 Sensitivity analyses

The WHO recommend that pulse oximetry should be universally available at first-level health facilities.[32,190] Although many barriers exist to realising this laudable goal, to account for the fact that in such contexts a severity score or prediction model would not be required to guide referral for children who are already hypoxaemic at the time of presentation, a sensitivity analysis excluding attendances with SpO_2 < 90% at presentation was conducted. Discrimination and calibration remained comparable (Table 4.3-5). Rule-in ability of the models (PLR) appeared to improve slightly, at a cost of poorer rule-out (NLR) performance across all referral thresholds (Table 4.3-6). This appeared to be driven by reductions in sensitivity, although fewer outcome events (n = 52) preclude firm inferences.

Table 4.3-5: Performance of the severity scores and clinical prediction models excluding attendances with SpO_2 < 90% at presentation. Sensitivity analysis includes 2,949 presentations, 52 of which met the primary outcome. *Pooled performance measure across multiply imputed datasets reported. †Optimism-adjusted performance measure reported. CI = confidence interval; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; mSIRS = modified Systemic Inflammatory Response Syndrome; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2.

	Analysis	LqSOFA	mSIRS	qPELOD-2
Severity scores (external validation)				
Discrimination (95% CI)*	Main	0.84 (0.79 to 0.89)	0.57 (0.51 to 0.63)	0.79 (0.74 to 0.85)
	Sensitivity	0.82 (0.74 to 0.89)	0.55 (0.47 to 0.63)	0.78 (0.71 to 0.86)
Prognostic models (original)				
Discrimination (95% CI)*†	Main	0.90 (0.86 to 0.94)	0.81 (0.76 to 0.86)	0.84 (0.79 to 0.89)
	Sensitivity	0.89 (0.83 to 0.95)	0.78 (0.70 to 0.85)	0.83 (0.76 to 0.90)
Calibration slope (95% CI)*†	Main	0.98 (0.96 to 0.99)	0.96 (0.95 to 0.97)	0.99 (0.98 to 1.00)
	Sensitivity	0.95 (0.94 to 0.97)	0.93 0.91 to 0.95)	0.98 (0.96 to 1.00)
Calibration intercept (95% CI)*†	Main	-0.05 (-0.09 to -0.01)	-0.09 (-0.13 to -0.05)	-0.01 (-0.05 to -0.03)
	Sensitivity	-0.12 (-0.17 to -0.07)	-0.23 (-0.32 to -0.16)	-0.04 (-0.12 to 0.03)
Prognostic models (updated)				
Discrimination (95% CI)†	Main	0.95 (0.93 to 0.96)	0.93 (0.90 to 0.95)	0.94 (0.93 to 0.96)
	Sensitivity	0.93 (0.91 to 0.96)	0.91 (0.87 to 0.93)	0.93 (0.89 to 0.95)
Calibration slope (95% CI)†	Main	1.05 (1.01 to 1.11)	1.02 (1.01 to 1.17)	1.04 (1.02 to 1.10)
	Sensitivity	1.08 (1.01 to 1.17)	1.05 (1.01 to 1.20)	1.05 (1.01 to 1.17)
Calibration intercept (95% CI)†	Main	0.10 (0.02 to 0.21)	0.03 (0.01 to 0.29)	0.07 (0.03 to 0.20)
	Sensitivity	0.20 (0.03 to 0.49)	0.14 (0.02 to 0.61)	0.14 (0.03 to 0.48)

Table 4.3-6: Classifications at different referral thresholds using the updated clinical prediction models excluding attendances with SpO_2 < 90% at presentation. Sensitivity analysis includes 2,949 presentations, 52 of which met the primary outcome. A referral threshold of 5% reflects a management strategy whereby any child with a predicted probability of requiring oxygen ≥ 5% is referred. CI = confidence interval; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; mSIRS = modified Systemic Inflammatory Response Syndrome; NLR = negative likelihood ratio; PLR = positive likelihood ratio; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2.

Model	Sensitivity (95% CI)	Specificity (95% CI)	NLR (95% CI)	PLR (95% CI)
Referral threshold = 1%				
LqSOFA	0.94 (0.86 to 1.00)	0.80 (0.76 to 0.85)	0.07 (0.02 to 0.17)	4.68 (3.97 to 6.30)
qPELOD-2	0.94 (0.86 to 0.98)	0.78 (0.77 to 0.85)	0.08 (0.02 to 0.17)	4.66 (4.05 to 6.53)
mSIRS	0.87 (0.78 to 0.94)	0.80 (0.70 to 0.85)	0.16 (0.07 to 0.29)	4.60 (3.06 to 6.35)
Referral threshold = 5%				
LqSOFA	0.69 (0.47 to 0.83)	0.94 (0.90 to 0.96)	0.33 (0.18 to 0.55)	11.22 (8.04 to 17.80)
qPELOD-2	0.63 (0.44 to 0.79)	0.94 (0.89 to 0.96)	0.39 (0.23 to 0.59)	10.77 (7.18 to 18.06)
mSIRS	0.71 (0.51 to 0.84)	0.90 (0.87 to 0.94)	0.32 (0.17 to 0.53)	7.49 (5.94 to 10.93)
Referral threshold = 10%				
LqSOFA	0.52 (0.34 to 0.68)	0.96 (0.95 to 0.97)	0.50 (0.33 to 0.68)	14.97 (10.86 to 21.45)
qPELOD-2	0.49 (0.32 to 0.63)	0.97 (0.96 to 0.97)	0.52 (0.38 to 0.70)	14.39 (10.42 to 21.35)
mSIRS	0.46 (0.25 to 0.64)	0.96 (0.93 to 0.98)	0.56 (0.38 to 0.77)	11.44 (8.14 to 17.71)
Referral threshold = 20%				
LqSOFA	0.31 (0.10 to 0.52)	0.98 (0.97 to 1.00)	0.70 (0.49 to 0.90)	Inf
qPELOD-2	0.32 (0.10 to 0.54)	0.98 (0.97 to 0.99)	0.69 (0.47 to 0.91)	21.96 (13.93 to 41.68)
mSIRS	0.18 (0.00 to 0.37)	0.99 (0.98 to 1.00)	0.83 (0.64 to 1.00)	Inf
Referral threshold = 40%				
LqSOFA	0.09 (0.00 to 0.25)	1.00 (0.99 to 1.00)	0.91 (0.76 to 1.00)	Inf
qPELOD-2	0.08 (0.00 to 0.27)	1.00 (0.99 to 1.00)	0.92 (0.74 to 1.00)	Inf
mSIRS	0.04 (0.00 to 0.14)	1.00 (1.00 to 1.00)	0.96 (0.86 to 1.00)	Inf

A limitation of many studies evaluating community-based triage tools is the lack of follow-up data for patients categorised as low risk;[123] in this study, 72.3% (2,175/3,010) of children were sent away from the clinic without admission. However, as acute illness visits were nested within the longitudinal birth cohort and the clinic was one of only two medical facilities with capacity for providing oxygen therapy in the camps, it was possible to confirm which attendances received supplemental

oxygen within the 28 days following presentation in 95.8% (2,083/2,175) of cases managed as outpatients. Amongst these, 1.4% (30/2,083) received supplemental oxygen within the next 28 days. Although it is uncertain whether this oxygen therapy related to the index ARI or a new illness, a sensitivity analysis conservatively classifying these 30 presentations as meeting the primary outcome (i.e., assuming the oxygen therapy related to the index ARI) resulted in a decrease in the discrimination and sensitivity of all three severity scores and models (Table 4.3-7 and Table 4.3-8).

Table 4.3-7: Performance of the severity scores and clinical prediction models when presentations sent away from the clinic but requiring oxygen within the next 28 days are classified as meeting the primary outcome. Sensitivity analysis includes 2,918 presentations, 134 of which met the primary outcome. *Pooled performance measure across multiply imputed datasets reported. †Optimism-adjusted performance measure reported. CI = confidence interval; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; mSIRS = modified Systemic Inflammatory Response Syndrome; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2.

	Analysis	LqSOFA	mSIRS	qPELOD-2
Severity scores (external validation)				
Discrimination (95% CI)*	Main	0.84 (0.79 to 0.89)	0.57 (0.51 to 0.63)	0.79 (0.74 to 0.85)
	Sensitivity	0.75 (0.70 to 0.80)	0.54 (0.49 to 0.60)	0.71 (0.66 to 0.76)
Prognostic models (original)				
Discrimination (95% CI)*†	Main	0.90 (0.86 to 0.94)	0.81 (0.76 to 0.86)	0.84 (0.79 to 0.89)
	Sensitivity	0.82 (0.78 to 0.86)	0.76 (0.71 to 0.81)	0.77 (0.72 to 0.81)
Calibration slope (95% CI)*†	Main	0.98 (0.96 to 0.99)	0.96 (0.95 to 0.97)	0.99 (0.98 to 1.00)
	Sensitivity	0.98 (0.96 to 0.99)	0.97 (0.95 to 0.98)	0.99 (0.98 to 1.01)
Calibration intercept (95% CI)*†	Main	-0.05 (-0.09 to -0.01)	-0.09 (-0.13 to -0.05)	-0.01 (-0.05 to -0.03)
	Sensitivity	-0.05 (-0.09 to -0.01)	-0.08 (-0.12 to -0.04)	-0.01 (-0.06 to 0.03)
Prognostic models (updated)				
Discrimination (95% CI)†	Main	0.95 (0.93 to 0.96)	0.93 (0.90 to 0.95)	0.94 (0.93 to 0.96)
	Sensitivity	0.86 (0.82 to 0.90)	0.85 (0.81 to 0.88)	0.85 (0.81 to 0.89)
Calibration slope (95% CI)†	Main	1.05 (1.01 to 1.11)	1.02 (1.01 to 1.17)	1.04 (1.02 to 1.10)
	Sensitivity	1.04 (1.01 to 1.09)	1.02 (1.00 to 1.11)	1.02 (1.01 to 1.09)
Calibration intercept (95% CI)†	Main	0.10 (0.02 to 0.21)	0.03 (0.01 to 0.29)	0.07 (0.03 to 0.20)
	Sensitivity	0.08 (0.02 to 0.20)	0.04 (0.01 to 0.26)	0.06 (0.03 to 0.20)

Table 4.3-8: Classifications at different referral thresholds using the updated clinical prediction models when presentations sent away from the clinic but requiring oxygen within the next 28 days are classified as meeting the primary outcome. Sensitivity analysis includes 2,918 presentations, 134 of which met the primary outcome. A referral threshold of 5% reflects a management strategy whereby any child with a predicted probability of requiring oxygen ≥ 5% is referred. CI = confidence interval; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; mSIRS = modified Systemic Inflammatory Response Syndrome; NLR = negative likelihood ratio; PLR = positive likelihood ratio; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2.

Model	Sensitivity (95% CI)	Specificity (95% CI)	NLR (95% CI)	PLR (95% CI)
Referral threshold = 1%				
LqSOFA	0.95 (0.90 to 0.99)	0.33 (0.11 to 0.58)	0.16 (0.06 to 0.35)	1.46 (1.14 to 2.37)
qPELOD-2	0.96 (0.88 to 0.98)	0.26 (0.06 to 0.62)	0.20 (0.08 to 1.16)	1.35 (1.07 to 2.67)
mSIRS	0.95 (0.90 to 0.99)	0.31 (0.09 to 0.50)	0.16 (0.03 to 0.34)	1.42 (1.11 to 2.00)
Referral threshold = 5%				
LqSOFA	0.75 (0.67 to 0.80)	0.84 (0.82 to 0.86)	0.30 (0.24 to 0.39)	4.57 (3.85 to 5.58)
qPELOD-2	0.75 (0.69 to 0.81)	0.83 (0.81 to 0.86)	0.30 (0.23 to 0.38)	4.49 (3.77 to 5.36)
mSIRS	0.75 (0.68 to 0.81)	0.84 (0.81 to 0.86)	0.30 (0.23 to 0.38)	4.76 (3.93 to 5.61)
Referral threshold = 10%				
LqSOFA	0.65 (0.55 to 0.73)	0.91 (0.88 to 0.93)	0.38 (0.30 to 0.49)	7.01 (5.63 to 9.12)
qPELOD-2	0.65 (0.54 to 0.72)	0.90 (0.87 to 0.93)	0.39 (0.31 to 0.51)	6.62 (5.45 to 9.40)
mSIRS	0.60 (0.54 to 0.75)	0.89 (0.86 to 0.91)	0.39 (0.29 to 0.50)	5.87 (4.93 to 7.16)
Referral threshold = 20%				
LqSOFA	0.48 (0.35 to 0.58)	0.96 (0.95 to 0.97)	0.54 (0.44 to 0.67)	12.46 (9.50 to 16.28)
qPELOD-2	0.46 (0.33 to 0.55)	0.96 (0.95 to 0.97)	0.56 (0.47 to 0.69)	12.46 (9.65 to 16.12)
mSIRS	0.41 (0.28 to 0.52)	0.95 (0.94 to 0.97)	0.61 (0.51 to 0.75)	9.31 (6.81 to 13.04)
Referral threshold = 40%				
LqSOFA	0.18 (0.09 to 0.29)	0.99 (0.99 to 1.00)	0.82 (0.71 to 0.91)	Inf
qPELOD-2	0.18 (0.06 to 0.27)	0.99 (0.99 to 1.00)	0.83 (0.73 to 0.94)	Inf
mSIRS	0.15 (0.04 to 0.26)	1.00 (0.99 to 1.00)	0.85 (0.75 to 0.96)	Inf

4.3.8 Clinical characteristics of the pneumonia cohort

Based on the results from the first phase of the analysis, the points-based LqSOFA score was taken forward for evaluation within the pneumonia cohort. Amongst the 900 attendances included in

the primary analyses, median age was 10.7 months (IQR = 6.0 to 16.3 months) and 50.7% of pneumonia attendances were for male children. Children had been symptomatic for a median of three days (IQR = 2 to 5 days) and fewer than 3% (2.8%; 25/900) had received antibiotics prior to presentation (Table 4.3-9). Admission rate to the clinic was 28.4% (256/900) and one quarter of pneumonia episodes (26.2%; 236/900) met WHO criteria for severe pneumonia at presentation.[152]

Forty-nine (5.4%; 49/900) presentations received supplemental oxygen during their illness visit (met the primary outcome). Unlike the main ARI cohort, abnormal lung auscultation was not associated with requirement of supplemental oxygen therapy (43/48 [90%] vs. 676/839 [81%]), consistent with the fact that all participants in this subgroup had pneumonia. Age-adjusted tachycardia and tachypnoea were less strongly associated with the primary outcome compared to the main ARI cohort and prior antibiotic consumption was not associated with supplemental oxygen therapy (1/49 [2.0%] vs. 24/851 [2.8%]). The LqSOFA score and certain biomarkers reflecting endothelial injury (Ang-2 and sFlt-1), immune activation (IL-8, IL-6, and IL-1ra), and inflammation (PCT) were associated with the primary outcome (Table 4.3-10; Figure 4.3-9).

Table 4.3-9: Baseline characteristics of the pneumonia cohort, stratified by primary outcome status. [#]Respiratory distress defined as head bobbing, tracheal tug, grunting, and/or chest indrawing; [†]abnormal chest auscultation defined as crepitations and/or wheeze; [‡]rectal temperature converted to axillary temperature for neonates and infants;[32] [§]median interval between anthropometric measurement and illness presentation: length = 8 days (IQR = 4 to 12 days); MUAC = 9 days (IQR = 4 to 13 days); weight = 4 days (IQR = 0 to 9 days). [*]Missing data: gestation = 3; birthweight = 7; comorbidity = 4; symptom duration = 3; abnormal lung auscultation = 13; lung crepitations = 15; wheeze = 22; heart rate = 2; respiratory rate = 1; temperature = 1; oxygen saturation = 138; capillary refill time = 63; mental status = 12; weight-for-length z-score = 47; weight-for-age z-score = 46; MUAC-for-age z-score = 98; length-for-age z-score = 3. Bpm = beats / breaths per minute; IQR = interquartile range; MUAC = mid-upper arm circumference.

Baseline characteristic	Overall N = 900 Median (IQR); n/N (%)	Supplemental oxygen		p-value[1]
		No N = 851 Median (IQR); n/N (%)	Yes N = 49 Median (IQR); n/N (%)	
Demographics				
Age (months)	10.7 (6.0, 16.3)	10.8 (6.0, 16.4)	7.0 (3.4, 14.9)	0.010
Male sex	456 / 900 (51%)	432 / 851 (51%)	24 / 49 (49%)	0.80
Birth history				
Gestation (weeks)[*]	39.1 (38.1, 40.0)	39.1 (38.1, 40.0)	38.5 (37.3, 39.5)	0.011
Birthweight (kg)[*]	2.9 (2.6, 3.2)	2.9 (2.6, 3.2)	2.6 (2.0, 3.1)	0.006
Previous medical history				
Number of previous illness visits	4.0 (2.0, 7.0)	4.0 (2.0, 7.0)	3.0 (2.0, 8.0)	0.40
Time since last illness visit (days)	45.0 (15.0, 106.2)	47.0 (17.0, 109.0)	18.0 (1.0, 69.0)	0.002
Known comorbidity[*]	11 / 896 (1.2%)	7 / 849 (0.8%)	4 / 47 (8.5%)	0.002
History of current illness				
Duration of symptoms (days)[*]	3.0 (2.0, 5.0)	3.0 (2.0, 5.0)	3.0 (2.0, 5.0)	0.060
Antibiotics prior to presentation	25 / 900 (2.8%)	24 / 851 (2.8%)	1 / 49 (2.0%)	> 0.90

Baseline characteristic	Overall N = 900 Median (IQR); n/N (%)	Supplemental oxygen		p-value[1]
		No N = 851 Median (IQR); n/N (%)	Yes N = 49 Median (IQR); n/N (%)	
Presenting symptoms and signs				
Fever	780 / 900 (87%)	743 / 851 (87%)	37 / 49 (76%)	0.018
Cough	895 / 900 (99%)	846 / 851 (99%)	49 / 49 (100%)	> 0.90
Respiratory distress[#]	278 / 900 (31%)	233 / 851 (27%)	45 / 49 (92%)	< 0.001
Head bobbing	33 / 900 (3.7%)	15 / 851 (1.8%)	18 / 49 (37%)	< 0.001
Tracheal tug	76 / 900 (8.4%)	53 / 851 (6.2%)	23 / 49 (47%)	< 0.001
Grunting	14 / 900 (1.6%)	5 / 851 (0.6%)	9 / 49 (18%)	< 0.001
Chest indrawing	273 / 900 (30%)	228 / 851 (27%)	45 / 49 (92%)	< 0.001
Abnormal lung auscultation[†*]	719 / 887 (81%)	676 / 839 (81%)	43 / 48 (90%)	0.12
Crepitations[*]	659 / 885 (74%)	621 / 838 (74%)	38 / 47 (81%)	0.30
Wheeze[*]	338 / 878 (38%)	311 / 830 (37%)	27 / 48 (56%)	0.009
Vital signs				
Heart rate (bpm) [*]				
Neonate	149.0 (143.8, 160.0)	154.0 (147.2, 160.0)	145.0 (140.0, 153.5)	0.70
Infant	140.0 (130.0, 148.0)	140.0 (130.0, 148.0)	144.0 (134.0, 158.0)	0.019
Child	132.0 (124.0, 140.0)	132.0 (124.0, 140.0)	136.0 (128.0, 141.5)	0.20
Respiratory rate (bpm) [*]				
Neonate	67.0 (63.5, 77.0)	63.0 (61.0, 65.5)	78.0 (73.0, 80.5)	0.081

Baseline characteristic	Overall N = 900 Median (IQR); n/N (%)	Supplemental oxygen		p-value[1]
		No N = 851 Median (IQR); n/N (%)	Yes N = 49 Median (IQR); n/N (%)	
Infant	58.0 (54.0, 60.8)	58.0 (54.0, 60.0)	60.0 (56.0, 64.0)	0.036
Child	50.0 (46.0, 56.0)	50.0 (46.0, 56.0)	58.0 (46.5, 61.5)	0.046
Axillary temperature (°C)[‡*]	37.1 (36.4, 37.7)	37.1 (36.4, 37.7)	37.3 (36.5, 38.1)	0.13
Oxygen saturation (%)[*]	95.0 (93.0, 96.0)	95.0 (93.0, 96.0)	88.0 (86.0, 92.0)	< 0.001
Capillary refill time > 2 secs[*]	12 / 837 (1.4%)	8 / 792 (1.0%)	4 / 45 (8.9%)	0.003
Not alert[*]	116 / 888 (13%)	87 / 842 (10%)	29 / 46 (63%)	< 0.001
Anthropometrics				
Weight-for-length z-score[§*]	-0.2 (-0.9, 0.6)	-0.2 (-0.9, 0.6)	-0.4 (-1.7, 0.7)	0.067
Weight-for-age z-score[§*]	-1.0 (-1.9, -0.4)	-1.0 (-1.8, -0.4)	-1.9 (-3.4, -0.4)	< 0.001
MUAC-for-age z-score[§*]	0.1 (-0.6, 0.7)	0.1 (-0.5, 0.7)	-1.2 (-2.0, -0.2)	< 0.001
Length-for-age z-score[§*]	-1.6 (-2.5, -0.9)	-1.6 (-2.4, -0.8)	-2.5 (-3.0, -1.3)	0.003

[1]Wilcoxon rank sum test; Pearson's Chi-squared test; Fisher's exact test

Table 4.3-10: Baseline LqSOFA scores and host biomarker concentrations in the pneumonia cohort, stratified by primary outcome status. Ang-1 = angiopoietin-1; Ang-2 = angiopoietin-2; CHI3L1 = chitinase-3-like protein-1; CRP = C-reactive protein; IL-1ra = interleukin-1 receptor antagonist; IL-6 = interleukin-6; IL-8 = interleukin-8; IL-10 = interleukin-10; IP-10 = interferon-γ induced protein-10; IQR = interquartile range; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; PCT = procalcitonin; sFlt-1 = soluble fms-like tyrosine kinase-1; sTREM-1 = soluble triggering receptor expressed on myeloid cells-1; sTNFR-1 = soluble tumour necrosis factor receptor-1.

Baseline characteristic	Overall N = 900 Median (IQR); n/N (%)	Supplemental oxygen		p-value[1]
		No N = 851 Median (IQR); n/N (%)	Yes N = 49 Median (IQR); n/N (%)	
LqSOFA score				
0	713 / 900 (79%)	703 / 851 (83%)	10 / 49 (20%)	< 0.001
1	156 / 900 (17%)	128 / 851 (15%)	28 / 49 (57%)	
2	30 / 900 (3.3%)	20 / 851 (2.4%)	10 / 49 (20%)	
3	1 / 900 (0.1%)	0 / 851 (0%)	1 / 49 (2.0%)	
Host biomarkers				
Endothelial injury				
Ang-1 (pg/ml)	33,913.5 (21,938.0, 46,836.5)	33,897.0 (21,849.0, 46,572.0)	35,880.0 (23,177.0, 50,341.0)	0.50
Ang-2 (pg/ml)	1,538.5 (1,187.0, 1,984.8)	1,485.0 (1,164.5, 1,915.5)	2,261.0 (1,792.0, 3,412.0)	< 0.001
sFlt-1 (pg/ml)	170.0 (145.0, 196.0)	168.0 (144.0, 193.0)	197.0 (171.0, 231.0)	< 0.001
Immune activation				
CHI3L1 (ng/ml)	43.2 (31.0, 61.4)	43.2 (31.1, 60.8)	44.0 (30.0, 67.8)	0.70
IL-1ra (pg/ml)	1,929.5 (1,217.0, 2,949.8)	1,890.0 (1,210.5, 2,794.0)	3,519.0 (1,858.0, 5,136.0)	< 0.001

Baseline characteristic	Overall N = 900 Median (IQR); n/N (%)	Supplemental oxygen		p-value[1]
		No N = 851 Median (IQR); n/N (%)	Yes N = 49 Median (IQR); n/N (%)	
IL-6 (pg/ml)	17.4 (8.9, 36.4)	16.9 (8.7, 34.6)	38.9 (13.8, 70.7)	< 0.001
IL-8 (pg/ml)	35.8 (23.8, 52.0)	34.9 (23.1, 51.0)	53.7 (37.0, 82.1)	< 0.001
IL-10 (pg/ml)	17.0 (11.2, 28.0)	16.9 (11.0, 27.2)	22.5 (13.5, 38.7)	0.015
IP-10 (pg/ml)	745.0 (441.0, 1,371.0)	742.0 (438.5, 1,350.5)	886.0 (551.0, 1,834.0)	0.064
sTNFR-1 (pg/ml)	1,583.5 (1,345.8, 1,906.0)	1,575.0 (1,341.0, 1,894.0)	1,821.0 (1,545.0, 2,196.0)	0.001
sTREM-1 (pg/ml)	399.5 (315.0, 522.0)	398.0 (313.5, 520.5)	423.0 (345.0, 533.0)	0.20
Acute phase proteins				
CRP (mg/l)	20.4 (7.4, 44.9)	20.1 (7.2, 43.8)	31.0 (7.5, 67.9)	0.20
PCT (pg/ml)	240.0 (176.0, 391.2)	236.0 (173.0, 372.5)	417.0 (248.0, 722.0)	< 0.001

[1]Wilcoxon rank sum test; Pearson's Chi-squared test; Fisher's exact test

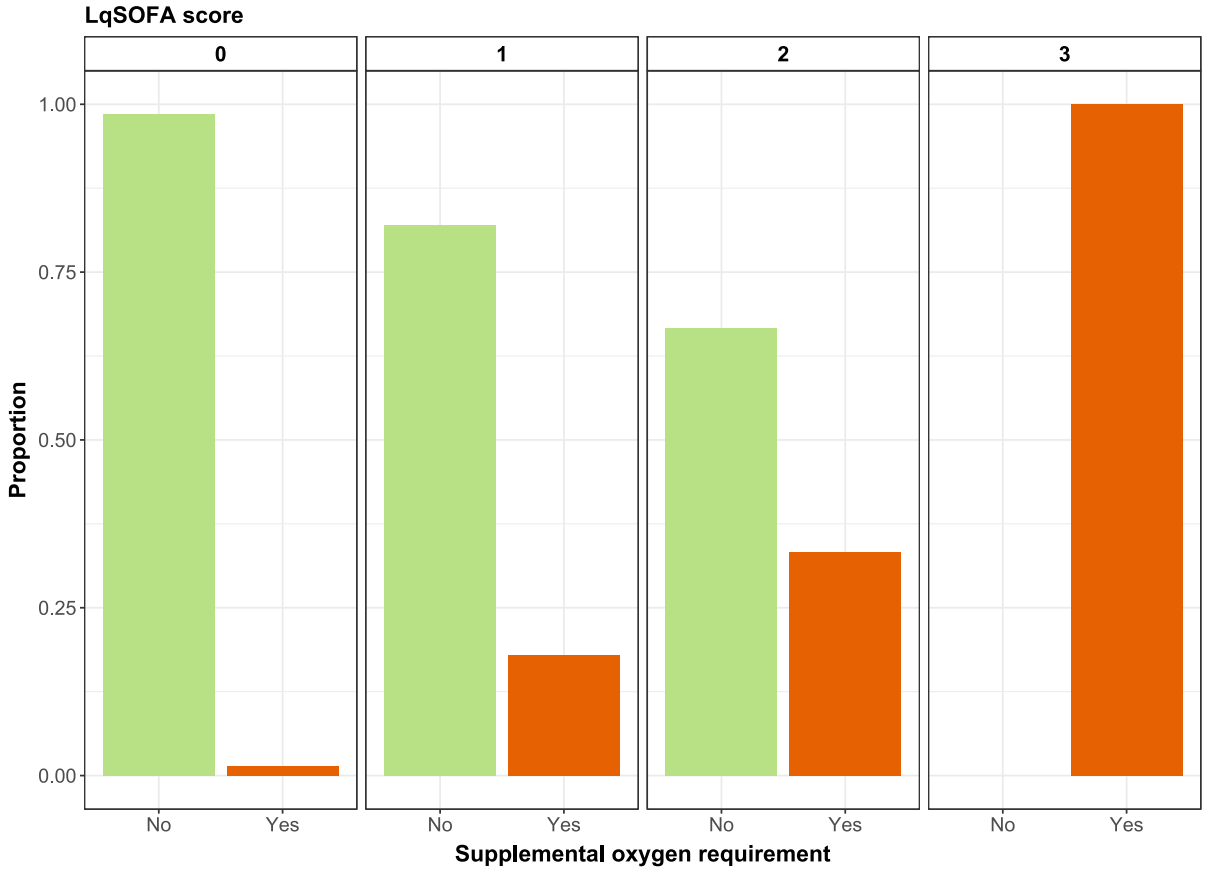


Figure 4.3-9: Relationship between baseline LqSOFA score and observed outcome proportions. LqSOFA = Liverpool quick Sequential Organ Failure Assessment.

4.3.9 Ang-2, sFlt-1, and IL-8 improve discrimination of the LqSOFA score

Ang-2 demonstrated substantially better discrimination than any other biomarker and comparable discrimination to the LqSOFA score (Table 4.3-11; AUC = 0.81 [95% CI = 0.74 to 0.87] vs. AUC = 0.82 [95% CI = 0.76 to 0.88]; p = 0.74). The LqSOFA score outperformed all other biomarkers. Increases in the baseline concentrations of Ang-2, sFlt-1, IL-1ra, and IL-8 corresponded to increases in the probability of supplemental oxygen therapy in a log-linear manner, over the concentration ranges where the majority of the data lay (Figure 4.3-10).

Table 4.3-11: Discrimination of the LqSOFA score and host biomarkers, and improvement in discrimination of the LqSOFA score when adding one of the five most discriminatory biomarkers. Ang-1 = angiopoietin-1; Ang-2 = angiopoietin-2; AUC = area under the receiver operating characteristic curve; CHI3L1 = chitinase-3-like protein-1; CI = confidence interval; CRP = C-reactive protein; IL-1ra = interleukin-1 receptor antagonist; IL-6 = interleukin-6; IL-8 = interleukin-8; IL-10 = interleukin-10; IP-10 = interferon-γ induced protein-10; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; PCT = procalcitonin; sFlt-1 = soluble fms-like tyrosine kinase-1; sTREM-1 = soluble triggering receptor expressed on myeloid cells-1; sTNFR-1 = soluble tumour necrosis factor receptor-1.

Predictor	**AUC** (95% CI)			
	Univariate	**p-value**[1]	**+ LqSOFA**	**p-value**[2]
LqSOFA	0.82 (0.76 to 0.88)	-	NA	–
Ang-2	0.81 (0.74 to 0.87)	0.74	0.91 (0.87 to 0.94)	< 0.001
IL-8	0.72 (0.65 to 0.79)	0.04	0.88 (0.85 to 0.92)	0.001
sFlt-1	0.69 (0.61 to 0.77)	0.02	0.89 (0.86 to 0.92)	< 0.001
PCT	0.69 (0.62 to 0.77)	0.01	0.78 (0.69 to 0.86)	0.015
IL-1ra	0.68 (0.59 to 0.77)	0.004	0.80 (0.72 to 0.88)	0.242
IL-6	0.65 (0.56 to 0.74)	0.001		
sTNFR-1	0.64 (0.55 to 0.72)	0.001		
IL-10	0.60 (0.52 to 0.69)	< 0.001		
IP-10	0.58 (0.49 to 0.66)	< 0.001		
sTREM-1	0.56 (0.49 to 0.63)	< 0.001		
CRP	0.55 (0.46 to 0.64)	< 0.001		
Ang-1	0.53 (0.44 to 0.62)	< 0.001		
CHI3L1	0.52 (0.43 to 0.61)	< 0.001		

[1]DeLong method to compare AUC of LqSOFA vs. AUC of biomarker; [2]DeLong method to compare AUC of LqSOFA vs. AUC of LqSOFA + biomarker

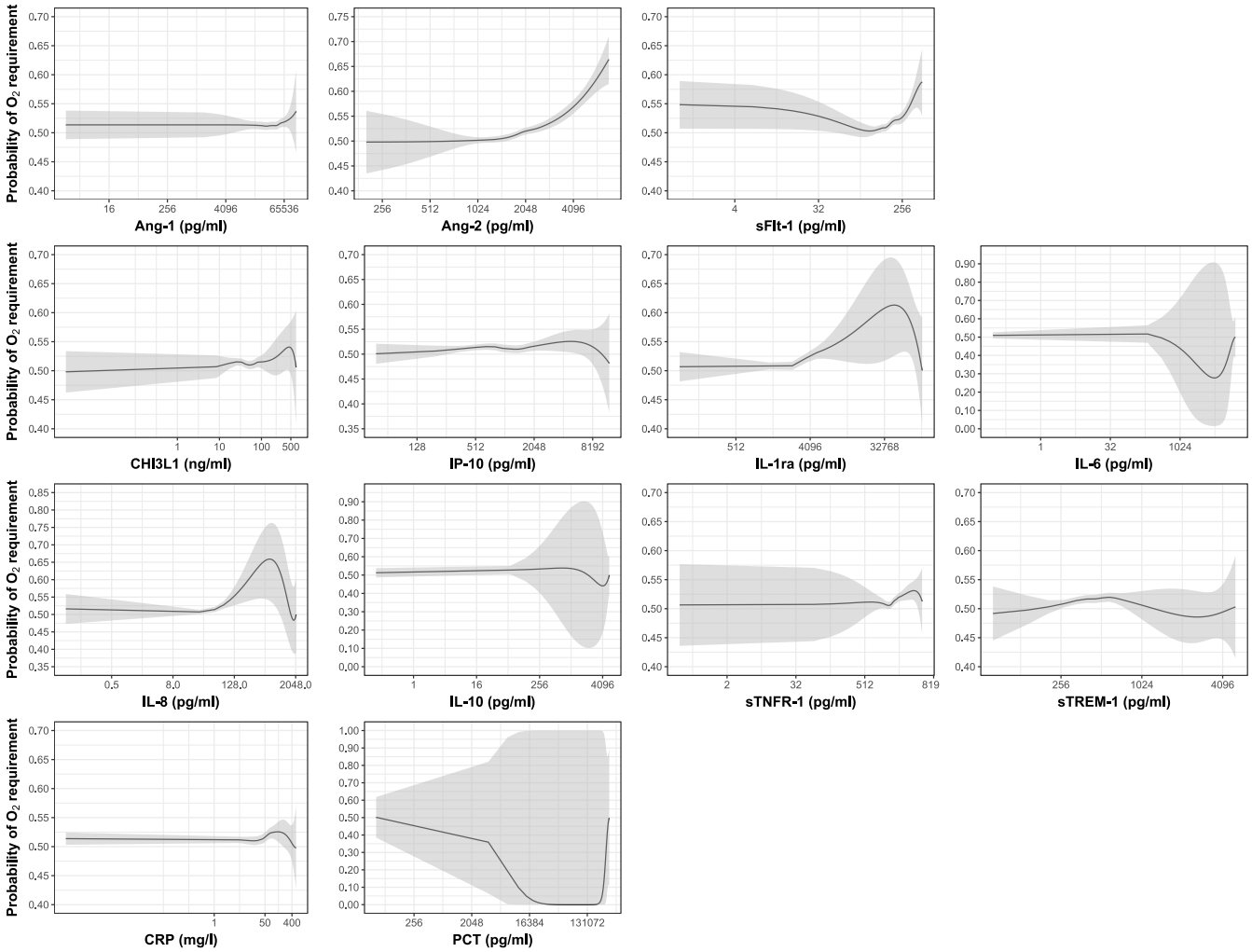


Figure 4.3-10: Association between baseline host biomarker concentrations and the primary outcome. Biomarker concentrations plotted on $\log_2$ scale. Black line = probability of oxygen requirement. Grey shaded areas = 95% confidence interval. Ang-1 = angiopoietin-1; Ang-2 = angiopoietin-2; CHI3L1 = chitinase-3-like protein-1; CRP = C-reactive protein; IL-1ra = interleukin-1 receptor antagonist; IL-6 = interleukin-6; IL-8 = interleukin-8; IL-10 = interleukin-10; IP-10 = interferon-gamma-inducible protein-10; PCT = procalcitonin; sFlt-1 = soluble fms-like tyrosine kinase-1; sTNFR-1 = soluble tumour necrosis factor receptor-1; sTREM-1 = soluble triggering receptor expressed on myeloid cells-1.

Cross-tabulations and box-and-whisker plots of baseline LqSOFA scores and biomarker concentrations suggested that most biomarkers were correlated with the clinical score (Figure 4.3-11 and Table 4.3-12), but polyserial correlation analysis indicated that this association was not strong, with correlation coefficients ranging from 0.01 to 0.41 (Table 4.3-13). Therefore, the five most discriminatory biomarkers (Ang-2, sFlt-1, IL-8, PCT, and IL-1ra) were selected for the next phase of the analysis to investigate whether a combinatorial clinical-biomarker approach, including the LqSOFA

score and one biomarker in a bivariate model, might improve performance. Discrimination improved when either Ang-2 (AUC = 0.91; 95% CI = 0.87 to 0.94; p < 0.001), sFlt-1 (AUC = 0.89; 95% CI = 0.86 to 0.92; p < 0.001), or IL-8 (AUC = 0.88; 95% CI = 0.85 to 0.92; p = 0.001) were combined with the LqSOFA score (Table 4.3-11).

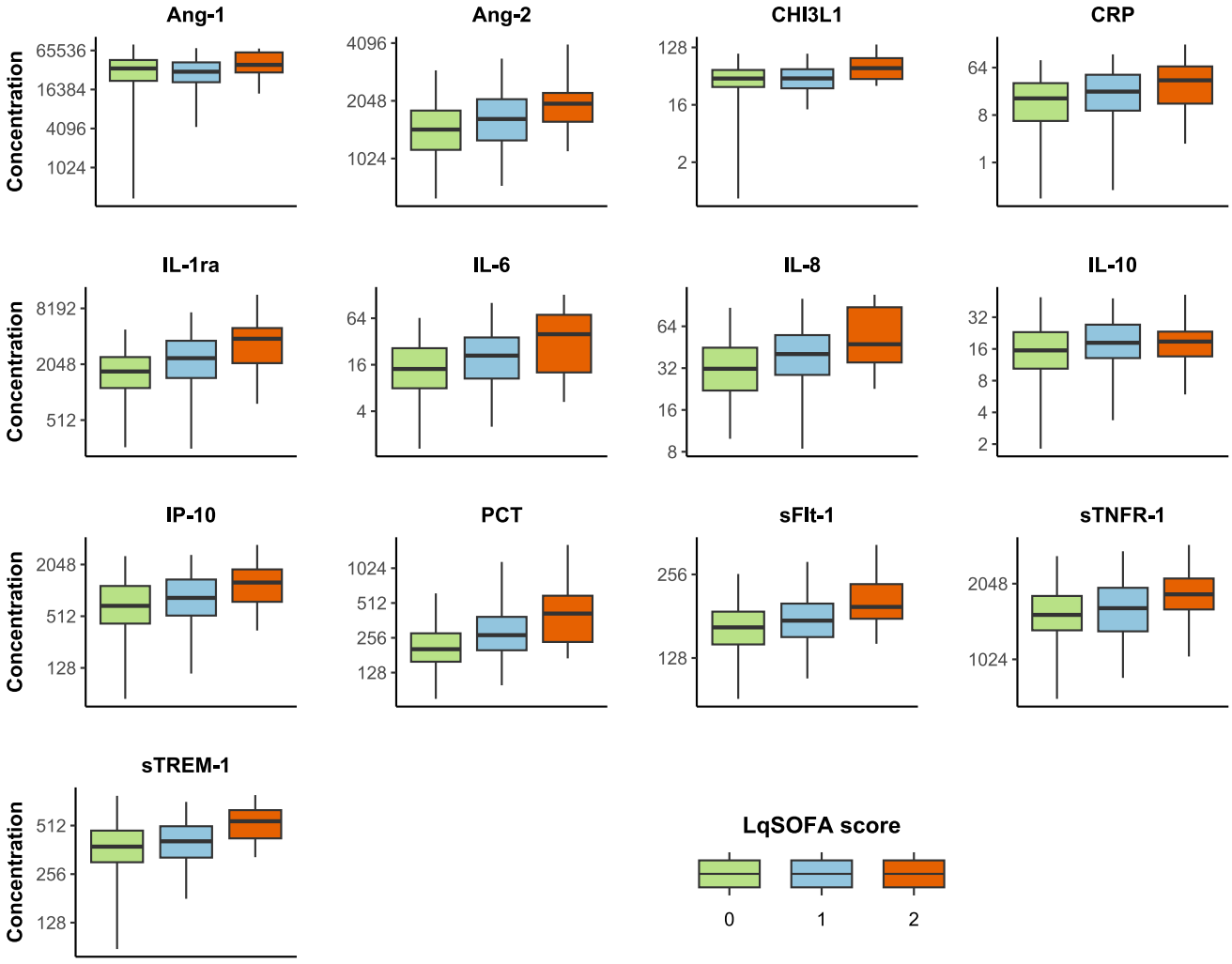


Figure 4.3-11: Baseline biomarker concentrations stratified by baseline LqSOFA score. Log_2 scales used to plot concentrations in pg/ml (Ang-1, Ang-2, IL-1ra, IL-6, Il-8, IL-10, IP-10, PCT, sFlt-1, sTNFR-1, sTREM-1), ng/ml (CHI3L1), and mg/l (CRP). Ang-1 = angiopoietin-1; Ang-2 = angiopoietin-2; CHI3L1 = chitinase-3-like protein-1; CRP = C-reactive protein; IL-1ra = interleukin-1 receptor antagonist; IL-6 = interleukin-6; IL-8 = interleukin-8; IL-10 = interleukin-10; IP-10 = interferon-gamma-inducible protein-10; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; PCT = procalcitonin; sFlt-1 = soluble fms-like tyrosine kinase-1; sTNFR-1 = soluble tumour necrosis factor receptor-1; sTREM-1 = soluble triggering receptor expressed on myeloid cells-1.

Table 4.3-12: Baseline biomarker concentrations stratified by baseline LqSOFA score. Ang-1 = angiopoietin-1; Ang-2 = angiopoietin-2; CHI3L1 = chitinase-3-like protein-1; CRP = C-reactive protein; IL-1ra = interleukin-1 receptor antagonist; IL-6 = interleukin-6; IL-8 = interleukin-8; IL-10 = interleukin-10; IP-10 = interferon-gamma-inducible protein-10; IQR = interquartile range; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; PCT = procalcitonin; sFlt-1 = soluble fms-like tyrosine kinase-1; sTNFR-1 = soluble tumour necrosis factor receptor-1; sTREM-1 = soluble triggering receptor expressed on myeloid cells-1.

Biomarker	LqSOFA score			
	0 **N = 713** Median (IQR)	**1** **N = 156** Median (IQR)	**2** **N = 30** Median (IQR)	**3** **N = 1**
Endothelial injury				
Ang1 (pg/ml)	34,661.0 (22,376.0, 47,143.0)	30,963.0 (21,157.5, 43,494.0)	39,321.5 (29,893.8, 63,535.2)	16,686.0
Ang2 (pg/ml)	1,480.0 (1,150.0, 1,889.0)	1,703.5 (1,288.5, 2,241.2)	2,031.0 (1,592.8, 2,782.0)	3,412.0
sFlt-1 (pg/ml)	166.0 (143.0, 190.0)	174.5 (150.0, 205.2)	197.5 (177.2, 243.2)	337.0
Immune activation				
CHI3L1 (ng/ml)	43.0 (30.9, 59.8)	42.6 (29.3, 60.8)	61.5 (40.8, 99.5)	208.3
IL-1ra (pg/ml)	1,807.0 (1,177.0, 2,673.0)	2,601.0 (1,497.5, 4,113.5)	4,109.0 (2,242.2, 6,739.0)	59,522.0
IL-6 (pg/ml)	15.8 (8.3, 31.4)	26.5 (11.2, 49.9)	57.0 (13.9, 105.5)	180.0
IL-8 (pg/ml)	33.5 (22.2, 49.0)	43.5 (29.8, 58.8)	53.4 (35.1, 97.4)	72.2
IL-10 (pg/ml)	16.4 (10.6, 26.6)	19.8 (13.5, 29.2)	20.9 (13.9, 36.8)	232.0
IP-10 (pg/ml)	706.0 (425.0, 1,289.0)	937.0 (557.8, 1,492.0)	1,398.0 (754.0, 1,882.0)	4,081.0
sTNFR-1 (pg/ml)	1,557.0 (1,340.0, 1,865.0)	1,682.5 (1,340.5, 2,048.2)	1,868.0 (1,617.8, 2,233.8)	4,555.0
sTREM-1 (pg/ml)	390.0 (310.0, 509.0)	427.0 (326.5, 530.2)	550.0 (426.5, 675.8)	652.0
Acute phase proteins				
CRP (mg/l)	18.4 (6.7, 40.0)	27.9 (10.2, 56.5)	40.5 (13.5, 85.1)	186.0
PCT (pg/ml)	222.0 (168.0, 350.0)	301.0 (208.2, 694.5)	460.5 (258.2, 1,342.5)	19,459.0

Table 4.3-13: Correlation between baseline biomarker concentrations and baseline LqSOFA score. Ang-1 = angiopoietin-1; Ang-2 = angiopoietin-2; CHI3L1 = chitinase-3-like protein-1; CRP = C-reactive protein; IL-1ra = interleukin-1 receptor antagonist; IL-6 = interleukin-6; IL-8 = interleukin-8; IL-10 = interleukin-10; IP-10 = interferon-gamma-inducible protein-10; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; PCT = procalcitonin; sFlt-1 = soluble fms-like tyrosine kinase-1; sTNFR-1 = soluble tumour necrosis factor receptor-1; sTREM-1 = soluble triggering receptor expressed on myeloid cells-1.

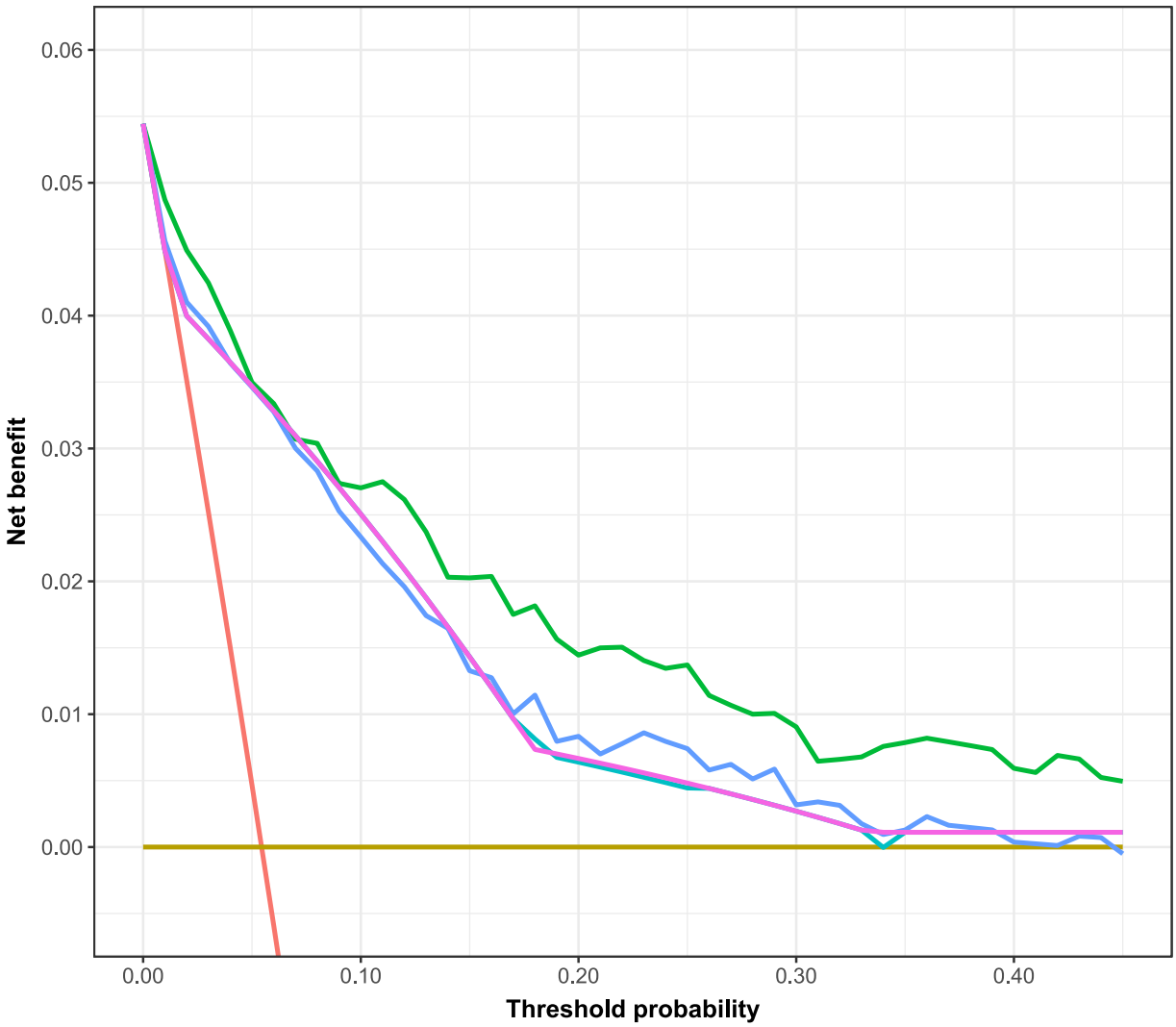
Biomarker	Polyserial correlation coefficient
Endothelial injury	
Ang-1	0.007
Ang-2	0.322
sFlt-1	0.280
Immune activation	
CHI3L1	0.187
IL-1ra	0.409
IL-6	0.145
IL-8	0.177
IL-10	0.093
IP-10	0.210
sTNFR-1	0.279
sTREM-1	0.065
Acute phase proteins	
PCT	0.163
CRP	0.196

4.3.10 Combining Ang-2 and LqSOFA improves sensitivity of community triage for pneumonia

Recognising that better discrimination does not necessarily translate into greater clinical utility, decision curve analyses were used to account for the fact that the relative value of a true negative (correctly identifying a child with pneumonia who could be safely managed in the community) and a false negative (misclassifying a child with pneumonia who would require supplemental oxygen) is context-dependent.[64,65] The net benefit of the LqSOFA score alone was compared to that of the LqSOFA score combined with either Ang-2, sFlt-1, or IL-8, over a range of clinically-plausible referral

thresholds.[191] At referral thresholds beyond ~8% (a strategy equivalent to referring any child in whom the predicted risk of requiring supplemental oxygen is ≥ 8%), addition of Ang-2 to the LqSOFA score provided greater utility than the LqSOFA score alone (Figure 4.3-12).

Figure 4.3-12: Utility of the LqSOFA score alone and combined with Ang-2, sFlt-1, or IL-8 for the identification of children suitable for community-based management of pneumonia. The net benefit of the LqSOFA score (pink line) is compared to the LqSOFA score combined with either Ang-2 (green line), IL-8 (turquoise line), or sFlt-1 (blue line), and a 'refer-all' (red line) and 'refer-none' (brown line) approach. A threshold probability of 5% is equivalent to a management strategy whereby any child with a predicted risk of oxygen requirement ≥ 5% is referred (i.e., a scenario where the value of one correct referral is equivalent to 19 incorrect referrals or a NNR of 20). Moving from left to right along the x-axis (increasing referral threshold) reflects increasing 'penalisation' of an incorrect referral (false positive), indicative of contexts in which referrals may be more challenging or costly. Ang-2 = angiopoietin-2; IL-8 = interleukin-8; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; NNR = number needed to refer; sFlt-1 = soluble fms-like tyrosine kinase-1.



Examining predicted classifications across these referral thresholds suggested that a combinatorial approach utilising Ang-2 and LqSOFA could reduce the number of children incorrectly identified as safe for community-based management by ~10-30% compared to the LqSOFA score alone (ratios of 18 to 71:1 vs. 23 to 77:1 for correct to incorrect cases managed in the community), at the cost of a small increase in the proportion of unnecessary referrals (Table 4.3-14). Addition of neither sFlt-1 nor IL-8 provided greater net benefit than the LqSOFA score alone at any referral threshold (Figure 4.3-12).

Table 4.3-14: Classifications at different referral thresholds using the LqSOFA score alone and the LqSOFA score combined with Ang-2. A referral threshold of 5% reflects a management strategy whereby any child with a predicted probability of requiring oxygen ≥ 5% is referred. *LqSOFA scores converted to predicted probabilities to facilitate comparison with the combinatorial approach (LqSOFA score + Ang-2); referral thresholds (predicted probabilities) approximate to the following LqSOFA scores: 1% ≥ 0; 5% ≥ 1; 20% ≥ 2; 40% ≥ 3. Ang-2 = angiopoietin-2; CI = confidence interval; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; NLR = negative likelihood ratio; PLR = positive likelihood ratio.

Referral threshold	Sensitivity (95% CI)	Specificity (95% CI)	NLR (95% CI)	PLR (95% CI)	Cases referred (%)	Cases managed in community (%)	Ratio of Incorrect to Correct referrals	Ratio of Correct to Incorrect cases managed in community
LqSOFA score*								
1%	1.00 (NA)	0.00 (NA)	0.00 (NA)	1.00 (NA)	900 (100%)	0 (0%)	17 to 1	NA
5%	0.79 (0.67 to 0.89)	0.83 (0.80 to 0.85)	0.25 (0.13 to 0.40)	4.56 (3.63 to 5.54)	184 (20.4%)	716 (79.6%)	4 to 1	71 to 1
20%	0.23 (0.13 to 0.35)	0.98 (0.97 to 0.99)	0.79 (0.67 to 0.90)	9.95 (4.62 to 16.91)	28 (3.1%)	872 (96.9%)	2 to 1	22 to 1
40%	0.20 (0.00 to 0.33)	0.98 (0.97 to 1.00)	0.82 (0.68 to 1.00)	Inf (NA)	1 (0.1%)	899 (99.9%)	0 to 1	18 to 1
LqSOFA score + Ang-2								
1%	1.00 (0.91 to 1.00)	0.20 (0.00 to 0.59)	0.00 (0.00 to 0.13)	1.32 (1.00 to 2.66)	559 (62.1%)	341 (37.9%)	10 to 1	Inf to 1
5%	0.85 (0.69 to 0.93)	0.82 (0.76 to 0.87)	0.19 (0.09 to 0.37)	4.73 (3.74 to 7.50)	202 (22.4%)	698 (77.6%)	4 to 1	77 to 1
20%	0.43 (0.27 to 0.56)	0.96 (0.95 to 0.97)	0.59 (0.45 to 0.76)	11.73 (7.26 to 19.28)	58 (6.4%)	842 (93.6%)	2 to 1	30 to 1
40%	0.29 (0.08 to 0.46)	0.99 (0.97 to 0.99)	0.72 (0.55 to 0.93)	21.28 (7.56 to 37.98)	22 (2.4%)	878 (97.6%)	1 to 1	23 to 1

4.3.11 Head-to-head comparison suggests Ang-2 cannot replace LqSOFA

Although policy makers and healthcare workers are most likely to adopt biomarker tests as adjuncts to clinical risk stratification, they have been proposed as standalone replacements in settings with limited capacity for collection of clinical data.[192] Therefore, in light of the fact that Ang-2 and LqSOFA had demonstrated comparable discrimination, but recognising that discrimination does not equate to clinical utility,[64] the net benefits and predicted classifications of the LqSOFA score and Ang-2 were compared using decision curve analyses. These analyses indicated LqSOFA to have better clinical utility than Ang-2 and confirmed the overall superiority of the combinatorial approach (Figure 4.3-13; Table 4.3-15).

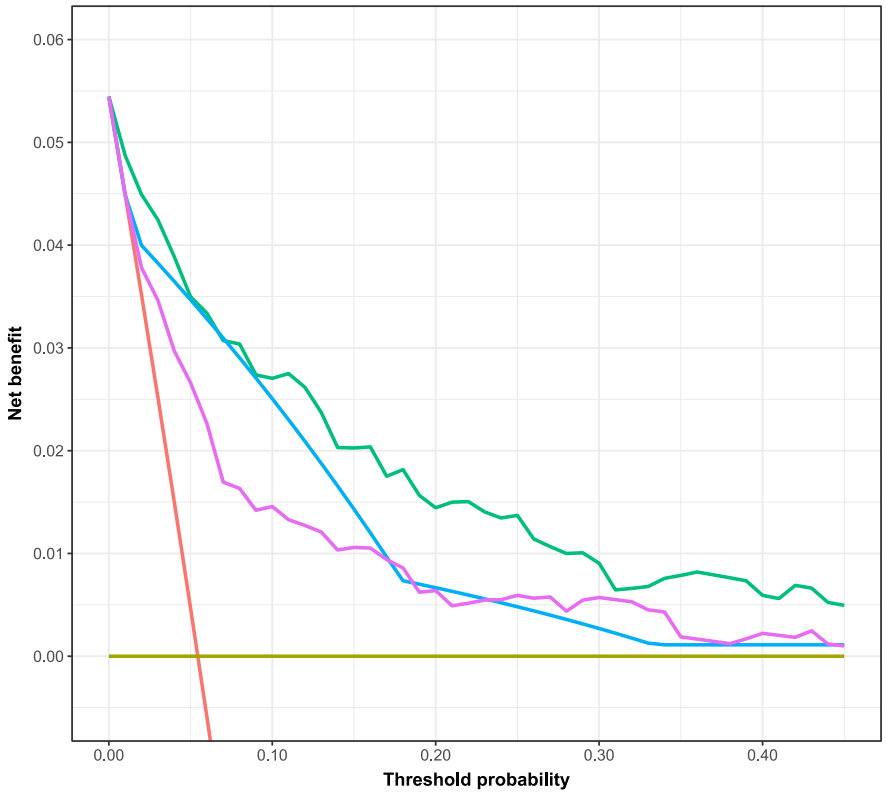


Figure 4.3-13: Utility of the LqSOFA score alone, Ang-2 alone, and a combinatorial approach for the identification of children suitable for community-based management of pneumonia. The net benefit of Ang-2 (pink line) is compared to the LqSOFA score (blue line), the combination of Ang-2 and LqSOFA (green line), and a 'refer-all' (red line) and 'refer-none' (brown line) approach. A threshold probability of 5% is equivalent to a management strategy whereby any child with a predicted risk of oxygen requirement ≥ 5% is referred (i.e., a scenario where the value of one correct referral is equivalent to 19 incorrect referrals or a number-needed-to-refer of 20). Moving from left to right along the x-axis (increasing referral threshold) reflects increasing 'penalisation' of an incorrect referral (false positive), indicative of contexts in which referrals may be more challenging or costly. Ang-2 = angiopoietin-2; LqSOFA = Liverpool quick Sequential Organ Failure Assessment.

Table 4.3-15: Classifications at different referral thresholds using the LqSOFA score alone and Ang-2 alone. A referral threshold of 5% reflects a management strategy whereby any child with a predicted probability of requiring oxygen ≥ 5% is referred. *LqSOFA scores converted to predicted probabilities to facilitate comparison with Ang-2; referral thresholds (predicted probabilities) approximate to the following LqSOFA scores: 1% ≥ 0; 5% ≥ 1; 20% ≥ 2; 40% ≥ 3. Ang-2 = angiopoietin-2; CI = confidence interval; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; NLR = negative likelihood ratio; PLR = positive likelihood ratio.

Referral threshold	Sensitivity (95% CI)	Specificity (95% CI)	NLR (95% CI)	PLR (95% CI)	Cases referred (%)	Cases managed in community (%)	Ratio of Incorrect to Correct referrals	Ratio of Correct to Incorrect cases managed in community
LqSOFA score*								
1%	1.00 (NA)	0.00 (NA)	0.00 (NA)	1.00 (NA)	900 (100%)	0 (0%)	17 to 1	NA
5%	0.79 (0.67 to 0.89)	0.83 (0.80 to 0.85)	0.25 (0.13 to 0.40)	4.56 (3.63 to 5.54)	184 (20.4%)	716 (79.6%)	4 to 1	71 to 1
20%	0.23 (0.13 to 0.35)	0.98 (0.97 to 0.99)	0.79 (0.67 to 0.90)	9.95 (4.62 to 16.91)	28 (3.1%)	872 (96.9%)	2 to 1	22 to 1
40%	0.20 (0.00 to 0.33)	0.98 (0.97 to 1.00)	0.82 (0.68 to 1.00)	Inf (NA)	1 (0.1%)	899 (99.9%)	0 to 1	18 to 1
Ang-2								
1%	1.00 (NA)	0.00 (0.00 to 0.31)	0.00 (NA)	1.00 (1.00 to 1.45)	899 (99.9%)	1 (0.1%)	17 to 1	Inf to 1
5%	0.67 (0.45 to 0.82)	0.77 (0.59 to 0.87)	0.43 (0.26 to 0.66)	3.07 (2.04 to 4.92)	225 (25.0%)	675 (75.0%)	6 to 1	44 to 1
20%	0.24 (0.10 to 0.42)	0.97 (0.96 to 0.99)	0.78 (0.60 to 0.92)	9.65 (5.18 to 18.48)	32 (3.6%)	868 (96.4%)	2 to 1	22 to 1
40%	0.13 (0.03 to 0.29)	0.99 (0.99 to 1.00)	0.88 (0.72 to 0.98)	18.34 (5.72 to 56.08)	12 (1.3%)	888 (98.7%)	1 to 1	20 to 1

4.3.12 Differing roles for host biomarkers in diagnosis and prognosis of pneumonia severity

To determine how and why Ang-2 improved performance of LqSOFA, the differential contribution of clinical assessment and host biomarkers to diagnosis and prognosis of the severity of pneumonia was explored. The ability of the five most discriminatory biomarkers and the LqSOFA score to distinguish children who were: (a) hypoxaemic at presentation (SpO_2 < 90%; diagnostic outcome); (b) not hypoxaemic at presentation but who required supplemental oxygen during their illness visit (prognostic outcome 1); and (c) not hypoxaemic at presentation but who required supplemental oxygen at any time in the next 28 days (prognostic outcome 2), was evaluated. Discrimination of the LqSOFA score deteriorated (AUC = 0.85 to 0.66) as the time horizon for prediction became more distant (moved from diagnosis to prognosis), whereas discrimination of the biomarkers appeared stable (Table 4.3-16). Decision curve analyses confirmed Ang-2 to have greater prognostic utility (net benefit) than either IL-8 or the LqSOFA score for the most distal outcome (prognostic outcome 2; Figure 4.3-14).

Table 4.3-16: Discrimination of the LqSOFA score and host biomarkers for diagnosis and prognosis of the severity of childhood pneumonia. Children with hypoxaemia (SpO_2 < 90%) at presentation excluded for assessment of prognostic outcomes. Attendances with missing outcome status excluded: diagnostic outcome assessed in 766 presentations (734 controls and 32 cases); prognostic outcome 1 assessed in 869 presentations (846 controls and 23 cases); prognostic outcome 2 assessed in 827 presentations (789 controls and 38 cases). Ang-2 = angiopoietin-2; AUC = area under the receiver operating characteristic curve; CI = confidence interval; IL-1ra = interleukin-1 receptor antagonist; IL-8 = interleukin-8; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; PCT = procalcitonin; sFlt-1 = soluble fms-like tyrosine kinase-1.

DIAGNOSTIC OUTCOME SpO_2 **< 90% at presentation**		**PROGNOSTIC OUTCOME 1** Supplemental O_2 **during illness visit**		**PROGNOSTIC OUTCOME 2** Supplemental $O_2 \leq$ **28 days after presentation**	
Predictor	**AUC** (95% CI)	**Predictor**	**AUC** (95% CI)	**Predictor**	**AUC** (95% CI)
LqSOFA	0.85 (0.78 to 0.91)	**Ang-2**	0.80 (0.70 to 0.90)	**IL-8**	0.75 (0.67 to 0.82)
Ang-2	0.74 (0.65 to 0.84)	**IL-8**	0.75 (0.64 to 0.85)	**Ang-2**	0.73 (0.64 to 0.82)
sFlt-1	0.71 (0.62 to 0.81)	**LqSOFA**	0.75 (0.64 to 0.85)	**IL-1ra**	0.70 (0.61 to 0.80)
IL-1ra	0.71 (0.61 to 0.81)	**PCT**	0.68 (0.56 to 0.80)	**PCT**	0.67 (0.59 to 0.76)
IL-8	0.70 (0.62 to 0.77)	**IL-1ra**	0.66 (0.53 to 0.79)	**LqSOFA**	0.66 (0.57 to 0.74)
PCT	0.69 (0.59 to 0.78)	**sFlt-1**	0.60 (0.47 to 0.73)	**sFlt-1**	0.65 (0.56 to 0.75)

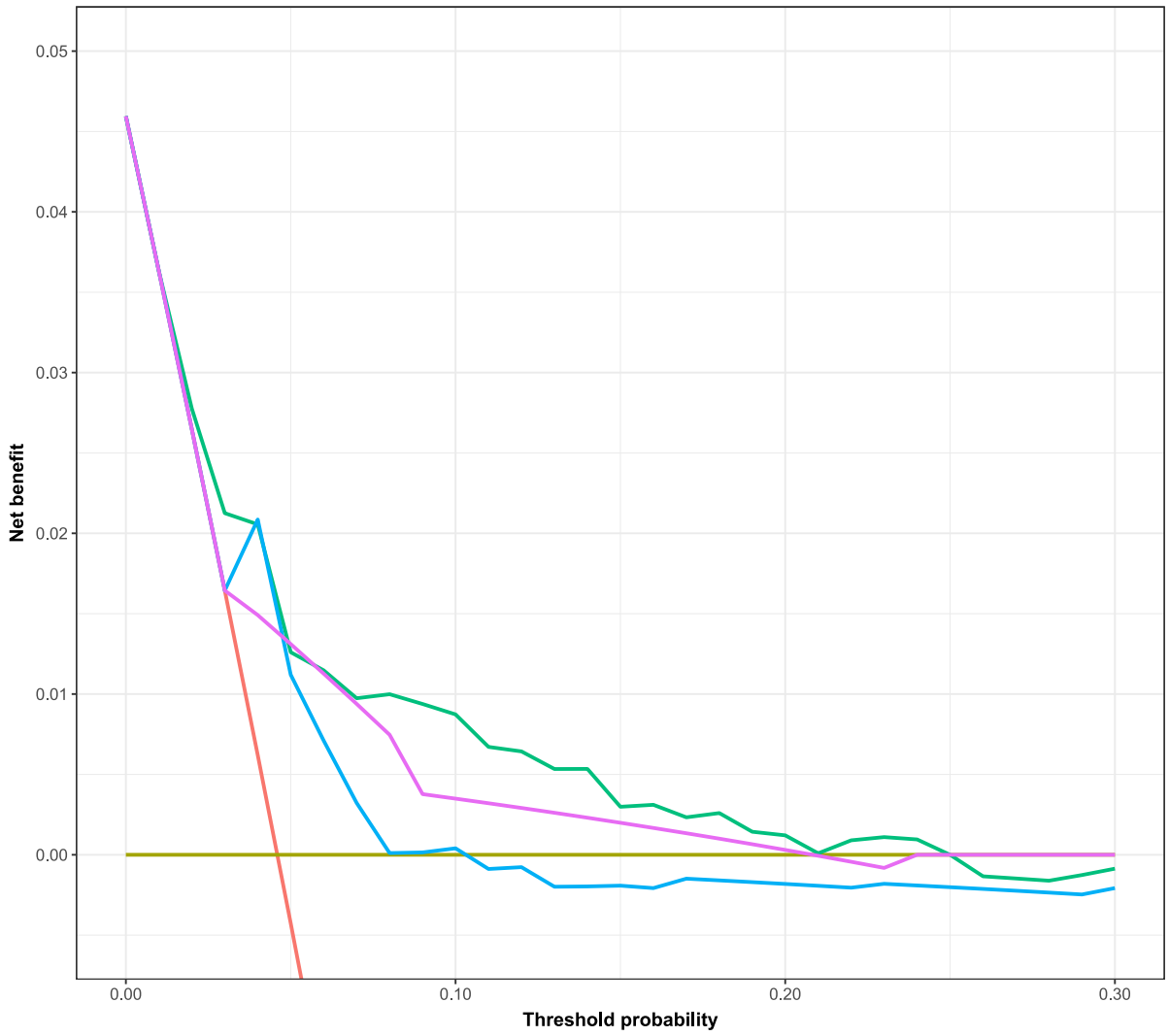


Figure 4.3-14: Prognostic utility of Ang-2, IL-8, and the LqSOFA score to predict requirement of supplemental oxygen therapy within the next 28 days for children presenting with pneumonia. The net benefit of the LqSOFA score (pink line) is compared to Ang-2 (green line) and IL-8 (blue line), and a 'refer-all' (red line) and 'refer-none' (brown line) approach. A threshold probability of 5% is equivalent to a management strategy whereby any child with a predicted risk of oxygen requirement ≥ 5% is referred (i.e., a scenario where the value of one correct referral is equivalent to 19 incorrect referrals or a number-needed-to-refer of 20). Moving from left to right along the x-axis (increasing referral threshold) reflects increasing 'penalisation' of an incorrect referral (false positive), indicative of contexts in which referrals may be more challenging or costly. Ang-2 = angiopoietin-2; IL-8 = interleukin-8; LqSOFA = Liverpool quick Sequential Organ Failure Assessment.

4.3.13 An algorithm for the safe outpatient management of childhood pneumonia

Finally, recognising that it might not be practical to measure a biomarker in all children presenting with pneumonia at the community level, the LqSOFA score and measurements of Ang-2 were combined in a sequential manner to generate a simple proof-of-concept algorithm for triage of

all children presenting with pneumonia (Figure 4.3-15). Since sensitivity would usually be prioritised for community-based triage, the cost of misclassifying a child who would require supplemental oxygen was prespecified as 10 times that of the cost of misclassifying a child who would not, reflecting a pragmatic approximation for the upper limit of the NNR from a typical resource-limited primary care setting. The algorithm demonstrated modest but potentially important rule-out (NLR = 0.28; sensitivity = 0.78) and rule-in potential (PLR = 3.66; specificity = 0.79), for the identification of children suitable for home-based management of pneumonia.[87]

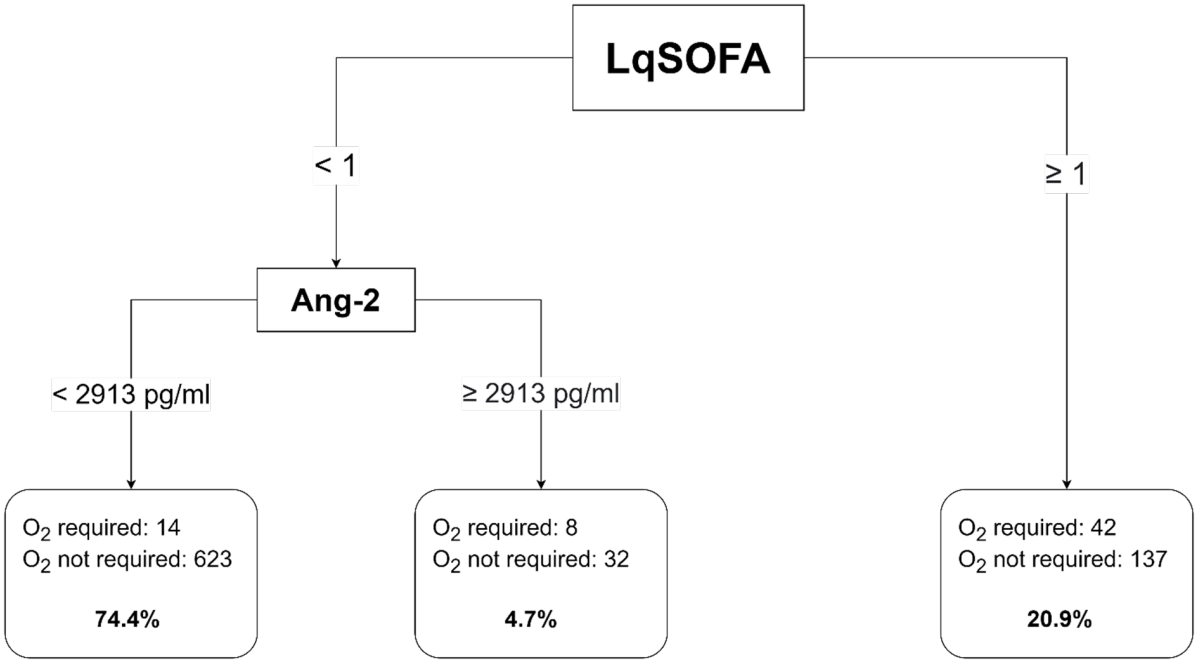


Figure 4.3-15: Algorithm for the triage of childhood pneumonia in resource-limited primary care contexts. The cost of misclassifying a child who would require supplemental oxygen was prespecified as 10 times that of misclassifying a child who would not require supplemental oxygen. Cut-points selected by recursive partitioning. The maximum level of tree depth was set at two and the minimum number of observations per node was set at 20. Percentages indicate proportion of cohort in each terminal node. Ang-2 = angiopoietin-2; LqSOFA = Liverpool quick Sequential Organ Failure Assessment.

4.3.14 Sensitivity analyses

Samples collected on Saturday evenings and Sundays were centrifuged and serum stored at 2-8°C for up to 48 hours prior to being transferred to definitive -80°C storage on Monday evenings. Although most biomarkers are stable at refrigeration temperatures for short periods following

centrifugation,[193] sensitivity analyses excluding weekend presentations were performed. Similar results were obtained (Table 4.3-17).

Table 4.3-17: Ability of host biomarkers to discriminate children who required supplemental oxygen, excluding presentations on a Saturday or Sunday. Primary analysis includes 900 presentations, 49 of which met the primary outcome. Sensitivity analysis includes 696 presentations, 34 of which met the primary outcome. Ang-1 = angiopoietin-1; Ang-2 = angiopoietin-2; AUC = area under the receiver operating characteristic curve; CHI3L1 = chitinase-3-like protein-1; CI = confidence interval; CRP = C-reactive protein; IL-1ra = interleukin-1 receptor antagonist; IL-6 = interleukin-6; IL-8 = interleukin-8; IL-10 = interleukin-10; IP-10 = interferon-gamma-inducible protein-10; PCT = procalcitonin; sFlt-1 = soluble fms-like tyrosine kinase-1; sTNFR-1 = soluble tumour necrosis factor receptor-1; sTREM-1 = soluble triggering receptor expressed on myeloid cells-1.

Biomarker	**AUC** (95% CI)	
	Primary analysis	**Sensitivity analysis**
Ang-2	0.81 (0.74 to 0.87)	0.81 (0.75 to 0.88)
IL-8	0.72 (0.65 to 0.79)	0.70 (0.62 to 0.79)
sFlt-1	0.69 (0.61 to 0.77)	0.66 (0.56 to 0.76)
PCT	0.69 (0.62 to 0.77)	0.67 (0.57 to 0.76)
IL-1ra	0.68 (0.59 to 0.77)	0.63 (0.52 to 0.74)
IL-6	0.65 (0.56 to 0.74)	0.64 (0.53 to 0.74)
sTNFR-1	0.64 (0.55 to 0.72)	0.59 (0.49 to 0.69)
IL-10	0.60 (0.52 to 0.69)	0.56 (0.46 to 0.67)
IP-10	0.58 (0.49 to 0.66)	0.54 (0.44 to 0.64)
sTREM-1	0.56 (0.49 to 0.63)	0.55 (0.47 to 0.63)
CRP	0.55 (0.46 to 0.64)	0.55 (0.44 to 0.66)
Ang-1	0.53 (0.44 to 0.62)	0.56 (0.45 to 0.66)
CHI3L1	0.52 (0.43 to 0.61)	0.53 (0.42 to 0.64)

For the diagnostic outcome, 15.3% (139/905) of pneumonia presentations were excluded as baseline SpO_2 measurements were missing. Missingness was unlikely to be at random as measurement of SpO_2 was a prerequisite for children considered for supplemental oxygen therapy. In support of this, 89.2% (124/139) of missing values were in outpatients and no presentations missing baseline SpO_2

received supplemental oxygen. A sensitivity analysis assuming that all presentations missing baseline SpO_2 measurements were not hypoxaemic (i.e., had $SpO_2 \geq 90\%$) produced almost identical results (Table 4.3-18).

Table 4.3-18: Discrimination of the LqSOFA score and host biomarkers for the diagnosis of severe pneumonia, assuming all presentations with missing baseline SpO_2 were not hypoxaemic. Main analysis includes 766 presentations, 32 of which met the primary outcome. Sensitivity analysis includes 905 presentations, 32 of which met the primary outcome. Only the five top-performing biomarkers from the primary analysis were evaluated for the secondary diagnostic and prognostic outcomes. Ang-2 = angiopoietin-2; AUC = area under the receiver operating characteristic curve; CI = confidence interval; IL-1ra = interleukin-1 receptor antagonist; IL-8 = interleukin-8; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; PCT = procalcitonin; sFlt-1 = soluble fms-like tyrosine kinase-1.

Predictor	**AUC** (95% CI)	
	Main analysis	**Sensitivity analysis**
LqSOFA	0.85 (0.78 to 0.91)	0.84 (0.77 to 0.91)
Ang-2	0.74 (0.65 to 0.84)	0.75 (0.65 to 0.85)
sFlt-1	0.71 (0.62 to 0.81)	0.72 (0.63 to 0.81)
IL-1ra	0.71 (0.61 to 0.81)	0.72 (0.61 to 0.82)
IL-8	0.70 (0.62 to 0.77)	0.70 (0.62 to 0.77)
PCT	0.69 (0.59 to 0.78)	0.69 (0.60 to 0.78)

Serum samples were only available for 77.8% (905/1,164) of pneumonia presentations. A comparison of the clinical characteristics of those with and without serum available indicated more severe presentations amongst those without serum available (Table 4.3-19).

Table 4.3-19: Comparison of baseline characteristics amongst pneumonia presentations with and without a serum sample available. Presentations without missingness for primary outcome analysed. [#]Respiratory distress defined as head bobbing, tracheal tug, grunting and/or chest indrawing; [†]abnormal chest auscultation defined as crepitations and/or wheeze; [‡]rectal temperature converted to axillary temperature for neonates and infants.[32] [*]Missing data: gestation = 3; birthweight = 7; comorbidity = 5; symptom duration = 8; fever = 1; abnormal lung auscultation = 23; lung crepitations = 27; wheeze = 36; heart rate = 5; respiratory rate = 2; temperature = 1; oxygen saturation = 187; capillary refill time = 100; mental status = 18; weight-for-length z-score = 59 weight-for-age z-score = 55; MUAC-for-age z-score = 135; length-for-age z-score = 5. IQR = interquartile range; MUAC = mid-upper arm circumference.

Baseline characteristic	Overall N = 900 Median (IQR); n/N (%)	Serum sample available		p-value[1]
		No N = 256 Median (IQR); n/N (%)	Yes N = 900 Median (IQR); n/N (%)	
Demographics				
Age (months)	10.3 (5.9, 15.6)	9.4 (5.7, 13.9)	10.7 (6.0, 16.3)	0.022
Male sex	605 / 1,156 (52%)	149 / 256 (58%)	456 / 900 (51%)	0.033
Birth history				
Gestation (weeks)[*]	39.1 (38.1, 40.0)	39.1 (37.6, 40.0)	39.1 (38.1, 40.0)	0.077
Birthweight (kg)[*]	2.9 (2.6, 3.2)	2.8 (2.5, 3.2)	2.9 (2.6, 3.2)	0.40
Previous medical history				
Number of previous illness visits	4.0 (2.0, 7.0)	5.0 (3.0, 7.0)	4.0 (2.0, 7.0)	0.059
Time since last illness visit (days)	35.5 (7.0, 96.2)	12.5 (3.0, 47.2)	45.0 (15.0, 106.2)	< 0.001
Known comorbidity[*]	17 / 1,151 (1.5%)	6 / 255 (2.4%)	11 / 896 (1.2%)	0.20
History of current illness				
Duration of symptoms (days)[*]	3.0 (2.0, 5.0)	3.0 (2.0, 6.0)	3.0 (2.0, 5.0)	0.70

Baseline characteristic	Overall N = 900 Median (IQR); n/N (%)	Serum sample available		p-value[1]
		No N = 256 Median (IQR); n/N (%)	Yes N = 900 Median (IQR); n/N (%)	
Antibiotics prior to presentation	79 / 1,156 (6.8%)	54 / 256 (21%)	25 / 900 (2.8%)	< 0.001
Presenting symptoms and signs				
Fever[*]	983 / 1,155 (85%)	203 / 255 (80%)	780 / 900 (87%)	0.005
Cough	1,148 / 1,156 (99%)	253 / 256 (99%)	895 / 900 (99%)	0.40
Respiratory distress[#]	403 / 1,156 (35%)	125 / 256 (49%)	278 / 900 (31%)	< 0.001
Head bobbing	46 / 1,156 (4.0%)	13 / 256 (5.1%)	33 / 900 (3.7%)	0.30
Tracheal tug	113 / 1,156 (9.8%)	37 / 256 (14%)	76 / 900 (8.4%)	0.004
Grunting	21 / 1,156 (1.8%)	7 / 256 (2.7%)	14 / 900 (1.6%)	0.30
Chest indrawing	395 / 1,156 (34%)	122 / 256 (48%)	273 / 900 (30%)	< 0.001
Abnormal lung auscultation[†*]	916 / 1,133 (81%)	197 / 246 (80%)	719 / 887 (81%)	0.70
Crepitations[*]	833 / 1,129 (74%)	174 / 244 (71%)	659 / 885 (74%)	0.30
Wheeze[*]	432 / 1,120 (39%)	94 / 242 (39%)	338 / 878 (38%)	> 0.9
Vital signs				
Heart rate (bpm)[*]				
Neonate	150.0 (143.8, 161.0)	159.0 (147.5, 168.5)	149.0 (143.8, 160.0)	0.40
Infant	140.0 (130.0, 148.0)	140.0 (130.0, 150.0)	140.0 (130.0, 148.0)	0.30
Child	132.0 (124.0, 140.0)	136.0 (126.0, 148.0)	132.0 (124.0, 140.0)	0.024
Respiratory rate (bpm)[*]				
Neonate	64.0 (57.5, 71.5)	52.0 (42.5, 58.2)	67.0 (63.5, 77.0)	0.050

Baseline characteristic	Overall N = 900 Median (IQR); n/N (%)	Serum sample available		p-value[1]
		No N = 256 Median (IQR); n/N (%)	Yes N = 900 Median (IQR); n/N (%)	
Infant	56.0 (54.0, 60.0)	56.0 (52.0, 60.0)	58.0 (54.0, 60.8)	< 0.001
Child	50.0 (46.0, 56.0)	52.0 (47.0, 58.0)	50.0 (46.0, 56.0)	0.054
Axillary temperature (°C) ‡*	37.1 (36.4, 37.8)	37.3 (36.4, 37.8)	37.1 (36.4, 37.7)	0.064
Oxygen saturation (%)*	94.0 (92.0, 96.0)	94.0 (92.0, 95.5)	95.0 (93.0, 96.0)	< 0.001
Capillary refill time > 2 secs*	18 / 1,056 (1.7%)	6 / 219 (2.7%)	12 / 837 (1.4%)	0.20
Not alert*	183 / 1,138 (16%)	67 / 250 (27%)	116 / 888 (13%)	< 0.001
Anthropometrics				
Weight-for-length z-score*	-0.2 (-0.9, 0.6)	-0.2 (-0.8, 0.8)	-0.2 (-0.9, 0.6)	0.30
Weight-for-age z-score*	-1.0 (-1.9, -0.4)	-1.0 (-1.9, -0.3)	-1.0 (-1.9, -0.4)	0.80
MUAC-for-age z-score*	0.1 (-0.6, 0.7)	0.3 (-0.5, 0.9)	0.1 (-0.6, 0.7)	0.025
Length-for-age z-score*	-1.6 (-2.5, -0.9)	-1.6 (-2.4, -0.9)	-1.6 (-2.5, -0.9)	0.90
Primary outcome				
Received supplemental oxygen	80 / 1,156 (6.9%)	31 / 256 (12%)	49 / 900 (5.4%)	< 0.001

[1]Wilcoxon rank sum test; Pearson's Chi-squared test; Fisher's exact test

4.4 Discussion

This study reports the external validation of three clinical severity scores amongst young children presenting with ARIs to a medical clinic on the Thailand-Myanmar border, quantifies the added value of biomarkers of endothelial injury, immune activation, and inflammation amongst the subset of children meeting WHO-pneumonia criteria, and explores ways of integrating clinical and biomarker measurements into patient assessment. Unlike other studies which investigate the prognostic accuracy of the scores and biomarkers in hospital settings,[106,145,146,161] this study evaluated their performance at the community level. Crucially, the unique circumstances of the cohort permitted follow-up of the children managed as outpatients, something that has been recognised as critical to advancing understanding of the performance of community-based triage tools.[123]

The results indicate that the LqSOFA and qPELOD-2 scores could support early recognition of children requiring referral or closer follow-up in settings with limited resources, and that Ang-2, a marker of endothelial injury, improves sensitivity of the best-performing severity score (LqSOFA), resulting in safer identification of children suitable for home-based management of pneumonia. These findings were consistent across a range of clinically-plausible referral thresholds. Ang-2 appeared to add particular value for the prognostication of childhood pneumonia, with baseline Ang-2 concentrations able to identify children who were not hypoxaemic at presentation but whose illnesses progressed over the next 28 days, outperforming other biomarkers and the LqSOFA score.

4.4.1 Clinical severity scores and ARI risk stratification

An LqSOFA score ≥ 1 yielded a sensitivity and specificity > 0.80 for identifying children with ARIs who would require supplemental oxygen. Encouragingly, this is remarkably consistent with the performance reported in the original LqSOFA development study and may reflect similarities in the use-case (febrile children presenting from the community) and severity of the cohorts (admission rate 12.1% vs. 27.7%), albeit despite obvious demographic differences.[133] In contrast to qPELOD-2, LqSOFA

contains age-adjusted tachypnoea, which may have improved performance in children with respiratory illnesses. Furthermore, the performance of LqSOFA (or qSOFA) has been shown to improve outside of the intensive care unit, when used to predict more proximal outcomes (for example, critical care admission rather than mortality), and if the AVPU scale (vs. GCS) is employed to assess mental status.[194] These all apply to this cohort. In keeping with previous literature, the mSIRS score was poorly discriminative, not well calibrated, and led to substantial misclassification.[106]

Performance of the severity scores improved when they were deployed as clinical prediction models and when nutritional status and respiratory distress were included as additional predictors. Whilst discrimination of all three updated models was good, the AUC is a summary measure of model performance and does not necessarily reflect clinical utility.[64,65,191] Decision curve analyses illustrate the superiority of the LqSOFA and qPELOD-2 models compared with the mSIRS model across a range of clinically-relevant referral thresholds.

With growing access to smartphones there may be contexts where the increased accuracy afforded by a clinical prediction model outweighs the simplicity and practicality of points-based scoring systems. At a 5% referral threshold, the updated LqSOFA model identified a similar proportion of presentations for referral as the LqSOFA score at a cut-off of ≥ 1 (14.3% vs. 16.1%), however use of the model would have resulted in ~20% fewer incorrect referrals and a ~30% decrease in the number of presentations incorrectly recommended for community-based management.

In addition to greater accuracy, prediction models permit more stable and nuanced evaluation of risk. One limitation of discrete points-based scores is that small changes in predictor variables can result in a patient being assigned a different level of a score, and as changes in score cut-off often have large impacts on performance (sensitivity, specificity, NLR, and PLR), this can result in substantially different management recommendations for patients with similar baseline risks. Many clinical predictors are easily influenced by a child's activity level (for example, heart rate, respiratory rate, etc.) contributing to variable and unstable predictions. This issue is partly mitigated for prediction models,

which provide a continuous output (predicted probability), particularly if dichotomisation of predictors is avoided.[195] Furthermore, prediction models allow decision (referral) thresholds to be adjusted to the needs of an individual patient and/or health system. This flexibility may be particularly important in the heterogeneous environments commonplace in many LMIC primary care contexts. For example, in locations where community follow-up is feasible (for example, via a telephone call or return clinic visit) and/or referral carries great cost (to the patient or system), a higher referral threshold (lower NNR) may be acceptable, compared with settings where safety-netting is impractical and/or access to secondary care is less challenging.

4.4.2 Combinatorial clinical-biomarker approaches to childhood pneumonia

The performance of the LqSOFA score in the pneumonia subgroup analyses was consistent with the broader analysis of the score in children with ARIs. The LqSOFA score is an age-adapted version of the widely-endorsed qSOFA score for adults,[194] and was specifically designed for triaging children presenting from the community. Unlike other paediatric pneumonia risk scores it uses routinely collected data, which facilitated external validation in this resource-limited primary care setting.[155]

The results illustrate the critical importance of considering clinical context when evaluating potential incremental value of biomarkers, rather than relying on summary measures such as the AUC alone.[64] Although measurement of some biomarkers of endothelial injury (Ang-2 and sFlt-1) and immune activation (IL-8) improved the ability of the LqSOFA score to discriminate children who required supplemental oxygen, only for Ang-2 did inclusion of a biomarker translate into superior clinical utility (net benefit).

Combining measurements of Ang-2 alongside the LqSOFA score could make triage of paediatric pneumonia safer. Sensitivity improved such that ~10-30% fewer children would be incorrectly identified for community-based management, without substantial increases in the proportion of

inappropriate referrals. However, laboratory tests carry an opportunity cost, especially in settings with limited resources. It should be noted that this strategy would require the measurement of Ang-2 in all children presenting with pneumonia. Whether this is feasible in routine practice would depend on many factors, including the availability, durability, turnaround time, and cost of a point-of-care test for Ang-2. Should such a test become available, cost-effectiveness analyses accounting for differing scenarios would be required before it could be recommended for use.

An alternative strategy, perhaps more compatible with the clinical workflow and resources available in some busy LMIC primary care settings, could be to use the easily practicable LqSOFA score as a screening tool to identify high risk children with pneumonia, and measure Ang-2 concentrations only in the remaining subset of children not readily identified as requiring referral to hospital by the LqSOFA score. This approach demonstrated encouraging rule-out performance (NLR = 0.28; sensitivity = 0.78 [50/64]) for identifying children who would require supplemental oxygen over the 28 days following presentation, whilst maintaining an incorrect to correct referral ratio of 3:1 (i.e., an NNR of four; PLR = 3.66; specificity = 0.79 [623/792]), and reducing the number of Ang-2 tests required by more than 20% (179/856). Further efficiencies could be achieved by converting the points-based LqSOFA score into a clinical prediction model, which would permit the identification of both low- and high-risk groups who could be adequately risk stratified, further restricting the number of measurements of Ang-2 to the group with intermediate risk.

The association between higher concentrations of Ang-2 and supplemental oxygen requirement has biological plausibility. Ang-2 destabilises the endothelium, increases microvascular permeability, and is implicated in the pathogenesis of acute lung injury and sepsis.[37,148,165,196] Previous work has illustrated the prognostic role of Ang-2 in adults with pneumonia and in hospitalised children with hypoxaemic pneumonia.[144,148] Although this is the first study to investigate the role of Ang-2 in paediatric pneumonia at the community-level, endothelial dysfunction has been documented in ambulatory children with mild ARIs.[150]

Recently, the immune activation marker sTREM-1 has been shown to be prognostic in hospitalised children with pneumonia and proposed as a possible risk stratification tool.[145-147] In this study, baseline sTREM-1 concentrations were similar in children who did and did not progress to require supplemental oxygen. As McDonald et al. note, the results of hospital-based studies cannot be generalised to community settings; in this study most children with pneumonia (71.6%; 644/900) were managed in the community and only a quarter (26.2%; 236/900) had severe pneumonia at presentation, compared to over three-quarters of children who had severe pneumonia at the time sTREM-1 levels were measured in previous hospital-based studies.[145,146] Furthermore, studies of adults with Covid-19 suggest that sTREM-1 may be useful for predicting mortality but less well-suited for predicting proximal outcomes such as supplemental oxygen requirement.[171,197]

4.4.3 Strengths and limitations

The local circumstances of the cohort enabled recruitment of children at the first point of contact with the formal health system and follow-up of those managed as outpatients, aspects that are critical for robust evaluation of clinical scores and biomarkers in primary care.[123] The study followed the latest guidelines in prediction model building and used bootstrap internal validation, penalised regression, placed knots at predefined locations, and limited the number of candidate predictors to avoid overfitting the models.[66,186,198,199]

The pneumonia subgroup analyses constitute the largest investigation of the role of markers of endothelial injury and immune activation in paediatric pneumonia, and the only study to date conducted at the community-level. A biomarker panel with mechanistic links to severe respiratory disease was prespecified and the value added to a validated clinical score quantified. The analytical approach acknowledged that the threshold for home-based management of pneumonia would vary in different healthcare settings and that there is an inherent difference between recognising a child who

is acutely unwell at the time of presentation (diagnosis) and identifying a child who appears clinically stable but is at risk of subsequent deterioration (prognosis).[200]

Although best-practice methodology was followed, the models still require validation on new data to assess generalisability and provide a fairer comparison with the pre-existing points-based scores. The full model equations are provided to encourage independent validation. The performance of the models deteriorated when children sent home directly from the clinic who required supplemental oxygen within the next 28 days were classified as meeting the primary outcome. Although this was a conservative assumption, it serves to highlight the potential importance of a problem that has been long suspected but hitherto relatively unexplored with regards the validity of existing community-based triage tools.[123] Prospective research with dedicated outpatient follow-up is ongoing to investigate this issue further.[201]

Supplemental oxygen therapy was selected as the primary outcome as this reflects a clinically-meaningful endpoint for respiratory illnesses and a pragmatic referral threshold for many resource-limited primary care settings. Oxygen was a scarce resource during the study (cylinders were transported in each week from ~60km away) and oxygen therapy was protocolised; hence the risk of outcome misclassification is low.

For those who met the primary outcome, the time of oxygen initiation was not available in the study database. Although no patient had met the outcome when baseline predictors were measured, some may have done so shortly after. Nevertheless, sensitivity analyses of the performance of the clinical scores excluding presentations with baseline SpO_2 < 90% (the qualifying criterion for supplemental oxygen) produced similar results.

Scores including SpO_2 were excluded at the shortlisting stage due to feasibility concerns in resource-limited community settings. Obtaining a reliable pulse oximetry measurement in every child, particularly those under the age of two years, could be challenging due to movement, crying, and/or dirty extremities.[202,203] It would also be time-consuming: half the respondents indicated a mean

measurement time of over two minutes in one survey, which may not be practical in busy primary care settings.[204] However, pulse oximeters are increasingly available in community settings, especially since the Covid-19 pandemic,[205] and hence some scores including SpO_2 may warrant reconsideration.

The pneumonia subgroup analyses were limited to presentations with stored serum samples. Presentations without serum samples available were more likely to have respiratory distress, altered mental status, and receive supplemental oxygen, and thus future studies should assess whether the findings are generalisable to more severe pneumonia presentations in the community.

The WHO pneumonia definition is recognised as prioritising sensitivity over specificity.[206] It is possible that some children who did not receive supplemental oxygen may have had upper respiratory tract infections and hence the discrimination demonstrated by Ang-2 and LqSOFA may partly reflect misclassification of the study population. However, using WHO pneumonia criteria is pragmatic and likely reflects the approach that would be taken if these triage strategies were to be implemented on the field.

Results of the recursive partitioning analysis inherently reflect the 10:1 trade-off between false negatives (missed referrals) and false positives (inappropriate referrals) specified in the loss-matrix. Whilst this was informed by clinical experience of working in resource-limited settings, and is comparable to approaches taken by other groups,[39,207] it will not apply in all contexts. The relatively few outcome events meant that it was not possible to cross-validate the decision trees and hence the results will be optimistic and should be viewed as an indicative framework for the integration of Ang-2 and LqSOFA for the triage of childhood pneumonia at the community level.

Finally, given the exploratory nature of this study the analyses were set within a simplified framework reflective of contexts in which a health worker is faced with a binary decision to manage a child in the community or refer them to hospital. In reality, strategies for delivery of primary care are often complex and heterogeneous. Ongoing prospective work will evaluate different triage strategies

including whether a 'watch-and-wait' approach for children at intermediate risk of disease progression could result in further gains.[201]

4.4.4 Conclusions

Simple 'pathogen agnostic' algorithms could be particularly impactful in resource-limited primary care settings where patient management is often syndromic and the infecting microbe is usually unknown at the time of initial assessment. Performance of the LqSOFA score was encouraging and comparable to that in the original derivation setting.[133] Converting the points-based severity scores into clinical prediction models and including additional variables relevant to resource-constrained LMIC settings improved accuracy and, if validated, might permit application across a wider range of contexts with differing referral thresholds. Measurements of Ang-2, a biomarker of endothelial injury, improved the sensitivity of the LqSOFA score and might enable safer community-based triage of childhood pneumonia, primarily through better identification of children who deteriorate later in their illness course.

Clinical risk scores and host biomarker measurements can act synergistically such that combinatorial approaches may assist health workers identify children who are acutely unwell at presentation and those who will deteriorate later, enabling earlier and more appropriate referrals to higher-level care. Future prospective work should focus on validating these findings and developing durable and affordable point-of-care tests for the most promising biomarkers. Clinical utility and cost-effectiveness of different strategies for integrating biomarker measurements into patient assessment and triage should be explored.

5 A prognostic model for critically ill children in Cambodia

This chapter is based upon work published in: **Chandna A**, Keang S, Vorlark M, et al. *Derivation of a prognostic model for critically ill children in locations with limited resources.* Pediatr Crit Care Med (in press). 2023.

5.1 Introduction

Historically, paediatric critical care has often been perceived as too complex, expensive, or unethical to provide in settings where resources are scarce.[208] These presumptions are increasingly countered by data which suggest that simple, low-cost interventions can result in substantial improvements in health outcomes and the premise of comprehensive healthcare as a universal human right.[12,209-211] Consequently, demand and capacities for paediatric critical care services are growing in many resource-limited settings.[212,213]

Notwithstanding these welcome developments, need for critical care often outstrips supply.[50] Evidence-based approaches to support resource stewardship are essential to promote equitable and sustainable critical care services. This is especially true in rural regions of many low- and middle-income countries (LMICs) where maldistribution of healthcare professionals and resources results in considerable disparities in access to paediatric critical care.[12,208,210,214-217]

Risk stratification tools can help target scarce resources optimally. However, tools developed for use in paediatric intensive care units (PICUs) in high-income settings are time-consuming to compute and often require diagnostic tests not routinely available in resource-constrained regions of LMICs.[218,219] Furthermore, prognosis is influenced by the level of care available and underlying host susceptibility states, and hence locally-developed and/or validated tools are required to support context-specific clinical decision making.[214,220,221] Consequently, there have been calls both to evaluate

existing severity scores in resource-constrained PICUs and to develop new risk stratification tools appropriate for use in these settings.[214,222] Unfortunately, most studies from LMIC PICUs to date have been limited to urban centres, hampered by small sample sizes, and used methods incompatible with development of robust clinical severity scores or prediction models.[66,223-226]

Using data from children admitted to the PICU at Angkor Hospital for Children (AHC) in Siem Reap, Cambodia, this study reports the external validation of nine existing paediatric severity scores in a resource-limited PICU setting. Secondly, it presents the development of a bespoke clinical prediction model, derived specifically to support risk stratification in resource-constrained PICU contexts.

5.2 Methods

5.2.1 Study population and setting

This retrospective cohort study screened consecutive admissions to the PICU at AHC between 1st January 2018 and 1st January 2020. All non-elective admissions of children aged > 28 days and ≤ 16 years were included.

AHC is a non-governmental paediatric healthcare organisation with a nationwide catchment area providing comprehensive primary-to-tertiary care. The hospital, located in Siem Reap, northern Cambodia, has 89 inpatient beds situated on two medical wards, a surgical ward, a special care baby unit, and neonatal and paediatric intensive care units. The 14-bedded PICU has approximately 1,000 annual admissions and is staffed by a team of 30 nurses, four senior doctors, five doctors in training, and receives medical and nursing students from Cambodia's three main medical schools. The unit provides the only critical care service for children in the north of the country and is the only PICU located outside of the capital city, Phnom Penh. Clinical staff have completed or are undertaking training in paediatric intensive care medicine. The unit (Level II or Community PICU)[216,227] provides mechanical and non-invasive ventilation (oxygen cylinders are delivered fortnightly), inotropic therapy,

peritoneal dialysis, and specialist nursing care (minimum 1:3 nurse-patient ratio) for critically ill children at AHC and those transferred from other health facilities. A backup generator ensures continuity of electrical supply during infrequent power outages.

5.2.2 Data collection

PICU admissions were identified from the electronic Hospital Information System (HIS) and cross-checked against the admission logbook located on the unit. Clinical records were retrieved and data extracted onto structured case report forms (CRFs; Appendix 9.9) by a team of trained research nurses. Data extraction occurred between 27th November 2020 and 14th December 2021. The hospital admission and PICU vital sign proforma (Appendix 9.10 and Appendix 9.11) helped standardise data extraction and all variables were prospectively defined in a data dictionary to ensure consistency of interpretation across the research team. Each CRF was reviewed by one of two study doctors in consultation with the clinical records, with particular focus on explanatory and outcome variables. Data were entered into an electronic study database and 10% of CRFs underwent review by the study Data Manager to ensure a data entry error rate of < 0.5%. Data profiling was conducted to identify missing and implausible values.

5.2.3 Primary and secondary outcomes

The primary outcome was death during PICU admission. Participants who were discharged from PICU to die at home were classified as meeting the primary outcome. Sensitivity analyses were conducted excluding these participants as well as those whose death was judged by either of two study doctors to have been related to a separate illness acquired during the PICU stay. It was not possible to blind the research team to outcome status during data extraction.

The secondary outcome was death in the 12 months following a PICU discharge. Caretakers of participants for whom post-discharge outcomes could not be determined from the clinical records or HIS were telephoned to ascertain vital status 12 months after PICU discharge.

5.2.4 Shortlisting of existing severity scores

The results of two recent systematic reviews were supplemented by searching MEDLINE using synonyms of "paediatric" AND ["severity score OR prediction model"].[154,228] Forty-nine scores or models were longlisted and assessed for suitability for external validation (Table 5.2-1): 15 were excluded as they required diagnostic tests unlikely to be available in resource-limited PICU settings (for example, arterial blood gases, creatinine, or serum electrolyte estimations) and 14 were excluded as they contained variables not relevant to the intended-use context and/or population (for example, arrival via emergency medical services, presence of an indwelling central venous catheter, or malaria rapid diagnostic test result). A further five were excluded as the information required to calculate the score was not provided in the original manuscript, and another eight were excluded as they contained variables not available in the AHC clinical records and no suitable proxy variables could be identified. Ultimately, nine severity scores were shortlisted for external validation (Table 5.2-2). Neither the setting, population, nor outcome used for the derivation were prerequisites for selection of a score for external validation.

Table 5.2-1: Severity scores excluded at longlisting. Severity scores identified from two recent systematic reviews and PubMed search.[154,228] Scores were excluded if they contained advanced diagnostic tests unlikely to be available in resource-constrained contexts, they included variables that were not relevant for the intended setting of use, the information required to calculate the score/model was not provided in the original manuscript, and/or the required variables were not available in the routine clinical records at the study site (and no suitable proxy variable could be identified). APTT = activated partial thromboplastin clotting time; ARI = acute respiratory infection; BUN = blood urea nitrogen; GE = gastroenteritis; HIV = human immunodeficiency virus; LDH = lactate dehydrogenase; IPSCC = international pediatric sepsis consensus conference; MUAC = mid-upper arm circumference; $paCO_2$ = partial pressure of carbon dioxide in arterial blood; paO_2 = partial pressure of oxygen in arterial blood; PCT = procalcitonin; PICU = paediatric intensive care unit; PT = prothrombin time; RA = room air; SpO_2 = oxygen saturation.

Name of score	Advanced diagnostic test required	Inappropriate for setting and/or population	Data not available	Reasons for exclusion
AQUAMAT	Y	N	N	BUN and base deficit required
BITWE MODEL	Y	N	Y	MUAC, infectious diagnosis (ARI, GE, malaria, bacteraemia, other) required
BITWE SCORE	Y	N	Y	MUAC, infectious diagnosis (ARI, GE, malaria, bacteraemia, other) required
DRAMAIX	Y	N	Y	Albumin, transthyretin, oedema, and MUAC required
ELSHOUT	N	Y	N	Sore throat, palpable lymphadenopathy not suitable for PICU population
ERDMAN	Y	N	N	Host biomarker tests required
FEAST-PETaL	Y	N	N	BUN, pH, and lactate required
ITAT	N	Y	N	SpO_2 on RA not relevant for PICU population
KWIZERA 1	N	N	N	Information not available for construction of the score/model
KWIZERA 2	N	N	N	Information not available for construction of the score/model
KWIZERA 3	N	N	N	Information not available for construction of the score/model
KWIZERA 4	N	N	N	Information not available for construction of the score/model
KWIZERA 5	N	N	N	Information not available for construction of the score/model
LIN NOMOGRAM	Y	N	N	Blood culture, albumin, and LDH required
LODS	N	N	Y	Deep breathing and prostration required
LOWLAAVAR 1	N	Y	N	HIV not relevant (low endemicity)
LOWLAAVAR 2	N	Y	Y	HIV not relevant (low endemicity), MUAC required
LOWLAAVAR 3	N	N	Y	MUAC required - and as this is a model cannot substitute a proxy variable

Name of score	Advanced diagnostic test required	Inappropriate for setting and/or population	Data not available	Reasons for exclusion
MPIMBAZA	N	N	Y	Prostration, jaundice, deep breathing, and meningitic signs required
mPRIO	Y	Y	Y	SpO_2 on RA, PCT, organ dysfunction as per IPSCC definition
mRISC	N	Y	Y	Malaria not relevant (low endemicity), dehydration, prostration, night sweats, and historical loss of consciousness required
PCIS	Y	N	Y	BUN, creatinine, K^+, Na^+, pH, paO_2, and gastrointestinal bleeding required
PEDIA-e	N	N	Y	Prostration, jaundice, and kwashiorkor required
PEDIA-i	N	N	Y	Deep breathing, prostration, and jaundice required
PEDIA-l	N	N	Y	Prostration and kwashiorkor required
PELOD-2	Y	N	Y	Lactate, creatinine, paO_2, $paCO_2$, and pupillary reaction required
PERCH	N	Y	Y	SpO_2 on RA not relevant, deep breathing, cough, and grunting required
PEWS BCH	N	N	Y	Skin colour, frequency of nebulisation prior to assessment, nurse concern, and family concern required
PIM III	Y	Y	Y	Base excess, paO_2, and pupillary reaction required, along with many other high-income country contextual variables
PIRO	Y	Y	Y	paO_2, BUN, transaminases, PT, blood culture, SpO_2 on RA not relevant for PICU population, signs of liver failure
pMODS	Y	N	N	Lactate, bilirubin, paO_2, fibrinogen, and BUN required
PRISM III	Y	N	Y	pCO_2, paO_2, pH, acidosis, total CO_2, K^+, BUN, creatinine, PT/APTT, and pupillary reaction required
pSOFA	Y	N	N	Bilirubin and creatinine required
qSOFA-L	Y	N	N	Lactate required
RISC	N	Y	Y	SpO_2 on RA not relevant for PICU population and HIV not relevant (low endemicity), prostration, and wheezing required
RISC-Malawi	N	Y	Y	SpO_2 on RA not relevant for PICU population, MUAC, and wheezing required
SCOTT	N	Y	Y	Arrival via emergency medical services, indwelling central line, and hospitalised within last year required
SICK	N	Y	N	SpO_2 on RA not relevant for PICU population
TORPS	N	Y	N	SpO_2 on RA not relevant for PICU population
YOS	N	N	Y	Quality of cry, reaction to parent stimulation, state variation, colour, and response to social overtures required

Table 5.2-2: Severity scores selected for external validation. Scores were selected for external validation irrespective of the setting, population, and outcome used for the original derivation study. The only prerequisites were that the score had to be calculable with the available data (with the exception that systolic blood pressure could be dropped if CRT was included),[133] relevant to the study population, and feasible for implementation in a resource-limited PICU context. AVPU = Alert Voice Pain Unresponsive scale; CRT = capillary refill time; ED = emergency department; FEAST-PET = Fluid Expansion as Supportive Therapy-Pediatric Emergency Triage; GCS = Glasgow Coma Scale; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; PAWS = Paediatric Advanced Warning Score; PEWS = Pediatric Early Warning Score; PEWS-RL = PEWS-Resource Limited; PICU = paediatric intensive care unit; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2; qSOFA = quick Sequential Organ Failure Assessment; SIRS = Systemic Inflammatory Response Syndrome; SpO_2 = oxygen saturation.

Score	Range	Predictors	Adjustments	Original setting, population, and outcome
FEAST-PET	0-10	Heart rate, temperature, CRT, pulse character, work of breathing, lung crepitations, mental status, pallor	Cut-off for CRT increased to ≥ 3 seconds; deep breathing omitted from work of breathing; age-adjusted WHO criteria for severe anaemia used as a proxy for pallor;[229] mental status dichotomised and assessed using AVPU or GCS, reducing the maximum possible score to 9	Score to predict 48-hour mortality on admission to secondary and tertiary-care hospitals in East Africa in children with severe febrile illness[98]
LqSOFA	0-4	Respiratory rate, heart rate, CRT, mental status	Mental status assessed using AVPU or GCS	Score to predict PICU admission or death in febrile children presenting to ED in the United Kingdom[133]
PAWS	0-21	Respiratory rate, heart rate, temperature, CRT, mental status, SpO_2, work of breathing	Mental status assessed using AVPU or GCS; CRT, mental status, and work of breathing dichotomised, reducing the maximum possible score to 17	Score to predict need for PICU admission in children presenting to ED in the United Kingdom[159]
PEWS	0-26	Respiratory rate, heart rate, systolic blood pressure, CRT, SpO_2, supplemental oxygen, work of breathing	Systolic blood pressure omitted; work of breathing dichotomised, reducing the maximum possible score to 20	Score to predict need for PICU admission in children hospitalised on a general paediatric ward in a tertiary-care hospital in Canada[113]
PEWS-IRISH	0-21	Respiratory rate, heart rate, systolic blood pressure, CRT, mental status, SpO_2, supplemental oxygen, work of breathing	Systolic blood pressure omitted; work of breathing and mental status dichotomised, reducing the maximum possible score to 15	Adaptation of the PEWS score by the Irish National Clinical Effectiveness Committee[230]

Score	Range	Predictors	Adjustments	Original setting, population, and outcome
PEWS-RL	0-6	Respiratory rate, heart rate, temperature, mental status, supplemental oxygen, work of breathing		Score to predict clinical deterioration in children hospitalised on a general paediatric ward in a tertiary-care hospital in Rwanda[231]
qPELOD-2	0-3	Heart rate, systolic blood pressure, mental status	CRT used as a proxy for systolic blood pressure; mental status assessed using AVPU or GCS	Score to predict mortality in children with suspected infection on admission to nine European PICUs[161]
qSOFA	0-3	Respiratory rate, systolic blood pressure, mental status	CRT used as a proxy for systolic blood pressure; mental status assessed using AVPU or GCS	Adult sepsis score to predict mortality adapted for children with suspected infection on admission to PICUs in Australia and New Zealand[161]
SIRS	0-4	Respiratory rate, heart rate, temperature, white cell count		Expert consensus definition for the diagnosis of paediatric sepsis[109]

5.2.5 Candidate predictors

Baseline variables at the time of PICU admission were extracted from the clinical records. For admissions occurring from the AHC Emergency Room (ER) the first set of vital signs was abstracted. For inter- and intra-hospital transfers the vital signs recorded at the time the decision to transfer was made were abstracted. If weight or height were not recorded at the time of PICU admission the closest values during the same hospital stay were used. Laboratory parameters measured within 24 hours of PICU admission were considered available on admission. Sensitivity analyses restricting this period to between two hours prior and up to four hours after admission were performed.[232]

For derivation of the new model, candidate predictors were selected *a priori* based on existing literature, expert knowledge, feasibility for implementation, and availability of data in the clinical records. Rather than solely focussing on parameters reflecting acute vital organ dysfunction, variables reflecting the background of the child and their illness journey were also considered to ensure holistic assessment of critical illness and inclusion of important contextual determinants of outcome often omitted by clinical risk scores developed in high-income settings. Variables were divided into five 'domains': background, illness journey, cardiovascular, respiratory, and neurological, and candidate predictors were selected across all domains. The 11 selected predictors were age, comorbidity status, weight-for-age z-score (WAZ), estimated travel time to hospital, route of admission to PICU, heart rate, capillary refill time (CRT), respiratory rate, peripheral oxygen saturation (SpO_2), receipt of supplemental oxygen, and mental status.

Comorbidity status was assessed by a study paediatrician blinded to outcome status using the following adapted working definition: *any previous health condition known to be present at PICU admission severe enough to require specialty paediatric care and probably a period of hospitalisation over 12 months*.[233] Both chronic conditions (for example, congenital heart disease) and conditions for which the child was known to be receiving treatment prior to the illness resulting in PICU admission (for example, active tuberculosis) were considered as comorbidities. For descriptive purposes, all

comorbidities were assigned to one of 11 categories reflecting the primary organ or body system affected (cardiac, endocrine, gastrointestinal, haematological, immune, neurological, oncological, renal, respiratory, other congenital abnormality, and other).

Table 5.2-3: Candidate variables selected across five clinical domains for derivation of the new prognostic model. AHC = Angkor Hospital for Children; AVPU = Alert Voice Pain Unresponsive scale; US-CDC = United States Centers for Disease Control; GCS = Glasgow Coma Scale; WHO = World Health Organization.

Predictor	Units	Variable type	Notes
Background			
Age	Months	Continuous	
Presence of comorbidity	NA	Binary	Adapted from Feudtner et al.[233]
Weight-for-age z-score	NA	Continuous	Reference ranges: WHO < 10 years;[234,235] US-CDC ≥ 10 years[236]
Illness journey			
Travel time	Minutes	Continuous	Estimated travel time via road using GoogleMaps
Intra-hospital transfer	NA	Binary	Transfer from AHC acute ward
Cardiovascular			
Heart rate	Beats per minute	Continuous	
Capillary refill time	Seconds	Binary	Dichotomised at > 2 seconds
Respiratory			
Respiratory rate	Breaths per minute	Continuous	
Oxygen saturation	Percentage	Continuous	
Supplemental oxygen	NA	Binary	Any supplemental oxygen therapy at time SpO_2 measured
Neurological			
Mental status	NA	Binary	Abnormal mental status: < A on AVPU or GCS < 15

The World Health Organization (WHO) reference ranges were used to calculate WAZ for children up to the age of 10 years.[234,235] For older children (10 to 16 years) reference ranges from the United States Centers for Disease Control (US-CDC) were used to calculate z-scores.[236] This was a pragmatic decision due to the absence of WHO reference ranges in children aged ≥ 10 years, which

was taken after confirming (a) good agreement between WHO and US-CDC reference ranges in children with WAZ less than zero (Figure 5.2-1); and (b) that the majority of children aged ≥ 10 years in the study population (90.0%; 188/209) had WAZ less than zero. All z-scores were calculated using the R package *zscorer*.[153]

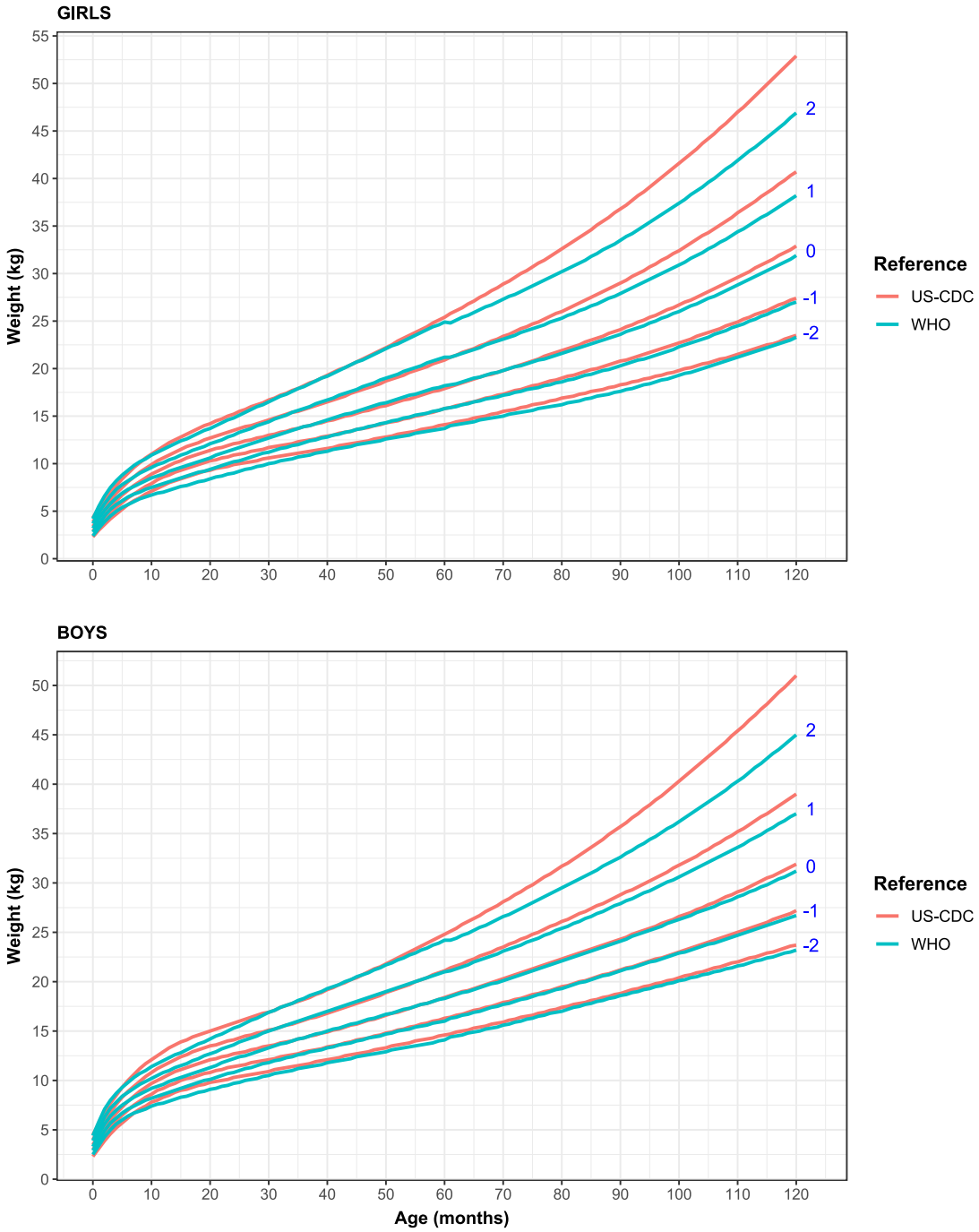


Figure 5.2-1: Comparison of WHO and US-CDC WAZ reference ranges for children up to age 10 years. Blue numbers indicate WAZ score pertaining to each pair of lines. CDC = United States Centre for Disease Control; WAZ = weight-for-age z-score; WHO = World Health Organization.

5.2.6 Missing data

Missing data were summarised for the existing scores and for each candidate predictor in the new model (Figure 5.2-2; R package: *naniar*).[237] For the existing scores, missingness ranged from 5.4% (qPELOD-2) to 15.6% (FEAST-PET). Amongst the 11 candidate predictors for the new model missingness ranged from 1.1% for heart rate and prolonged CRT to 6.9% for respiratory rate, whilst four predictors had no missing data. Given the relatively low proportion of missing data, single (median) imputation conditional on outcome status was proposed to address missingness. Sensitivity analyses comparing this to a full-case approach, as well as best- and worst-case imputation, produced similar results (Table 5.2-4), confirming that single (median) imputation was appropriate for the primary analysis.

Table 5.2-4: Sensitivity analyses comparing different methods for handling missing data. Complete case analyses were performed using pairwise deletion. Median imputation was performed conditional on outcome status. For best case imputation, missing values amongst admissions that met the primary outcome were assigned the most extreme values in the dataset, whilst missing values amongst admissions that did not meet the primary outcome were assigned a normal value (for example, median heart rate, 100% SpO_2, etc.) The opposite approach was taken for worst case imputation. AUC = area under the receiver operating characteristic curve; CI = confidence interval; FEAST-PET = Fluid Expansion as Supportive Therapy-Pediatric Emergency Triage; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; PAWS = Paediatric Advanced Warning Score; PEWS = Pediatric Early Warning Score; PEWS-RL = PEWS-Resource Limited; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2; qSOFA = quick Sequential Organ Failure Assessment; SIRS = Systemic Inflammatory Response Syndrome.

Model / Score	**AUC** (95% CI)			
	Complete Case	**Median Imputation**	**Best Case**	**Worst Case**
NEW MODEL	0.83 (0.79 to 0.88)	0.84 (0.80 to 0.88)	0.86 (0.83 to 0.90)	0.78 (0.74 to 0.83)
FEAST-PET	0.67 (0.60 to 0.74)	0.72 (0.66 to 0.78)	0.74 (0.68 to 0.80)	0.67 (0.61 to 0.73)
LqSOFA	0.75 (0.69 to 0.80)	0.76 (0.71 to 0.81)	0.78 (0.73 to 0.83)	0.71 (0.66 to 0.76)
PAWS	0.73 (0.67 to 0.78)	0.76 (0.71 to 0.81)	0.78 (0.73 to 0.83)	0.70 (0.65 to 0.75)
PEWS	0.67 (0.61 to 0.73)	0.71 (0.65 to 0.76)	0.73 (0.67 to 0.78)	0.67 (0.62 to 0.73)
PEWS-IRISH	0.69 (0.63 to 0.75)	0.74 (0.69 to 0.79)	0.76 (0.71 to 0.81)	0.69 (0.64 to 0.74)
PEWS-RL	0.68 (0.62 to 0.74)	0.72 (0.67 to 0.77)	0.73 (0.68 to 0.78)	0.66 (0.61 to 0.71)
qPELOD-2	0.73 (0.68 to 0.79)	0.75 (0.70 to 0.80)	0.75 (0.70 to 0.80)	0.70 (0.64 to 0.75)
qSOFA	0.72 (0.66 to 0.77)	0.74 (0.69 to 0.79)	0.74 (0.69 to 0.79)	0.70 (0.64 to 0.75)
SIRS	0.62 (0.55 to 0.68)	0.59 (0.53 to 0.65)	0.63 (0.57 to 0.68)	0.56 (0.50 to 0.62)

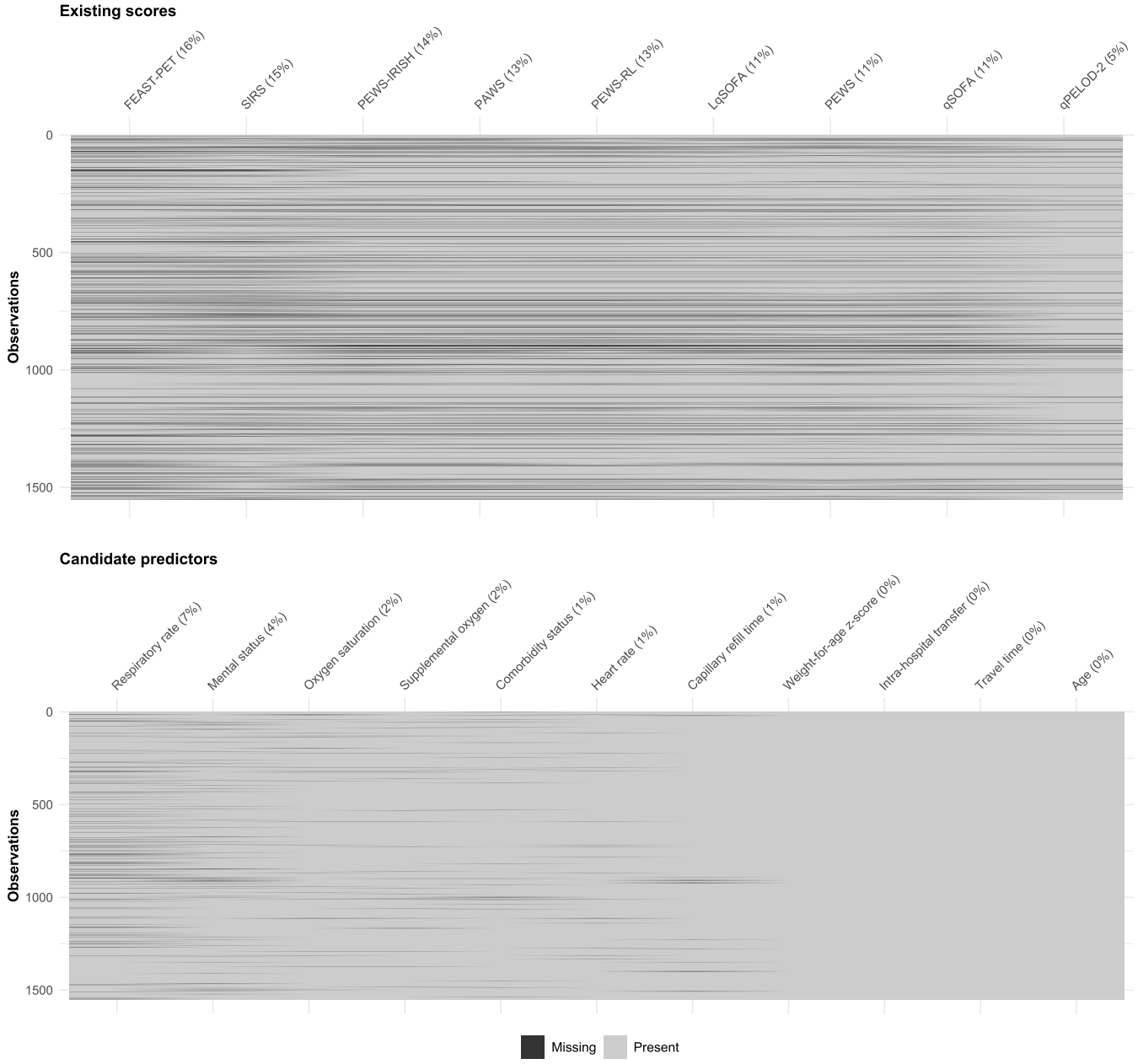


Figure 5.2-2: Missing data patterns. Top panel: missingness in existing severity scores. Bottom panel: missingness in candidate predictors included in new clinical prediction model. Variables are ordered from left to right by proportion of missingness. FEAST-PET = Fluid Expansion as Supportive Therapy-Pediatric Emergency Triage; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; PAWS = Paediatric Advanced Warning Score; PEWS = Pediatric Early Warning Score; PEWS-RL = PEWS-Resource Limited; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2; qSOFA = quick Sequential Organ Failure Assessment; SIRS = Systemic Inflammatory Response Syndrome.

5.2.7 Statistical methods

Discrimination and calibration of each existing score were assessed by quantifying the area under the receiver operating characteristic curve (AUC; R package: *pROC*)[182] and plotting the proportion of admissions that met the primary outcome at each level of a score. Positive and negative likelihood ratios (PLRs and NLRs) were reported at each of the scores' cut-offs to quantify the change in pre-test probability that a PICU admission would result in death. As a rule-of-thumb, a PLR > 10 or NLR < 0.1 is often deemed conclusive, a PLR between 5-10 or NLR between 0.1-0.2 considered substantial, a PLR between 2-5 or NLR between 0.2-0.5 regarded as small but important, and a PLR between 1-2 or NLR between 0.5-1 likely clinically insignificant.[87]

Prior to model building the relationship between continuous predictors and PICU survival status was examined using locally weighted scatterplot smoothing (LOWESS) to determine if transformations were required. Age-specific relationships for heart rate and respiratory rate (< 12 months; 12-59 months; 5-12 years; > 12 years) were explored to account for known changes in these parameters associated with physiological maturation. Stratum-specific odds ratios (ORs) and likelihood ratio tests (LRTs) were used to identify important interactions between age and each of heart rate and respiratory rate, as well as between SpO_2 and receipt of supplemental oxygen. Penalised (ridge) logistic regression was used to derive the model and adjust for optimism (R package: *ridge*).[238] All predictors were prespecified and no predictor selection was performed during model development.

Discrimination (AUC), calibration (calibration intercept, slope, and plots), and classification indices (sensitivity, specificity, NLR, and PLR) at clinically-relevant decision thresholds (R package: *reportROC*)[239] were reported to summarise model performance. Recognising that the relative value of a true positive (TP; PICU admission correctly identified as at high-risk of death) and false positive (FP; PICU admission incorrectly identified as at high-risk of death) will be context-dependent (for example, depending on the human and material capacities of a high-acuity area that at-risk PICU admissions might be triaged to), the clinical utility of the new model was compared to the best-performing existing

scores using decision curves to visualise their net benefits over a range of clinically-plausible decision thresholds (R package: *dcurves*).[59,180]

All analyses were done in R, version 4.2.2.[89]

5.2.8 Sample size

Routinely collected data from the hospital indicated that approximately 100 deaths on the PICU were expected over two calendar years (mortality rate of ~5%), which would ensure sufficient outcome events for external validation of the existing severity scores.[185] At this prevalence, and assuming a conservative Nagelkerke R^2 of 0.15 and shrinkage factor of 0.9, up to 10 candidate predictors (events per parameter [EPP] = 9.7) could be used to build the new prediction model (R package: *pmsampsize*).[186,187] In order to allow for inclusion of interaction terms between age and heart rate and age and respiratory rate, penalisation was used to shrink regression coefficients and permit inclusion of up to 13 parameters whilst still minimising the risk of overfitting.

5.2.9 Ethics and reporting

The study was approved by the AHC Research Committee (AHC 0656/20), Cambodian National Ethics Committee for Health Research (NECHR 257), and the Oxford Tropical Research Ethics Committee (OxTREC 565-20), and is reported in accordance with the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) guidelines (Appendix 9.12).[66]

5.3 Results

5.3.1 Study population

Between 1st January 2018 and 1st January 2020 there were 2,066 admissions to the hospital's PICU, of which case notes were located for 2,021 (97.8%; 2,021/2,066). In total, 1,550 non-elective admissions were eligible for inclusion in the study (eligibility rate 76.7%; 1,550/2,021; Figure 5.3-1). There were 1,366 individual children in the cohort, with 91.1% (1,245/1,366) admitted to the PICU only once during the study period. Median age at PICU admission was 14.0 months (interquartile range [IQR] = 4.0 to 73.0 months) and 59.8% of admissions (927/1,550) were for male children (Table 5.3-1). Nearly one in five admissions had a WAZ < -3 (271/1,550; 17.5%) with a similar proportion having a WAZ between -3 and -2 (261/1,550; 16.8%).

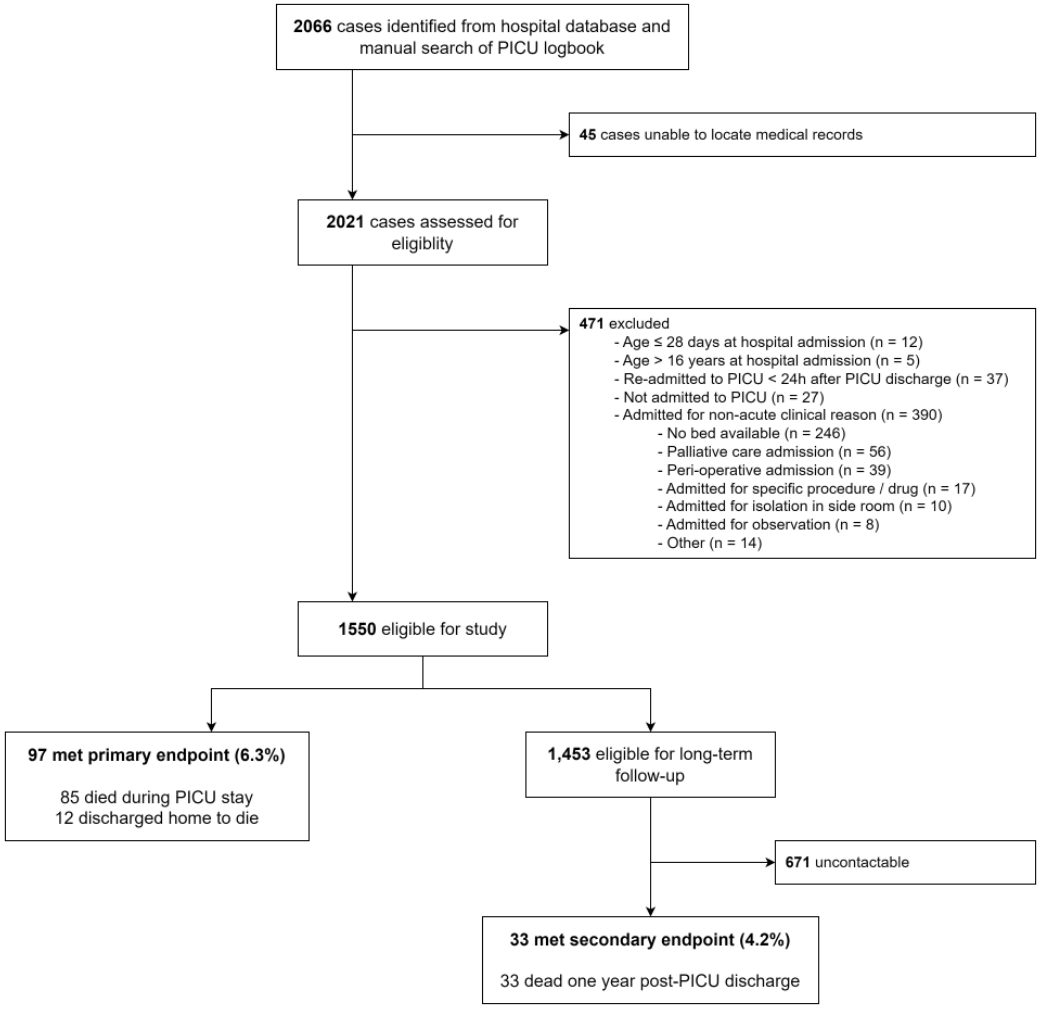


Figure 5.3-1: Inclusion of cases in study. PICU = paediatric intensive care unit.

Table 5.3-1: Baseline characteristics. Baseline demographic, background, illness history, anthropometric, clinical, and laboratory characteristics of the cohort, stratified by primary outcome status. [$]Baseline SpO_2 amongst those not receiving supplemental oxygen at the time of PICU admission confirmed a similar relationship (96.5% vs. 98.0%; p < 0.001; n = 798). [†]Not alert = GCS < 15 or AVPU < A; prolonged CRT defined as > 2 seconds; respiratory distress = chest indrawing, tracheal tug, or nasal flaring; laboratory parameters included if measured within 24 hours of PICU admission. [*]Missing data: comorbidity = 23; preterm birth = 84; low birthweight = 238; illness duration = 2; number of previous care encounters = 391; axillary temperature = 33; heart rate = 17; respiratory rate = 107; oxygen saturation = 26; supplemental oxygen = 24; mental status = 55; CRT = 17; pulse character = 42; cool extremities = 10; respiratory distress = 22; lung crackles = 10; white cell count = 117; neutrophil count = 117; lymphocyte count = 118; haemoglobin = 114; platelet count = 115; C-reactive protein = 411; glucose = 386; LqSOFA = 177; qSOFA = 170; qPELOD-2 = 83; SIRS = 235; PEWS = 175; PEWS-RL = 206; PEWS-IRISH = 220; PAWS = 209; FEAST-PET = 242. AVPU = Alert Voice Pain Unresponsiveness scale; Bpm = beats/breaths per minute; CRT = capillary refill time; FEAST-PET = Fluid Expansion as Supportive Therapy-Pediatric Emergency Triage; GCS = Glasgow Coma Scale; IQR = interquartile range; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; PAWS = Paediatric Advanced Warning Score; PEWS = Pediatric Early Warning Score; PEWS-RL = PEWS-Resource Limited; PICU = paediatric intensive care unit; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2; qSOFA = quick Sequential Organ Failure Assessment; SIRS = Systemic Inflammatory Response Syndrome.

Baseline characteristic	**Overall** N = 1,550 Median (IQR); n/N (%)	**PICU survival**		**p-value**[1]
		Yes N = 1,453 Median (IQR); n/N (%)	**No** N = 97 Median (IQR); n/N (%)	
Demographics				
Age (months)	14.0 (4.0, 73.0)	13.0 (4.0, 73.0)	18.0 (6.0, 60.0)	0.30
Male sex	927 / 1,550 (60%)	873 / 1,453 (60%)	54 / 97 (56%)	0.40
Known comorbidity[*]	266 / 1,527 (17%)	239 / 1,433 (17%)	27 / 94 (29%)	0.003
Perinatal history				
Reported preterm birth[*]	111 / 1,466 (7.6%)	104 / 1,379 (7.5%)	7 / 87 (8.0%)	0.90
Reported low birthweight[*]	193 / 1,312 (15%)	180 / 1,237 (15%)	13 / 75 (17%)	0.50
Location of residence				
Travel time to hospital (minutes)	69.0 (27.0, 156.0)	69.0 (27.0, 156.0)	88.0 (50.0, 197.0)	0.008
Distance to hospital (kilometres)	60.8 (14.8, 147.0)	60.8 (14.8, 139.0)	80.8 (31.4, 178.0)	0.006
Illness history				
Duration of illness (days)[*]	3.0 (2.0, 5.0)	3.0 (2.0, 5.0)	3.0 (2.0, 7.0)	0.07

Baseline characteristic	Overall N = 1,550 Median (IQR); n/N (%)	PICU survival		p-value[1]
		Yes N = 1,453 Median (IQR); n/N (%)	No N = 97 Median (IQR); n/N (%)	
Care prior to admission at study site	1,160 / 1,550 (75%)	1,082 / 1,453 (74%)	78 / 97 (80%)	0.20
Traditional healer	30 / 1,160 (2.6%)	27 / 1,082 (2.5%)	3 / 78 (3.8%)	0.40
Government primary health centre	296 / 1,160 (26%)	279 / 1,082 (26%)	17 / 78 (22%)	0.40
Private pharmacy	142 / 1,160 (12%)	131 / 1,082 (12%)	11 / 78 (14%)	0.60
Government hospital	114 / 1,160 (9.8%)	101 / 1,082 (9.3%)	13 / 78 (17%)	0.04
Non-governmental healthcare provider	133 / 1,160 (11%)	118 / 1,082 (11%)	15 / 78 (19%)	0.03
Private hospital/clinic	449 / 1,160 (39%)	415 / 1,082 (38%)	34 / 78 (44%)	0.40
Study site	171 / 1,160 (15%)	158 / 1,082 (15%)	13 / 78 (17%)	0.60
Satellite clinic of study site	195 / 1,160 (17%)	188 / 1,082 (17%)	7 / 78 (9.0%)	0.06
Other healthcare provider	49 / 1,160 (4.2%)	46 / 1,082 (4.3%)	3 / 78 (3.8%)	>0.90
Number of previous care encounters*	1.0 (1.0, 2.0)	1.0 (1.0, 2.0)	1.0 (1.0, 2.0)	0.30
Overnight admission prior to presentation	268 / 1,550 (17%)	239 / 1,453 (16%)	29 / 97 (30%)	< 0.001
Inter-hospital transfer to study site	318 / 1,550 (21%)	292 / 1,453 (20%)	26 / 97 (27%)	0.11
Satellite clinic of study site	194 / 318 (61%)	186 / 292 (64%)	8 / 26 (31%)	< 0.001
Other	108 / 318 (34%)	96 / 292 (33%)	12 / 26 (46%)	0.20
Self-transfer	16 / 318 (5.0%)	10 / 292 (3.4%)	6 / 26 (23%)	< 0.001
Intra-hospital admission to PICU	294 / 1,550 (19%)	262 / 1,453 (18%)	32 / 97 (33%)	< 0.001
Anthropometrics				
Weight-for-age z-score	-1.40 (-2.41, -0.53)	-1.39 (-2.37, -0.51)	-1.81 (-3.37, -0.97)	0.003
Vital signs				
Axillary temperature (°C)*	36.9 (36.4, 37.6)	36.9 (36.4, 37.6)	37.0 (36.4, 37.7)	0.90
Fever (≥ 37.5°C)	462 / 1,517 (30%)	432 / 1,423 (30%)	30 / 94 (32%)	0.80

Baseline characteristic	Overall N = 1,550 Median (IQR); n/N (%)	PICU survival		p-value[1]
		Yes N = 1,453 Median (IQR); n/N (%)	No N = 97 Median (IQR); n/N (%)	
Hypothermia (< 35.5°C)	46 / 1,517 (3.0%)	37 / 1,423 (2.6%)	9 / 94 (9.6%)	0.001
Heart rate (bpm)[*]				
< 12 months	165.0 (148.0, 176.0)	165.0 (148.0, 176.0)	160.0 (145.0, 176.8)	0.50
12 to 59 months	152.0 (132.0, 169.5)	153.0 (134.0, 170.0)	142.0 (124.0, 161.0)	0.03
60 months to 12 years	115.0 (98.0, 132.0)	113.0 (98.0, 129.0)	150.0 (132.0, 164.0)	< 0.001
> 12 years	99.0 (83.0, 114.0)	98.0 (81.0, 112.0)	128.0 (100.8, 151.5)	0.05
Respiratory rate (bpm)[*]				
< 12 months	56.0 (48.0, 62.0)	56.0 (48.0, 62.0)	56.0 (48.5, 68.0)	0.60
12 to 59 months	48.0 (38.0, 60.0)	50.0 (38.0, 60.0)	46.0 (38.0, 56.5)	0.50
60 months to 12 years	30.0 (26.0, 36.0)	30.0 (26.0, 36.0)	42.0 (32.0, 52.0)	< 0.001
> 12 years	26.0 (24.0, 32.0)	26.0 (24.0, 32.0)	39.0 (35.8, 40.0)	0.002
Oxygen saturation (%)[$*]	98.0 (96.0, 99.0)	98.0 (97.0, 99.0)	98.0 (89.2, 99.0)	< 0.001
On supplemental oxygen[*]	767 / 1,526 (50%)	710 / 1,431 (50%)	57 / 95 (60%)	0.05
Clinical assessment				
Not alert[†*]	351 / 1,495 (23%)	295 / 1,406 (21%)	56 / 89 (63%)	< 0.001
Prolonged central capillary refill time[†*]	126 / 1,533 (8.2%)	95 / 1,436 (6.6%)	31 / 97 (32%)	< 0.001
Weak pulse[*]	183 / 1,508 (12%)	155 / 1,416 (11%)	28 / 92 (30%)	< 0.001
Cool extremities[*]	406 / 1,540 (26%)	363 / 1,443 (25%)	43 / 97 (44%)	< 0.001
Respiratory distress[†*]	900 / 1,528 (59%)	833 / 1,436 (58%)	67 / 92 (73%)	0.01
Lung crepitations[*]	561 / 1,540 (36%)	527 / 1,445 (36%)	34 / 95 (36%)	0.90
Laboratory parameters				
White cell count ($x10^9$ cells/l)[†*]	12.4 (8.2, 17.2)	12.4 (8.4, 17.2)	12.1 (6.2, 17.8)	0.20

Baseline characteristic	Overall N = 1,550 Median (IQR); n/N (%)	PICU survival		p-value[1]
		Yes N = 1,453 Median (IQR); n/N (%)	No N = 97 Median (IQR); n/N (%)	
Neutrophil count (x10^9 cells/l)[†*]	6.0 (3.2, 10.1)	6.0 (3.2, 10.1)	5.8 (2.5, 9.5)	0.40
Lymphocyte count (x10^9 cells/l)[†*]	3.7 (2.1, 6.2)	3.7 (2.1, 6.3)	3.1 (1.5, 6.0)	0.05
Haemoglobin (g/dl)[†*]	108.0 (95.0, 122.0)	109.0 (97.0, 123.0)	96.0 (79.0, 108.0)	< 0.001
Platelet count (x10^9 cells/l)[†*]	376.0 (214.5, 506.5)	384.0 (229.2, 509.5)	242.0 (76.0, 371.0)	< 0.001
C-reactive protein (mg/l)[†*]	7.0 (2.0, 33.0)	7.0 (2.0, 30.0)	13.5 (2.0, 74.5)	0.03
Glucose (mg/dl)[†*]	104.0 (86.0, 129.0)	103.5 (86.0, 128.0)	111.5 (86.5, 159.8)	0.20
Severity scores				
FEAST-PET[*]	2.0 (1.0, 3.0)	2.0 (1.0, 3.0)	3.0 (2.0, 4.0)	< 0.001
LqSOFA[*]	0.0 (0.0, 1.0)	0.0 (0.0, 1.0)	1.0 (1.0, 2.0)	< 0.001
PAWS[*]	3.0 (2.0, 5.0)	3.0 (2.0, 5.0)	5.0 (3.0, 7.0)	< 0.001
PEWS[*]	6.0 (3.0, 8.0)	5.0 (3.0, 8.0)	7.5 (6.0, 11.0)	< 0.001
PEWS-IRISH[*]	4.0 (2.0, 7.0)	4.0 (2.0, 6.0)	6.0 (4.2, 8.8)	< 0.001
PEWS-RL[*]	3.0 (2.0, 3.0)	3.0 (1.0, 3.0)	3.0 (3.0, 4.0)	< 0.001
qPELOD-2[*]	0.0 (0.0, 1.0)	0.0 (0.0, 1.0)	1.0 (0.0, 2.0)	< 0.001
qSOFA[*]	1.0 (1.0, 1.0)	1.0 (1.0, 1.0)	2.0 (1.0, 2.0)	< 0.001
SIRS[*]	2.0 (1.0, 2.0)	2.0 (1.0, 2.0)	2.0 (1.0, 3.0)	< 0.001

[1]Wilcoxon rank sum test; Pearson's Chi-squared test; Fisher's exact test

5.3.2 Illness journeys

Admissions originated from 23 of Cambodia's 25 provinces (Figure 5.3-2) and median travel time from a child's residence to the hospital was 69 minutes (IQR = 27 to 156 minutes). Children had been sick for a median of three days (IQR = 2 to 5 days) prior to admission to the study site, with the majority (1,160/1,550; 74.8%) seeking care from at least one other healthcare provider prior to presentation. Approximately one in five children (268/1,550; 17.3%) had been admitted overnight at another healthcare facility. A similar proportion were inter-hospital transfers (318/1,550; 20.5%), the majority referred from the hospital's satellite clinic (194/318; 61.0%) located approximately 45 minutes from the main site.

Over two-thirds of PICU admissions originated from the hospital's ER (1,067/1,550; 68.8%), whilst 294 (294/1,550; 19.0%) were intra-hospital transfers from one of the hospital's three acute wards. Most children were admitted with febrile illnesses (history of fever and/or a documented axillary temperature ≥ 37.5°C; 1,128/1,540; 79.1%). Common reasons for PICU admission included respiratory distress (989/1,550; 63.8%), circulatory instability (545/1,550; 35.2%), and impaired consciousness (352/1,550; 22.7%).

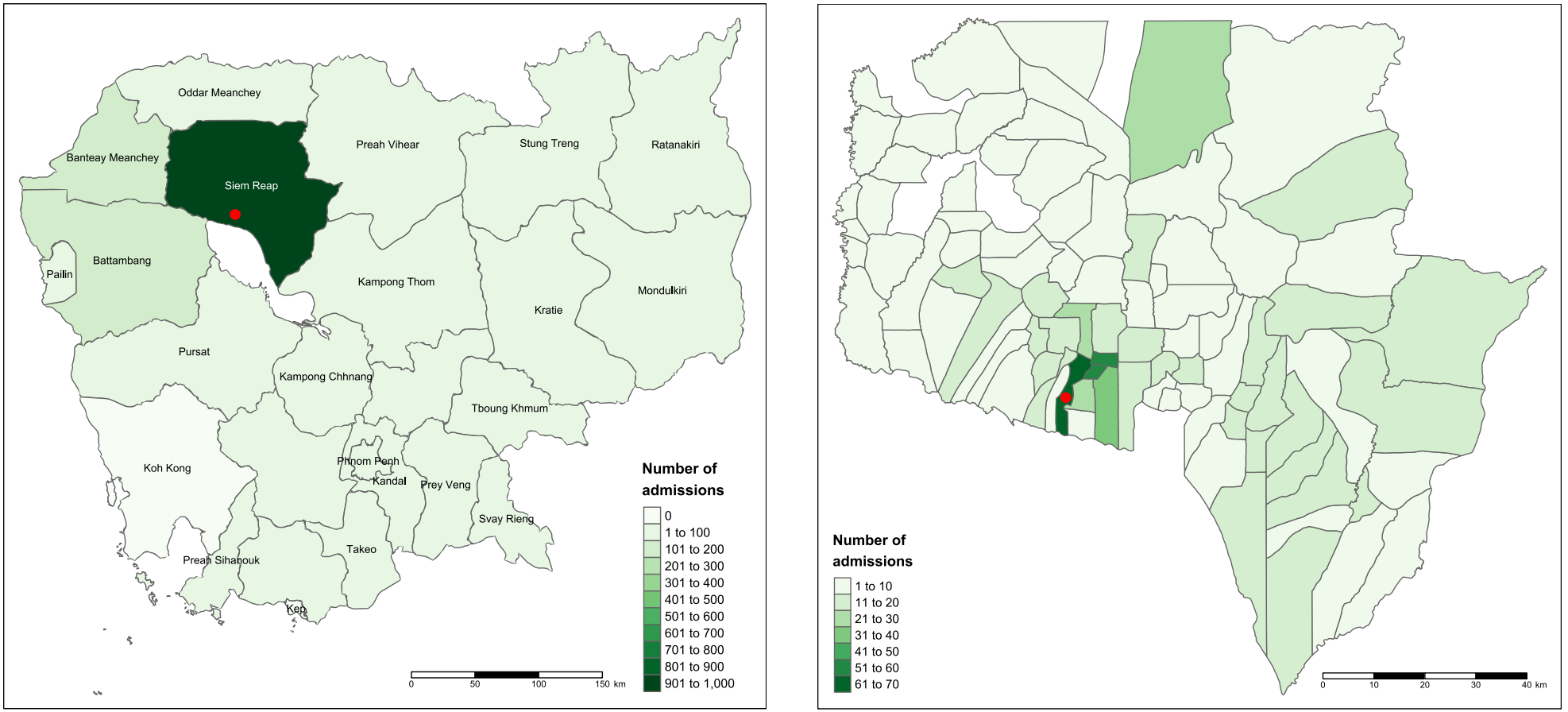


Figure 5.3-2: Maps depicting locations of residence for study participants. Left panel: Distribution of admissions across Cambodian provinces. Right panel: Distribution of admissions across communes in Siem Reap province. The study site is indicated by the red dot.

5.3.3 Baseline characteristics

Children with known comorbidities (27/94 [29%] vs. 239/1,433 [17%]; as summarised in Figure 5.3-3) and lower WAZ (-1.81 vs. -1.39), as well as those sick for longer (3 [IQR = 2 to 7] days vs. 3 [IQR = 2 to 5] days], admitted elsewhere prior to presentation (29/97 [30%] vs. 239/1,453 [16%]), residing further from the hospital (80.8 km vs. 60.8 km), and intra-hospital transfers (32/97 [33%] vs. 262/1,453 [18%]) were all less likely to survive their PICU admission. Presentations with hypothermia (9/94 [9.6%] vs. 37/1,423 [2.6%]), lower SpO_2 (98.0% [IQR = 89.2 to 99.0] vs. 98.0% [IQR = 97.0 to 99.0]), impaired consciousness (56/89 [63%] vs. 295/1,406 [21%]), and signs of cardiovascular (prolonged CRT [31/97 {32%} vs. 95/1,436 {6.6%}], weak pulses [28/92 {30%} vs. 155/1,416 {11%}], or cool extremities [43/97 {44%} vs. 363/1,443 {25%}]) or respiratory compromise (67/92 [73%] vs. 833/1,436 [58%]) were all more likely to meet the primary outcome. Baseline tachycardia (150 bpm vs. 113 bpm and 128 bpm vs. 98 bpm) and tachypnoea (42 bpm vs. 30 bpm and 39 bpm vs. 26 bpm) were associated with death during PICU stay for children aged five years and older but this association was not observed for younger children. Statistically significant differences between admissions that did and did not meet the primary outcome were observed for a number of laboratory parameters, however only for haemoglobin was this association compatible with a potentially clinically important difference (96.0 vs. 109.0 g/dL). All nine clinical severity scores were higher in admissions that resulted in death.

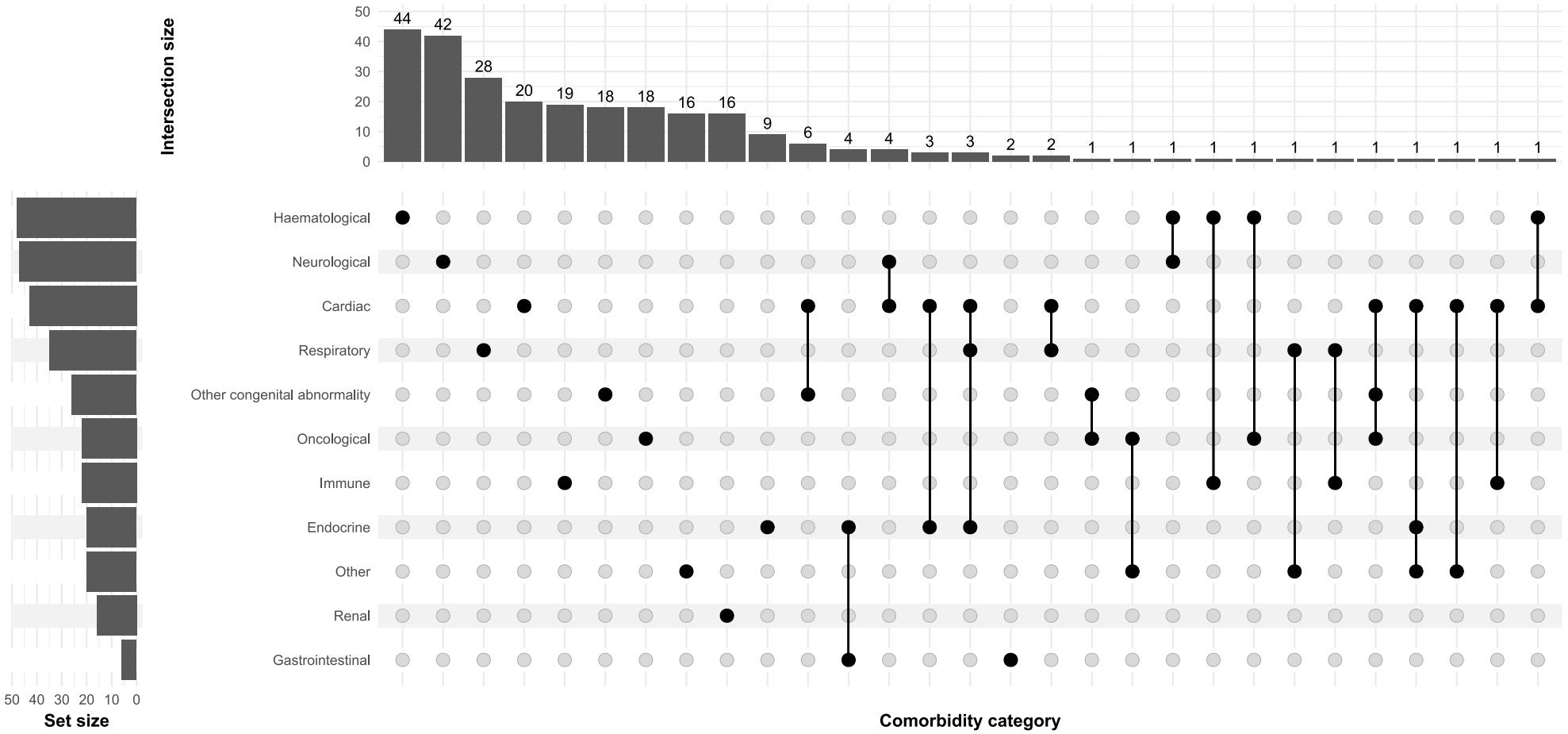


Figure 5.3-3: UpSet plot illustrating categories of comorbidities amongst study population. Sets contain children with a particular category of comorbidity. Intersections contain children with a particular combination of comorbidity categories.

5.3.4 Clinical outcomes

Vital organ support was provided to 41.7% (647/1,550) of PICU admissions: 516/1,550 (33.3%) were non-invasively ventilated, 354/1,550 (22.8%) mechanically ventilated, 98/1,550 (6.3%) received inotropic therapy, and 4/1,550 (0.3%) received peritoneal dialysis. Median length of stay on the unit was two days (IQR = 1 to 4 days). The most frequent clinical discharge diagnoses included pneumonia (501/1,550; 32.3%), bronchiolitis (319/1,550; 20.6%), and dengue (221/1,550; 14.3%).

The PICU mortality rate was 6.3% (97/1,550), with 85 children dying during their PICU stay and a further 12 discharged to die at home. The most common causes of death were pneumonia (52/97; 53.6%), undifferentiated sepsis (15/97; 15.5%), and cardiac failure (9/97; 9.3%), although 28 children (28/97; 28.9%) had more than one cause of death implicated (Figure 5.3-4). Median time to death was four days (IQR = 1 to 7 days; Figure 5.3-5) post-PICU admission. Amongst the 12 children discharged to die at home, death was confirmed in two (via the 12-month post-PICU discharge telephone follow-up), whilst caretakers of the remaining 10 children were uncontactable.

Of the 1,453 admissions surviving to leave PICU, vital status 12 months post-PICU discharge could be ascertained for 782 (53.8%; 782/1,453). Of these, 33/782 admissions (4.2%) had resulted in death (25/672 [3.7%] individual children). The median time to death was one month (IQR = 1 to 4 months) post-PICU discharge (Figure 5.3-5).

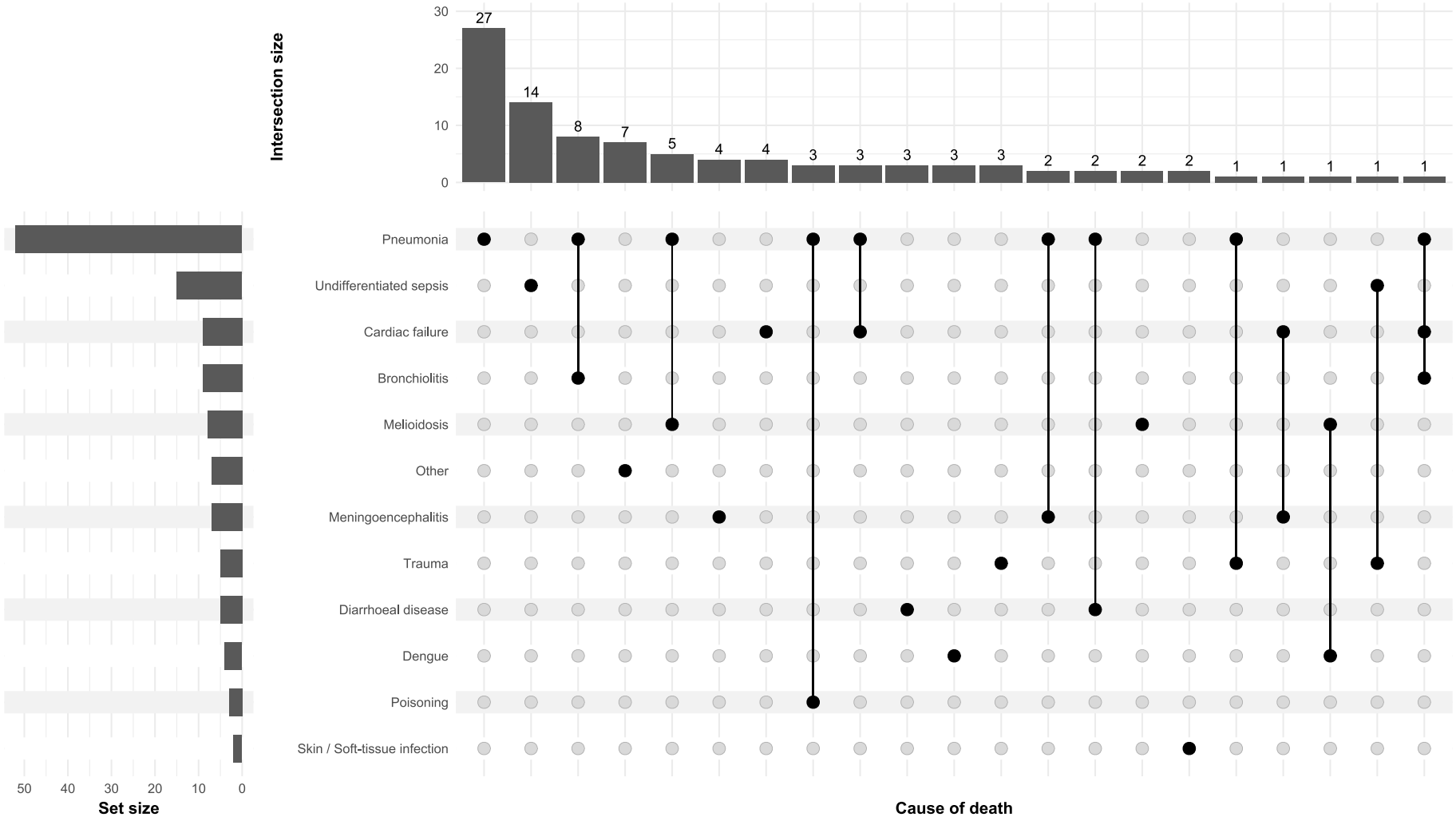


Figure 5.3-4: UpSet plot illustrating clinical diagnoses amongst study participants that died. Sets contain children with a particular diagnosis. Intersections contain children with a particular combination of diagnoses.

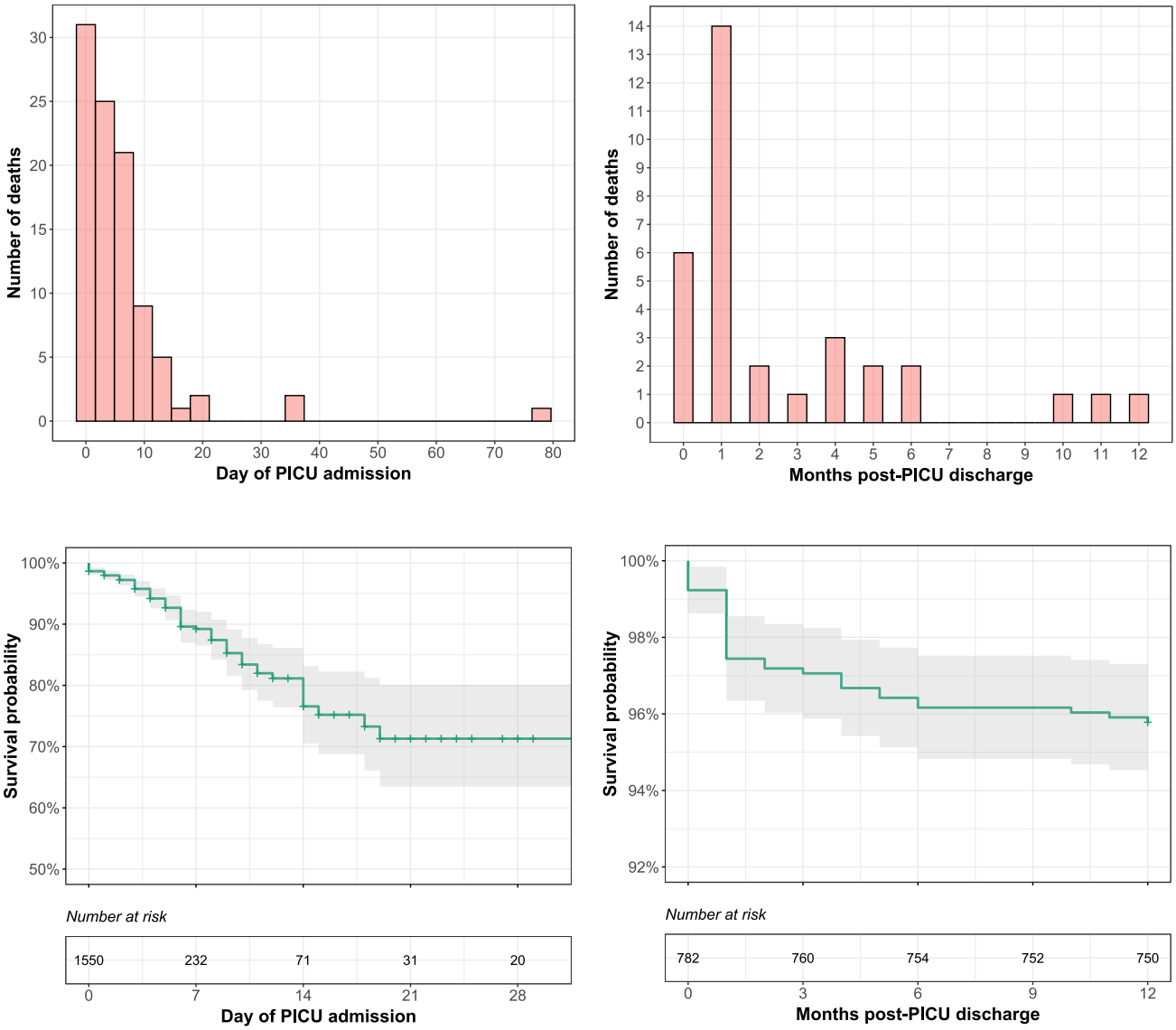


Figure 5.3-5: Time to meeting the primary and secondary outcomes. Left panel: days to death after PICU admission. Right panel: months to death after PICU discharge. Grey shaded areas = 95% confidence intervals. PICU = paediatric intensive care unit.

5.3.5 Moderate utility of PAWS, qPELOD-2, and qSOFA for triage of critically ill children

All scores achieved comparable discrimination (Figure 5.3-6; AUCs = 0.71 to 0.76) with the exception of SIRS, for which discrimination was poorer (AUC = 0.59; 95% CI 0.53 to 0.65). Across all scores, admissions with higher baseline scores were more likely to progress to meet the primary outcome, although this association was less pronounced for PEWS-RL and SIRS (Figure 5.3-7). For scores with multiple levels (PAWS, PEWS, and PEWS-IRISH) the increase in proportion of admissions

meeting the primary outcome across lower levels was modest, indicating redundancy in these more granular scoring systems.

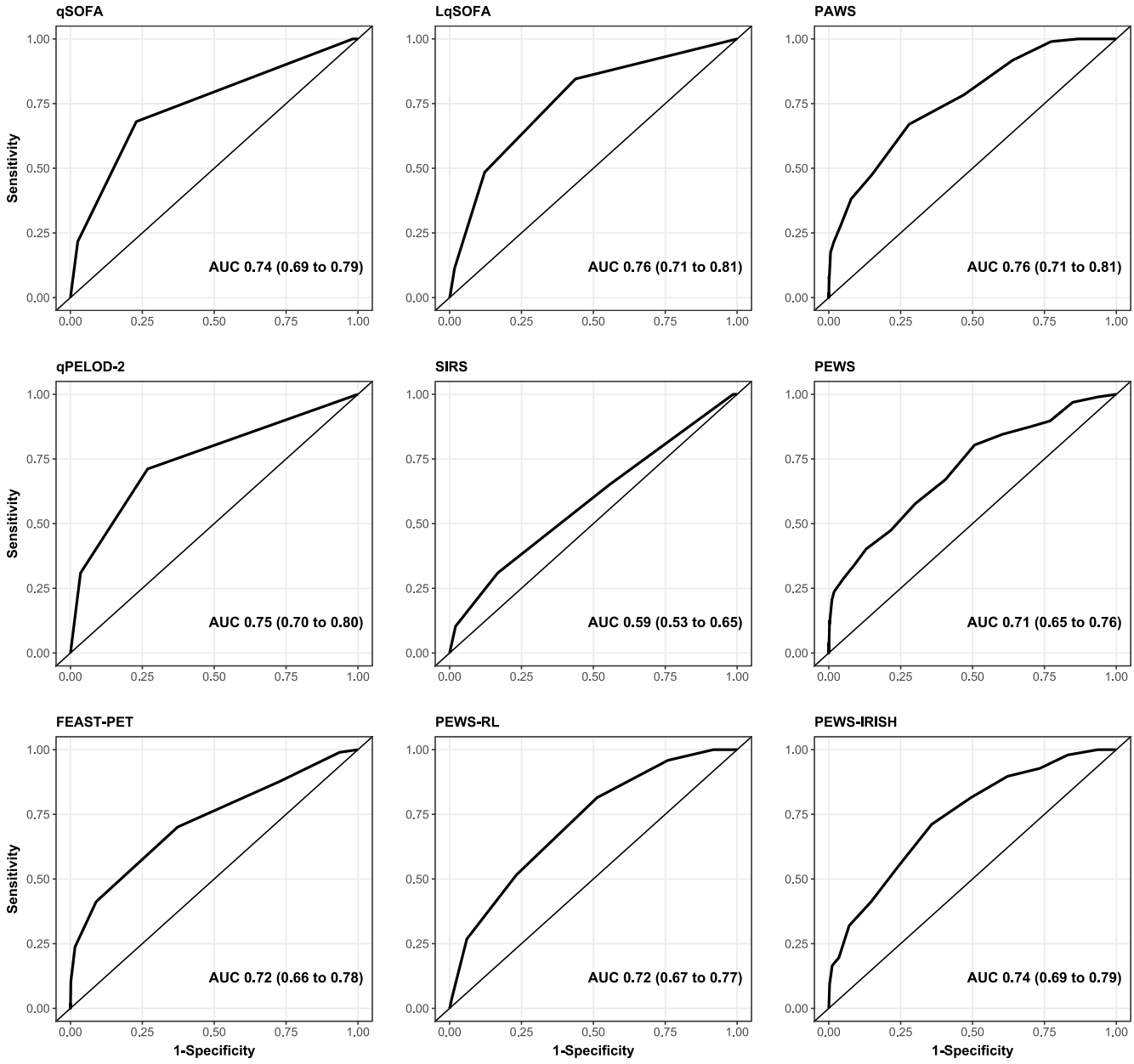


Figure 5.3-6: Discrimination of existing severity scores. AUC = area under the receiver operating characteristic curve. FEAST-PET = Fluid Expansion as Supportive Therapy-Pediatric Emergency Triage; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; PAWS = Paediatric Advanced Warning Score; PEWS = Pediatric Early Warning Score; PEWS-RL = PEWS-Resource Limited; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2; qSOFA = quick Sequential Organ Failure Assessment; SIRS = Systemic Inflammatory Response Syndrome.

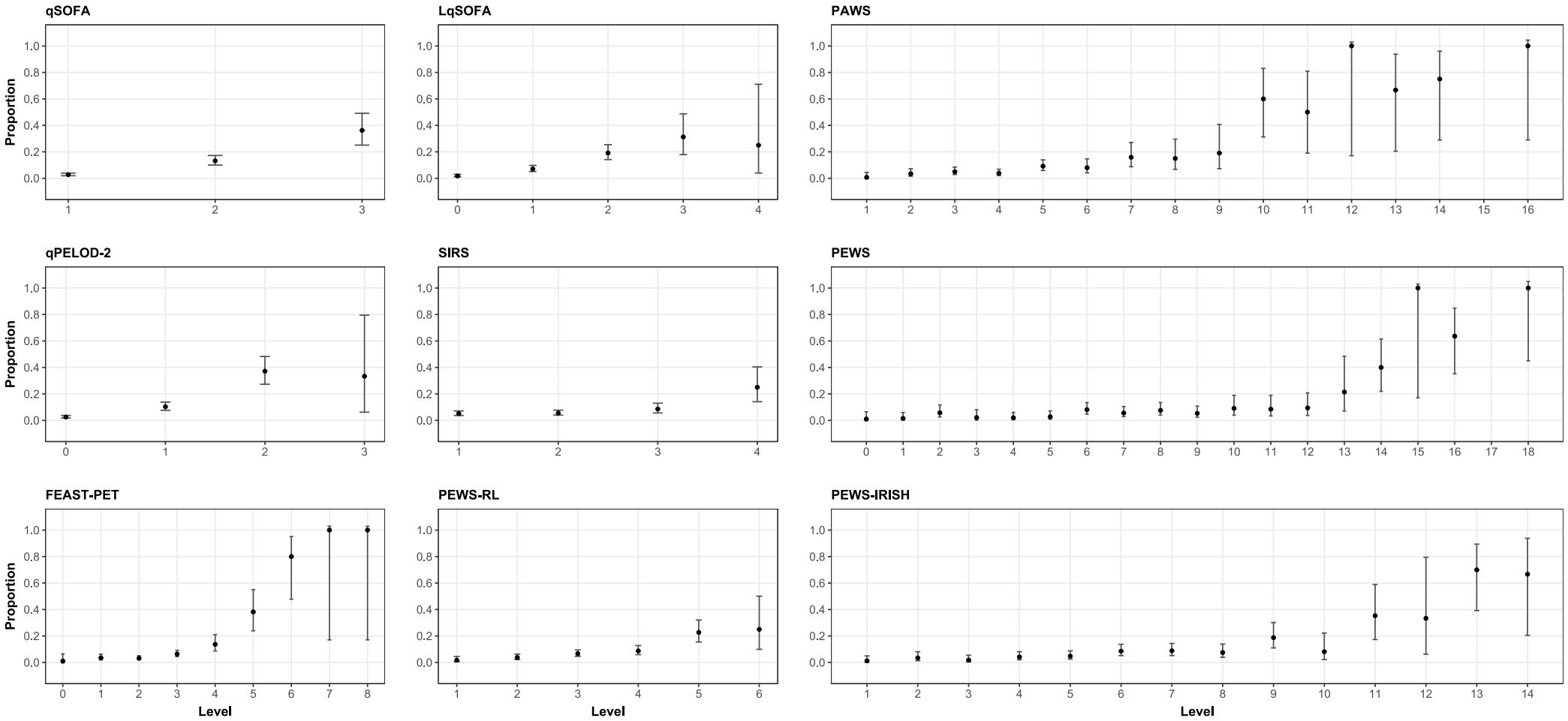


Figure 5.3-7: Calibration of existing severity scores. Proportion of admissions at each level of each score that died during their PICU stay. Error bars indicate Wilson 95% confidence intervals. PICU = paediatric intensive care unit. FEAST-PET = Fluid Expansion as Supportive Therapy-Pediatric Emergency Triage; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; PAWS = Paediatric Advanced Warning Score; PEWS = Pediatric Early Warning Score; PEWS-RL = PEWS-Resource Limited; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2; qSOFA = quick Sequential Organ Failure Assessment; SIRS = Systemic Inflammatory Response Syndrome.

At a cut-off of ≥ 1, the qPELOD-2 score demonstrated a sensitivity of 0.71 (95% CI = 0.62 to 0.80) and specificity of 0.73 (95% CI = 0.71 to 0.75). No other score achieved a sensitivity and specificity of > 0.70 at any cut-off (Figure 5.3-8). Three scores demonstrated potential for stratifying PICU admissions into low- and high-risk groups (Figure 5.3-9), achieving small but important changes in pre-test probability using a single cut-off: qPELOD-2 at a cut-off of ≥ 1 (PLR = 2.65; 95% CI = 2.28 to 3.09 and NLR = 0.40; 95% CI = 0.29 to 0.54), qSOFA at a cut-off of ≥ 2 (PLR = 2.97; 95% CI = 2.52 to 3.50 and NLR = 0.42; 95% CI = 0.31 to 0.56), and PAWS at a cut-off of ≥ 5 (PLR = 2.40; 95% CI = 2.04 to 2.82 and NLR = 0.46; 95% CI = 0.34 to 0.61).

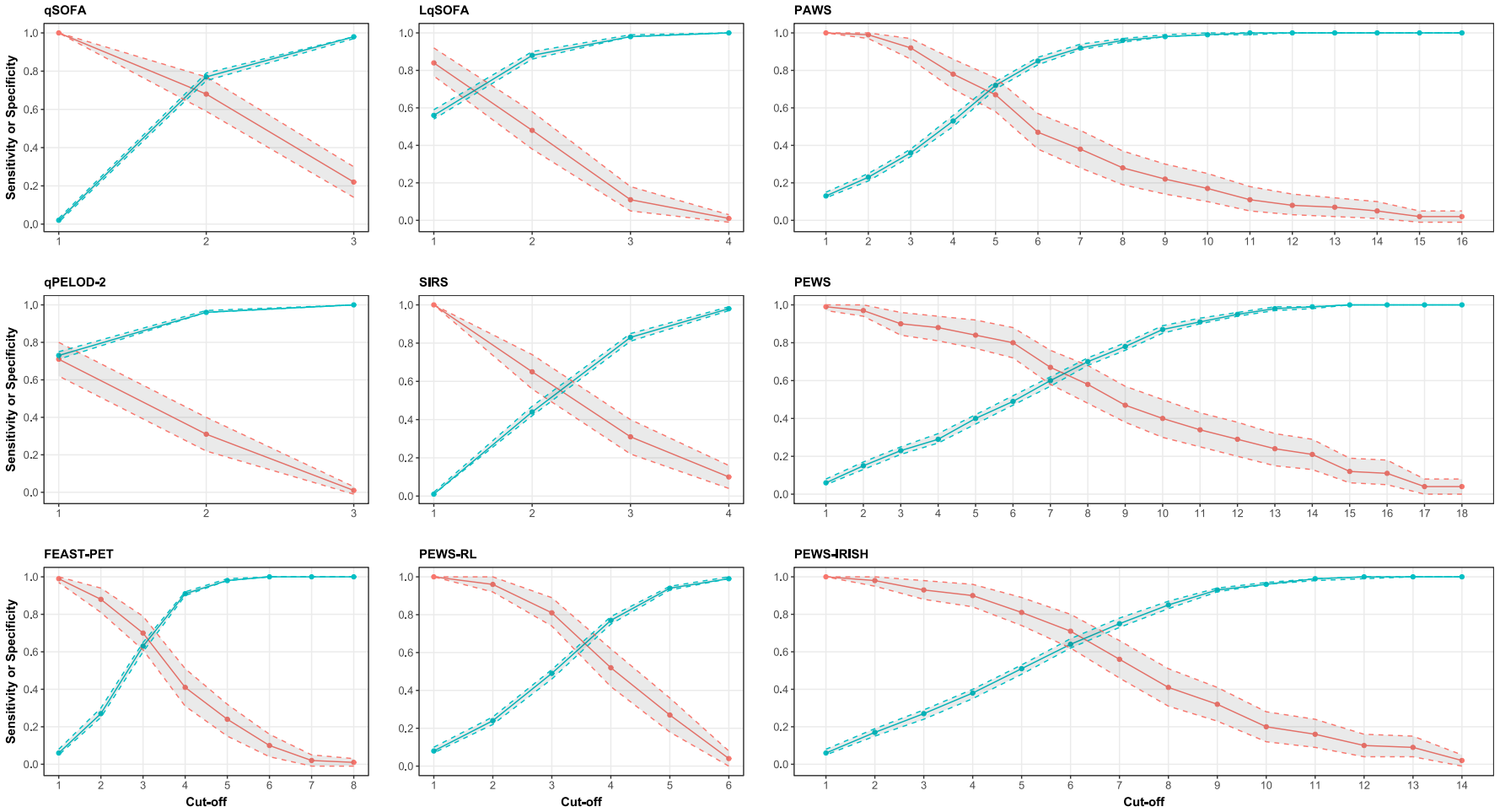


Figure 5.3-8: Sensitivity and specificity of existing severity scores. Change in sensitivity (red line) and specificity (blue line) at increasing cut-offs of the severity scores. Grey shaded areas = 95% confidence intervals. FEAST-PET = Fluid Expansion as Supportive Therapy-Pediatric Emergency Triage; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; PAWS = Paediatric Advanced Warning Score; PEWS = Pediatric Early Warning Score; PEWS-RL = PEWS-Resource Limited; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2; qSOFA = quick Sequential Organ Failure Assessment; SIRS = Systemic Inflammatory Response Syndrome.

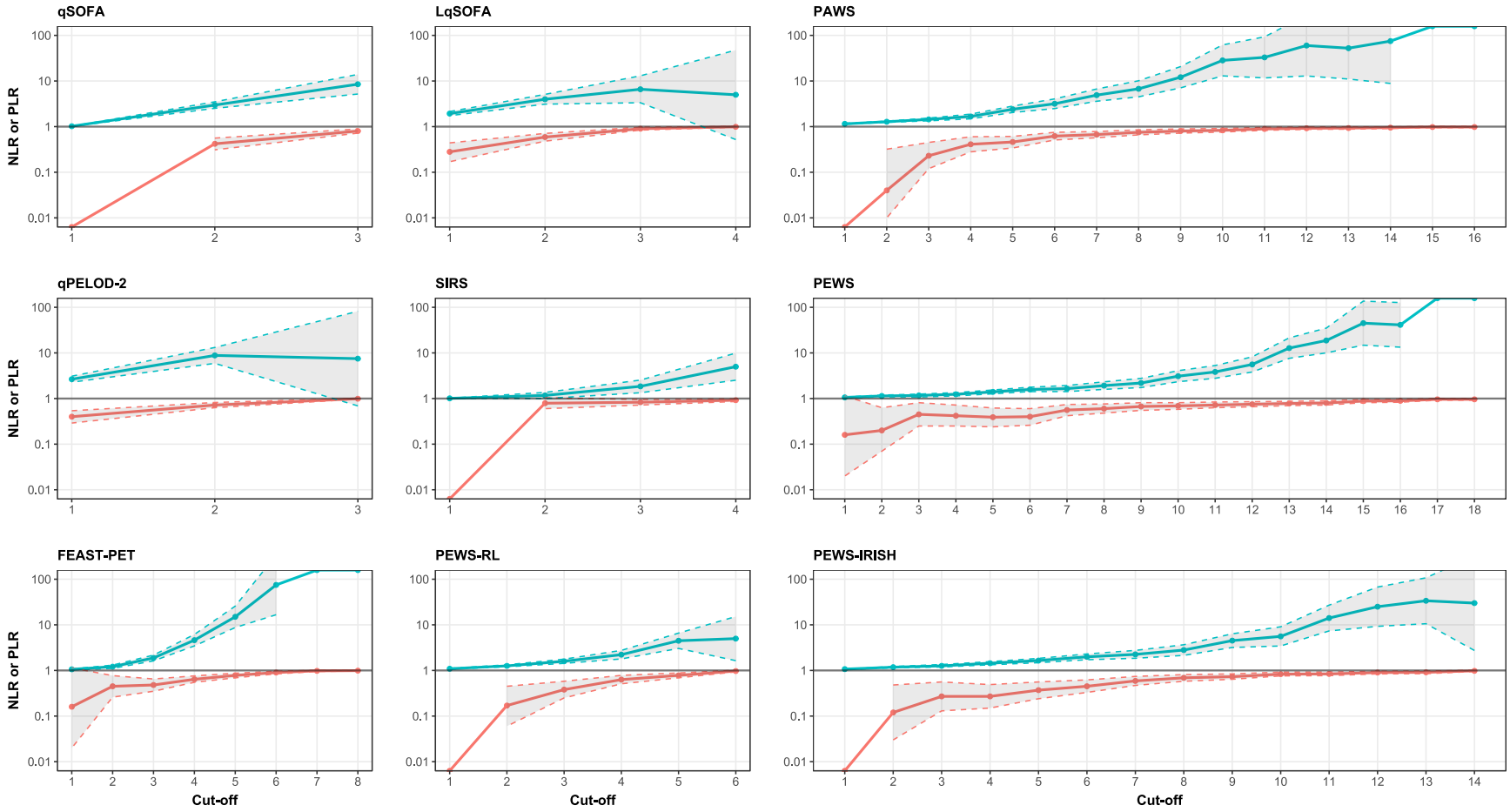


Figure 5.3-9: Negative and positive likelihood ratios of existing severity scores. Change in NLR (red line) and PLR (blue line) at increasing cut-offs of the severity scores, illustrated on a $\log_{10}$ scale. Grey shaded areas = 95% confidence intervals. Confidence intervals not calculable for certain cut-offs of some scores as sensitivity or specificity were 100%. FEAST-PET = Fluid Expansion as Supportive Therapy-Pediatric Emergency Triage; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; NLR = negative likelihood ratio; PAWS = Paediatric Advanced Warning Score; PEWS = Pediatric Early Warning Score; PEWS-RL = PEWS-Resource Limited; PLR = positive likelihood ratio; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2; qSOFA = quick Sequential Organ Failure Assessment; SIRS = Systemic Inflammatory Response Syndrome.

5.3.6 Superior performance of a new prognostic model

Assessment of the relationship between continuous candidate predictors and the primary outcome did not identify serious violations of linearity (Figure 5.3-10). Age-dependent relationships between the primary outcome and heart rate and respiratory rate were evident and this was confirmed via examination of stratum-specific ORs and LRTs for interaction ($p < 0.001$). Examination of stratum-specific ORs and an LRT for interaction between SpO_2 and supplemental oxygen status did not indicate evidence of interaction ($p = 0.92$) and thus only the main effects for these parameters were included in the model. The full model, including the formulae to calculate the probability that a PICU admission will result in death, is presented (Table 5.3-2).

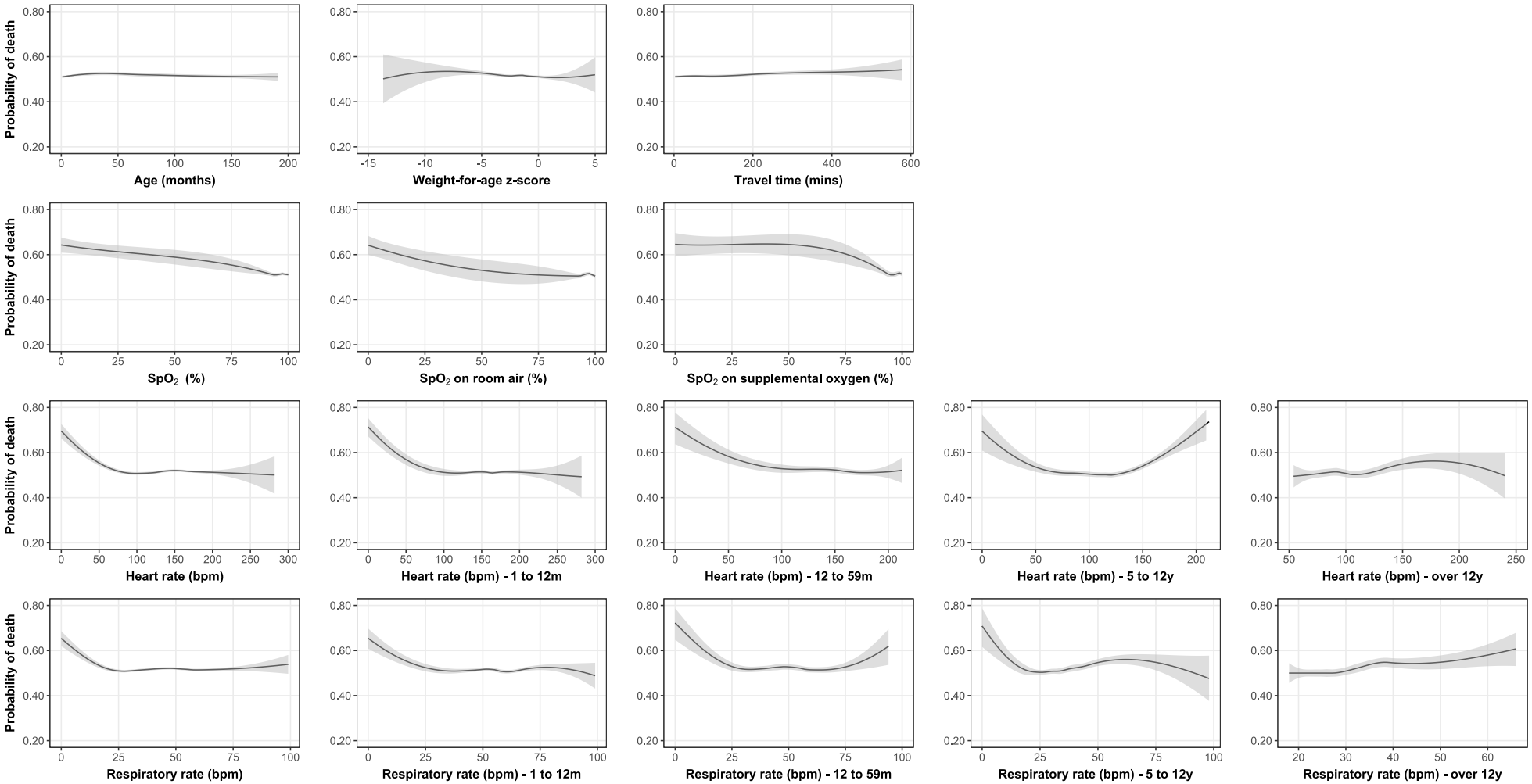


Figure 5.3-10: Relationship between continuous candidate predictors and the primary outcome. Black line = probability of death during PICU admission. Grey shaded areas = 95% confidence intervals. Bpm = beats / breaths per minute; PICU = paediatric intensive care unit.

Table 5.3-2: Clinical prediction model to estimate the probability that a PICU admission will end in death. Predictors spanning the five clinical domains are presented along with their regression coefficients and the formulae required to calculate the probability that a PICU admission will end in death. [*]Assessed by a study paediatrician blinded to outcome status using the following adapted working definition: *any previous health condition known to be present at PICU admission severe enough to require specialty paediatric care and probably a period of hospitalisation over 12 months;*[233] [#]calculated (R package: *zscorer*)[153] using WHO (children < 10 years)[234,235] and US-CDC (children ≥ 10 years)[236] reference ranges; [$]travel by road estimated using GoogleMaps; [†]admission from acute medical or surgical ward; [‡]CRT > 2 seconds; [¶]GCS < 15 and/or AVPU < A. AVPU = Alert Voice Pain Unresponsiveness scale; CRT = capillary refill time; GCS = Glasgow Coma Scale; PICU = paediatric intensive care unit; US-CDC = United States Centers for Disease Control; WHO = World Health Organization.

Clinical domain	Predictors	Ridge regression coefficient	p-value
	Intercept	-1.8525	
Background	**Age** (months)	-0.0013	0.02
	Presence of comorbidity[*]	0.2201	0.12
	Weight-for-age z-score[#]	-0.1108	< 0.001
Illness journey	**Estimated travel time** (minutes)[$]	0.0013	0.01
	Intra-hospital transfer[†]	0.4626	< 0.001
Cardiovascular	**Heart rate** (beats per minute)	-0.0002	0.89
	Heart rate (beats per minute) x **Age** (months)	8.6514×10^{-6}	0.14
	Prolonged capillary refill time[‡]	1.0244	< 0.001
Respiratory	**Respiratory rate** (breaths per minute)	0.0090	0.01
	Respiratory rate (breaths per minute) x **Age** (months)	4.8939×10^{-5}	0.03
	Oxygen saturation (%)	-0.0253	< 0.001
	Receipt of supplemental oxygen	0.1808	0.10
Neurological	**Abnormal mental status**[¶]	1.0313	< 0.001

The model estimates the log odds of death during a PICU admission, using the sum of the intercept and the predictors multiplied by their coefficients, according to the following equation:

$$-1.8525 - 0.0013 * age + \begin{cases} 0 \; if \; no \; comorbidity \\ 0.2201 \; if \; comorbidity \end{cases} - 0.1108 * waz + 0.0013 * travel\ time +$$

$$\begin{cases} 0 \; if \; no \; intra\text{-}hopital \; transfer \\ 0.4626 \; if \; intra\text{-}hospital \; transfer \end{cases} - 0.0002 * heart\ rate\ + 8.6514x10^{-6} * age * heart\ rate + \begin{cases} 0 \; if \; no \; prolonged \; CRT \\ 1.0244 \; if \; prolonged \; CRT \end{cases} +$$

$$0.0090 * respiratory\ rate + 4.8939x10^{-5} * age * respiratory\ rate - 0.0253 * oxygen\ saturation +$$

$$\begin{cases} 0 \; if \; no \; supplemental \; oxygen \\ 0.1808 \; if \; supplemental \; oxygen \end{cases} + \begin{cases} 0 \; if \; normal \; mental \; status \\ 1.0313 \; if \; abnormal \; mental \; status \end{cases}$$

To support clinical decision making, the output of the model (log odds) is converted into the probability that a PICU admission will result in death using the following transformation:

$$\Pr(death\ during\ PICU\ admission) = \frac{e^{logodds}}{1 + e^{logodds}}$$

Discrimination of the new model (Figure 5.3-11; AUC = 0.84; 95% CI = 0.80 to 0.88) was significantly better than all of the existing scores (DeLong test; $p < 0.001$), although this is in part expected for a comparison between a newly derived (internally-validated) model and external validation of existing scores. Calibration of the new model appeared best at lower predicted probabilities (Figure 5.3-11), with the model underestimating risk for admissions with observed probabilities of death > 20-25%.

The ability of the model to triage PICU admissions into high- and low-acuity groups at cut-offs of 2.5%, 5%, 7.5%, 10%, and 15% is presented (Table 5.3-3). A cut-off of 10% reflects a triage strategy whereby all PICU admissions with a predicted probability of death ≥ 10% are directed to a high-acuity area (where human and material resources are concentrated) and all other PICU admissions are managed on the main unit. At this cut-off, PICU admissions triaged to the high-acuity area would have a probability of death almost five times that of the general PICU population (PLR = 5.75; 95% CI = 4.57 to 7.23), whereas the probability amongst those triaged to the low-acuity area would be less than half that of the general PICU population (NLR = 0.47; 95% CI = 0.37 to 0.59), and almost a tenth of those triaged to the high-acuity area. At the 10% cut-off, approximately 13.0% of all PICU admissions would be triaged to the high-acuity area, resulting in a ratio of 3:1 incorrect to correct (FP:TP) high-acuity triages.

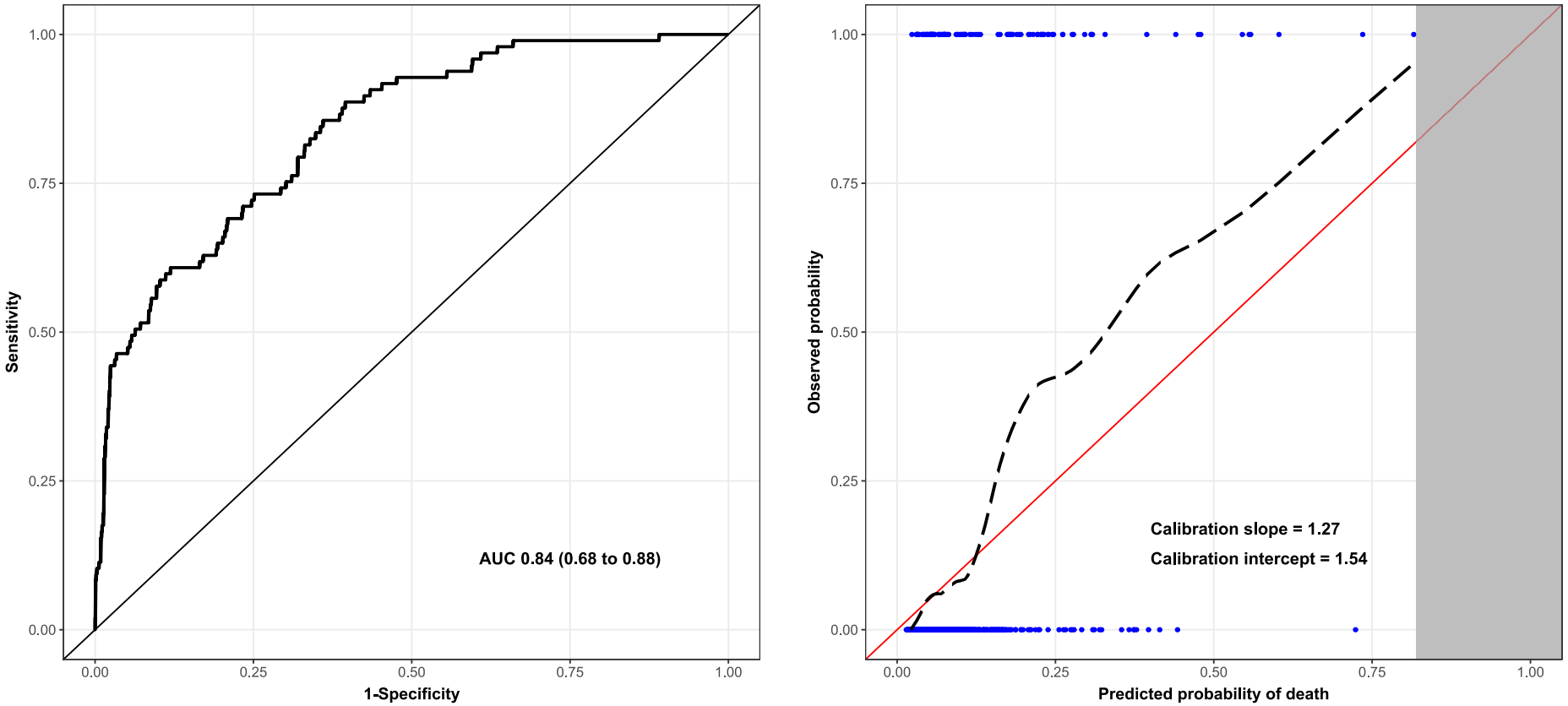


Figure 5.3-11: Discrimination and calibration of the new model. Left panel: discrimination of the new model. Right panel: calibration of the new model. Red line indicates perfect calibration. Dashed line indicates calibration of the model. Blue rug plots indicate distribution of predicted risks for participants who did (top) and did not (bottom) meet the primary outcome. AUC = area under the receiver operating characteristic curve.

Table 5.3-3: Ability of the model to triage PICU admissions. Performance of the model at five cut-offs (decision thresholds or threshold probabilities). A cut-off of 10% reflects a strategy whereby all PICU admissions with a predicted probability of death ≥ 10% are directed to a high-acuity area and all other PICU admissions managed on the main unit. A decrease in threshold probability (cut-off) is associated with an increase in the sensitivity of the triage strategy for identifying high-risk PICU admissions, at the cost of a greater proportion of admissions being directed to the high-acuity area. CI = confidence interval; FN = false negative (high-risk PICU admission triaged to low-acuity area); FP = false positive (low-risk PICU admission triaged to high-acuity area); NLR = negative likelihood ratio; PICU = paediatric intensive care unit; PLR = positive likelihood ratio; TN = true negative (low-risk PICU admission triaged to low-acuity area); TP = true positive (high-risk PICU admission triaged to high-acuity area).

Predicted probability of death	Sensitivity (95% CI)	Specificity (95% CI)	PLR (95% CI)	NLR (95% CI)	Per 1,000 PICU admissions (~63 of which would die)				Percentage of PICU admissions triaged as high-acuity	Ratio of incorrect to correct high-acuity triages
					TP	FP	TN	FN		
2.5%	0.99 (0.97 to 1.00)	0.16 (0.14 to 0.18)	1.18 (1.14 to 1.21)	0.07 (0.01 to 0.46)	62	788	150	1	85.0%	13:1
5%	0.86 (0.79 to 0.93)	0.63 (0.61 to 0.66)	2.31 (2.08 to 2.57)	0.23 (0.14 to 0.37)	54	347	590	9	40.1%	6:1
7.5%	0.65 (0.56 to 0.74)	0.80 (0.78 to 0.82)	3.27 (2.73 to 3.91)	0.44 (0.33 to 0.57)	41	186	751	22	22.7%	5:1
10%	0.58 (0.48 to 0.68)	0.90 (0.88 to 0.92)	5.75 (4.57 to 7.23)	0.47 (0.37 to 0.59)	36	94	844	26	13.0%	3:1
15%	0.46 (0.37 to 0.56)	0.96 (0.95 to 0.97)	11.43 (8.22 to 15.88)	0.56 (0.46 to 0.67)	29	39	899	34	12.3%	1:1

5.3.7 New model could support triage across a range of settings

There is great heterogeneity in the approach to critical care provision across different resource-constrained contexts, with the relative value of a TP and FP depending on the available human and material resources. Decision curve analyses accounting for these differing contexts indicate that using the model to support triage decisions could provide utility at cut-offs ≥ 7.5% (Figure 5.3-12), or simply put, in contexts where it might be desirable and feasible to manage up to a quarter (22.7%) of critical care admissions in a high-acuity area and tolerate up to 5:1 incorrect to correct (FP:TP) high-acuity triages.

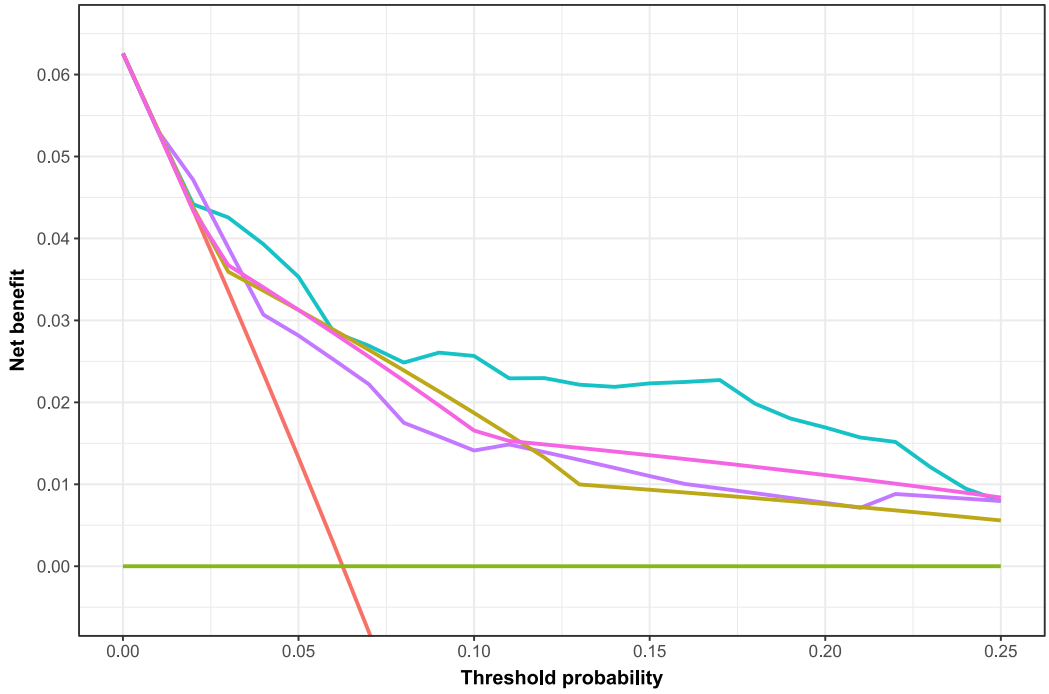


Figure 5.3-12: Clinical utility of the new model across a range of plausible decision thresholds. A cut-off (decision threshold or threshold probability) of 10% reflects a triage strategy whereby all PICU admissions with a predicted probability of death ≥ 10% are directed to a high-acuity area and all other PICU admissions managed on the main unit. The net benefit of the new model (turquoise line) is compared to the three existing scores that demonstrated potential for stratifying PICU admissions into low- and high-risk groups from the external validation (qPELOD-2 [pink line]; qSOFA [brown line]; PAWS [purple line]), and a 'triage-all' (red line) and 'triage-none' (green line) approach. Above a cut-off of 7.5% (i.e., scenarios where up to 5 incorrect high-risk triages can be tolerated for every one correct high-risk triage) using the new model to triage PICU admissions appears to be the optimal strategy. Moving from left to right along the x-axis (increasing threshold for triage to the high acuity area) reflects increasing 'penalisation' of an incorrect high-acuity triage (false positive), indicative of contexts in which resources in the high-acuity area may be more constrained. PAWS = Paediatric Advanced Warning Score; PICU = paediatric intensive care unit; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2; qSOFA = quick Sequential Organ Failure Assessment.

5.3.8 Sensitivity analyses

In resource-constrained settings, where hospital admission is often associated with substantial out-of-pocket expenditure for patients and families, children with terminal illnesses are not infrequently 'discharged home to die' at the request of their caretaker. Although the expectation of the treating healthcare worker in these cases is that the child will not survive their current illness, to account for the fact that these children had not died prior to leaving PICU, sensitivity analyses were conducted excluding these children (n = 12), as well as those children whose death was judged to have been related to a second illness acquired during their PICU stay (n = 9) by either of two study doctors. The results of these analyses produced similar results to the main analyses (Table 5.3-4).

Table 5.3-4: Sensitivity analyses for primary outcome excluding children 'discharged home to die' or in whom a second illness acquired during their PICU stay was judged to have contributed to their death. Comparison between the primary analysis (n = 1,550; outcome events = 97) and a sensitivity analysis (n = 1529; outcome events = 76) where admissions which met the primary outcome but in which the death was judged to have been related to a second illness acquired during the PICU stay (n = 9) and admissions which were discharged to die at home (n = 12) were excluded. AUC = area under the receiver operating characteristic curve; CI = confidence interval; FEAST-PET = Fluid Expansion as Supportive Therapy-Pediatric Emergency Triage; LqSOFA = Liverpool quick Sequential Organ Failure Assessment; PAWS = Paediatric Advanced Warning Score; PEWS = Pediatric Early Warning Score; PEWS-RL = PEWS-Resource Limited; PICU = paediatric intensive care unit; qPELOD-2 = quick Pediatric Logistic Organ Dysfunction-2; qSOFA = quick Sequential Organ Failure Assessment; SIRS = Systemic Inflammatory Response Syndrome.

Model / Score	**AUC** (95% CI)	
	Primary Analysis	**Sensitivity Analysis**
NEW MODEL	0.84 (0.80 to 0.88)	0.83 (0.78 to 0.88)
FEAST-PET	0.72 (0.66 to 0.78)	0.74 (0.67 to 0.80)
LqSOFA	0.76 (0.71 to 0.81)	0.79 (0.74 to 0.84)
PAWS	0.76 (0.71 to 0.81)	0.77 (0.72 to 0.82)
PEWS	0.71 (0.65 to 0.76)	0.72 (0.66 to 0.79)
PEWS-IRISH	0.74 (0.69 to 0.79)	0.74 (0.69 to 0.80)
PEWS-RL	0.72 (0.67 to 0.77)	0.72 (0.67 to 0.78)
qPELOD-2	0.75 (0.70 to 0.80)	0.76 (0.70 to 0.81)
qSOFA	0.74 (0.69 to 0.79)	0.74 (0.69 to 0.80)
SIRS	0.59 (0.53 to 0.65)	0.62 (0.55 to 0.68)

Sensitivity analyses were also conducted for the two existing scores that contained laboratory parameters (FEAST-PET and SIRS), in which more restrictive criteria (between two hours prior and up to four hours after)[232] were used to determine if these variables were considered available on PICU admission. These analyses also produced similar results (Table 5.3-5).

Table 5.3-5: Sensitivity analyses for primary outcome only including laboratory tests between two hours prior and up to four hours after PICU admission. Comparison between the primary analysis and a sensitivity analysis in which laboratory parameters were restricted to those available between two hours prior and four hours after PICU admission.[232] Results are presented for the two scores which included laboratory parameters: FEAST-PET (haemoglobin as a proxy for pallor) and SIRS (white cell count). In the sensitivity analysis, with the more restrictive criteria for inclusion of laboratory parameters, missingness (addressed by median imputation conditional on outcome status as for the primary analysis) increased to 30.2% for FEAST-PET and 31.0% for SIRS. AUC = area under the receiver operating characteristic curve; CI = confidence interval; FEAST-PET = Fluid Expansion as Supportive Therapy-Pediatric Emergency Triage; PICU = paediatric intensive care unit; SIRS = Systemic Inflammatory Response Syndrome.

Model / Score	**AUC** (95% CI)	
	Primary Analysis	**Sensitivity Analysis**
FEAST-PET	0.72 (0.66 to 0.78)	0.72 (0.66 to 0.78)
SIRS	0.59 (0.53 to 0.65)	0.58 (0.52 to 0.63)

5.4 Discussion

This study reports the ability of nine paediatric severity scores to risk stratify children on admission to an intensive care unit in Cambodia and compares their performance to that of a novel clinical prediction model derived specifically for locations where critical care resources are scarce. Whilst three scores (qPELOD-2, qSOFA, and PAWS) appear to have moderate clinical utility, the new model proved superior and, if validated, could be a practical and flexible tool to support risk stratification of critically ill children in a variety of resource-constrained contexts.

5.4.1 Existing severity scores

With the exception of SIRS, which is known to be non-specific and perform poorly for risk assessment of acutely unwell children,[106,161] the other eight existing scores demonstrated comparable discrimination. However, discrimination is a poor indicator of clinical utility.[64] Only three scores were associated with meaningful changes in the pre-test probability that a PICU admission might result in death, such that a single cut-off could be used to triage children into high- and low-risk groups. Whilst separate cut-offs could conceivably identify high- and low-risk PICU admissions, in settings where resources are scarce it is unclear how a middle or 'indeterminate' group might be managed, and dividing admissions into multiple risk categories may not be practical on the ground.

Discrimination and classification of FEAST-PET, qSOFA, and qPELOD-2 were comparable to their original development studies,[98,115,161] which may reflect similarities in the population (critically ill children), outcome (mortality), and for the FEAST-PET study, contextual factors (access to care etc.). Performance of LqSOFA was inferior to the original development study,[133] which is not unexpected as LqSOFA is known to perform best as a screening tool outside of PICU.[133,194] It is notable that discrimination and classification of PAWS, PEWS, and PEWS-RL were considerably worse in this study.[113,133,159,231] These scores are *diagnostic* scores (aiming to predict events occurring very shortly [< 24 hours] after the time of calculation) and it is therefore perhaps not surprising that their ability to prognosticate more distant outcomes is sub-optimal.

5.4.2 The AHC PICU prognostic model

The new model developed in this study considers the background of a child, their illness journey, and vital organ status to provide a contextualised assessment of critical illness and estimate the probability that the child will not survive to PICU discharge, given the resources available in a typical Level II PICU located outside of a major urban centre in Southeast Asia. Discrimination of the model was considerably better than all existing severity scores evaluated. It provided good prognostic

value and was well calibrated over the threshold probabilities (cut-offs) of interest. In contexts where it might be feasible to resource a particular clinical area to manage up to a quarter of the highest risk PICU admissions, the model, if validated, could provide a readily implementable mechanism to identify children whom might benefit most from being cared for in such an area. Importantly, as the output of such a model is continuous (as opposed to discrete as is the case for points-based scores), the cut-off (threshold probability) for triage to the high-acuity area could be better tailored to account for unit capacity, seasonal bed-pressures, hospital policy, and other dynamic contextual factors.

It is essential that risk stratification tools do not inadvertently concentrate all available resources on patients with untreatable and terminal illnesses.[208] This is particularly important in contexts where resources are at a premium and prolonged hospitalisation can be associated with catastrophic expenditure for patients' families. The PICU is often an environment where more resources are available to manage the end of life:[240] data-driven risk stratification can help reduce pressure on individual doctors and provide a framework for discussions related to dignified withdrawal of care.[50]

In addition to patient triage, accurate severity assessment tools, such as the one developed in this study, offer ancillary benefits. Severity-adjusted mortality rates can help standardise inter-unit comparisons and interpret impact of new interventions such as training programmes, therapeutics, or organisational changes.[58] Risk scores and models can also improve transparency and consistency in the way policies are applied, and increase focus on care pathways to promote equitable and practical delivery of critical care.[241] It is important to note that this risk prediction model is not intended to replace clinical assessment but rather to provide an additional data point to assist busy healthcare professionals plan and organise care for children who are critically ill.

This study provides one of few descriptions of paediatric critical care delivery in regions of LMICs where critical care demand and services are growing. Although preceding illness duration was reportedly short, children often had protracted journeys, consisting of multiple care encounters

involving both the private and public health systems. The short average length of stay on PICU is striking and likely reflects the fact that in many resource-limited settings effective intensive care consists of providing simple, life-saving interventions for critically ill children with readily reversible conditions.[12,208] The ~4% post-PICU discharge mortality rate is in keeping with other studies and is likely an underestimate due to considerable losses to follow-up.[242,243]

5.4.3 Strengths and limitations

Amongst this study's strengths are its relatively unique setting in a PICU outside of a major urban tertiary centre in a country with high under-five mortality. Best-practice methods were followed for external validation of the existing scores and derivation of the new model, with particular care taken to prespecify and limit the number of candidate predictors and use penalised regression to avoid overfitting.[66] Important contextual determinants of outcome absent in tools developed in high-income settings were included in the model and contributed to its promising performance. It should be noted that certain contextual factors may be more or less relevant in other locations, emphasising the need for setting-specific derivation, validation, and/or updating of models, rather than expecting a single model to perform well in multiple locations.[214,220]

The major limitation of this study is the lack of external validation of the new model. Although steps were taken to avoid overfitting, assessment of the model's performance with new data is required before it can be recommended for clinical use and to enable a fairer comparison with the pre-existing severity scores that were assessed in this study. A prospective validation study is underway in which the out-of-sample performance of the model will be compared to the best-performing of the existing severity scores.

The study was conducted in a single centre and hence findings may not be applicable to other resource-limited PICU settings, particularly those where the prevalent causes of critical illness differ. Due to the retrospective nature of this study, travel time had to be estimated based on a child's

location of residence. This may not reflect actual travel time, which will be influenced by their mode of transport, seasonal variation in road conditions, and interruptions to their journey. It was not possible to evaluate the performance of all longlisted scores. In particular, only two of the nine included scores were developed in LMICs and it is disappointing that seven longlisted LMIC-derived scores had to be excluded.[98,231]

Almost 60% of the cohort were male children, which may reflect their predisposition to severe infections or gender biases in care seeking, although the latter is not known to be prevalent in Cambodia.[244,245] Nevertheless, the findings may be biased towards males. The use of routine records means that clinical parameters will not have been measured in a standardised manner. However, the existence of structured admission and vital signs proforma partially mitigate this issue and helped keep missingness low. Use of clinical records did ensure that the new model contains predictors feasible for collection under routine circumstances and will hopefully increase the likelihood of successful out-of-sample validation.

5.4.4 Conclusions

This study presents a new clinical prediction model for estimating the probability that a child admitted to a PICU in a resource-constrained context will not survive to discharge. The model contains predictors from multiple domains to ensure holistic assessment of critical illness. It outperformed nine existing paediatric severity scores and, if validated, offers a readily implementable and flexible mechanism to support risk stratification of critically ill children across a variety of contexts with emerging critical care capacities.

6 Predicting disease progression in moderate Covid-19 in India

This chapter is based upon work published in: **Chandna A**, Mahajan R, Gautam P, et al. *Facilitating safe discharge through predicting disease progression in moderate Covid-19: a prospective cohort study to develop and validate a clinical prediction model in resource-limited settings.* Clin Infect Dis. 75(1), 2022; **Chandna A**, Mahajan R, Gautam P, et al. *Host biomarkers reflect prognosis in patients presenting with moderate Covid-19: a prospective cohort study.* Open Forum Infect Dis. 9(10), 2022; and **Chandna A,** Mahajan R, Gautam P, et al. *Point-of-care prognostication in moderate Covid-19: analytical validation and prognostic accuracy of a soluble urokinase plasminogen activator receptor (suPAR) rapid test.* PLOS Glob Public Health. Aug 21;3(8): e0001538, 2023.

6.1 Global context at the time of the study

By the end of 2021, two years after the first SARS-CoV-2 infection was detected,[246] almost 300 million cases of Covid-19 and 5.5 million deaths had been reported.[247] Unprecedented demand had overwhelmed triage rooms, exhausted oxygen supplies, and led to rationing of intensive care.[248,249] However, rollout of effective vaccines and treatments meant that global attention was moving inexorably towards the 'safe re-opening' of societies and borders.[250-255]

In many low- and middle-income countries (LMICs) however, the pandemic was unfolding rather differently; fewer than 5% of individuals had been vaccinated;[247] dysfunctional global supply chains were compromising scale-up of diagnostics, with less than 0.5% of all Covid-19 tests performed in LMICs;[256] and despite benefiting from enormous public funding, restrictive licensing agreements between pharmaceutical corporations and local manufacturers were limiting development and distribution of generic drugs.[257] Furthermore, inequitable access to key components of supportive care such as ventilators, medical oxygen, and personal protective equipment was hindering pandemic response:[258] in India, prior to the emergence of the SARS-CoV-2 delta variant of concern (VOC;

B.1.617.2), 50,000 ventilators (predominantly in the private sector) had been estimated as available to use for the projected one million Covid-19 patients requiring ventilation.[259]

All this was occurring as a new, more transmissible, VOC (B.1.1.529; omicron) was detected. Impact on disease severity was as yet unknown but hospitalisation rates were rising and the World Health Organization (WHO) issued a recommendation for countries to increase public health and medical capacities.[260] Fragile health systems remained vulnerable to being overwhelmed by a surge in Covid-19 cases and optimal resource allocation was paramount.[261-263]

6.2 Introduction

Only a small proportion of patients with Covid-19 require admission to hospital. Oxygen is the most important supportive treatment and the practical ceiling of care in many LMICs.[264] The WHO estimates that 15% of patients with symptomatic Covid-19 will require supplemental oxygen.[265] Effective identification of patients unlikely to become hypoxaemic and who are suitable for community-based management would have considerable potential to decompress health systems and help healthcare providers allocate resources more efficiently.[266] Efficient and accurate allocation becomes increasingly important as resources are stretched by high case numbers.

Numerous prognostic models for Covid-19 have been developed.[267,268] Almost all predict critical illness or mortality and thus cannot inform whether a patient might be safely managed in the community. Of the few that focus on patients with moderate disease, most rely on retrospective or registry-based data,[269-273] lack external validation,[274,275] and are not feasible for use in resource-limited settings.[268,276] Moreover, most existing studies did not follow best-practice guidelines for model building and reporting,[66] are at high risk of bias,[267] and the resulting models are neither suitable nor recommended for use in LMIC contexts.[268]

Biochemical biomarkers of the host response to infection, including those reflecting endothelial injury, immune activation, inflammation, and coagulation have been shown to be prognostic in a variety of febrile illnesses,[40,135,164] including SARS-CoV-2.[24,169,171,277-282] This has led to them being proposed as risk stratification tools (alone or as adjuncts to clinical risk scores) to help health workers identify patients with a poor prognosis and guide resource allocation.[39,134]

Some of these biomarkers are measurable using commercially-available point-of-care tests. However, studies developing triage tools commonly use laboratory-based platforms to quantify biomarker concentrations, the accuracy of which may differ from point-of-care tests required to inform timely management of individual patients in settings without established laboratory capacity.[283] One such biomarker is the soluble version of the urokinase plasminogen activator receptor (suPAR), which is upregulated during the host response to infection.[38] Measurements of suPAR have been shown to be useful in the diagnosis and prognosis of a wide range of infections and infectious syndromes,[282,284-288] including Covid-19.[278,282] Although suPAR is measurable using a commercially-available rapid diagnostic test (RDT), these studies quantified suPAR using a laboratory-based enzyme-linked immunosorbent assay (ELISA).

The primary objective of this study (PRIORITISE; Prognostication of Oxygen Requirement in Patients with Non-severe SARS-CoV-2 Infection) was to develop and validate a clinical prediction model to rule-out progression to supplemental oxygen requirement in patients presenting with moderate Covid-19. The hypothesis was that combining simple clinical parameters with host biomarkers implicated in the pathogenesis of Covid-19 would improve prognostication. The study was motivated by a recognition that a tool to inform this critical clinical decision could support health systems in under-resourced settings during surges in SARS-CoV-2 infections.[261] Thus, only biomarkers measurable using commercially-available point-of-care tests were considered for the primary analysis.

A nested secondary objective was to evaluate the analytical performance and prognostic accuracy of a commercially-available RDT for suPAR, by comparing the predictive performance of

suPAR concentrations quantified using the RDT and an ELISA. The aim was to investigate whether the suPAR RDT might prove useful for triage of patients presenting with moderate Covid-19 irrespective of its analytical performance when compared with the reference test.

The final objective was to evaluate the prognostic utility of a broader panel of host response biomarkers, previously found to reflect final common pathways to severe febrile illness and sepsis, for predicting a range of disease severities in patients presenting with moderate Covid-19. The hypothesis was that different biomarkers might predict distinct clinical outcomes and be better suited to particular clinical use-cases.[200] The primary aim of this objective was to inform variable selection for future studies investigating biomarker-guided prognosis of acute febrile illnesses and hence a broader range of biomarkers was considered for these analyses.

6.3 Methods

6.3.1 Study population

PRIORITISE was a prospective observational cohort study. Consecutive patients aged ≥ 18 years with clinically-suspected SARS-CoV-2 infection presenting with moderate symptoms to the All India Institute of Medical Sciences (AIIMS) Hospital in Patna, Bihar, India and the Christian Medical College (CMC) Hospital in Vellore, Tamil Nadu, India were screened (daytime hours, Monday to Saturday). AIIMS is a 1,000-bed hospital and the largest medical facility providing primary-to-tertiary healthcare in the state of Bihar. CMC is a 3,000-bed not-for-profit medical college hospital that provided care for ~1,500 patients with Covid-19 each day during the peak of the pandemic.

The case definitions in the WHO Clinical Management Guideline (moderate disease)[265] and WHO Clinical Progression Scale (WHO-CPS; scores 2, 3, or 4)[289] were adapted to define moderate disease as follows: a peripheral oxygen saturation (SpO_2) ≥ 94% and respiratory rate < 30 bpm, in the context of systemic symptoms (breathlessness or fever and chest pain, abdominal pain, diarrhoea, or

severe myalgia). These adaptations were made in view of the fact that the threshold for hospitalisation varies throughout a pandemic and hence would be a sub-optimal surrogate for the severity of a patient's disease,[290] and that a sensitive cut-off for hypoxaemia (i.e., 94% vs. 90%) would be desirable in a tool to inform safe community-based management.[289]

Patients with a previous history of laboratory-confirmed SARS-CoV-2 infection (prior to the current presenting illness) were excluded. Towards the end of recruitment (March 2021 in AIIMS and May 2021 in CMC) vaccines against Covid-19 began to become available in the study areas and a decision was made to exclude vaccinated individuals as the study would not be powered to determine whether the prediction models were valid in this subgroup.

6.3.2 Data collection

Structured electronic case report forms (CRFs; Open Data Kit [ODK], preloaded onto Android tablets; Appendix 9.13) were completed at enrolment, day 7, and day 14 for all participants, and daily during hospitalisation for participants admitted to the study facilities. Anthropometrics and vital signs were measured by the study staff at enrolment and demographics, clinical symptoms, comorbidities, and medication history collected via brief interview with the participant. Venous blood samples were collected at enrolment in ethylenediaminetetraacetic acid (EDTA) tubes. Study staff were familiarised with the study's Standard Operating Procedures for clinical data and specimen collection via remote Site Initiation Visits prior to the start of recruitment. All variables were defined in a data dictionary to ensure consistency of interpretation across both sites.

For participants admitted at the study sites, follow-up was conducted in-person on each day of hospitalisation until day 14. For those not admitted or discharged prior to day 14, follow-up was conducted via telephone on days 7 and 14, with those who reported worsening (day 7) and/or persistent (day 14) symptoms recalled to have their SpO_2 and respiratory rate measured.

6.3.3 Primary and secondary outcomes

The primary endpoint was development of a supplemental oxygen requirement within 14 days of enrolment, defined as any of the following: SpO_2 < 94%; respiratory rate > 30 bpm; SpO_2/fraction of inspired oxygen (FiO_2) < 400;[291,292] or death, aligning closely with a WHO-CPS score of ≥ 5.[289] Patients who received supplemental oxygen outside the study facilities were classified as meeting the primary endpoint if it was not possible to retrieve their case notes, provided the oxygen was prescribed in a licensed medical facility. The site study teams were unaware of which baseline variables had been preselected as candidate predictors when determining outcome status.

The secondary endpoint was an ordinal outcome reflecting the maximum level of respiratory support a participant required in the 14 days following enrolment. Outcome categories were defined as: (i) no supplemental oxygen required (SpO_2 ≥ 94% *and* RR < 30 *and* SpO_2/FiO_2 ≥ 400; WHO-CPS ≤ 4); (ii) supplemental oxygen required (SpO_2 < 94% *or* RR ≥ 30 *or* SpO_2/FiO_2 < 400; WHO-CPS = 5); (iii) non-invasive ventilation (NIV; WHO-CPS = 6); and (iv) mechanical ventilation (MV) and/or death (WHO-CPS ≥ 7).

The outcomes for the nested evaluation of the suPAR RDT included: (a) the comparative predictive performance of the RDT and the ELISA assessed using the area under the receiver operating characteristic curve (AUC) of suPAR concentrations alone or as part of a multivariable clinical prediction model to predict development of a supplemental oxygen requirement (prognostic accuracy); and (b) the agreement between the RDT and the ELISA (analytical performance).

6.3.4 Selection of host biomarkers

Host biomarkers were shortlisted following a review of the literature conducted by searching MEDLINE (1st June 2020) using synonyms of "SARS-CoV-2" AND ["biomarker" OR "prognosis"]. In consultation with FIND, the global alliance for diagnostics (Geneva, Switzerland), biomarkers that had demonstrated promising prognostic utility in Covid-19 were divided into two groups: a primary

translational panel and a secondary exploratory panel (Table 6.3-1). To qualify for inclusion in the primary panel, biomarkers had to be quantifiable using near-patient tests that were either already in clinical use or in late-stage development (Technology Readiness Level ≥ 4; Table 6.3-2).[293]

The primary panel included seven host response biomarkers: C-reactive protein (CRP), D-dimer, interleukin-6 (IL-6), neutrophil-to-lymphocyte ratio (NLR), procalcitonin (PCT), soluble triggering receptor expressed on myeloid cells 1 (sTREM-1), and suPAR.[24,171,278,280-282] The secondary panel comprised an additional six biomarkers, including those implicated in endothelial injury (angiopoietin-2 [Ang-2]), immune activation (interferon-gamma-inducible protein-10 [IP-10; CXCL-10], interleukin-1 receptor antagonist [IL-1ra], interleukin-8 [IL-8], interleukin-10 [IL-10]), and coagulation (platelets).[24,169,277,279,280]

Table 6.3-1: Supportive data for the primary and secondary panels of host biomarkers selected for inclusion in the study. Ang-2 = angiopoietin-2; ARDS = acute respiratory distress syndrome; CXCL-10 = C-X-C motif chemokine-10; CRP = C-reactive protein; ICU = intensive care unit; IL-1ra = interleukin-1 receptor antagonist; IL-6 = interleukin-6; IL-8 = interleukin-8; IL-10 = interleukin-10; IP-10 = interferon-gamma-inducible protein-10; ISRCTN = International Standard Randomised Controlled Trial Number; MHRA = Medicines and Healthcare products Regulatory Agency; NLR = neutrophil-to-lymphocyte ratio; PCT = procalcitonin; RECOVERY = Randomised Evaluation of Covid-19 Therapy; sTREM-1 = soluble triggering receptor expressed on myeloid cells-1; suPAR = soluble urokinase plasminogen activator receptor; UK = United Kingdom; USA = United States of America.

Biomarker	Overview of supportive data in Covid-19
Primary panel	
CRP	Elevated CRP has been found to be an indicator of poor prognosis in patients in China and Germany,[281,294-296] and is being used to enrich study populations in trials of novel therapeutics (for example, RECOVERY, ISRCTN 50189673) and to inform access to newly licensed treatments (for example, Remdesivir, UK MHRA).
D-dimer	Coagulopathy, including multi-system arterial and venous thromboses, is emerging as an important component of Covid-19 pathophysiology. Several studies have established raised D-dimer levels as predictive of poor outcomes in hospitalised patients and measurement of D-dimer is recommended by the International Society of Thrombosis and Haemostasis in the triage of Covid-19 patients.[24,297-300]
IL-6	Several studies, including a recent meta-analysis, have demonstrated that elevated IL-6 levels predict the risk of severe Covid-19, ARDS, and respiratory failure.[296,299,301,302] In addition, IL-6 was identified as the best predictor of supplemental oxygen requirement in a cohort of Swiss patients.[171] IL-6 is a therapeutic target under investigation in many clinical trials (for example. RECOVERY, ISRCTN 50189673).

Biomarker	Overview of supportive data in Covid-19
NLR	Multiple studies have identified elevated neutrophil counts and/or decreased lymphocyte counts as poor prognostic indicators.[281,294,299,303]
PCT	A number of studies, including a recent meta-analysis, have indicated that patients with elevated PCT levels at admission are more likely to have severe Covid-19 and require ICU admission.[24,302,304,305]
sTREM-1	Increased sTREM-1 has been shown to predict respiratory failure in hospitalised patients.[171]
suPAR	suPAR has demonstrated prognostic utility in the UK, Denmark, and Greece. High suPAR levels have been shown to identify patients at high risk of respiratory failure and those safe for discharge from hospital.[278,282,306]
Secondary panel	
Ang-2	Elevated Ang-2 predicted need for ICU admission in hospitalised patients in France.[277]
IL-1ra	IL-1ra is produced early in the course of a Covid-19 illness and elevated measurements in the first week predict clinical outcomes in hospitalised patients in China.[169]
IL-8	High admission IL-8 levels predicted survival amongst hospitalised patients in the USA.[280]
IL-10	IL-10 is produced early in the course of a Covid-19 illness and elevated measurements in the first week predict clinical outcomes in hospitalised patients in China.[169]
IP-10 / CXCL-10	Elevated IP-10 was associated with disease severity and predicted disease progression amongst hospitalised patients in China.[279]
Platelets	Decreased platelet counts were associated with progression from mild to moderate Covid-19 amongst hospitalised patients in China.[24]

Table 6.3-2: Commercially-available or late-stage development quantitative near-patient tests for the biomarkers included in the primary panel. CRP = C-reactive protein; IL-6 = interleukin-6; NLR = neutrophil-to-lymphocyte ratio; PCT = procalcitonin; sTREM-1 = soluble triggering receptor expressed on myeloid cells-1; suPAR = soluble urokinase plasminogen activator receptor.

Biomarker	Device, Manufacturer	Test characteristics		
		Sample type	Sample volume	Turn-around-time
CRP	NycoCard, Abbott	Whole blood, serum or plasma	5uL	3 minutes
D-dimer	RAMP, Response Biomedical	Whole blood	75uL	15 minutes
IL-6	IL-6, Hotgen	Serum	100uL	15 minutes
NLR	WBC DIFF, HemoCue	Whole blood	10uL	5 minutes
PCT	BRAHMS PCT direct, Roche	Whole blood	20uL	20 minutes
sTREM-1	FIND, personal communication	Whole blood	TBC	TBC
suPAR	suPARnostic, ViroGates	Plasma	10uL	20 minutes

6.3.5 Laboratory procedures

Complete blood counts (XP-300-Hematology-Analyzer, Sysmex, Illinois, USA) were measured on site. Within four hours of collection the remaining enrolment blood sample was centrifuged and aliquots of EDTA-plasma stored at ≤ -20°C. Frozen plasma aliquots were transported on dry ice to the study laboratories and thawed at 2-8°C overnight prior to analysis. Remaining plasma was biobanked on site. In patients without a laboratory-confirmed diagnosis of SARS-CoV-2 infection in the three days prior to presentation, combined oral and/or nasopharyngeal swabs were tested via RT-PCR (Cepheid Xpert Xpress SARS-CoV-2, California, USA or Altona RealStar SARS-CoV-2 rRT-PCR, Hamburg, Germany).

Plasma concentrations of biomarkers of endothelial injury, immune activation, inflammation, and coagulation were quantified using the suPARnostic ELISA and Quick Triage test (ViroGates, Lyngby, Denmark) and the Ella microfluidic platform (ProteinSimple, San Jose, California, USA; Appendix 9.14). SARS-CoV-2 IgM and IgG antibodies were measured using the SCoV-2 Detect ELISA (InBios,

Washington, USA). Custom built Simple Plex multi-analyte assays were performed on the Ella platform at the laboratories of Médecins Sans Frontières (MSF), Patna, India. The suPARnostic ELISAs were performed at the laboratories of the Rajendra Memorial Research Institute of Medical Sciences (RMRI), Patna, India. The suPARnostic Quick Triage tests and the SCoV-2 Detect ELISAs were performed at the CMC laboratories in Vellore, India. For biomarkers quantified on the multi-analyte Simple Plex Ella platform (Ang-2, CRP, IP-10, D-dimer, IL-1ra, IL-6, IL-8, IL-10, PCT, and sTREM-1), analytes outside the limits of quantification (LOQ) were assigned a value extrapolated from the standard curve if available, and a value half the lower LOQ if an extrapolated value was not available (Table 6.3-3).

Table 6.3-3: Proportion of host biomarkers quantified on the multi-analyte platform outside the LOQs. Values extrapolated from the standard curve were available for IP-10 and D-dimer. Values for IL-10 were assigned a concentration half the lower LOQ. No other biomarkers had samples outside the LOQs. IL-10 = interleukin-10; IP-10 = interferon-gamma-inducible protein-10; LOQ = limit of quantification.

Biomarker	Number outside the limits of quantification % (n/N)
D-dimer	0.7% (3/423)
IL-10	1.4% (6/423)
IP-10	1.2% (5/423)

The suPARnostic ELISA is a simplified double monoclonal antibody sandwich assay which requires 15µl of venous plasma. Samples were run in duplicate and the mean concentration reported. The suPARnostic Quick Triage test is a RDT based on lateral flow principles. It requires 10µl of venous plasma and has a dynamic range of 2-15 ng/ml. Paired with a centrifuge and automated lateral flow optical reader it has a time-to-result of 20 minutes. Both tests were performed as per the manufacturer's instructions using the same aliquot of thawed plasma,[307,308] and the operators who performed each test were blinded to the results of the other test. Values outside the dynamic range of the RDT were assigned to the LOQs for the assessment of prognostic accuracy and excluded for the evaluation of assay analytical performance.

6.3.6 Candidate predictors

Based on previous field experience of MSF it was decided *a priori* that a model using four predictors would be practical for use in high-patient-throughput resource-limited settings. Following review of the evidence and considering reliability, validity, feasibility (ease of collection, time-to-decision, etc.), and biological plausibility it was prespecified that each model would contain age, sex, SpO_2, and one host response biomarker from the primary panel.[269,276,309] Clinical predictors were measured at enrolment and all biomarkers except NLR were measured retrospectively from samples obtained at enrolment. NLR was measured on site and was not repeated if it had been measured at the site within 24 hours prior to recruitment. All predictors were measured blinded to outcome status.

6.3.7 Missing data

Due to the small fraction of missing data (< 3% for any single predictor), missing observations were replaced with the median value, conditional on outcome status (Table 6.3-4).

Table 6.3-4: Proportion of participants with missing data for each candidate predictor, stratified by cohort. CRP = C-reactive protein; IL-6 = interleukin-6; NLR = neutrophil-to-lymphocyte ratio; PCT = procalcitonin; sTREM-1 = soluble triggering receptor expressed on myeloid cells-1; suPAR = soluble urokinase plasminogen activator receptor.

Predictor	Proportion missing % (n/N)		
	Overall (N = 423)	**Development cohort** (N = 257)	**Validation cohort** (N = 166)
Age	0%	0%	0%
Sex	0%	0%	0%
SpO_2	0%	0%	0%
CRP	1.9% (8/423)	0/257	4.8% (8/166)
D-dimer	0.7% (3/423)	0.4% (1/257)	1.2% (2/166)
IL-6	0.5% (2/423)	0%	1.2% (2/166)
NLR	2.6% (11/423)	3.9% (10/257)	0.6% (1/166)
PCT	0.5% (2/423)	0%	1.2% (2/166)
sTREM-1	0.5% (2/423)	0%	1.2% (2/166)
suPAR	0%	0%	0%

6.3.8 Statistical methods

The relationship between each continuous predictor (age, SpO_2, and all biomarkers in the primary panel) and the primary outcome was examined in the development cohort using locally weighted scatterplot smoothing (LOWESS). Deviations from linearity were observed for CRP, D-dimer, and IL-6. Multivariable fractional polynomials (R package: *mfp*)[310] were used to select transformations ($\log_{10}$ for CRP and D-dimer; $\frac{1}{\sqrt{IL-6/100}}$ for IL-6) which best predicted the outcome.

Penalised (ridge) logistic regression was used to develop the prediction models and shrink regression coefficients to minimise model optimism (R package: *glmnet*).[179] All predictors were prespecified and no predictor selection was performed during model development. The coefficients derived from the development cohort were applied to the external (temporal) validation cohort and the discrimination (AUC) and calibration (calibration intercepts, slopes, and plots) of each model reported.

Given that the aim was to develop a model to rule-out progression to supplemental oxygen requirement, classifications (true positives [TP], false positives [FP], true negatives [TN], and false negatives [FN]) were examined at clinically-relevant admission thresholds (predicted probabilities). Recognising that the relative value of a TP and FP would vary at different stages of the pandemic (for example, reflecting bed pressures and/or capacity for outpatient follow-up),[290] clinical utility of the models was evaluated using decision curve analyses to quantify the net benefit between correctly identified TP or TN and incorrectly identified FP or FN at a range of plausible admission thresholds (R package: *dcurves*).[59,180] Of particular interest were the threshold probabilities at which the prediction models might offer benefit over the 'admit-all' approach and the added value of the models containing each of the candidate biomarkers compared to the clinical model.

Prognostic accuracy of the suPAR RDT and ELISA, both alone and as constituents of a multivariable clinical prediction model, were compared using the DeLong method for comparing AUCs.[181] Analytical performance of the RDT was evaluated using a Bland-Altman plot to estimate the bias and limits of agreement between the RDT and the ELISA (R package: *blandr*).[311,312] Assessment of agreement was limited to samples within the dynamic range of the RDT (2-15 ng/ml). Sensitivity analyses were conducted whereby samples quantified on the ELISA but outside the dynamic range of the RDT were set to the LOQs of the RDT.

Finally, logistic regression (R package: *pROC*)[182] was used to quantify the ability of each biomarker in the primary and secondary panels to discriminate (AUC) participants who developed increasingly severe pulmonary dysfunction, in accordance with the secondary endpoint's ordinal outcome categories.

All analyses were performed in R, versions 4.0.2, 4.0.3, and 4.1.2.[89]

6.3.9 Sample size

Following the recommendations of Riley et al., a conservative R^2 Nagelkerke of 0.15 and shrinkage factor of 0.9 were assumed.[186] Anticipating that ~8% of participants would meet the primary endpoint, it was estimated that 44 outcome events would be required to derive a prediction model comprising four candidate predictors (events per parameter [EPP] = 11) and minimise the risk of overfitting (R package: *pmsampsize*).[187] Allowing for 5% attrition, the initial recruitment target was 600 participants in the model development cohort.

Given the uncertainty around deterioration rates amongst patients with moderate Covid-19 at the time of study inception, an interim review after the first 100 participants were recruited was prespecified in the study protocol. At this review, the proportion of participants meeting the primary endpoint was higher than anticipated (20% vs. 8%). At this higher prevalence, and using R^2 values from 0.20-0.15, between 52-68 outcome events (EPP = 13-17) would be required to develop the prediction models.[186] Recognising that (i) the range of R^2 estimates was conservative, (ii) penalised regression methods would reduce the risk of overfitting, and (iii) the external validation cohort would allow assessment of model optimism, and following the advice of the Study Management Group and External Advisory Panel, a decision was made to use the first 50 outcome events to derive the models. Participants recruited after that point were entered into the external temporal validation cohort.

For the purposes of an analytical validation, the Clinical and Laboratory Standards Institute (CLSI) recommend a minimum sample size of 100 to evaluate agreement between a candidate and reference test.[313] To maximise precision of the results, all available samples (n = 425) were used for the evaluation of the suPAR RDT.

6.3.10 Ethics and reporting

PRIORITISE was a prospectively-registered investigator-initiated study (ClinicalTrials.gov; NCT04441372), with protocol and statistical analysis plan uploaded to the Open Science Framework

platform (DOI: 10.17605/OSF.IO/DXQ43). Ethical approval was given by the AIIMS, Patna Ethics Committee (AIIMS/Pat/IEC 2020/590; CMC Ethics Committee (CMC IRB 13465); Oxford Tropical Research Ethics Committee (OxTREC 47-20); and MSF Ethical Review Board (MSF ERB 2060). The study is reported in accordance with the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) and Standards for Reporting Diagnostic accuracy studies (STARD) guidelines (Appendix 9.15 and Appendix 9.16).[66,314]

6.4 Results

6.4.1 Study population

Between 22nd October 2020 and 3rd July 2021, 2,808 patients with clinically-suspected Covid-19 were screened, of whom 446 were eligible (446/2,808; 15.9%) and 426 were recruited (20/446; 4.5% refusal rate). All participants had laboratory-confirmed SARS-CoV-2 infection (425/426 [99.8%] via RT-PCR and one via rapid antigen test).

Three participants were lost-to-follow-up (3/426; 0.7%) and excluded from all prognostic analyses, leaving 423 participants for the primary analysis (Figure 6.4-1). The first 257 participants comprised the development cohort (22nd October 2020 to 26th April 2021) and the remaining 166 participants comprised the temporal validation cohort (26th April 2021 to 3rd July 2021). Recruitment closed in AIIMS, Patna in March 2021 after enrolment of 125 participants (48% [124/257] of the development cohort and one participant lost-to-follow-up). A total of 301 participants were recruited at CMC, Vellore up until July 2021 (52% [133/257] of the development cohort, 100% [166/166] of the validation cohort, and two participants lost-to-follow-up).

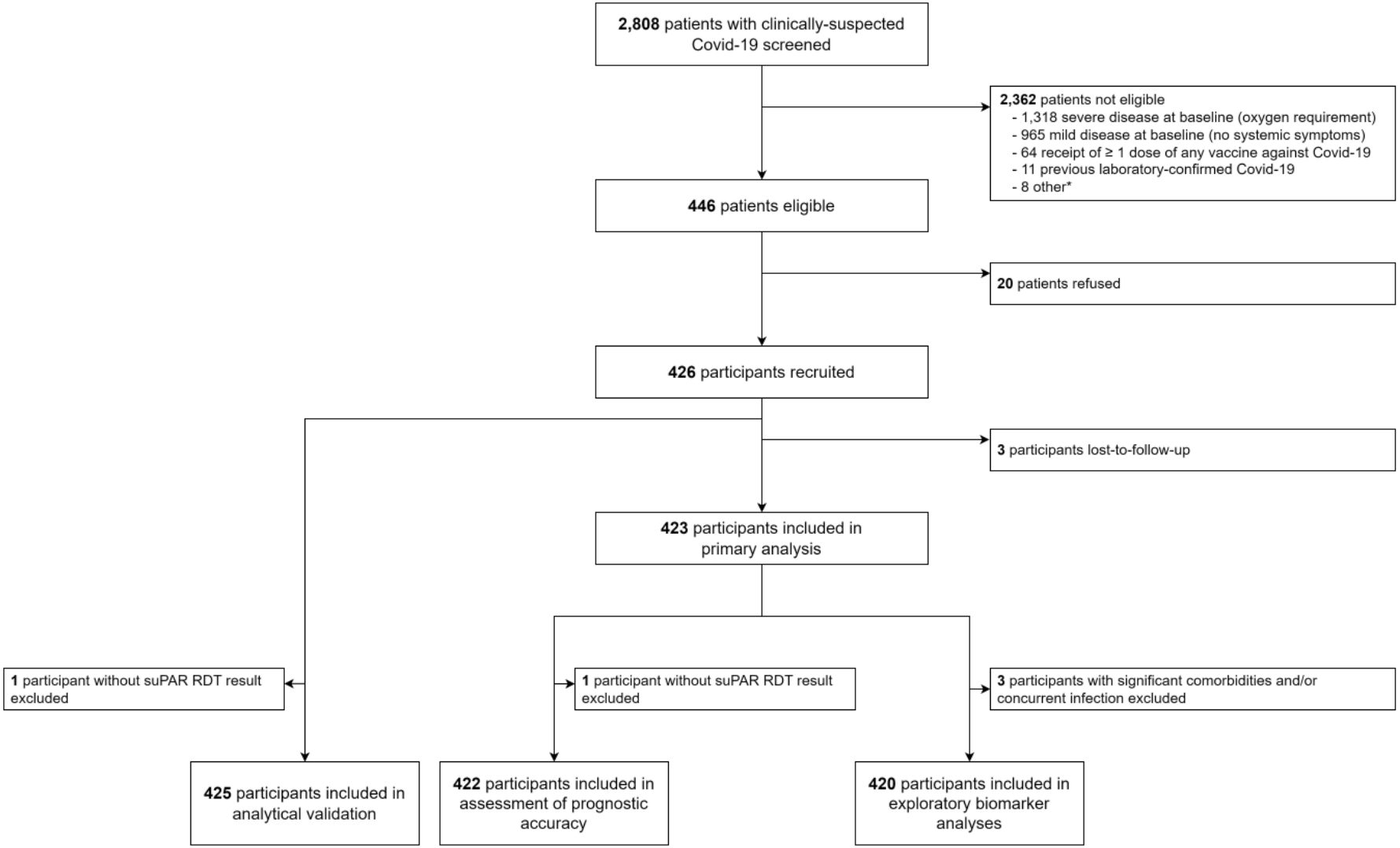


Figure 6.4-1: Screening and enrolment of participants. *Reasons for exclusion: 3 = unable to provide informed consent; 5 = not documented.

Three participants with significant comorbidities or coinfections (two active malignancies with neutropenia and one acute pyelonephritis [3/423; 0.7%]) were excluded from the secondary biomarker analyses. One participant did not have suPAR concentrations quantified using the RDT (1/426; 0.2%) and was excluded from these analyses.

6.4.2 Baseline characteristics

Median age of the cohort was 53.0 years (interquartile range [IQR] = 41.0 to 62.0 years) and 67.6% of the cohort were male (286/423). Most participants presented after six days of symptoms (IQR 4 to 8 days) with an elevated respiratory rate (median 22 bpm; IQR = 22 to 24 bpm) but other vital signs (heart rate, SpO_2, temperature, blood pressure, and mental status) were largely within normal physiological ranges. Two-thirds of participants reported at least one comorbidity (282/423; 66.7%), with diabetes (156/421; 37.1%), hypertension (153/422; 36.3%), and cardiovascular disease (40/418; 9.6%) predominating. Few participants were current tobacco smokers (14/423; 3.3%).

Development and validation cohorts were largely balanced with respect to baseline characteristics (Table 6.4-1). There was a higher proportion of males in the development cohort (72% [185/257] vs. 61% [101/166]; p = 0.017). In the validation cohort, more participants presented with a qSOFA score ≥ 2 (16/166 [9.6%] vs. 13/257 [5.1%]; p = 0.069), and the validation cohort had higher median CRP (58.1 mg/l vs. 24.4 mg/l; p <0.001) and IL-6 (31.6 pg/ml vs. 11.0 pg/ml; p < 0.001) concentrations.

Table 6.4-1: Baseline characteristics, stratified by cohort and primary outcome status. [#]Details for presenting symptoms and comorbidities with prevalence ≥ 5% in any of the outcome groups for either the development or validation cohorts are reported. Data on other comorbidities (chronic neurological disorder, liver disease, active tuberculosis, malignant neoplasm, HIV, and other immunosuppression) and presenting symptoms (otalgia, wheeze, lower chest indrawing, altered consciousness, seizures, conjunctivitis, skin rash, skin ulcers, and lymphadenopathy) not reported; [§]different specimen collection procedures and PCR assays were used at each site; [†]seronegative defined as negative for both SARS-CoV-2 IgM and IgG antibodies; [‡]comparison between participants who did and did not meet the primary outcome within each of the cohorts; [¶]comparison of overall values between cohorts, independent of primary outcome status. [*]Missing data: BMI = 1; CRP = 8; D-dimer = 3; IL-6 = 2; NLR = 12; PCT = 2; sTREM-1 = 2; Ct value = 181; serostatus = 11. BMI = body mass index; BP = blood pressure; bpm = beats / breaths per minute; CRP = C-reactive protein; Ct = cycle threshold; HIV = human immunodeficiency virus; IL-6 = interleukin-6; IQR = interquartile range; NLR = neutrophil-to-lymphocyte ratio; PCT = procalcitonin; qSOFA = quick Sequential Organ Failure Assessment; sTREM-1 = soluble triggering receptor expressed on myeloid cells-1; suPAR = soluble urokinase plasminogen activator receptor.

Baseline characteristic	DEVELOPMENT COHORT				VALIDATION COHORT				
	Overall (n = 257) Median (IQR); n / N (%)	Developed oxygen requirement		p-value[1,‡]	Overall (n = 166) Median (IQR); n / N (%)	Developed oxygen requirement		p-value[1,‡]	p-value[1,¶]
		No (n = 207) Median (IQR); n / N (%)	Yes (n = 50) Median (IQR); n / N (%)			No (n = 127) Median (IQR); n / N (%)	Yes (n = 39) Median (IQR); n / N (%)		
Demographics									
Age (years)	52.0 (40.0 to 61.0)	52.0 (40.0 to 60.0)	54.0 (42.2 to 62.0)	0.30	54.0 (41.2 to 63.0)	55.0 (41.5 to 63.0)	54.0 (41.0 to 66.0)	0.80	0.20
Male sex	185 / 257 (72%)	144 / 207 (70%)	41 / 50 (82%)	0.079	101 / 166 (61%)	76 / 127 (60%)	25 / 39 (64%)	0.60	0.017
BMI (kg/m^2)[*]	26.0 (23.6 to 28.7)	26.2 (23.8 to 28.8)	25.8 (22.4 to 28.3)	0.40	24.9 (23.4 to 27.6)	24.8 (23.4 to 27.6)	26.1 (23.7 to 27.6)	0.70	0.071
Vital signs									
Heart rate (bpm)	88.0 (80.0 to 97.0)	86.0 (79.0 to 96.0)	90.0 (86.0 to 99.5)	0.010	84.0 (74.0 to 92.0)	84.0 (74.0 to 90.0)	84.0 (77.0 to 94.0)	0.40	< 0.001
Respiratory rate (bpm)	22.0 (22.0 to 24.0)	22.0 (22.0 to 24.0)	22.0 (22.0 to 24.0)	0.30	24.0 (22.0 to 24.0)	22.0 (22.0 to 24.0)	24.0 (22.0 to 24.0)	0.12	0.006
Oxygen saturation (%)	98.0 (96.0 to 99.0)	98.0 (97.0 to 99.0)	96.0 (95.2 to 98.0)	< 0.001	98.0 (96.0 to 99.0)	98.0 (96.0 to 99.0)	96.0 (95.5 to 98.0)	0.002	0.50
Axillary temperature (°C)	36.8 (36.4 to 37.1)	36.7 (36.4 to 37.0)	36.9 (36.5 to 37.2)	0.045	36.9 (36.7 to 37.2)	36.9 (36.7 to 37.2)	37.0 (36.9 to 37.2)	0.053	< 0.001

Baseline characteristic	DEVELOPMENT COHORT				VALIDATION COHORT				
	Overall (n = 257) Median (IQR); n / N (%)	Developed oxygen requirement		p-value[1,‡]	Overall (n = 166) Median (IQR); n / N (%)	Developed oxygen requirement		p-value[1,‡]	p-value[1,¶]
		No (n = 207) Median (IQR); n / N (%)	Yes (n = 50) Median (IQR); n / N (%)			No (n = 127) Median (IQR); n / N (%)	Yes (n = 39) Median (IQR); n / N (%)		
Systolic BP (mmHg)	128.0 (116.0 to 138.0)	128.0 (116.0 to 140.0)	126.0 (118.0 to 134.8)	0.70	121.0 (110.0 to 130.0)	120.0 (110.0 to 130.0)	122.0 (110.0 to 131.0)	> 0.9	< 0.001
Diastolic BP (mmHg)	80.0 (72.0 to 88.0)	80.0 (72.0 to 88.0)	79.0 (70.0 to 88.0)	0.30	76.0 (70.0 to 82.0)	76.0 (70.0 to 82.0)	74.0 (67.0 to 80.0)	0.30	< 0.001
qSOFA score ≥ 2	13 / 257 (5.1%)	9 / 207 (4.3%)	4 / 50 (8.0%)	0.30	16 / 166 (9.6%)	10 / 127 (7.9%)	6 / 39 (15%)	0.20	0.069
Comorbidities									
Current smokers	10 / 257 (3.9%)	8 / 207 (3.9%)	2 / 50 (4.0%)	> 0.9	4 / 166 (2.4%)	3 / 127 (2.4%)	1 / 39 (2.6%)	> 0.90	0.40
Reported comorbidity[#]	165 / 257 (64%)	128 / 207 (62%)	37 / 50 (74%)	0.11	117 / 166 (70%)	91 / 127 (72%)	26 / 39 (67%)	0.60	0.20
Cardiovascular disease	21 / 252 (8.3%)	16 / 203 (7.9%)	5 / 49 (10%)	0.60	19 / 166 (11%)	15 / 127 (12%)	4 / 39 (10%)	> 0.90	0.30
Diabetes	92 / 255 (36%)	73 / 205 (36%)	19 / 50 (38%)	0.80	64 / 166 (39%)	55 / 127 (43%)	9 / 39 (23%)	0.023	0.60
Hypertension	89 / 256 (35%)	73 / 206 (35%)	16 / 50 (32%)	0.60	64 / 166 (39%)	47 / 127 (37%)	17 / 39 (44%)	0.50	0.40
Chronic lung disease	3 / 252 (1.2%)	3 / 203 (1.5%)	0 / 49 (0%)	> 0.90	3 / 166 (1.8%)	1 / 127 (0.8%)	2 / 39 (5.1%)	0.14	0.70
Asthma	16 / 252 (6.3%)	13 / 203 (6.4%)	3 / 49 (6.1%)	> 0.90	11 / 166 (6.6%)	9 / 127 (7.1%)	2 / 39 (5.1%)	> 0.90	> 0.90
Chronic kidney disease	10 / 252 (4.0%)	8 / 203 (3.9%)	2 / 49 (4.1%)	> 0.90	9 / 166 (5.4%)	5 / 127 (3.9%)	4 / 39 (10%)	0.20	0.50
Presenting illness									
Symptom duration (days)	6.0 (4.0 to 8.0)	6.0 (4.0 to 8.0)	5.5 (5.0 to 7.0)	0.80	6.0 (4.0 to 8.0)	6.0 (3.5 to 8.0)	5.0 (4.0 to 7.0)	0.50	0.10

Baseline characteristic	DEVELOPMENT COHORT				VALIDATION COHORT				
	Overall (n = 257) Median (IQR); n / N (%)	Developed oxygen requirement		p-value[1,‡]	Overall (n = 166) Median (IQR); n / N (%)	Developed oxygen requirement		p-value[1,‡]	p-value[1,¶]
		No (n = 207) Median (IQR); n / N (%)	Yes (n = 50) Median (IQR); n / N (%)			No (n = 127) Median (IQR); n / N (%)	Yes (n = 39) Median (IQR); n / N (%)		
History of fever	243 / 257 (95%)	196 / 207 (95%)	47 / 50 (94%)	0.70	155 / 166 (93%)	118 / 127 (93%)	37 / 39 (95%)	> 0.90	0.60
Breathlessness	154 / 257 (60%)	119 / 207 (57%)	35 / 50 (70%)	0.11	90 / 166 (54%)	65 / 127 (51%)	25 / 39 (64%)	0.20	0.20
Chest pain	59 / 257 (23%)	48 / 207 (23%)	11 / 50 (22%)	0.90	15 / 166 (9.0%)	9 / 127 (7.1%)	6 / 39 (15%)	0.12	< 0.001
Abdominal pain	35 / 257 (14%)	32 / 207 (15%)	3 / 50 (6.0%)	0.080	15 / 166 (9.0%)	12 / 127 (9.4%)	3 / 39 (7.7%)	> 0.90	0.20
Diarrhoea	80 / 257 (31%)	65 / 207 (31%)	15 / 50 (30%)	0.80	47 / 166 (28%)	33 / 127 (26%)	14 / 39 (36%)	0.20	0.50
Severe myalgia	140 / 257 (54%)	110 / 207 (53%)	30 / 50 (60%)	0.40	75 / 166 (45%)	65 / 127 (51%)	10 / 39 (26%)	0.005	0.062
Sore throat	87 / 257 (34%)	69 / 207 (33%)	18 / 50 (36%)	0.70	31 / 166 (19%)	23 / 127 (18%)	8 / 39 (21%)	0.70	< 0.001
Rhinorrhoea	49 / 257 (19%)	42 / 207 (20%)	7 / 50 (14%)	0.30	19 / 166 (11%)	14 / 127 (11%)	5 / 39 (13%)	0.80	0.037
Cough	205 / 257 (80%)	160 / 207 (77%)	45 / 50 (90%)	0.045	144 / 166 (87%)	109 / 127 (86%)	35 / 39 (90%)	0.50	0.065
Arthralgia	49 / 257 (19%)	37 / 207 (18%)	12 / 50 (24%)	> 0.90	2 / 166 (1.2%)	2 / 127 (1.6%)	0 / 39 (0%)	> 0.90	< 0.001
Fatigue	190 / 257 (74%)	153 / 207 (74%)	37 / 50 (74%)	> 0.90	97 / 166 (58%)	76 / 127 (60%)	21 / 39 (54%)	0.50	< 0.001
Anorexia	69 / 257 (27%)	52 / 207 (25%)	17 / 50 (34%)	0.20	14 / 166 (8.4%)	12 / 127 (9.4%)	2 / 39 (5.1%)	0.50	< 0.001
Nausea or vomiting	55 / 257 (21%)	42 / 207 (20%)	13 / 50 (26%)	0.40	23 / 166 (14%)	17 / 127 (13%)	6 / 39 (15%)	0.80	0.051
Headache	75 / 257 (29%)	65 / 207 (31%)	10 / 50 (20%)	0.11	41 / 166 (25%)	32 / 127 (25%)	9 / 39 (23%)	0.80	0.30

Baseline characteristic	DEVELOPMENT COHORT				VALIDATION COHORT				
	Overall (n = 257) Median (IQR); n / N (%)	Developed oxygen requirement		p-value[1,‡]	Overall (n = 166) Median (IQR); n / N (%)	Developed oxygen requirement		p-value[1,‡]	p-value[1,¶]
		No (n = 207) Median (IQR); n / N (%)	Yes (n = 50) Median (IQR); n / N (%)			No (n = 127) Median (IQR); n / N (%)	Yes (n = 39) Median (IQR); n / N (%)		
Anosmia	60 / 257 (23%)	52 / 207 (25%)	8 / 50 (16%)	0.20	13 / 166 (7.8%)	9 / 127 (7.1%)	4 / 39 (10%)	0.50	< 0.001
Ageusia	60 / 257 (23%)	51 / 207 (25%)	9 / 50 (18%)	0.30	19 / 166 (11%)	14 / 127 (11%)	5 / 39 (13%)	0.80	0.002
Host biomarkers									
CRP (mg/l)[*]	24.4 (3.9 to 88.9)	17.9 (2.8 to 85.4)	62.5 (19.7 to 134.4)	< 0.001	58.1 (17.2 to 147.1)	42.5 (12.3 to 111.9)	95.8 (52.8 to 176.9)	< 0.001	< 0.001
D-dimer (ng/ml)[*]	725.0 (382.4 to 1,466.4)	640.6 (329.7 to 1,234.9)	1,201.7 (679.9 to 2,307.0)	< 0.001	968.2 (620.7 to 1,599.0)	918.8 (579.0 to 1,454.9)	1,148.1 (829.5 to 3,200.2)	0.009	< 0.001
IL-6 (pg/ml)[*]	11.0 (4.9 to 36.2)	8.7 (4.2 to 27.9)	36.4 (18.4 to 70.7)	< 0.001	31.6 (13.9 to 63.0)	24.4 (11.4 to 47.2)	71.1 (39.4 to 98.9)	< 0.001	< 0.001
NLR[*]	3.2 (1.9 to 4.9)	2.9 (1.7 to 4.5)	4.4 (3.2 to 7.2)	< 0.001	2.8 (1.8 to 5.4)	2.5 (1.6 to 4.2)	5.3 (2.7 to 7.0)	< 0.001	0.60
PCT (ng/ml)[*]	0.1 (0.1 to 0.2)	0.1 (0.1 to 0.1)	0.1 (0.1 to 0.2)	< 0.001	0.1 (0.1 to 0.2)	0.1 (0.1 to 0.2)	0.1 (0.1 to 0.3)	0.30	0.004
sTREM-1 (pg/ml)[*]	378.0 (265.0 to 537.0)	362.0 (259.0 to 522.0)	424.5 (306.8 to 649.5)	0.055	419.0 (285.0 to 596.8)	389.0 (282.0 to 562.0)	437.0 (349.0 to 660.8)	0.13	0.20
suPAR (ng/ml)	4.2 (3.1 to 5.8)	4.0 (2.9 to 5.5)	5.4 (4.0 to 6.8)	< 0.001	4.1 (3.1 to 5.6)	3.8 (2.9 to 5.1)	5.5 (3.9 to 6.7)	< 0.001	0.90
Viral markers									
Ct value[§*]	26.0 (20.7 to 30.8)	26.0 (20.6 to 30.1)	26.4 (22.0 to 31.4)	0.60	32.1 (28.3 to 36.2)	32.8 (28.1 to 36.2)	31.5 (28.4 to 36.0)	> 0.9	< 0.001
Seronegative[†*]	117 / 252 (46%)	90 / 203 (44%)	27 / 49 (55%)	0.20	73 / 160 (46%)	51 / 123 (41%)	22 / 37 (59%)	0.054	0.90

[1]Wilcoxon rank sum test; Pearson's Chi-squared test; Fisher's exact test

6.4.3 Clinical outcomes

Eighty-nine participants met the primary outcome (89/423; 21.0%): 50 in the development cohort (50/257; 19.5%) and 39 in the validation cohort (39/166; 23.5%). The median time to oxygen requirement was one day (IQR = 1 to 3 days); 11 participants died; two were mechanically ventilated; 15 received NIV; 49 received oxygen via a face mask and/or nasal cannula (one outside of the study facilities); and 12 had an SpO_2 < 94% but did not receive oxygen supplementation (Table 6.4-2).

Table 6.4-2: Maximum level of supplemental oxygen support received by participants who met the primary outcome, stratified by cohort. One participant (not included in this table) received supplemental oxygen via nasal cannula but did not meet the primary outcome as their SpO_2/FiO_2 remained above 400. *One participant met the endpoint on the basis of being prescribed oxygen at another licensed medical facility; all other participants who met the endpoint did so on the basis of an SpO_2 < 94% and/or SpO_2/FiO_2 < 400 and/or death. FiO_2 = fraction of inspired oxygen.

	Development (n = 257)	**Validation** (n = 166)	**Overall** (n = 423)
Number meeting primary outcome	50	39	89
Deaths	2	9	11
Mechanical ventilation	1	1	2
Non-invasive ventilation	5	10	15
Supplemental oxygen via face mask and/or nasal cannula*	32	17	49
No supplemental oxygen received	10	2	12

6.4.4 Promising performance of models containing IL-6, NLR, or suPAR

Associations between candidate predictors and the probability of developing a supplemental oxygen requirement are presented in Figure 6.4-2. Linear relationships with the primary outcome were observed for SpO_2, NLR, sTREM-1, and suPAR, with deviations from linearity for CRP, D-dimer, and IL-6 being substantial enough to warrant transformations prior to logistic regression modelling.

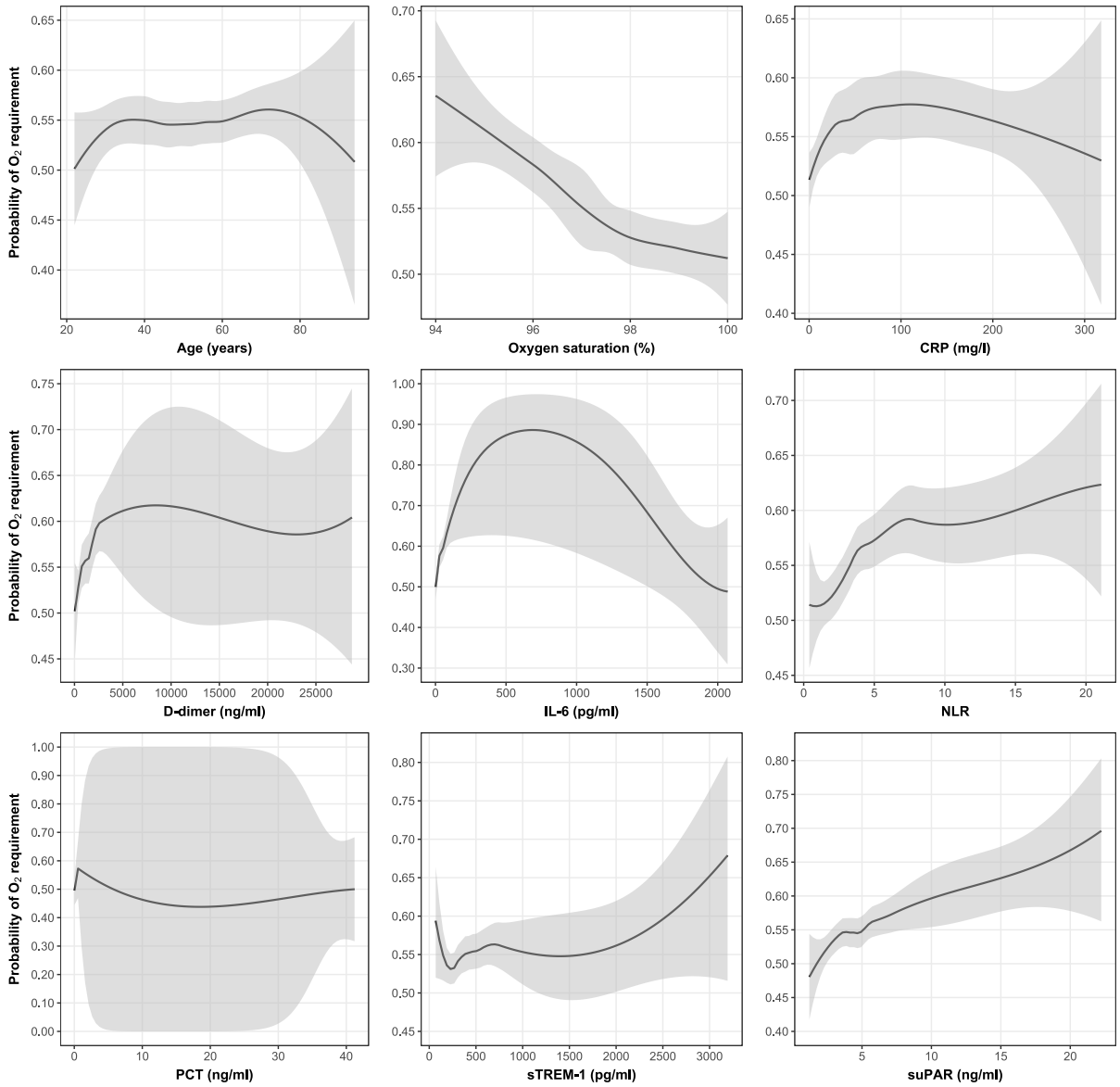


Figure 6.4-2: Exploration of the relationship between continuous candidate predictors and the primary outcome. Black line = probability of oxygen requirement. Grey shaded areas = 95% confidence intervals. CRP = C-reactive protein; IL-6 = interleukin-6; NLR = neutrophil-to-lymphocyte ratio; PCT = procalcitonin; sTREM-1 = soluble triggering receptor expressed on myeloid cells-1; suPAR = soluble urokinase plasminogen activator receptor.

The full models, including intercepts, ridge regression coefficients, and adjusted odds ratios, are presented in Tables 6.4-3 and Equation 6.4-1. After adjustment for the three clinical variables (age, sex, and SpO_2), baseline concentrations of five biomarkers (CRP, D-dimer, IL-6, NLR, and suPAR) were independently associated with progression to a supplemental oxygen requirement in the development cohort.

Table 6.4-3: Adjusted associations between candidate predictors and primary outcome for each model in the development cohort. *Transformations used due to non-linear relationship between predictor and outcome: $\log_{10}$ used for CRP and D-dimer; $\frac{1}{\sqrt{IL-6/100}}$ used for IL-6. As a result, odds ratios are not comparable between continuous predictors, as scales cannot be standardised. Odds ratios are expressed for a one unit increase in non-transformed continuous predictors (age, NLR, PCT, sTREM-1, and suPAR) and a ten-fold increase in CRP or D-dimer. CI = confidence interval; CRP = C-reactive protein; IL-6 = interleukin-6; NLR = neutrophil-to-lymphocyte ratio; PCT = procalcitonin; sTREM-1 = soluble triggering receptor expressed on myeloid cells-1; suPAR = soluble urokinase plasminogen activator receptor.

Variable	Ridge regression coefficient (95% CI)	Adjusted odds ratio (95% CI)
Clinical model		
Intercept	41.89 (23.14 to 67.65)	-
Male sex	0.31 (-0.30 to 1.22)	1.37 (0.74 to 3.37)
Age (years)	0.00 (-0.02 to 0.02)	1.00 (0.98 to 1.02)
SpO_2 (%)	-0.45 (-0.71 to -0.26)	0.64 (0.49 to 0.77)
CRP model		
Intercept	40.47 (21.51 to 66.25)	-
Male sex	0.16 (-0.56 to 1.03)	1.18 (0.57 to 2.79)
Age (years)	-0.00 (-0.03 to 0.02)	1.00 (0.97 to 1.02)
SpO_2 (%)	-0.44 (-0.71 to -0.24)	0.64 (0.49 to 0.78)
CRP*	0.54 (0.23 to 0.98)	1.71 (1.26 to 2.65)
D-dimer model		
Intercept	38.15 (22.19 to 67.18)	-
Male sex	0.34 (-0.29 to 1.26)	1.40 (0.75 to 3.54)
Age (years)	-0.00 (-0.03 to 0.02)	1.00 (0.97 to 1.02)
SpO_2 (%)	-0.44 (-0.75 to -0.28)	0.65 (0.47 to 0.76)
D-dimer*	0.98 (0.47 to 1.88)	2.67 (1.60 to 6.54)
IL-6 model		
Intercept	39.92 (24.84 to 64.91)	-
Male sex	0.11 (-0.64 to 1.02)	1.11 (0.53 to 2.76)
Age (years)	-0.01 (-0.03 to 0.02)	0.99 (0.97 to 1.02)
SpO_2 (%)	-0.41 (-0.68 to -0.25)	0.66 (0.51 to 0.78)
IL-6*	-0.44 (-0.66 to -0.25)	0.65 (0.52 to 0.78)
NLR model		
Intercept	38.27 (19.93 to 66.20)	-
Male sex	0.21 (-0.45 to 0.98)	1.24 (0.63 to 2.65)
Age (years)	-0.00 (-0.03 to 0.02)	1.00 (0.97 to 1.02)

Variable	Ridge regression coefficient (95% CI)	Adjusted odds ratio (95% CI)
SpO_2 (%)	-0.41 (-0.71 to -0.22)	0.66 (0.49 to 0.80)
NLR	0.10 (0.04 to 0.20)	1.10 (1.04 to 1.22)
PCT model		
Intercept	41.03 (1.03 to 66.59)	-
Male sex	0.32 (-0.29 to 1.19)	1.37 (0.75 to 3.28)
Age (years)	0.00 (-0.02 to 0.02)	1.00 (0.98 to 1.02)
SpO_2 (%)	-0.44 (-0.71 to -0.03)	0.64 (0.49 to 0.97)
PCT (ng/ml)	-0.03 (-0.03 to 1.28)	0.97 (0.97 to 3.60)
sTREM-1 model		
Intercept	42.55 (22.26 to 70.09)	-
Male sex	0.28 (-0.35 to 1.11)	1.33 (0.71 to 3.02)
Age (years)	0.00 (-0.02 to 0.02)	1.00 (0.98 to 1.02)
SpO_2 (%)	-0.46 (-0.75 to -0.25)	0.63 (0.48 to 0.78)
sTREM-1 (pg/ml)	0.00 (-0.00 to 0.00)	1.00 (1.00 to 1.00)
suPAR model		
Intercept	38.82 (21.24 to 66.61)	-
Male sex	0.38 (-0.24 to 1.32)	1.46 (0.79 to 3.74)
Age (years)	-0.00 (-0.03 to 0.02)	1.00 (0.98 to 1.02)
SpO_2 (%)	-0.42 (-0.71 to -0.24)	0.65 (0.49 to 0.79)
suPAR (ng/ml)	0.14 (0.06 to 0.25)	1.15 (1.06 to 1.29)

$$\Pr(\boldsymbol{Oxygen\ requirement}) = \frac{\boldsymbol{e^{LP}}}{\boldsymbol{1 + e^{LP}}}$$

where LP is the linear predictor

The LP predictor should be estimated for each model using the following equations:

Clinical model

$$LP(\text{Clinical model}) = 41.89 + \begin{cases} 0 \; if \; female \\ 0.31 \; if \; male \end{cases} + \; 0.003 * age - 0.45 * SpO_2$$

CRP model

$$LP(\text{CRP model}) = 40.47 + \begin{cases} 0 \; if \; female \\ 0.16 \; if \; male \end{cases} - \; 0.002 * age - 0.44 * SpO_2 \; + 0.54 * log_{10}(CRP)$$

D-dimer model

$$LP(D\ dimer) = 38.15 + \begin{cases} 0 \; if \; female \\ 0.34 \; if \; male \end{cases} - \; 0.002 * age - 0.44 * SpO_2 \; + 0.98 * log_{10}(D\ dimer)$$

IL-6 model

$$LP(\text{IL6 model}) = 39.92 + \begin{cases} 0 \; if \; female \\ 0.11 \; if \; male \end{cases} - \; 0.007 * age - 0.41 * SpO_2 \; - 0.44 * \frac{1}{\sqrt{(IL6)/100}}$$

NLR model

$$LP(\text{NLR model}) = 38.27 + \begin{cases} 0 \; if \; female \\ 0.21 \; if \; male \end{cases} - \; 0.002 * age - 0.41 * SpO_2 \; + 0.10 * NLR$$

PCT model

$$LP(\text{PCT model}) = 41.03 + \begin{cases} 0 \; if \; female \\ 0.32 \; if \; male \end{cases} + \; 0.003 * age - 0.44 * SpO_2 \; - 0.03 * PCT$$

sTREM-1 model

$$LP(\text{sTREM1 model}) = 42.55 + \begin{cases} 0 \; if \; female \\ 0.28 \; if \; male \end{cases} + \; 0.0003 * age - 0.46 * SpO_2 \; + 0.0006 * sTREM\ 1$$

suPAR model

$$LP(\text{suPAR model}) = 38.82 + \begin{cases} 0 \; if \; female \\ 0.38 \; if \; male \end{cases} - \; 0.002 * age - 0.42 * SpO_2 \; + 0.14 * suPAR$$

Note:
exp is the exponential function

Equation 6.4-1: Model equations to illustrate how the predicted probability of supplemental oxygen requirement can be calculated for patients presenting with moderate Covid-19. CRP = C-reactive protein; IL-6 = interleukin-6; NLR = neutrophil-to-lymphocyte ratio; PCT = procalcitonin; sTREM-1 = soluble triggering receptor expressed on myeloid cells-1; suPAR = soluble urokinase plasminogen activator receptor.

Discrimination and calibration of each model in the temporal validation cohort are presented in Table 6.4-4. AUCs ranged from 0.66 for the clinical model and the model containing PCT to 0.74 for the model containing IL-6. Calibration slopes ranged from 0.62 for the model containing PCT to 1.01 for the model containing suPAR. Calibration plots indicated that most models were well calibrated up until a predicted probability of supplemental oxygen requirement of 40%, beyond which some models began to overestimate risk (Figure 6.4-3).

Table 6.4-4: Discrimination and calibration of the clinical prediction models in the validation cohort. AUC = area under receiver operating characteristic curve; CI = confidence interval; CRP = C-reactive protein; IL-6 = interleukin-6; NLR = neutrophil-to-lymphocyte ratio; PCT = procalcitonin; sTREM-1 = soluble triggering receptor expressed on myeloid cells-1; suPAR = soluble urokinase plasminogen activator receptor.

Model	AUC (95% CI)	Calibration intercept (95% CI)	Calibration slope (95% CI)
Clinical model	0.66 (0.56 to 0.76)	-0.21 (-1.02 to 0.56)	0.68 (0.27 to 1.27)
CRP model	0.70 (0.61 to 0.79)	-0.11 (-0.74 to 0.61)	0.85 (0.49 to 1.40)
D-dimer model	0.70 (0.61 to 0.79)	-0.04 (-0.68 to 0.64)	0.86 (0.50 to 1.42)
IL-6 model	0.74 (0.66 to 0.82)	-0.10 (-0.64 to 0.49)	1.05 (0.70 to 1.65)
NLR model	0.72 (0.64 to 0.82)	0.23 (-0.44 to 1.04)	1.03 (0.67 to 1.69)
PCT model	0.66 (0.55 to 0.75)	-0.28 (-1.07 to 0.53)	0.62 (0.16 to 1.22)
sTREM-1 model	0.67 (0.58 to 0.77)	-0.25 (-0.91 to 0.56)	0.65 (0.27 to 1.21)
suPAR model	0.72 (0.64 to 0.81)	0.31 (-0.35 to 1.13)	1.01 (0.65 to 1.60)

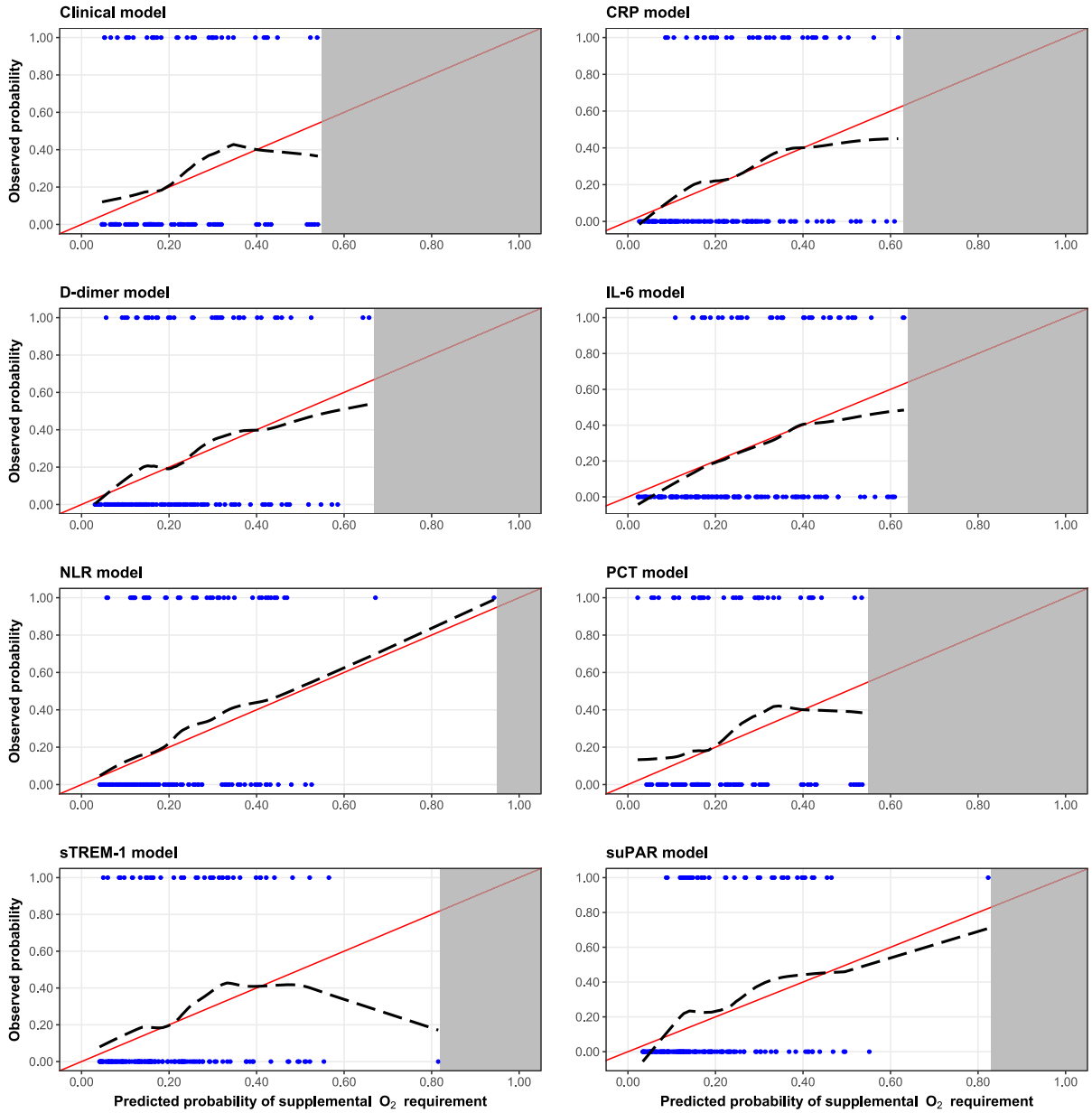


Figure 6.4-3: Calibration plots of the clinical prediction models in the validation cohort. Red line indicates perfect calibration; black dashed line indicates calibration slope for that particular model; blue rug plots indicate distribution of predicted risk for participants who did (top) and did not (bottom) meet the primary outcome. CRP = C-reactive protein; IL-6 = interleukin-6; NLR = neutrophil-to-lymphocyte ratio; PCT = procalcitonin; sTREM-1 = soluble triggering receptor expressed on myeloid cells-1; suPAR = soluble urokinase plasminogen activator receptor.

6.4.5 Models containing IL-6, NLR, or suPAR facilitate safe discharge across varied contexts

The ability of each model to rule-out progression to oxygen requirement amongst patients with moderate Covid-19 at predicted probabilities (cut-offs) of 10%, 15%, and 20% is shown in Table

6.4-5 and Table 6.4-6. A cut-off of 10% reflects a management strategy equivalent to discharging any patient in whom the predicted risk of developing an oxygen requirement is < 10%. At this cut-off, the results suggest that a model containing the three clinical parameters (age, sex, and SpO_2) without any biomarkers could facilitate correctly sending home ~25% of patients with moderate Covid-19 who would not subsequently require supplemental oxygen, at the cost of also sending home ~9% of moderate patients who would deteriorate and require supplemental oxygen, i.e., a ratio of correctly to incorrectly discharged patients of 10:1.

The inclusion of either NLR or suPAR improved the predictive performance such that the ratio of correctly to incorrectly discharged patients increased to 23:1 or 25:1 respectively, whilst a model containing IL-6 resulted in slightly fewer (~21%) correctly discharged patients compared to the clinical model but without missing any patients who would subsequently deteriorate and require supplemental oxygen. Inclusion of any of the other candidate biomarkers (CRP, D-dimer, PCT, or sTREM-1) did not improve the ability of the clinical model to rule-out progression to supplemental oxygen requirement.

The relative value of a TP and FP, i.e., admitted patients who would and would not subsequently require supplemental oxygen, is not fixed and varies at different stages of the pandemic, reflecting bed pressures and/or capacity for community-based follow-up.[290] Decision curve analyses (Figure 6.4-4) accounting for this differential weighting suggest that the clinical model could provide utility (net benefit over an 'admit-all' approach) at a threshold probability above 15% (i.e., when the value of one TP is equal to ~6 FPs). Furthermore, the results indicate that models containing any one of IL-6, NLR, or suPAR could offer greater net benefit than the clinical model and extend the range of contexts in which a model might provide utility to include threshold probabilities above 5% (value of one TP is equal to 19 FPs; i.e., when bed pressures are less critical). For the model containing IL-6, this higher net benefit appeared to be maintained across a range of plausible threshold probabilities.

Table 6.4-5: Predicted classification of patients at different cut-offs for each model, using the prevalence of the primary outcome in the validation cohort. A cut-off (decision threshold or threshold probability) of 10% reflects a management strategy whereby any patient with a predicted risk of requiring oxygen < 10% is discharged. CRP = C-reactive protein; FN = false negative; FP = false positive; IL-6 = interleukin-6; NLR = neutrophil-to-lymphocyte ratio; PCT = procalcitonin; sTREM-1 = soluble triggering receptor expressed on myeloid cells-1; suPAR = soluble urokinase plasminogen activator receptor; TN = true negative; TP = true positive.

Predicted probability of oxygen requirement	Per 100 patients (23 patients who would require oxygen)				Ratio of incorrect to correct admissions (FP : TP)	Ratio of correct to incorrect discharges (TN : FN)
	Patients who would require oxygen admitted (TP)	Unnecessary hospital admissions (FP)	Patients who would require oxygen discharged (FN)	Patients correctly discharged (TN)		
Clinical model						
10%	21	58	2	19	3 to 1	10 to 1
15%	18	46	5	31	3 to 1	6 to 1
20%	14	29	9	48	2 to 1	5 to 1
IL-6 model						
10%	23	61	0	16	3 to 1	NA
15%	21	49	2	28	2 to 1	14 to 1
20%	19	38	4	39	2 to 1	10 to 1
NLR model						
10%	22	54	1	23	2 to 1	23 to 1
15%	17	39	6	38	2 to 1	6 to 1
20%	15	25	8	52	2 to 1	6 to 1
suPAR model						
10%	22	52	1	25	2 to 1	25 to 1
15%	16	34	7	43	2 to 1	6 to 1
20%	13	22	10	55	2 to 1	6 to 1
CRP model						
10%	21	54	2	23	3 to 1	12 to 1
15%	20	43	3	34	2 to 1	11 to 1
20%	16	36	7	41	2 to 1	6 to 1
D-dimer model						
10%	21	54	2	23	3 to 1	12 to 1
15%	19	39	4	38	2 to 1	10 to 1
20%	15	31	8	46	2 to 1	6 to 1
PCT model						
10%	21	57	2	20	3 to 1	10 to 1
15%	18	45	5	32	2 to 1	6 to 1
20%	14	27	9	50	2 to 1	6 to 1
sTREM-1 model						
10%	20	55	3	22	3 to 1	7 to 1
15%	17	41	6	36	2 to 1	6 to 1
20%	14	28	9	49	2 to 1	5 to 1

Table 6.4-6: Sensitivity, specificity, negative likelihood ratio, and positive likelihood ratio at different cut-offs for each model in the validation cohort. A cut-off (decision threshold or threshold probability) of 10% reflects a management strategy whereby any patient with a predicted risk of requiring oxygen < 10% is discharged. CI = confidence interval; CRP = C-reactive protein; IL-6 = interleukin-6; NLR = neutrophil-to-lymphocyte ratio; PCT = procalcitonin; sTREM-1 = soluble triggering receptor expressed on myeloid cells-1; suPAR = soluble urokinase plasminogen activator receptor.

Predicted probability of oxygen requirement	Sensitivity (95% CI)	Specificity (95% CI)	Negative likelihood ratio (95% CI)	Positive likelihood ratio (95% CI)
Clinical model				
10%	89.7 (75.8 to 97.1)	25.2 (17.9 to 33.7)	0.41 (0.15 to 1.08)	1.20 (1.04 to 1.39)
15%	76.9 (60.7 to 88.9)	40.9 (32.3 to 50.2)	0.56 (0.31 to 1.04)	1.30 (1.04 to 1.63)
20%	61.5 (44.6 to 76.6)	62.9 (53.9 to 71.4)	0.61 (0.40 to 0.93)	1.66 (1.19 to 2.33)
IL-6 model				
10%	100 (90.9 to 100)	21.3 (14.5 to 29.4)	0 (NA)	1.27 (1.16 to 1.39)
15%	92.3 (79.1 to 98.4)	36.2 (27.9 to 45.2)	0.21 (0.07 to 0.65)	1.45 (1.23 to 1.70)
20%	82.1 (66.5 to 92.5)	51.2 (42.2 to 60.2)	0.35 (0.18 to 0.70)	1.68 (1.33 to 2.12)
NLR model				
10%	95.0 (82.7 to 99.3)	29.9 (22.1 to 38.7)	0.17 (0.04 to 0.68)	1.35 (1.18 to 1.55)
15%	74.4 (57.9 to 86.9)	49.6 (40.6 to 58.6)	0.52 (0.29 to 0.91)	1.48 (1.15 to 1.90)
20%	66.7 (49.8 to 80.9)	67.7 (58.9 to 75.7)	0.49 (0.31 to 0.78)	2.10 (1.48 to 2.89)
suPAR model				
10%	95.0 (82.7 to 99.4)	33.1 (24.9 to 41.9)	0.16 (0.04 to 0.61)	1.42 (1.23 to 1.63)
15%	69.2 (52.4 to 82.9)	55.9 (46.8 to 64.7)	0.55 (0.34 to 0.90)	1.57 (1.18 to 2.10)
20%	56.4 (39.6 to 72.2)	70.9 (62.2 to 78.6)	0.62 (0.42 to 0.89)	1.94 (1.32 to 2.85)
CRP model				
10%	92.3 (79.1 to 98.4)	29.9 (22.1 to 38.7)	0.26 (0.08 to 0.79)	1.32 (1.14 to 1.52)
15%	87.2 (72.6 to 95.7)	44.1 (35.3 to 53.2)	0.29 (0.13 to 0.67)	1.56 (1.28 to 1.90)
20%	69.2 (52.4 to 82.9)	53.5 (44.5 to 62.4)	0.57 (0.35 to 0.95)	1.49 (1.13 to 1.97)
D-dimer model				
10%	92.3 (79.1 to 98.4)	29.9 (22.1 to 38.7)	0.26 (0.08 to 0.79)	1.32 (1.14 to 1.52)
15%	82.1 (66.5 to 92.5)	48.8 (39.8 to 57.8)	0.37 (0.18 to 0.74)	1.60 (1.28 to 2.01)
20%	64.1 (47.2 to 78.8)	59.8 (50.8 to 68.4)	0.6 (0.39 to 0.93)	1.60 (1.16 to 2.19)
PCT model				
10%	89.7 (75.8 to 97.1)	25.9 (18.6 to 34.5)	0.39 (0.15 to 1.05)	1.21 (1.05 to 1.41)
15%	76.9 (60.7 to 88.9)	41.7 (33.1 to 50.8)	0.55 (0.3 to 1.02)	1.32 (1.05 to 1.66)
20%	58.9 (42.1 to 74.4)	64.6 (55.6 to 72.6)	0.64 (0.43 to 0.95)	1.66 (1.17 to 2.37)
sTREM-1 model				
10%	87.2 (72.6 to 95.7)	29.1 (21.4 to 37.9)	0.44 (0.19 to 1.04)	1.23 (1.04 to 1.45)
15%	74.4 (57.9 to 86.9)	46.5 (37.6 to 55.5)	0.55 (0.31 to 0.97)	1.36 (1.09 to 1.77)
20%	61.5 (44.6 to 76.6)	63.8 (54.8 to 72.1)	0.60 (0.40 to 0.92)	1.7 (1.21 to 2.38)

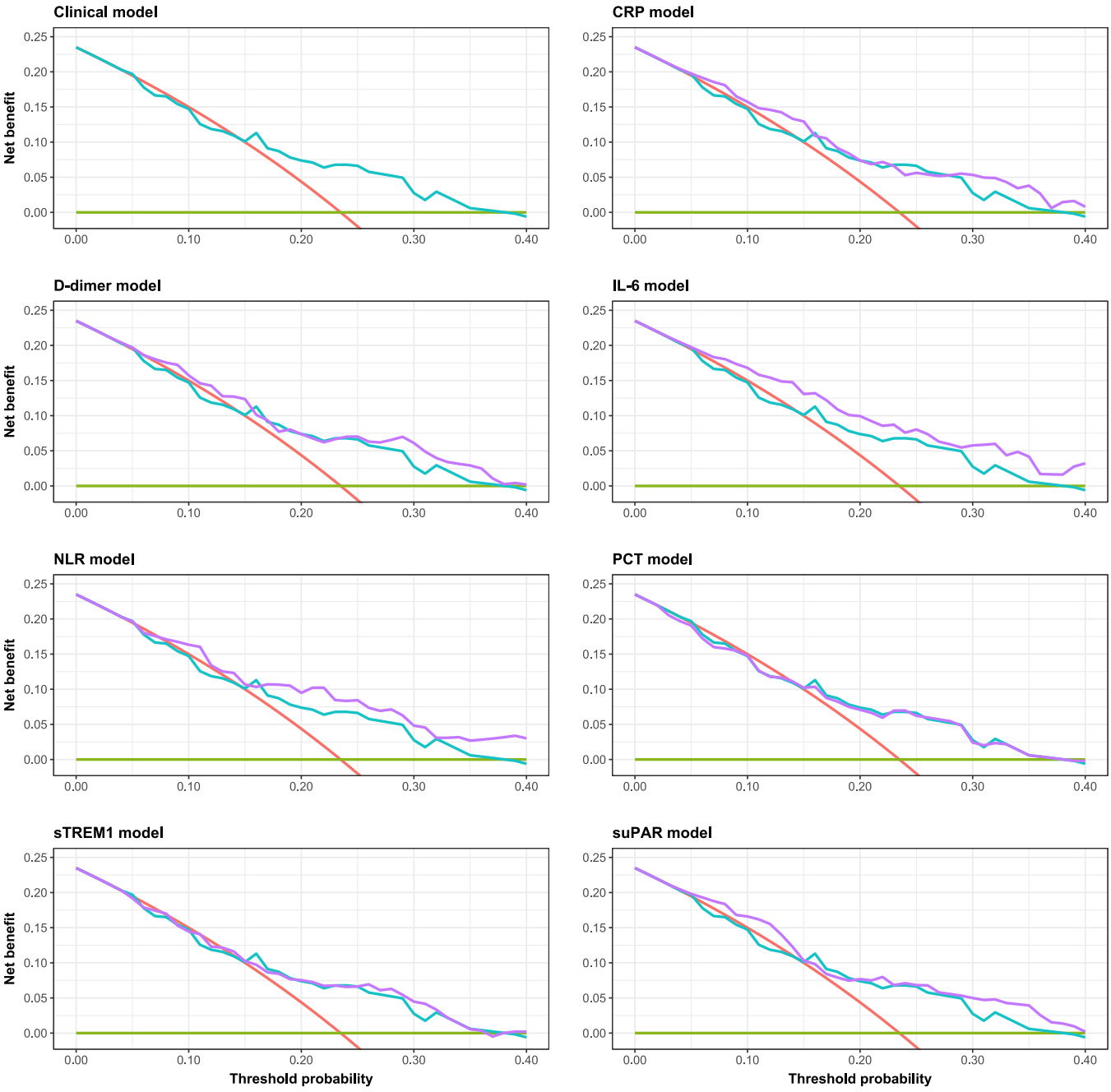


Figure 6.4-4: Decision curve analysis for each model in the validation cohort. The net benefit for each model is compared to an 'admit-all' (red line) and 'admit-none' (green line) approach, and each model containing a host response biomarker (purple line) is also compared to the model containing only clinical variables (blue line). A threshold probability of 5% indicates a scenario where the value of 1 TP (patient admitted who will subsequently require oxygen) is equivalent to 19 FPs (patients admitted who will not subsequently require oxygen). Moving from left to right along the x-axis (increasing threshold for admission) reflects increasing 'penalisation' of FP, indicative of contexts in which inpatient resources may be more constrained. CRP = C-reactive protein; FP = false positive; IL-6 = interleukin-6; NLR = neutrophil-to-lymphocyte ratio; PCT = procalcitonin; sTREM-1 = soluble triggering receptor expressed on myeloid cells-1; suPAR = soluble urokinase plasminogen activator receptor; TP = true positive.

6.4.6 Limited agreement between an RDT and ELISA for suPAR

Forty-four samples returned suPAR concentrations outside the dynamic range of the RDT on either the ELISA (n = 25), RDT (n = 13), or both (n = 6), leaving 381 paired samples for assessment of agreement. Median suPAR concentration was higher when quantified by the RDT (Table 6.4-7; 6.6 vs. 4.2 ng/ml; $p < 0.001$). The two tests were correlated (Pearson's correlation = 0.66; 95% CI = 0.60 to 0.71; $p < 0.001$) but there was limited agreement, with the RDT returning higher values than the ELISA on average. A Bland-Altman plot indicated a bias of -2.46 ng/ml (95% CI = -2.65 to -2.27 ng/ml) with upper and lower limits of agreement of 1.21 ng/ml (95% CI = 0.89 to 1.54 ng/ml) and -6.13 ng/ml (95% CI = -6.45 to -5.81 ng/ml) respectively (Figure 6.4-5).

Table 6.4-7: Baseline suPAR concentrations of the cohort, stratified by progression to supplemental oxygen requirement. *Three participants missing information about supplemental oxygen requirement excluded from table but included in assessment of agreement. ELISA = enzyme-linked immunosorbent assay; IQR = interquartile range; RDT = rapid diagnostic test; suPAR = soluble urokinase plasminogen activator receptor.

suPAR assay	Overall (n = 378)* Median (IQR)	Developed oxygen requirement		p-value[1]
		No (n = 297) Median (IQR)	Yes (n = 81) Median (IQR)	
suPAR ELISA (ng/ml)	4.2 (3.2, 5.6)	4.0 (3.1, 5.3)	5.2 (3.8, 6.4)	< 0.001
suPAR RDT (ng/ml)	6.6 (5.2, 8.5)	6.2 (5.0, 8.3)	8.0 (6.8, 9.4)	< 0.001

[1]Wilcoxon rank sum test

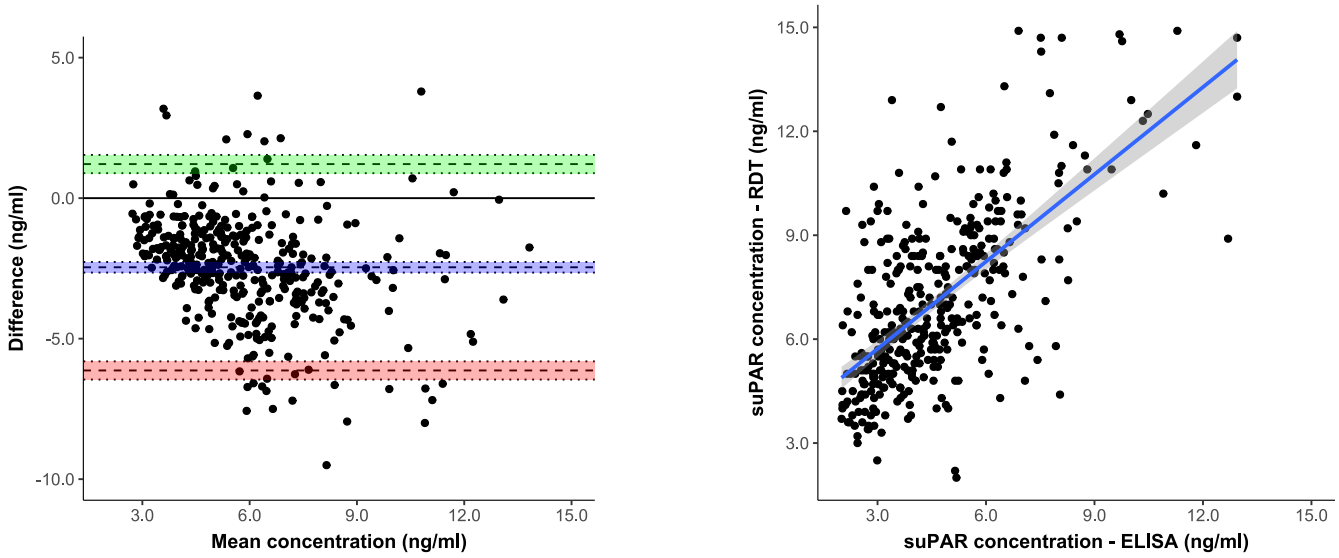


Figure 6.4-5: Agreement and correlation between the suPAR RDT and ELISA. Left panel: Bland-Altman plot indicating agreement between the two tests. Difference between RDT and ELISA measurement in ng/ml plotted on y-axis. Mean of the RDT and ELISA measurement in ng/ml plotted on x-axis. Limits of agreement defined by the concentrations within which 95% of the data fall. Blue line indicates bias, green line indicates upper limit of agreement, red line indicates lower limit of agreement, all with 95% confidence intervals. Bias = -2.46 ng/ml (95% CI = -2.65 to -2.27 ng/ml), upper limit of agreement = 1.21 ng/ml (95% CI = 0.89 to 1.54 ng/ml), lower limit of agreement = -6.13 ng/ml (95% CI = -6.45 to -5.81 ng/ml). Right panel: scatterplot indicating correlation between the two tests (R = 0.66 [95% CI = 0.60 to 0.71]; $p < 0.001$. ELISA = enzyme-linked immunosorbent assay; RDT = rapid diagnostic test; suPAR = soluble urokinase plasminogen activator receptor.

A sensitivity analysis in which samples outside the dynamic range of the RDT were set to the LOQs of the RDT returned similar results (Figure 6.4-6).

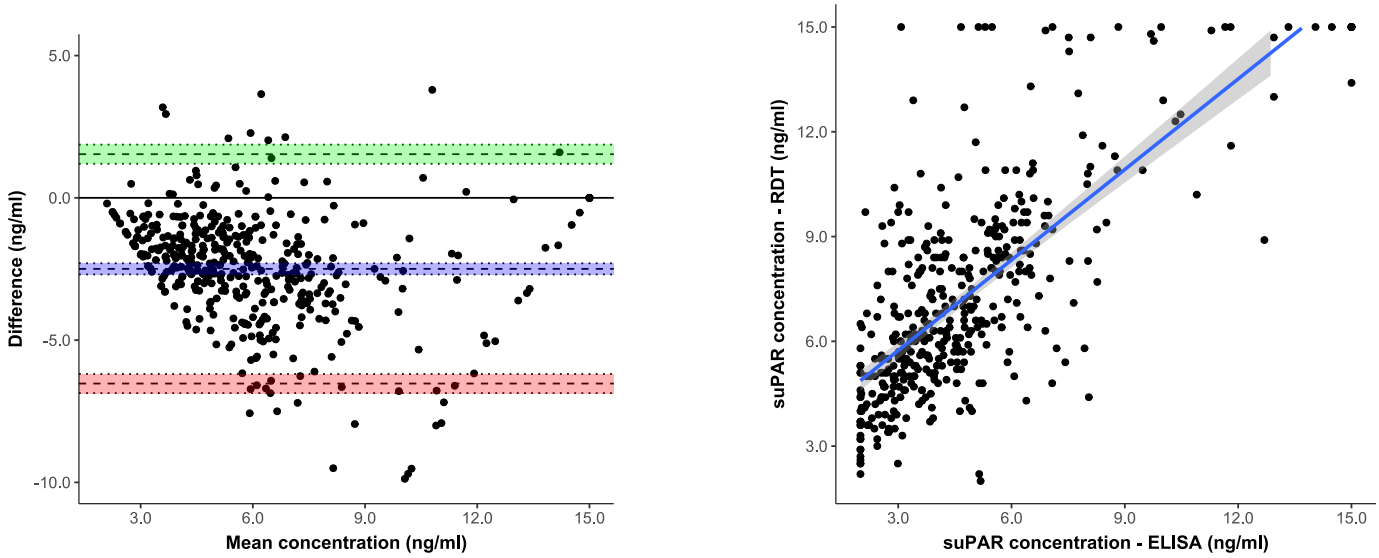


Figure 6.4-6: Agreement and correlation between the suPAR RDT and ELISA when samples outside the dynamic range of the RDT are set to the limits of quantification. Left panel: Bland-Altman plot indicating agreement between the two tests, where samples outside dynamic range of the RDT are set to the limits of detection of the RDT (n = 425). Difference between RDT and ELISA measurement in ng/ml plotted on y-axis. Mean of the RDT and ELISA measurement in ng/ml plotted on x-axis. Limits of agreement defined by the concentrations within which 95% of the data fall. Blue line indicates bias, green line indicates upper limit of agreement, red line indicates lower limit of agreement, all with 95% confidence intervals. Bias = -2.50 ng/ml (95% CI = -2.69 to -2.30 ng/ml), upper limit of agreement = 1.53 ng/ml (95% CI = 1.20 to 1.87 ng/ml), lower limit of agreement = -6.53 ng/ml (95% CI = -6.86 to -6.19 ng/ml). Right panel: scatterplot indicating correlation between the two tests (R = 0.73 [95% CI = 0.69 to 0.77]; p < 0.001). ELISA = enzyme-linked immunosorbent assay; RDT = rapid diagnostic test; suPAR = soluble urokinase plasminogen activator receptor.

Given the disagreement between the reference test and the RDT, the repeatability of the RDT results was investigated by repeating the measurements of all participants at one site using another batch of RDTs (Figure 6.4-7). A Bland-Altman plot indicated good agreement between the repeated RDT measurements: bias = 0.50 ng/ml (95% CI = 0.20 to 0.81 ng/ml), upper limit of agreement = 3.52 ng/ml (95% CI = 3.00 to 4.05 ng/ml), lower limit of agreement = -2.52 ng/ml (95% CI = -3.04 to -1.99 ng/ml).

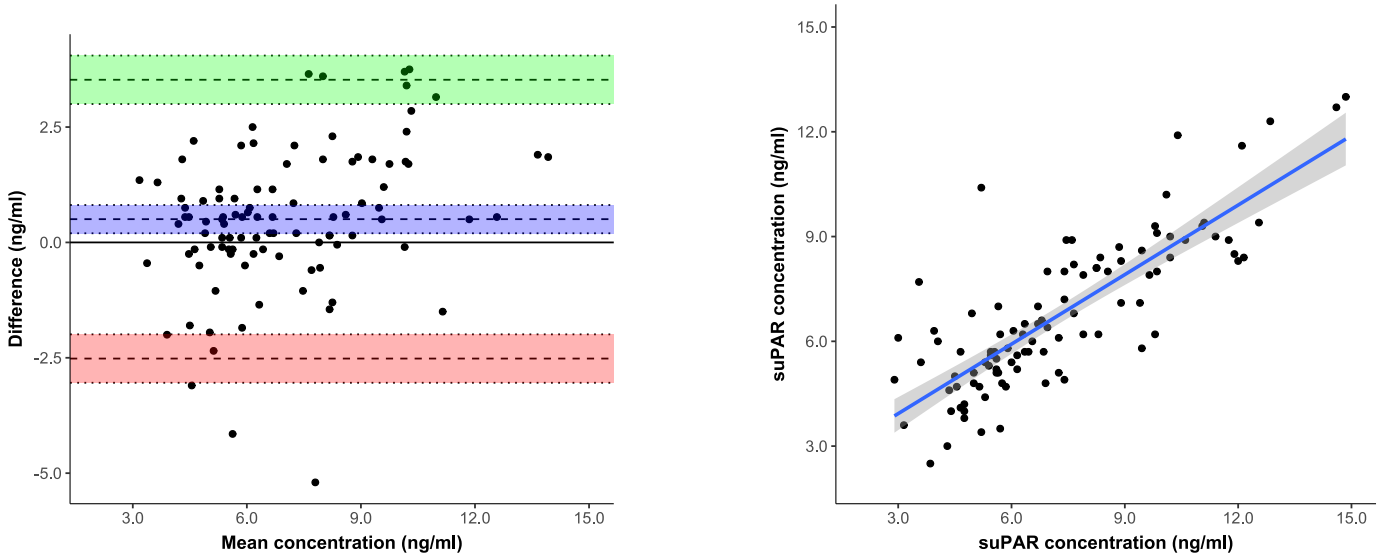


Figure 6.4-7: Agreement and correlation between repeated suPAR RDT measurements. Left panel: Bland-Altman plot indicating agreement between the two RDT measurements (n = 100). Samples outside the dynamic range of the RDT are excluded. Difference between the RDT measurements in ng/ml plotted on y-axis. Mean of the RDT measurements in ng/ml plotted on x-axis. Limits of agreement defined by the concentrations within which 95% of the data fall. Blue line indicates bias, green line indicates upper limit of agreement, red line indicates lower limit of agreement, all with 95% confidence intervals. Bias = 0.50 ng/ml (95% CI = 0.20 to 0.81 ng/ml), upper limit of agreement = 3.52 ng/ml (95% CI = 3.00 to 4.05 ng/ml), lower limit of agreement = -2.52 ng/ml (95% CI = -3.04 to -1.99 ng/ml). Right panel: scatterplot indicating correlation between the two measurements (R = 0.82 [95% CI = 0.74 to 0.87]; p < 0.001). RDT = rapid diagnostic test; suPAR = soluble urokinase plasminogen activator receptor.

6.4.7 Prognostic accuracy is maintained when using a RDT for suPAR

Participants who progressed to develop a supplemental oxygen requirement had higher median baseline suPAR levels, irrespective of the assay used for quantification (Table 6.4-7; RDT = 8.0 vs. 6.2 ng/ml, p < 0.001; ELISA = 5.2 vs. 4.0 ng/ml, p < 0.001). Discrimination of the RDT was comparable to the ELISA (AUC of RDT = 0.73 [95% CI = 0.68 to 0.79] vs. AUC of ELISA = 0.70 [95% CI = 0.63 to 0.76]; p = 0.12) for identifying participants who would progress to require supplemental oxygen (Figure 6.4-8). Consistent with the RDT returning higher readings on average when compared to the ELISA, cut-offs associated with particular sensitivities differed between the two assays (Table 6.4-8).

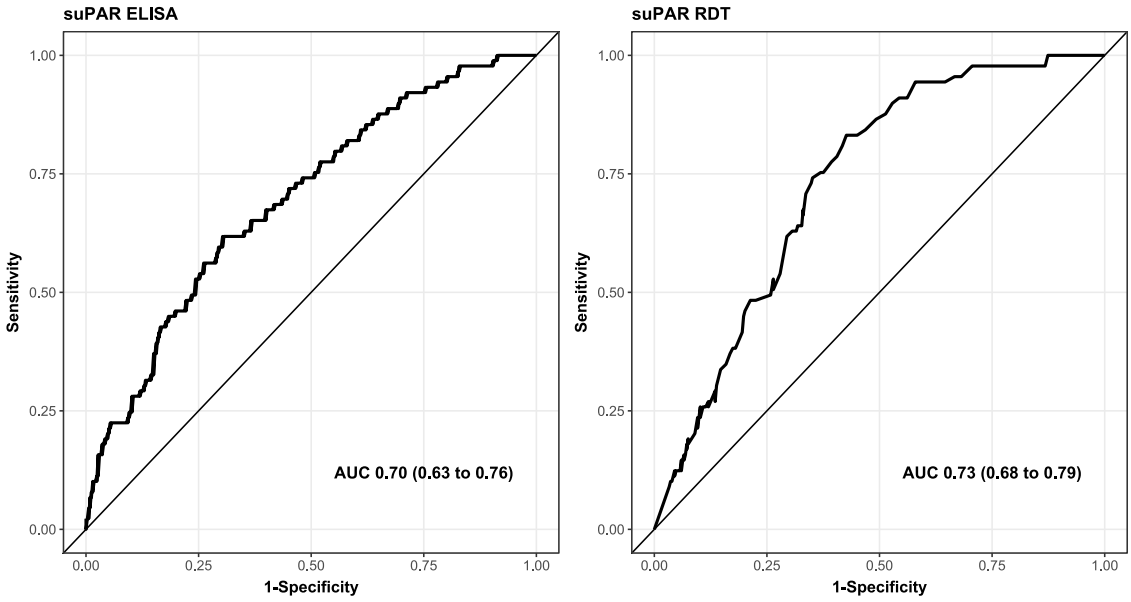


Figure 6.4-8: Ability of suPAR concentrations to discriminate patients who progress to require supplemental oxygen when quantified using the RDT or ELISA. AUC = area under the receiver operating characteristic curve; ELISA = enzyme-linked immunosorbent assay; RDT = rapid diagnostic test; suPAR = soluble urokinase plasminogen activator receptor.

Table 6.4-8: suPAR concentrations associated with different sensitivities for identifying patients who progress to require supplemental oxygen when quantified using the RDT or ELISA. ELISA = enzyme-linked immunosorbent assay; RDT = rapid diagnostic test; suPAR = soluble urokinase plasminogen activator receptor.

Sensitivity (%)	**suPAR concentration** (ng/ml)	
	RDT	**ELISA**
97.5	5.05	2.59
95.0	5.25	2.74
90.0	5.95	3.10
80.0	6.55	3.69

Evaluation of the prognostic accuracy of the RDT within the framework of the previously developed multivariable clinical prediction model confirmed comparable predictive performance of the model whether suPAR was quantified using the RDT or the ELISA. Participants who progressed to develop a supplemental oxygen requirement had higher suPAR levels, when quantified using the RDT or ELISA, in both the development and validation cohorts (Table 6.4-9).

Table 6.4-9: Baseline suPAR concentrations, stratified by cohort and primary outcome status. *Comparison between participants who did and did not meet the primary outcome within each of the cohorts; #comparison of overall values between cohorts, independent of primary outcome status. ELISA = enzyme-linked immunosorbent assay; IQR = interquartile range; RDT = rapid diagnostic test; suPAR = soluble urokinase plasminogen activator receptor.

suPAR assay	DEVELOPMENT COHORT				VALIDATION COHORT				
	Overall (n = 256) Median (IQR)	Developed oxygen requirement		p-value[1,*]	Overall (n = 166) Median (IQR)	Developed oxygen requirement		p-value[1,*]	p-value[1,#]
		No (n = 206) Median (IQR)	Yes (n = 50) Median (IQR)			No (n = 127) Median (IQR)	Yes (n = 39) Median (IQR)		
suPAR ELISA (ng/ml)	4.2 (3.1, 5.7)	4.0 (2.9, 5.5)	5.4 (4.0, 6.8)	< 0.001	4.1 (3.1, 5.6)	3.8 (2.9, 5.1)	5.5 (3.9, 6.7)	< 0.001	0.90
suPAR RDT (ng/ml)	6.4 (5.1, 8.8)	6.0 (4.8, 8.4)	8.2 (7.1, 9.4)	< 0.001	6.7 (5.2, 8.7)	6.2 (5.0, 8.2)	8.7 (6.9, 10.9)	< 0.001	0.70

[1]Wilcoxon rank sum test

suPAR concentrations measured on the RDT were transformed ($\frac{1}{(suPAR\ RDT/10)^2}$) prior to model building due to a non-linear relationship with supplemental oxygen requirement. After adjustment for the clinical variables (age, sex, and SpO_2), baseline suPAR concentrations were independently associated with progression to a supplemental oxygen requirement in the development cohort, irrespective of whether they were measured on the RDT or ELISA (Table 6.4-10). The weighting (regression coefficient) for suPAR within the multivariable prediction model varied depending on which assay was used to quantify suPAR concentrations.

Table 6.4-10: Regression coefficients and adjusted odds ratios for candidate predictors in the development cohort. *Transformation $\frac{1}{(\mathbf{suPAR\ RDT/10})^2}$ used due to non-linear relationship between suPAR concentrations and supplemental oxygen requirement. As a result, odds ratios and regression coefficients are not directly comparable. Odds ratios expressed for one unit increase in non-transformed continuous predictors (age, SpO2, and suPAR ELISA). CI = confidence interval; ELISA = enzyme-linked immunosorbent assay; RDT = rapid diagnostic test; suPAR = soluble urokinase plasminogen activator receptor.

Variable	Ridge regression coefficient (95% CI)	Adjusted odds ratio (95% CI)
Clinical model		
Intercept	42.60 (22.40 to 65.90)	-
Male sex	0.31 (-0.29 to 1.14)	1.37 (0.75 to 3.12)
Age (years)	0.02 (-0.02 to 0.02)	1.00 (0.98 to 1.02)
SpO_2 (%)	-0.46 (-0.69 to -0.24)	0.63 (0.50 to 0.79)
suPAR ELISA model		
Intercept	37.37 (19.63 to 62.53)	-
Male sex	0.37 (-0.21 to 1.29)	1.44 (0.81 to 3.63)
Age (years)	-0.00 (-0.02 to 0.02)	1.00 (0.98 to 1.02)
SpO_2 (%)	-0.41 (-0.67 to -0.23)	0.66 (0.51 to 0.80)
suPAR ELISA (ng/ml)	0.14 (0.07 to 0.25)	1.15 (1.07 to 1.29)
suPAR RDT model		
Intercept	40.52 (21.57 to 60.07)	-
Male sex	0.22 (-0.44 to 1.10)	1.24 (0.64 to 3.00)
Age (years)	-0.01 (-0.03 to 0.01)	0.99 (0.97 to 1.01)
SpO_2 (%)	-0.42 (-0.62 to -0.23)	0.66 (0.54 to 0.80)
suPAR RDT*	-0.37 (-0.53 to -0.22)	0.69 (0.59 to 0.80)

Discrimination and calibration of the models containing suPAR were superior to the clinical model and there was no appreciable difference between the ELISA-based and RDT-based suPAR models (Table; 6.4-11; Figure 6.4-9). The ability of the models to rule-out progression to supplemental oxygen requirement at predicted probabilities (discharge thresholds) of 10%, 15%, and 20% is shown (Table 6.4-12). The results suggest that both suPAR-containing models were comparable, with the RDT-based model appearing to have slightly greater utility for ruling out progression to supplemental oxygen requirement, achieving a ratio of correct to incorrect discharges of 30:1 vs. 25:1, whilst maintaining the same ratio (2:1) of incorrect to correct admissions.

Table 6.4-11: Discrimination and calibration of the clinical prediction models in the validation cohort, with and without suPAR quantified using the RDT or the ELISA. AUC = area under receiver operating characteristic curve; CI = confidence interval; ELISA = enzyme-linked immunosorbent assay; RDT = rapid diagnostic test; suPAR = soluble urokinase plasminogen activator receptor.

Model	AUC (95% CI)	Calibration intercept (95% CI)	Calibration slope (95% CI)
Clinical model	0.66 (0.56 to 0.76)	-0.23 (-0.98 to 0.53)	0.67 (0.23 to 1.24)
suPAR ELISA model	0.72 (0.64 to 0.81)	0.38 (-0.32 to 1.21)	1.07 (0.66 to 1.64)
suPAR RDT model	0.74 (0.66 to 0.83)	0.22 (-0.40 to 0.86)	1.00 (0.66 to 1.49)

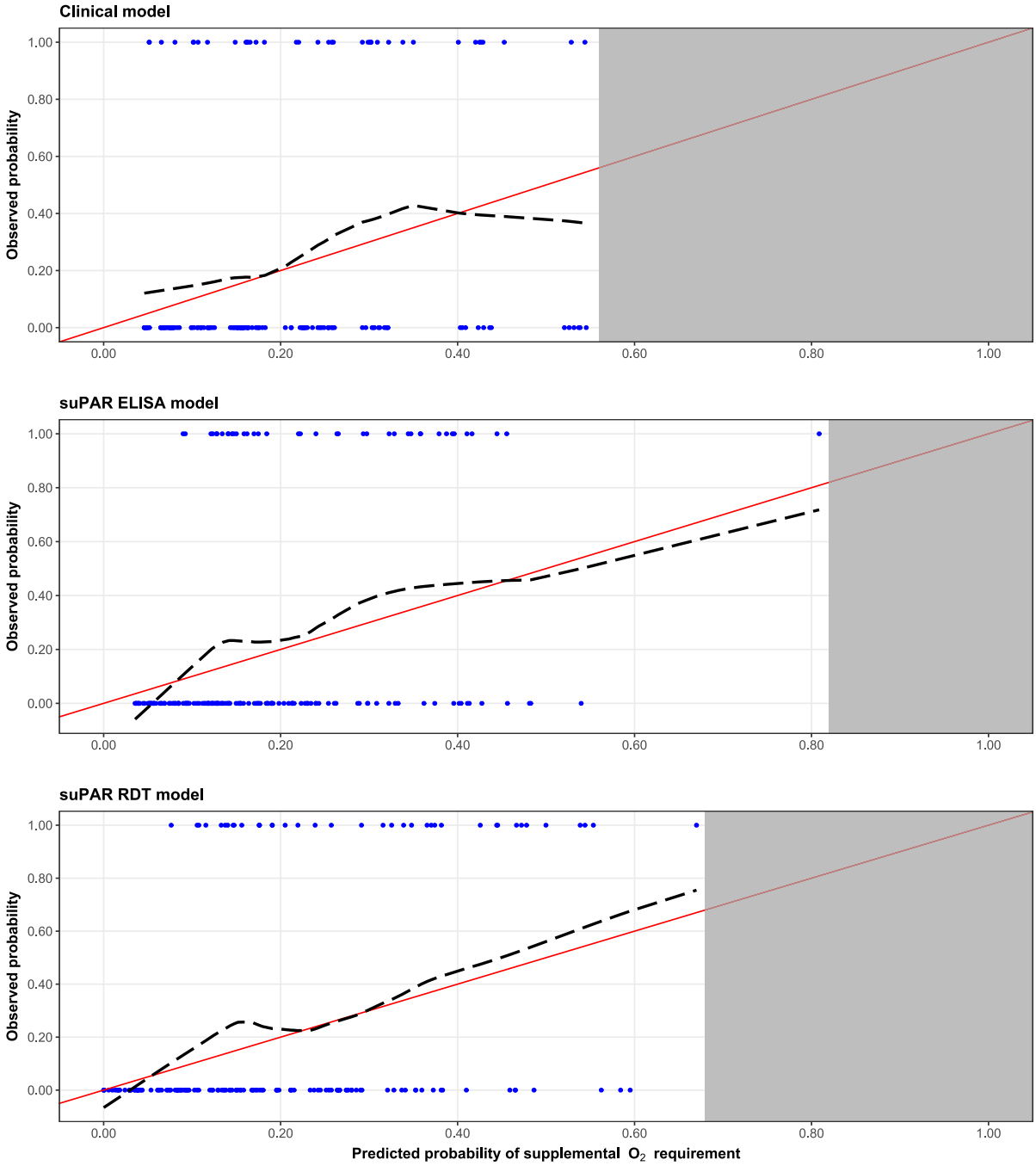


Figure 6.4-9: Calibration of the clinical prediction models in the validation cohort, with and without suPAR quantified using the RDT or the ELISA. Red line indicates perfect calibration; black dashed line indicates calibration slope for that particular model; blue rug plots indicate distribution of predicted risk for participants who did (top) and did not (bottom) meet the primary outcome. ELISA = enzyme-linked immunosorbent assay; RDT = rapid diagnostic test; suPAR = soluble urokinase plasminogen activator receptor.

Table 6.4-12: Predicted classifications of patients at different cut-offs for each clinical prediction model in the validation cohort. A cut-off (decision threshold or threshold probability) of 10% reflects a management strategy whereby any patient with a predicted risk of requiring oxygen < 10% is discharged. ELISA = enzyme-linked immunosorbent assay; FN = false negative; FP = false positive; RDT = rapid diagnostic test; suPAR = soluble urokinase plasminogen activator receptor; TN = true negative; TP = true positive.

Predicted probability of oxygen requirement	Per 100 patients (23 patients who would require oxygen)				Ratio of incorrect to correct admissions (FP : TP)	Ratio of correct to incorrect discharges (TN : FN)
	Patients who would require oxygen admitted (TP)	Unnecessary hospital admissions (FP)	Patients who would require oxygen discharged (FN)	Patients correctly discharged (TN)		
Clinical model						
10%	21	56	2	20	3 to 1	10 to 1
15%	18	45	5	31	3 to 1	6 to 1
20%	14	28	9	48	2 to 1	5 to 1
suPAR ELISA model						
10%	22	52	1	25	2 to 1	25 to 1
15%	16	34	7	43	2 to 1	6 to 1
20%	13	22	10	54	2 to 1	5 to 1
suPAR RDT model						
10%	23	46	1	30	2 to 1	30 to 1
15%	18	37	5	40	2 to 1	8 to 1
20%	14	27	9	50	2 to 1	6 to 1

6.4.8 Prognostic value of host biomarkers for different severities of Covid-19

For the secondary biomarker analyses, development and validation cohorts were combined and clinical outcomes were assessed on an ordinal scale. Most participants did not progress to require supplemental oxygen (Category 1 = 331/420; 78.8%). Of the 89 participants whose clinical condition deteriorated, 62 required supplemental oxygen (Category 2 = 62/420; 14.8%), another 15 received NIV (Category 3 = 15/420; 3.6%), and a further two were mechanically ventilated and 10 died (Category 4 = 12/420; 2.9%).

For all biomarkers, baseline concentrations varied across participants who progressed to different severities of Covid-19 (Figure 6.4-10; Table 6.4-13). Apart from for NLR, more deranged baseline biomarker values were observed in participants who progressed to more severe disease. Some biomarkers (for example, Ang-2, PCT, and sTREM-1) demonstrated baseline concentrations

which were only notably elevated in participants who developed the most severe clinical phenotypes (Category 4). For other biomarkers (for example, IP-10, suPAR, and CRP), increases in baseline concentrations were most substantial between participants in outcome categories 1 and 2, while further increases in participants who progressed to categories 3 (NIV) or 4 (MV and/or death) were less apparent.

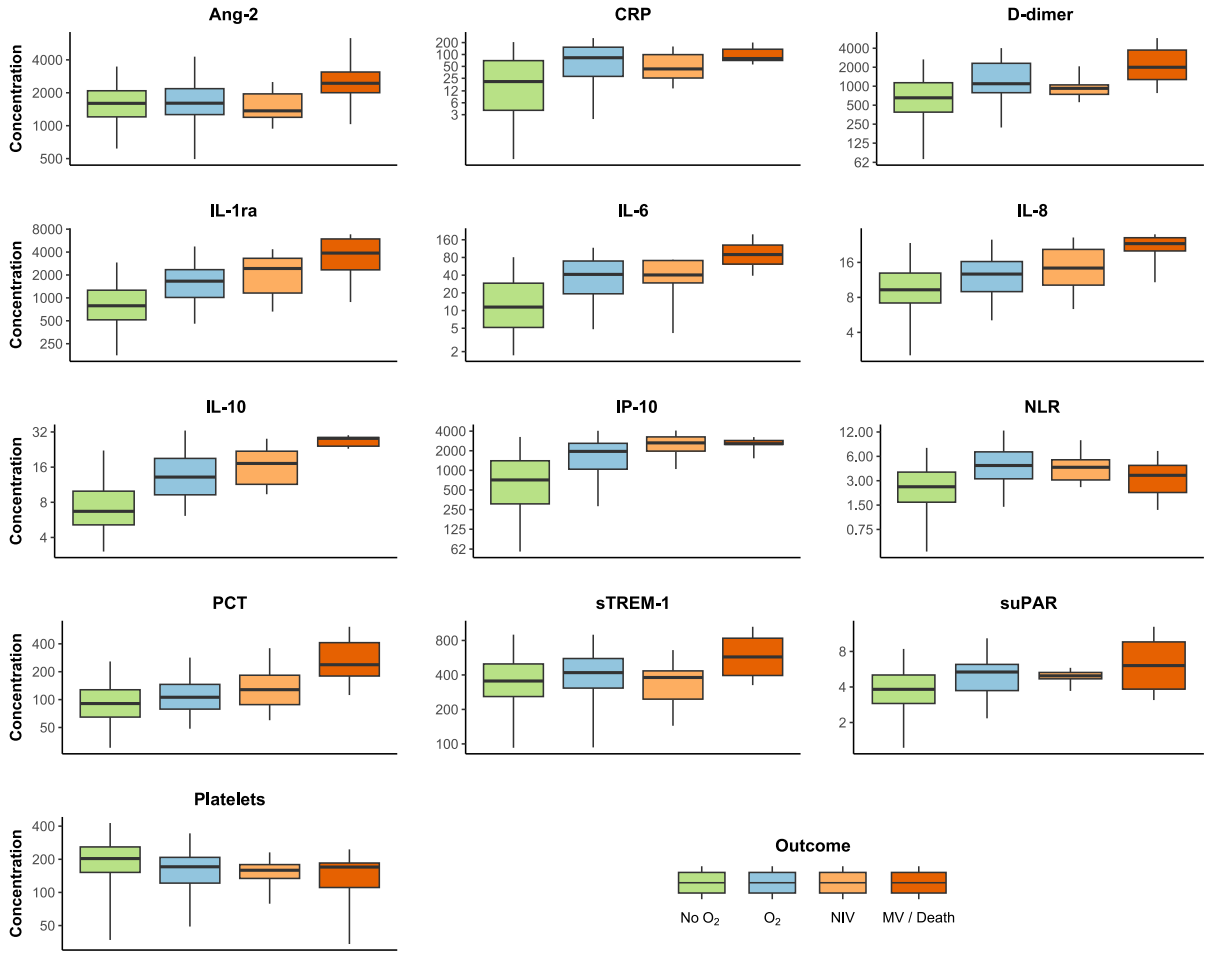


Figure 6.4-10: Baseline concentrations of host biomarkers in participants with moderate Covid-19, stratified by ordinal outcome status. Concentrations plotted on log_2 scale and expressed in pg/ml (Ang-2, IL-1ra, IL-6, IL-8, IL-10, IP-10, PCT, sTREM-1), ng/ml (D-dimer, suPAR), mg/l (CRP), or $x10^9$ cells/l (platelets). Ang-2 = angiopoietin-2; CRP = C-reactive protein; IL-1ra = interleukin-1 receptor antagonist; IL-6 = interleukin-6; IL-8 = interleukin-8; IL-10 = interleukin-10; IP-10 = interferon-gamma-inducible protein-10; MV = mechanical ventilation; NIV = non-invasive ventilation; NLR = neutrophil-to-lymphocyte ratio; PCT = procalcitonin; sTREM-1 = soluble triggering receptor expressed on myeloid cells-1; suPAR = soluble urokinase plasminogen activator receptor.

Table 6.4-13: Baseline characteristics of participants with moderate Covid-19, stratified by ordinal outcome status. †Seronegative defined as negative for both IgM and IgG antibodies against SARS-CoV-2. *Missing data: BMI = 1 (Category 1); platelet count, white cell count, neutrophil count, lymphocyte count, NLR = 10 (Category 1 = 7; Category 2 = 3); Ang-2, IL-1ra, IL-6, IL-8, IL-10, IP-10, PCT, sTREM-1 = 2 (Category 1 = 1; Category 4 = 1); D-dimer = 3 (Category 1 = 2; Category 4 = 1); CRP = 8 (Category 1); serostatus = 11 (Category 1 = 8; Category 2 = 3). Ang-2 = angiopoietin-2; BMI = body mass index; BP = blood pressure; bpm = beats/breaths per minute; CRP = C-reactive protein; IL-1ra = interleukin-1 receptor antagonist; IL-6 = interleukin-6; IL-8 = interleukin-8; IL-10 = interleukin-10; IP-10 = interferon-gamma-inducible protein-10; IQR = interquartile range; NLR = neutrophil-to-lymphocyte ratio; PCT = procalcitonin; qSOFA = quick Sequential Organ Failure Assessment; sTREM-1 = soluble triggering receptor expressed on myeloid cells-1; suPAR = soluble urokinase plasminogen activator receptor.

Baseline characteristic	Overall[1] N = 420 Median (IQR); n/N (%)	Outcome category				p-value[1]
		1 No Supplemental O_2 Requirement N = 331 Median (IQR); n/N (%)	2 Supplemental O_2 Requirement N = 62 Median (IQR); n/N (%)	3 Non-invasive ventilation N = 15 Median (IQR); n/N (%)	4 Mechanical ventilation and/or Death N = 12 Median (IQR); n/N (%)	
Demographics and background						
Age (years)	53.0 (41.0, 62.0)	53.0 (40.5, 61.5)	53.5 (42.0, 64.8)	51.0 (45.0, 65.5)	59.5 (38.8, 64.0)	0.70
Male sex	285 / 420 (68%)	218 / 331 (66%)	48 / 62 (77%)	13 / 15 (87%)	6 / 12 (50%)	0.058
BMI (kg/m^2)*	25.4 (23.5, 28.3)	25.4 (23.6, 28.6)	25.3 (22.3, 28.0)	26.8 (24.0, 28.2)	26.5 (24.4, 27.3)	0.80
Comorbidity	279 / 420 (66%)	216 / 331 (65%)	43 / 62 (69%)	10 / 15 (67%)	10 / 12 (83%)	0.60
Clinical features						
Duration of illness (days)	6.0 (4.0, 8.0)	6.0 (4.0, 8.0)	6.0 (4.2, 7.0)	5.0 (4.5, 7.5)	5.0 (4.0, 5.0)	0.30
Heart rate (bpm)	86.0 (78.0, 95.0)	86.0 (78.0, 94.0)	90.0 (84.0, 97.8)	82.0 (78.0, 89.0)	87.0 (73.0, 102.5)	0.017
Respiratory rate (bpm)	22.0 (22.0, 24.0)	22.0 (22.0, 24.0)	24.0 (22.0, 24.0)	22.0 (22.0, 24.0)	23.0 (22.0, 24.0)	0.20

Baseline characteristic	Overall[1] N = 420 Median (IQR); n/N (%)	Outcome category				p-value[1]
		1 No Supplemental O_2 Requirement N = 331 Median (IQR); n/N (%)	2 Supplemental O_2 Requirement N = 62 Median (IQR); n/N (%)	3 Non-invasive ventilation N = 15 Median (IQR); n/N (%)	4 Mechanical ventilation and/or Death N = 12 Median (IQR); n/N (%)	
Oxygen saturation (%)	98.0 (96.0, 99.0)	98.0 (97.0, 99.0)	96.0 (95.0, 97.0)	97.0 (96.0, 98.0)	96.0 (96.0, 98.2)	< 0.001
Systolic BP (mmHg)	126.0 (115.0, 134.0)	126.0 (115.5, 135.0)	123.0 (115.2, 132.8)	130.0 (117.0, 140.0)	125.0 (116.5, 130.8)	0.80
Axillary temperature (°C)	36.9 (36.6, 37.1)	36.8 (36.5, 37.1)	36.9 (36.5, 37.2)	37.0 (36.8, 37.1)	37.1 (36.9, 37.8)	0.005
qSOFA ≥ 2	26 / 420 (6.2%)	17 / 331 (5.1%)	7 / 62 (11%)	1 / 15 (6.7%)	1 / 12 (8.3%)	0.20
Biochemical biomarkers						
Platelet count (x10^9 cells/l)*	199.0 (147.0, 261.0)	207.5 (156.0, 268.2)	175.0 (126.0, 246.5)	159.0 (126.5, 184.5)	170.0 (88.8, 187.0)	0.002
White cell count (x10^9 cells/l)*	6.2 (4.6, 7.8)	6.2 (4.7, 7.6)	6.6 (4.6, 9.8)	5.0 (3.4, 5.9)	6.8 (3.9, 9.2)	0.024
Neutrophil count (x10^9 cells/l)*	4.0 (2.8, 5.7)	4.0 (2.8, 5.4)	4.9 (3.2, 7.5)	3.7 (2.6, 4.3)	5.8 (2.9, 6.4)	0.013
Lymphocyte count (x10^9 cells/l)*	1.3 (0.9, 1.9)	1.4 (1.0, 2.0)	1.1 (0.7, 1.4)	0.9 (0.6, 1.0)	1.1 (0.6, 1.4)	< 0.001
NLR*	3.1 (1.8, 5.1)	2.7 (1.7, 4.4)	4.7 (3.2, 7.4)	4.6 (3.0, 5.8)	4.0 (2.2, 9.0)	< 0.001
Ang-2 (pg/ml)*	1,688.0 (1,237.0, 2,306.8)	1,688.0 (1,237.0, 2,225.0)	1,645.0 (1,266.8, 2,529.5)	1,367.0 (1,166.5, 2,003.0)	3,095.0 (2,004.5, 6,996.5)	0.039
CRP (mg/l)*	37.7 (6.8, 107.8)	26.0 (5.5, 87.8)	83.0 (27.7, 158.0)	49.7 (23.2, 130.2)	79.1 (69.5, 159.4)	< 0.001

Baseline characteristic	Overall[1] N = 420 Median (IQR); n/N (%)	Outcome category				p-value[1]
		1 No Supplemental O_2 Requirement N = 331 Median (IQR); n/N (%)	2 Supplemental O_2 Requirement N = 62 Median (IQR); n/N (%)	3 Non-invasive ventilation N = 15 Median (IQR); n/N (%)	4 Mechanical ventilation and/or Death N = 12 Median (IQR); n/N (%)	
D-dimer (ng/ml)*	847.3 (467.0, 1,520.2)	735.7 (410.1, 1,358.8)	1,115.1 (788.4, 2,334.3)	970.1 (751.3, 1,674.2)	2,612.6 (1,153.5, 4,802.7)	< 0.001
IL-1ra (pg/ml)*	1,000.5 (591.0, 1,838.0)	841.5 (543.2, 1,535.0)	1,688.5 (998.8, 2,588.8)	2,433.0 (1,087.0, 3,405.5)	3,879.0 (2,016.0, 6,333.0)	< 0.001
IL-6 (pg/ml)*	19.5 (6.5, 47.0)	13.6 (5.2, 36.5)	43.0 (19.3, 81.6)	68.9 (29.6, 73.2)	89.6 (56.2, 148.0)	< 0.001
IL-8 (pg/ml)*	10.6 (7.8, 15.6)	9.8 (7.4, 13.9)	13.9 (9.2, 18.3)	14.3 (9.6, 22.1)	25.1 (20.2, 31.6)	< 0.001
IL-10 (pg/ml)*	8.4 (5.6, 15.1)	7.1 (5.2, 12.5)	13.8 (9.4, 20.2)	17.2 (11.2, 22.0)	28.1 (22.1, 29.4)	< 0.001
IP-10 (pg/ml)*	977.5 (377.5, 1,951.2)	766.5 (311.5, 1,532.5)	1,955.5 (989.0, 2,661.5)	2,641.0 (1,763.0, 3,445.5)	2,566.0 (2,009.5, 3,188.0)	< 0.001
PCT (pg/ml)*	103.5 (70.1, 164.0)	98.7 (68.2, 149.8)	122.0 (80.5, 188.2)	133.0 (88.3, 295.0)	319.0 (180.0, 961.0)	< 0.001
sTREM-1 (pg/ml)*	390.0 (271.0, 562.2)	376.0 (262.8, 536.0)	437.0 (316.5, 649.5)	380.0 (245.5, 470.0)	586.0 (395.0, 1,000.5)	0.009
suPAR (ng/ml)	4.2 (3.1, 5.7)	3.9 (2.9, 5.3)	5.7 (3.9, 6.7)	5.0 (4.5, 5.6)	6.1 (3.8, 10.0)	< 0.001
Seronegative†*	188 / 409 (46%)	139 / 323 (43%)	27 / 59 (46%)	12 / 15 (80%)	10 / 12 (83%)	0.002

[1]Kruskal-Wallis rank sum test; Fisher's exact test; Pearson's Chi-squared test

Discrimination of many of the biomarkers varied depending on the severity of disease predicted. Biomarkers that demonstrated dose-response relationships between their baseline concentrations and the maximum level of subsequent respiratory support (for example, IL-1ra, IL-6, and IL-10; Figure 6.4-10) appeared to have best discrimination (Table 6.4-14; Figure 6.4-11).

Table 6.4-14: Prognostic potential of host biomarkers in participants with moderate Covid-19. *Missing data (in baseline group unless otherwise stated): platelet count, NLR = 10 (Supplemental O_2 Requirement = 3/10); Ang-2, IL-1ra, IL-6, IL-8, IL-10, IP-10, PCT, sTREM-1 = 2 (MV and/or Death = 1/2); D-dimer = 3 (MV and/or Death = 1/3); CRP = 8. Ang-2 = angiopoietin-2; AUC = area under the receiver operating characteristic curve; CI = confidence interval; CRP = C-reactive protein; IL-1ra = interleukin-1 receptor antagonist; IL-6 = interleukin-6; IL-8 = interleukin-8; IL-10 = interleukin-10; IP-10 = interferon-gamma-inducible protein-10; NLR = neutrophil-to-lymphocyte ratio; PCT = procalcitonin; sTREM-1 = soluble triggering receptor expressed on myeloid cells-1; suPAR = soluble urokinase plasminogen activator receptor.

Biomarker	**AUC** (95% CI)		
	Supplemental O_2 Requirement N = 89	**Non-invasive ventilation** N = 27	**Mechanical ventilation and/or Death** N = 12
Ang-2 (pg/ml)*	0.52 (0.44 to 0.59)	0.54 (0.40 to 0.67)	0.73 (0.52 to 0.93)
CRP (mg/l)*	0.67 (0.61 to 0.73)	0.66 (0.57 to 0.74)	0.73 (0.65 to 0.80)
D-dimer (ng/ml)*	0.66 (0.60 to 0.72)	0.67 (0.58 to 0.77)	0.77 (0.62 to 0.92)
IL-1ra (pg/ml)*	0.73 (0.67 to 0.79)	0.77 (0.68 to 0.87)	0.83 (0.71 to 0.95)
IL-6 (pg/ml)*	0.75 (0.70 to 0.81)	0.77 (0.67 to 0.87)	0.86 (0.77 to 0.95)
IL-8 (pg/ml)*	0.68 (0.62 to 0.74)	0.72 (0.61 to 0.82)	0.83 (0.71 to 0.95)
IL-10 (pg/ml)*	0.77 (0.72 to 0.82)	0.80 (0.73 to 0.87)	0.86 (0.78 to 0.94)
IP-10 (pg/ml)*	0.77 (0.71 to 0.82)	0.80 (0.71 to 0.89)	0.77 (0.61 to 0.94)
NLR *	0.70 (0.64 to 0.76)	0.64 (0.53 to 0.75)	0.60 (0.40 to 0.80)
Platelet count ($x10^9$ cells/l)*	0.62 (0.55 to 0.69)	0.68 (0.59 to 0.78)	0.68 (0.52 to 0.84)
PCT (pg/ml)*	0.63 (0.56 to 0.69)	0.70 (0.59 to 0.82)	0.80 (0.65 to 0.95)
sTREM-1 (pg/ml)*	0.58 (0.51 to 0.65)	0.56 (0.44 to 0.68)	0.72 (0.57 to 0.88)
suPAR (ng/ml)	0.69 (0.63 to 0.75)	0.67 (0.58 to 0.76)	0.68 (0.51 to 0.86)

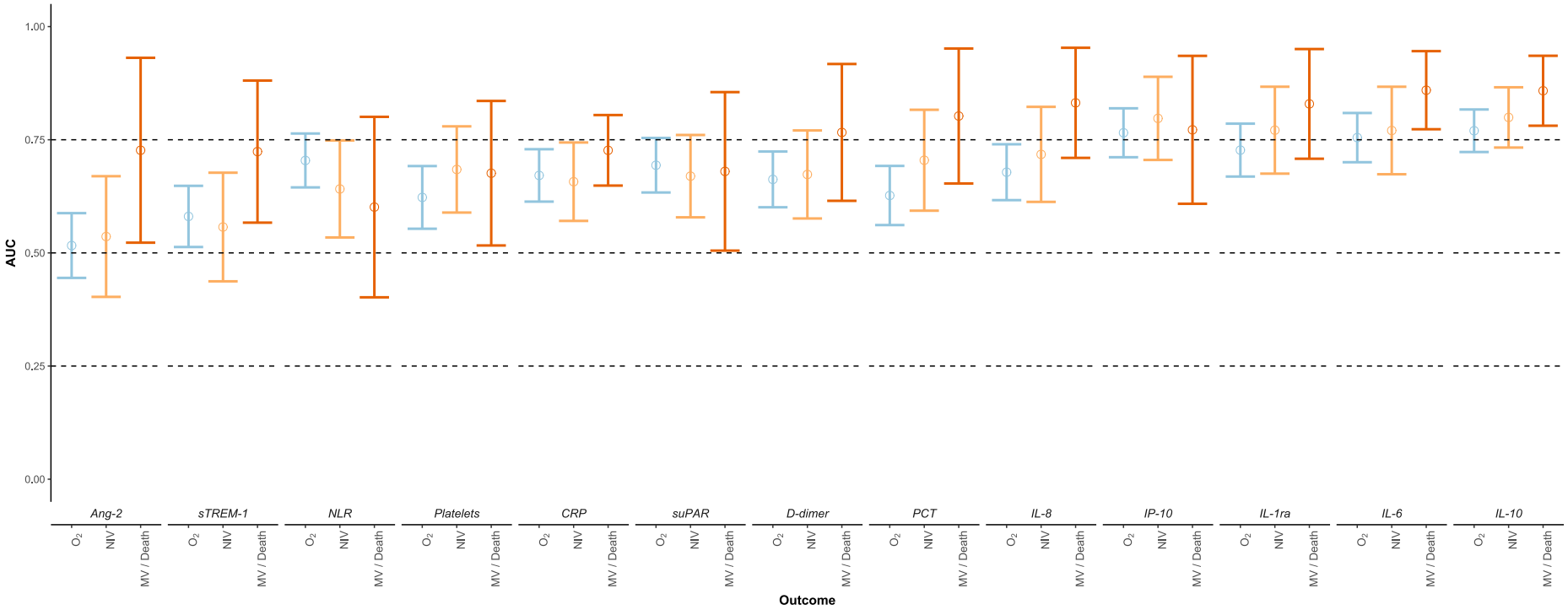


Figure 6.4-11: Prognostic potential of host biomarkers in participants with moderate Covid-19. Biomarkers ordered from left to right by ascending mean AUC across the three outcome categories. Ang-2 = angiopoietin-2; AUC = area under the receiver operating characteristic curve; CRP = C-reactive protein; IL-1ra = interleukin-1 receptor antagonist; IL-6 = interleukin-6; IL-8 = interleukin-8; IL-10 = interleukin-10; IP-10 = interferon-gamma-inducible protein-10; MV = mechanical ventilation; NIV = non-invasive ventilation; NLR = neutrophil-to-lymphocyte ratio; PCT = procalcitonin; sTREM-1 = soluble triggering receptor expressed on myeloid cells-1; suPAR = soluble urokinase plasminogen activator receptor.

6.4.9 Longitudinal changes in biomarker concentrations

Previous studies have illustrated longitudinal changes in biomarker concentrations during the natural history of a Covid-19 illness.[169] Comparison of participants who presented in the first vs. second week of their illness did not suggest clinically meaningful differences in baseline biomarker concentrations, although participants who presented after seven days of symptoms were more likely to be seropositive (Table 6.4-15). For participants who presented in their first week of illness, median symptom duration prior to enrolment was five days (IQR = 3 to 6 days) and for participants presenting in their second week of illness median symptom duration was 10 days (IQR = 8 to 11 days).

Table 6.4-15: Baseline concentrations of host biomarkers stratified by week of illness. †Seronegative defined as negative for both IgM and IgG antibodies against SARS-CoV-2. *Missing data: platelet count, white cell count, neutrophil count, lymphocyte count, NLR = 10 (Week 1 = 4; Week 2 = 6); Ang-2, IL-1ra, IL-6, IL-8, IL-10, IP-10, PCT, sTREM-1 = 2 (Week1 = 1; Week 2 = 1); D-dimer = 3 (Week 1 = 2; Week 2 = 1); CRP = 8 (Week 1); serostatus = 11 (Week 1 = 7; Week 2 = 4). Ang-2 = angiopoietin-2; CRP = C-reactive protein; IL-1ra = interleukin-1 receptor antagonist; IL-6 = interleukin-6; IL-8 = interleukin-8; IL-10 = interleukin-10; IP-10 = interferon-gamma-inducible protein-10; IQR = interquartile range; NLR = neutrophil-to-lymphocyte ratio; PCT = procalcitonin; sTREM-1 = soluble triggering receptor expressed on myeloid cells-1; suPAR = soluble urokinase plasminogen activator receptor.

Biomarker	Overall N = 420 Median (IQR); n/N (%)	Days of symptoms prior to enrolment		p-value[1]
		≤ 7 days N = 299 Median (IQR); n/N (%)	> 7 days N = 121 Median (IQR); n/N (%)	
Ang-2 (pg/ml)*	1,688.0 (1,237.0, 2,306.8)	1,603.5 (1,212.5, 2,195.5)	1,927.5 (1,325.8, 2,493.5)	0.03
CRP (mg/l)*	37.7 (6.8, 107.8)	36.4 (6.6, 96.0)	39.8 (9.3, 133.4)	0.40
D-dimer (ng/ml)*	847.3 (467.0, 1,520.2)	847.3 (457.5, 1,484.5)	861.5 (476.8, 1,598.1)	0.80
IL-1ra (pg/ml)*	1,000.5 (591.0, 1,838.0)	1,094.0 (625.5, 1,991.2)	826.0 (503.2, 1,487.8)	0.005
IL-6 (pg/ml)*	19.5 (6.5, 47.0)	20.3 (6.9, 51.2)	16.7 (5.2, 43.4)	0.13
IL-8 (pg/ml)*	10.6 (7.8, 15.6)	10.5 (7.9, 16.0)	10.6 (7.4, 13.9)	0.40
IL-10 (pg/ml)*	8.4 (5.6, 15.1)	9.1 (5.7, 16.1)	7.4 (5.3, 11.9)	0.041
IP-10 (pg/ml)*	977.5 (377.5, 1,951.2)	1,034.5 (418.8, 2,120.0)	766.5 (310.5, 1,558.0)	0.011
NLR*	3.1 (1.8, 5.1)	2.9 (1.7, 4.7)	3.3 (2.2, 5.9)	0.018

Biomarker	Overall N = 420 Median (IQR); n/N (%)	Days of symptoms prior to enrolment		p-value[1]
		≤ 7 days N = 299 Median (IQR); n/N (%)	> 7 days N = 121 Median (IQR); n/N (%)	
White cell count (x10^9 cells/l)*	6.2 (4.6, 7.8)	5.9 (4.5, 7.3)	6.8 (5.2, 9.2)	< 0.001
Neutrophil count (x10^9 cells/l)*	4.0 (2.8, 5.7)	3.7 (2.8, 5.3)	4.8 (3.4, 6.7)	< 0.001
Lymphocyte count (x10^9 cells/l)*	1.3 (0.9, 1.9)	1.3 (0.9, 1.9)	1.4 (1.0, 1.9)	0.60
Platelet count (x10^9 cells/l)*	199.0 (147.0, 261.0)	190.0 (143.0, 246.0)	230.0 (166.0, 309.0)	< 0.001
PCT (pg/ml)*	103.5 (70.1, 164.0)	106.0 (70.2, 168.5)	98.6 (70.0, 147.0)	0.30
sTREM-1 (pg/ml)*	390.0 (271.0, 562.2)	377.5 (259.2, 530.5)	441.5 (306.2, 633.2)	0.005
suPAR (ng/ml)	4.2 (3.1, 5.7)	3.9 (2.9, 5.5)	4.6 (3.4, 5.9)	0.003
Seronegative†*	46% (188 / 409)	50% (147 / 292)	35% (41 / 117)	0.005

[1]Wilcoxon rank sum test; Pearson's Chi-squared test

6.4.10 No suggestion of confounding by corticosteroid use

Corticosteroids were readily available in the study's contexts and were often self-prescribed or used off-license. Although they have only been shown to convey benefit in patients who have already developed a supplemental oxygen requirement,[251] it is possible that they may confer benefit (or cause harm) in a subset of patients with moderate disease. Notwithstanding that the primary purpose of the model was prediction (rather than causal inference), potential confounding by corticosteroid use was explored to account for the fact that patterns of corticosteroid use may vary in different settings and impact generalisability of the clinical prediction models. Steroid use was associated with some candidate predictors (Table 6.4-16) but was not associated with the primary outcome (Table 6.4-17) and is therefore unlikely to have confounded the observed association.

Table 6.4-16: Candidate predictor variables, stratified by corticosteroid use. Corticosteroid use is defined as use of oral or parenteral steroids, occurring at least one calendar day prior to developing an oxygen requirement for participants who met the primary outcome. *Missing data: CRP = 8, D-dimer = 3, IL-6 = 2, NLR = 12; PCT = 2; sTREM-1 = 2. CRP = C-reactive protein; IL-6 = interleukin-6; IQR = interquartile range; NLR = neutrophil-to-lymphocyte ratio; PCT = procalcitonin; sTREM-1 = soluble triggering receptor expressed on myeloid cells-1; suPAR = soluble urokinase plasminogen activator receptor.

Candidate predictor	Overall (n = 423) Median (IQR); n/N (%)	Corticosteroid use		p-value[1]
		No (n = 265) Median (IQR); n/N (%)	Yes (n = 158) Median (IQR); n/N (%)	
Age (years)	53.0 (41.0, 62.0)	53.0 (40.0, 62.0)	54.5 (42.0, 62.0)	0.40
Male sex	286 / 423 (68%)	175 / 265 (66%)	111 / 158 (70%)	0.40
Oxygen saturation (%)	98.0 (96.0, 99.0)	98.0 (96.0, 99.0)	97.0 (96.0, 98.0)	0.017
CRP (mg/l)*	36.7 (7.0, 108.6)	37.4 (6.8, 110.3)	35.5 (7.3, 101.0)	0.80
D-dimer (ng/ml)*	856.1 (470.8, 1,522.7)	846.7 (484.0, 1,484.5)	882.1 (458.4, 1,706.4)	0.60
IL-6 (pg/ml)*	19.8 (6.6, 47.6)	21.2 (7.2, 53.9)	15.4 (4.8, 43.2)	0.042
NLR*	3.1 (1.9, 4.9)	2.7 (1.7, 4.3)	3.7 (2.3, 6.2)	< 0.001
PCT (ng/ml)*	0.1 (0.1, 0.2)	0.1 (0.1, 0.2)	0.1 (0.1, 0.2)	0.80
sTREM-1 (pg/ml)*	394.0 (272.0, 563.0)	400.0 (269.0, 563.0)	381.0 (281.2, 565.2)	> 0.90
suPAR (ng/ml)	4.2 (3.1, 5.7)	4.1 (3.0, 5.6)	4.3 (3.2, 5.8)	0.20

[1]Wilcoxon rank sum test; Pearson's Chi-squared test

Table 6.4-17: Association of corticosteroid use with primary outcome, stratified by recruitment site. Corticosteroid use is defined as use of oral or parenteral steroids, occurring at least one calendar day prior to developing an oxygen requirement for participants who met the primary outcome. AIIMS = All India Institute of Medical Sciences; CI = confidence interval; CMC = Christian Medical College.

		Developed oxygen requirement		Total	Relative risk (95% CI)	p-value[1]
		No	Yes			
All participants (n = 423)						
Corticosteroid use	**Yes**	127	31	**158**	0.90 (0.61 to 1.32)	0.58
	No	207	58	**265**		
	Total	**334**	**89**	**423**		
AIIMS participants (n = 124)						
Corticosteroid use	**Yes**	53	12	**65**	0.73 (0.37 to 1.42)	0.35
	No	44	15	**59**		
	Total	**97**	**27**	**124**		
CMC participants (n = 299)						
Corticosteroid use	**Yes**	74	19	**93**	0.98 (0.60 to 1.58)	0.93
	No	163	43	**206**		
	Total	**237**	**62**	**299**		

[1]Pearson's Chi-squared test

6.4.11 Role of the virus and the host in influencing disease progression

As respiratory specimens were collected using different techniques and assayed on different PCR platforms at each site, it was not possible to accurately determine the roles of the virus and host in influencing the risk of deterioration. Furthermore, respiratory specimens were only collected in participants who did not have documented laboratory-confirmed SARS-CoV-2 infection within the three days prior to recruitment. Exploratory analyses, restricted to participants who had respiratory specimens collected within the 24 hours prior to recruitment, and stratified by collection technique and PCR platform, indicated that cycle threshold (Ct) value was not associated with the risk of deterioration (Table 6.4-18). As expected, combined nasopharyngeal and oropharyngeal swabs resulted in lower Ct values than nasopharyngeal swabs. Seronegative status was associated with an

increased risk of deterioration: 49/190 (25.8%) participants seronegative at enrolment vs. 37/222 (16.7%) participants seropositive at enrolment met the primary outcome (p = 0.023).

Table 6.4-18: Association of baseline Ct value with primary outcome, stratified by collection technique and PCR platform. NPS were collected at CMC and combined NPS and OPS were collected at AIIMS. Only swabs collected within the 24 hours prior to recruitment (n = 242/423; 57.2%) are included in these subgroup analyses. All swabs were tested using the Cepheid Xpert Xpress SARS-CoV-2 (California, USA) at AIIMS. Both the Cepheid Xpert Xpress SARS-CoV-2 (California, USA) and Altona RealStar SARS-CoV-2 rRT-PCR (Hamburg, Germany) were used at CMC. AIIMS = All India Institute of Medical Sciences; CMC = Christian Medical College; Ct = cycle threshold; IQR = interquartile range; NPS = nasopharyngeal swab; OPS = oropharyngeal swab.

Stratification factor	Ct value			p-value[1]
	Overall Median (IQR)	Developed oxygen requirement		
		No Median (IQR)	Yes Median (IQR)	
All participants (n = 242)	27.7 (23.1, 32.7)	27.4 (22.8, 33.0)	28.4 (24.1, 32.4)	0.40
Collection technique (n = 242)				
NPS (n = 118)	31.5 (26.7, 35.7)	31.5 (26.1, 35.8)	31.5 (28.0, 34.6)	0.90
Combined NPS and OPS (n = 124)	25.3 (20.5, 28.5)	25.5 (20.2, 28.8)	24.5 (21.4, 27.8)	> 0.90
PCR platform (n = 242)				
Cepheid Xpert Xpress (n = 173)	26.9 (21.0, 31.5)	26.6 (20.7, 30.9)	27.6 (24.0, 32.1)	0.20
Altona RealStar (n = 69)	31.6 (27.0, 35.4)	31.8 (27.0, 35.4)	31.2 (28.0, 35.4)	0.60
PCR platform at CMC site (n = 118)				
Cepheid Xpert Xpress (n = 49)	31.2 (26.3, 35.8)	30.7 (23.2, 36.1)	31.9 (28.4, 33.3)	0.40
Altona RealStar (n = 69)	31.6 (27.0, 35.4)	31.8 (27.0, 35.4)	31.2 (28.0, 35.4)	0.60

[1]Wilcoxon rank sum test

6.5 Discussion

This study reports the development and temporal validation of three promising clinical prediction models to assist with the assessment of patients presenting with moderate Covid-19 and demonstrates equivalent prognostic accuracy of a commercially-available RDT and ELISA for suPAR, a

leading biomarker candidate for the risk stratification of Covid-19 and other febrile illnesses. Unlike the majority of studies which focus on risk stratification of hospitalised patients,[267] PRIORITISE included patients in whom there is clinical uncertainty as to whether admission is warranted, and adopted an analytical framework which acknowledged that the trade-offs inherent in this decision will vary at different stages of a pandemic and in different healthcare settings. Clinically, all patients had moderate severity disease at the time of enrolment, suggesting that rather than being crude laboratory surrogates for bedside assessment, certain biomarkers might reflect subclinical pathophysiological changes and could add value to clinical risk scores.

6.5.1 Clinical prediction models

The study was motivated by the need for a practical tool to support health systems in under-resourced settings triage high volumes of patients during current and future surges in SARS-CoV-2 infections.[261] As such, the models combined three easily ascertainable clinical parameters (age, sex, and SpO_2) with measurement of a single host response biomarker (IL-6, NLR, or suPAR), quantifiable using commercially-available near-patient tests. Whilst the models would benefit from additional validation, if these results are confirmed they would be readily implementable and could help decompress overstretched health systems by supporting clinicians to identify which patients are most appropriate for community-based management.

Specific systemic symptoms were used to define moderate severity disease rather than hospitalisation as per the WHO-CPS, recognising, as did the scale's original authors, that the lower-end of the WHO-CPS is subjective.[289] Performance of any prediction model is influenced by the characteristics of the cohort within which it predicts and thus these more robust study entry criteria will hopefully better standardise patient characteristics and facilitate model transportability; symptoms were defined using the widely-adopted International Severe Acute Respiratory and emerging Infection Consortium (ISARIC) CRF in order to permit validation by other groups.[315]

In recognition of the fact that laboratory tests carry an opportunity cost, especially when resources are limited, the study focussed on quantifying the added value of host biomarkers. Although a model containing clinical parameters alone would be simpler to implement, the results indicate that inclusion of one biomarker test would provide greater utility and allow use of the models in a broader range of contexts, including when bed pressures are less acute early in a Covid-19 surge.

The models have face validity. All clinical and laboratory predictors had previously been implicated in the pathogenesis and risk stratification of Covid-19.[24,269,276,278,280,309] Reassuringly, the three host response biomarkers that demonstrate most promise for ruling out progression to supplemental oxygen requirement also have biological plausibility. In addition to being a therapeutic target,[250] previous studies indicate that raised IL-6 levels predict development of an oxygen requirement,[171,280] and along with an elevated NLR, form part of the Covid-19-associated hyperinflammatory syndrome (cHIS) diagnostic criteria.[316] Elevated suPAR levels are associated with disease severity and progression in both moderate and severe Covid-19,[278,282,306] and have been used to inform recruitment into trials of immunomodulatory agents.[317] Just as no silver bullet exists for any problem in healthcare,[318] no single biomarker will provide optimal risk stratification for all patients with Covid-19. However, the results suggest that when used in conjunction with simple clinical parameters, these biomarkers, which are measurable with commercially-available near-patient tests, could help clinicians risk stratify patients presenting with moderate Covid-19.

6.5.2 Importance of developing tools for specific clinical use-cases

In this study, age and sex were not strongly associated with disease progression, in contrast to their well-recognised association with Covid-19 mortality. This is similar to others' findings and underlines the importance of developing models for specific clinical use-cases.[309] Models developed to predict mortality amongst inpatients are not necessarily applicable to less severe disease and

ambulatory care, just as models developed for populations in well-resourced health systems may not generalise to individuals in LMIC settings.[221]

In keeping with this, results from the secondary exploratory biomarker analyses illustrate that the prognostic utility of a particular biomarker is not a standalone concept and is linked to the clinical outcome(s) being predicted. Different biomarkers predict different disease severities and may be suited to distinct clinical use-cases. Biomarkers which discriminate between patients who progress to require supplemental oxygen may be most helpful to support community-based triage (ruling-out need for hospitalisation), whereas those which predict more severe illness (NIV, MV, or death) may be better deployed to guide inpatient resource allocation (ruling-in need for frequent monitoring and/or admission to restricted-capacity high-dependency care areas). Biomarkers with dose-response relationships with subsequent disease progression (IL-1ra, IL-6, IL-8, and IL-10) demonstrated promising discrimination across a range of disease severities and are particularly attractive candidates for further exploration.

The apparent stepwise reduction in NLR in patients who progressed to more severe illness is difficult to reconcile. Although lower white cell and neutrophil counts are recognised in patients with overwhelming sepsis, this phenomenon is unlikely to explain this observation as all participants had moderate severity disease at the time of biomarker measurements. It may be a function of the small number of events in the more severe outcome categories and comparatively higher proportion (n = 10; 2.4%) of missing data for NLR. Nevertheless, if confirmed, it may limit the practical utility of NLR as a risk stratification marker for patients presenting with moderate Covid-19 and suggests that in the presence of a low NLR a second parameter (clinical or biomarker) would be required to identify patients at risk of disease progression.

6.5.3 Prognostic accuracy of a RDT for suPAR

Comparable prognostic performance of the suPAR RDT was achieved despite limited agreement with the reference test and was maintained whether suPAR concentrations were used alone or as a constituent parameter in the multivariable clinical prediction model. These results illustrate the importance of conducting both analytical and clinical validations of candidate point-of-care tests. The limited agreement between the two suPAR tests indicates that the RDT cannot be used interchangeably with the ELISA and results from studies utilising different suPAR assays should not be pooled, a conclusion that is likely to remain valid for other biomarkers and assays unless strong agreement has been demonstrated. If tests are adopted for clinical use, cut-offs proposed to inform particular decisions (for example, admission or discharge from the emergency department) or weightings within multivariable clinical prediction models must be assay-dependent.[319,320] Similarly, if measurements are used to inform participant recruitment into clinical trials, it is important that eligibility criteria are tailored to the assay used for enrolment.[317]

The lack of agreement between the RDT and ELISA is unexpected. The manufacturer reports that suPAR concentrations measured using the RDT should be within ± 10% of measurements made on the ELISA, further underlining the need for context-specific evaluation.[307,319] The impact of different detection methods for suPAR has been demonstrated,[321] however in this study both the ELISA and RDT used the same capture antibodies. Although RDT measurements were made after an additional freeze-thaw cycle, multiple studies have confirmed that suPAR concentrations are stable up to at least five repeated freeze-thaw cycles.[322,323]

In this multi-site Indian cohort, suPAR concentrations quantified using the RDT were higher than might be anticipated for non-severe patients attending an emergency department.[307] One other study has quantified suPAR concentrations using the same RDT in India: 66.3% (126/190) of patients attending an emergency department had suPAR concentrations > 5.5 ng/ml, which is broadly

consistent with the results described in this study.[324] It is possible that an unknown factor may lead to elevated suPAR concentrations in this population and this merits further investigation.

Nevertheless, as these results demonstrate, lack of agreement between inexpensive, quick, and practical RDTs and batched, quality-controlled, laboratory-based assays, does not necessarily preclude clinical utility of an index test on the field. Given these promising results, further exploration of the utility of suPAR-guided risk stratification of patients presenting with Covid-19 and other acute infections in settings with limited laboratory capacity is warranted.

6.5.4 Strengths and limitations

This study addressed the limitations identified in other Covid-19 prognostic models by following the TRIPOD guidelines,[66] and using a prospectively collected dataset with minimal loss-to-follow-up and missing data.[267] Biomarker concentrations were measured at the time of recruitment, before any patient had developed a supplemental oxygen requirement, allowing confident evaluation of the prognostic potential of the biomarkers.[325] It is the first study to evaluate the analytical performance and prognostic accuracy of a RDT for suPAR, head-to-head against a reference test, in tropical resource-limited environments which subjected the tests to wide temperature and humidity ranges.

Nevertheless, the small validation cohort (determined by the natural history of the pandemic in India) limits the ability to draw strong conclusions. Although the same models appeared superior in the different analyses performed, further external validation would be beneficial. Publication of the full models will hopefully encourage independent evaluation.

Vaccinated and previously infected individuals, as well as those with more recent VOCs, were not included and hence caution is required if findings are extrapolated to these populations. It is likely that similar pathophysiological pathways (and biomarkers) are implicated, although this requires

empirical testing and disease progression will be more frequent in unvaccinated and immune-naïve cohorts. Consequently, the models may require recalibration for use in vaccinated populations with lower baseline risk of progression to severe Covid-19. However, it is important to note that only 15/54 African countries met the WHO target of vaccinating 10% of their population by the end of September 2021,[326] and vaccination rates remain inadequate in many low-income settings.[247] An estimated 55-70% vaccination coverage is required to achieve herd immunity for a vaccine with 90% efficacy.[327] Unfortunately, the timelines for adequate vaccination coverage in many LMICs are likely to be long.

The models were developed and validated within the Indian healthcare context. Whilst the two sites are over 2,000 kilometres apart, data from another continent would be important to assess generalisability. There were no licensed treatments for moderate Covid-19 available in LMICs during the time the study was conducted, and hence the models will require updating if they are to be used in contexts where these become available: it may be that the models can help clinicians identify which patients are most likely to benefit from such therapies.[28] Although it is possible that immunomodulatory drugs that have demonstrated efficacy in later-stage disease may provide benefit in a subset of moderate patients with particular immunological signatures,[250,251,317,328] the analyses indicate that corticosteroid use is unlikely to have confounded the observed associations.

Oxygen requirement was selected as the primary outcome as this reflects a clinically meaningful endpoint and the practical ceiling of care in many LMIC settings. An $SpO_2/FiO_2 < 400$ was used for participants without documented hypoxaemia or tachypnoea prior to initiation of supplemental oxygen, as the threshold for oxygen therapy can be subjective and vary depending on available resources.[289,292] It is unlikely that the outcome lacked sensitivity; only one participant who received supplemental oxygen did not meet the primary outcome. It may have lacked specificity (12 participants who met the primary outcome did not receive supplemental oxygen and calculation of FiO_2 in non-ventilated patients can overestimate pulmonary dysfunction),[329] but sensitivity would always be prioritised in a tool to inform community-based management. Furthermore, any outcome

misclassification is likely to have reduced, rather than exaggerated, the prognostic performance of the candidate predictors and models.[330]

Follow-up was censored at day 14, however it was possible to retrospectively confirm that no further disease progression occurred beyond this point. MV and death were combined into a single ordinal category for the secondary outcome due to few surviving ventilated patients. This loss of granularity may have underestimated biomarker discrimination.[330] However, combining MV and death was preferable to combining NIV and MV: NIV may have had a less-consistent threshold for initiation as it could be provided outside the ICU setting.

It was not possible to comprehensively explore the roles of the virus and host in influencing the risk of deterioration, as respiratory specimens were collected using different techniques and assayed on different PCR platforms at each site. In keeping with other studies, seronegative status at enrolment was associated with an increased risk of deterioration.[328,331] As near-patient antibody tests are available this warrants further exploration, acknowledging that this is likely most relevant in patients without a history of previous Covid-19 illness or vaccination.

Limitations of the nested RDT evaluation include the fact that the RDTs were performed on frozen plasma by laboratory technicians. If the tests were to inform real-time clinical decisions, fresh plasma would be used and trained laboratory technicians may not be available, especially in contexts with limited laboratory capacity. Future research should extend these results to explore the field-based implementation of the RDT using unfrozen patient samples and evaluate reliability and usability amongst lesser-trained practitioners. Development of a test that is compatible with capillary (finger-prick) whole blood rather than venous plasma would increase the range of contexts in which it could be considered for use.

6.5.5 Conclusions

This study indicates that a number of host biomarkers implicated in the pathophysiology of other febrile illnesses may also play a role in the natural history of SARS-CoV-2 infections. The clinical prediction models presented address an unmet need in the Covid-19 care continuum and could help clinicians to identify patients who are suitable for community-based management. They are of particular relevance where resources are scarce and, if validated, would be practical for implementation in many LMIC contexts. Routinely collected data from MSF medical facilities across 26 LMICs indicate that 54.4% (18,400 / 33,780) of patients presenting with clinically-suspected Covid-19 between March 2020 and November 2021 whom might be considered for admission (moderate, severe, or critical disease), or 16.2% of all patients (18,400 / 113,455), would have been eligible for assessment using the models, illustrating the potential for widespread impact.

7 Discussion

Almost all clinical encounters involve health workers predicting the risk of a particular condition or future outcome in their patients in order to guide management. Predictions can either be diagnostic or prognostic in nature, and may be based on individual factors or integrate multiple pieces of data according to a specific framework, commonly referred to as a multivariable clinical prediction rule or model.

This thesis identified existing prognostic factors and clinical prediction rules relevant to the management of paediatric febrile illnesses in resource-limited settings and evaluated the performance of a number of different clinical prediction rules in children with acute respiratory infections on the Thailand-Myanmar border and critically ill children admitted to a paediatric intensive care unit in northern Cambodia. It also developed clinical prediction models to support resource allocation in settings with emerging paediatric critical care capacity and investigated the value added by host biomarker measurements to clinical assessment for community-based triage of childhood pneumonia and adults with moderate Covid-19.

7.1 Individual prognostic factors vs. multivariable clinical prediction

Predictions based on individual factors are prone to confounding unless the entity to be predicted can be measured directly. By definition, this is not possible for prognostic factors as the outcome of interest has not occurred at the time the prediction is made. As such, predictive performance of individual prognostic factors often varies across settings and multivariable prediction should be the preferred option for accurate prognostication.

Whilst some individual prognostic factors function reasonably effectively as red flags, they struggle to rule-out progression to serious illness. Only 3/200 prognostic factors identified in the

systematic review demonstrated rule-out potential [Chapter 3], indicating that a multivariate approach is required to accurately exclude severe disease. Ruling-out serious disease can be particularly valuable for common conditions such as febrile illnesses, especially in under-resourced health systems which are susceptible to overburdening.

Combinatorial approaches are also important when dealing with heterogeneous conditions such as febrile illnesses. Although final common pathways to severe febrile illness have been identified, no single clinical parameter or biomarker (no silver bullet) will adequately risk stratify all febrile patients. In the systematic review [Chapter 3], only 30/200 prognostic factors met the prespecified threshold for clinical relevance, illustrating the difficulties in identifying parsimonious predictors relevant for all febrile patients.

7.2 Clinical prediction rules and clinical prediction models

Historically, multivariable clinical prediction tools have taken the form of 'prediction rules' (also referred to as clinical risk scores), which allocate a certain number of 'points' to each constituent predictor dependent on its value. Points are then summed to give the overall 'score' for an individual patient and specific clinical actions are recommended in line with the obtained score. Although simple to compute, prediction rules constructed in this manner are inflexible: outputs are discrete (tied to levels of a score) and prognostic utility changes substantially between levels, hindering application across different settings [Chapter 4; Chapter 5]. This is particularly problematic for young febrile children whereby many clinical parameters included in clinical prediction rules (for example, heart rate and respiratory rate) are affected by other factors (for example, temperature and a child's activity level), small alterations in which can lead to large changes in predicted risk.

In this regard, clinical prediction models offer an advantage. Combining their constituent parameters to provide a continuous estimate of risk, predictions are more stable and decision

thresholds can be adjusted along a continuum and tailored to meet the needs of specific healthcare contexts. With the proliferation and accessibility of smartphones, application of clinical prediction models is now feasible in many resource-limited settings and could offer more nuanced and context-specific guidance for the management of febrile illnesses. Systems and actors responsible for healthcare vary greatly between and within countries, particularly in rural and conflict settings. Decision thresholds that can be adapted to suit different modes of care delivery and resource availability are essential. Further, thresholds that can be adjusted to account for the dynamic nature of febrile illnesses would ensure health facilities are not unduly burdened at times of high patient volumes or during disease epidemics [Chapter 5; Chapter 6].

7.3 Identifying patients at risk of disease progression

Prognostic performance often deteriorates as the time horizon for prediction becomes more distant, reflecting increasing complexity of the prediction problem as the clinical question moves from diagnosis to prognosis [Chapter 3; Chapter 4]. Different variables are suited to identifying patients sick at the time of presentation compared to those who might deteriorate later in their illness course.

Whilst many clinical prediction tools include variables that accurately identify patients who are already unwell (for example, vital signs and bedside coma scales), inclusion of parameters that reflect underlying host susceptibility states (for example, nutritional indices and comorbidities), as well as contextual determinants of outcome (for example, variables reflecting patients' illness journeys) could improve prognostication [Chapter 5]. Furthermore, certain host response biomarkers implicated in final common pathways to severe febrile illness and that reflect subclinical pathophysiological changes could add value to clinical assessment [Chapter 4; Chapter 6]. In essence, just as combinatorial approaches can address heterogeneity in the target condition (febrile illness), combining predictors from multiple domains can also help deal with heterogeneity in the outcome, ensuring algorithms are

able to identify both patients overtly sick at the time of assessment and those whose clinical condition will deteriorate later (Figure 7.3-1).

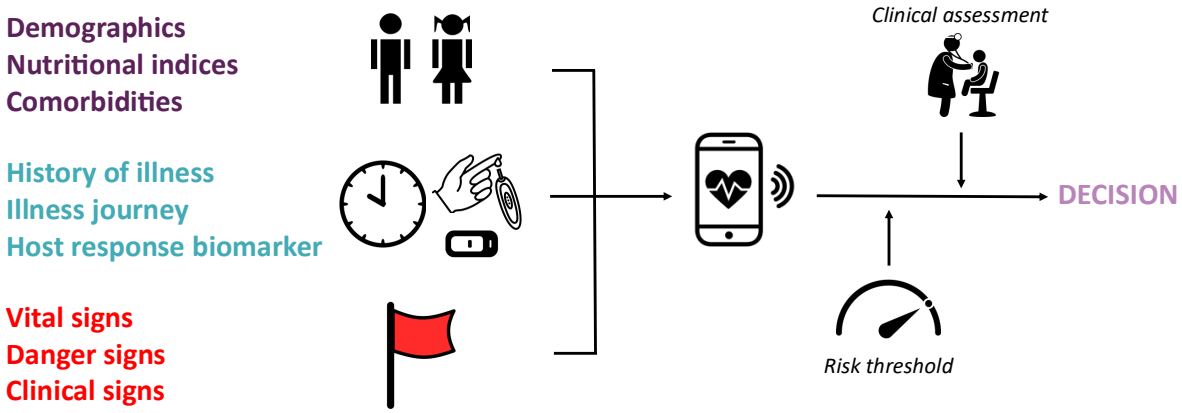


Figure 7.3-1: Schematic to illustrate important domains in the prediction of disease severity in febrile illnesses in resource-limited settings.

Biochemical biomarkers could be particularly useful adjuncts to clinical assessment in community and outpatient settings where a major challenge is identifying patients at risk of disease progression who lack clinical signs of severity as determined by existing management algorithms [Chapter 4; Chapter 6]. Biomarkers of endothelial injury and the host immune response are attractive candidates for these purposes: immune activation is a feature of all systemic infections and the microvasculature represents a highly conserved structure distributed throughout the human body, making it ideally positioned to sense physiological disturbances irrespective of the organ system in which they arise. Angiopoietin-2, a marker of endothelial injury studied in a number of hospital-based febrile illness studies, demonstrated utility for risk stratification of childhood pneumonia in a community setting on the Thailand-Myanmar border [Chapter 4]. However, these findings were not replicated in adults presenting with moderate Covid-19 in India, where other biomarkers of immune activation (soluble urokinase plasminogen activator receptor and interleukin-6) appeared more promising [Chapter 6]. As mentioned above, a silver bullet will not exist for risk stratification of all febrile illnesses nor for all outcomes of interest and further research is needed to determine which

(combinations of) biomarkers might be most useful in different contexts, populations, and diseases, and how best to integrate them within existing and changing clinical workflows.

7.4 Quality of clinical prediction research

Laboratory parameters, such as host biomarkers, always carry an opportunity cost, especially in settings where resources are scarce. It is crucial that the value added by new laboratory tests is considered. In order for this to be possible, reporting of clinical prediction research must improve. Most studies stop at the area under the receiver operating characteristic curve but this correlates poorly with clinical utility [Chapter 3]. Methods that allow different clinical prediction models to be compared and which acknowledge that the threshold for ruling in and ruling out serious disease is context dependent are becoming more widespread. In this thesis emphasis was placed on both the calibration and discrimination of prediction models and whilst analyses were always framed within a particular clinical use-case, decision curve analyses provided a mechanism to evaluate to what extent comparative clinical utility of different models was generalisable across various settings.

Clinical prediction models must be developed in populations that are representative of their intended use settings. For models hoping to contribute to the management of febrile illnesses in resource-limited settings this can be challenging. However, studies conducted in specialist urban centres or restricted to hospitalised patients will be subject to spectrum bias if their findings are applied in rural and/or community settings. In this thesis, scores developed in urban centres performed sub-optimally in a more rural paediatric intensive care unit [Chapter 5] and soluble triggering receptor expressed on myeloid cells-1 (sTREM-1), a host biomarker of severity which has demonstrated great promise in hospital-based febrile illness studies, failed to generalise to community settings [Chapter 4; Chapter 6]. Furthermore, as others have highlighted, studies that focus on hospitalised individuals exclude patients initially triaged as low-risk and suitable for community-based management: results from this thesis illustrate that if even a small proportion of these patients deteriorate, performance of

clinical prediction models is adversely impacted such that they may not be fit for purpose at the community level [Chapter 4].

7.5 Ongoing complementary studies and directions for future research

There are many reasons that few clinical prediction studies are conducted in resource-limited community settings [Chapter 3], from logistical constraints to structural inequities in funding mechanisms for health research. One issue relates to the selection of relevant clinical outcomes. Whilst 'hard' outcomes (for example, mortality) are used in most hospital-based studies, these occur at very low prevalence in community settings, making it impractical to adequately power studies using these endpoints. Furthermore, although mortality is a 'hard' and verifiable endpoint, it too has limitations, being influenced by the resources available in a particular context and susceptible to biasing findings towards patients presenting with fulminant disease and neglecting those with severe but in many instances treatable illnesses.

Whilst 'softer' outcomes (for example, hospitalisation or length of stay) may be more relevant and practical for community-based studies, these outcomes can be influenced by a multitude of contextual factors and come with an increased risk of outcome misclassification. Even small amounts of outcome misclassification can result in promising predictors being prematurely discounted.[330] An alternative approach could be to avoid dichotomisation and classify patient outcomes in a more granular way. Spot Sepsis (NCT04285021) is a multi-country observational prospective cohort study aiming to develop a clinical prediction model to guide referral of febrile children from resource-limited community settings and will explore different options for outcome classification (binary, categorical, and continuous).[201] The study enrolled both inpatients and outpatients and used a stratified recruitment strategy to capture a sufficient number of outcome events, whilst collecting the required data to weight the analysis to reflect the prevalence of severe disease at the community level and minimise spectrum bias.

Although the design of Spot Sepsis drew on recent methodological advances to mitigate the effects of spectrum bias, it remains a hospital-based study.[332] As such, if the results are promising they will still require verification in community settings. One very practical barrier to conducting biomarker research in these contexts relates to specimen collection, processing, and storage. Collection of venous blood, timely centrifugation, and maintenance of a cold chain is not feasible at the peripheral levels of many low- and middle-income country health systems. Initial work indicates that it may be possible to partially relax assumed requirements for accurate analyte quantification, with most protein biomarkers appearing stable at refrigeration temperatures on dried blood spots for up to three months.[193] Further work is required to verify these findings and explore interchangeability of capillary and venous blood samples but if these results are confirmed they could help facilitate research in rural and remote environments.

Finally, it is extremely important that impact on patient-centred outcomes is assessed. Cluster-randomised controlled trials and implementation studies are two possible methods. The South and Southeast Asian Community-based Trials Network (www.seactn.org) is a network of rural health facilities and providers spanning five countries which could provide a platform to validate and trial any clinical prediction model developed by the Spot Sepsis study.[333] If the model contains biomarkers not currently measurable using point-of-care tests, prototype devices would need to be developed. Research is ongoing to address this gap (www.isglobal.org/en/-/echilibrist-project). At Angkor Hospital for Children in Cambodia, a prospective study is under way to validate the clinical prediction model developed on the paediatric intensive care unit [Chapter 5]. Work on integrating the model within the electronic hospital information system is being planned. If validation of the model is successful this could provide a simple method to evaluate the impact of the model under routine conditions.

7.6 Conclusions

The classical prediction paradigm provides a framework to consider the diagnostic and prognostic functions that different predictors of severity can fulfil and to evaluate their roles in assessment of febrile patients. Individual prognostic factors are susceptible to confounding and perform poorly, particularly for ruling-out serious disease. Clinical prediction models combining predictors from multiple domains to provide an estimate of risk on a continuous scale could provide accurate and flexible mechanisms to support health workers and health system planners deliver equitable and efficient care for patients with febrile illnesses in resource-limited settings.

Various host response biomarkers implicated in final common pathways to severe febrile illness hold promise but further studies are needed to assess generalisability and investigate the clinical utility and cost-effectiveness of different strategies for integrating biomarker measurements into patient assessment and triage. High quality research that includes populations living in rural and remote settings and which utilises recent advances in clinical prediction science is essential if the potential for clinical prediction models to impact health outcomes for patients with febrile illnesses in resource-limited settings is to be realised.

8 References

1. World Health Organization. Informal consultation on fever management in peripheral health care settings: a global review of evidence and practice. Geneva, 2013.

2. Prasad N, Sharples KJ, Murdoch DR, Crump JA. Community prevalence of fever and relationship with malaria among infants and children in low-resource areas. *Am J Trop Med Hyg* 2015; **93**(1): 178-80.

3. McDonald CR, Weckman A, Richard-Greenblatt M, Leligdowicz A, Kain KC. Integrated fever management: disease severity markers to triage children with malaria and non-malarial febrile illness. *Malar J* 2018; **17**(1): 353.

4. Rudd KE, Johnson SC, Agesa KM, et al. Global, regional, and national sepsis incidence and mortality, 1990–2017: analysis for the Global Burden of Disease Study. *Lancet* 2020; **395**(10219): 200-11.

5. van den Berg M, van Beuningen FE, Ter Maaten JC, Bouma HR. Hospital-related costs of sepsis around the world: A systematic review exploring the economic burden of sepsis. *J Crit Care* 2022; **71**: 154096.

6. de Bont EG, Lepot JM, Hendrix DA, et al. Workload and management of childhood fever at general practice out-of-hours care: an observational cohort study. *BMJ Open* 2015; **5**(5): e007365.

7. Feikin DR, Olack B, Bigogo GM, et al. The burden of common infectious disease syndromes at the clinic and household level from population-based surveillance in rural and urban Kenya. *PLoS One* 2011; **6**(1): e16085.

8. O'Neill J. Tackling Drug-Resistant Infections Globally: Final Report and Recommendations, 2016.

9. Laxminarayan R, Matsoso P, Pant S, et al. Access to effective antimicrobials: a worldwide challenge. *Lancet* 2016; **387**(10014): 168-75.

10. Hossain MS, Siddiqee MH, Siddiqi UR, Raheem E, Akter R, Hu W. Dengue in a crowded megacity: Lessons learnt from 2019 outbreak in Dhaka, Bangladesh. *PLoS Negl Trop Dis* 2020; **14**(8): e0008349.

11. Kruk ME, Gage AD, Arsenault C, et al. High-quality health systems in the Sustainable Development Goals era: time for a revolution. *Lancet Glob Health* 2018; **6**(11): e1196-e252.

12. Firth P, Ttendo S. Intensive Care in Low-Income Countries - A Critical Need. *N Engl J Med* 2012; **367**(21): 1974-6.

13. Rudd KE, Kissoon N, Limmathurotsakul D, et al. The global burden of sepsis: barriers and potential solutions. *Crit Care* 2018; **22**(1): 232.

14. Prasad N, Murdoch DR, Reyburn H, Crump JA. Etiology of Severe Febrile Illness in Low- and Middle-Income Countries: A Systematic Review. *PLoS One* 2015; **10**(6): e0127962.

15. O'Brien KL, Baggett HC, Brooks WA, et al. Causes of severe pneumonia requiring hospital admission in children without HIV infection from Africa and Asia: the PERCH multi-country case-control study. *Lancet* 2019; **394**(10200): 757-79.

16. Hopkins H, Bassat Q, Chandler CI, et al. Febrile Illness Evaluation in a Broad Range of Endemicities (FIEBRE): protocol for a multisite prospective observational study of the causes of fever in Africa and Asia. *BMJ Open* 2020; **10**(7): e035632.

17. Kotloff KL, Nataro JP, Blackwelder WC, et al. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. *Lancet* 2013; **382**(9888): 209-22.

18. Saha SK, Schrag SJ, El Arifeen S, et al. Causes and incidence of community-acquired serious infections among young children in south Asia (ANISA): an observational cohort study. *Lancet* 2018; **392**(10142): 145-59.

19. Elven J, Dahal P, Ashley EA, et al. Non-malarial febrile illness: a systematic review of published aetiological studies and case reports from Africa, 1980-2015. *BMC Med* 2020; **18**(1): 279.

20. Shrestha P, Dahal P, Ogbonnaa-Njoku C, et al. Non-malarial febrile illness: a systematic review of published aetiological studies and case reports from Southern Asia and South-eastern Asia, 1980-2015. *BMC Med* 2020; **18**(1): 299.

21. Moreira J, Barros J, Lapouble O, et al. When fever is not malaria in Latin America: a systematic review. *BMC Med* 2020; **18**(1): 294.

22. Iroh Tam PY, Obaro SK, Storch G. Challenges in the Etiology and Diagnosis of Acute Febrile Illness in Children in Low- and Middle-Income Countries. *J Pediatric Infect Dis Soc* 2016; **5**(2): 190-205.

23. Yacoub S, Willis B. Predicting outcome from dengue. *BMC Med* 2014; **12**(147).

24. Cen Y, Chen X, Shen Y, et al. Risk factors for disease progression in patients with mild to moderate coronavirus disease 2019-a multi-centre observational study. *Clin Microbiol Infect* 2020; **26**(9): 1242-7.

25. Mayxay M, Castonguay-Vanier J, Chansamouth V, et al. Causes of non-malarial fever in Laos: a prospective study. *Lancet Glob Health* 2013; **1**(1): e46-e54.

26. Moons KG, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? *BMJ* 2009; **338**(b375).

27. Durack DT, Lukes AS, Bright DK, Duke Endocarditis Service. New Criteria for Diagnosis of Infective Endocarditis: Utilization of Specific Echocardiographic Findings. *Am J Med* 1994; **96**: 200-9.

28. Gupta RK, Calderwood CJ, Yavlinsky A, et al. Discovery and validation of a personalized risk predictor for incident tuberculosis in low transmission settings. *Nat Med* 2020; **26**(12): 1941-9.

29. World Health Organization. Pocket book of hospital care for children: guidelines for the management of common childhood illnesses. Geneva, Switzerland; 2013.

30. Sangkaew S, Ming D, Boonyasiri A, et al. Risk predictors of progression to severe disease during the febrile phase of dengue: a systematic review and meta-analysis. *Lancet Infect Dis* 2021; **21**(7): 1014-26.

31. Vuong NL, Le Duyen HT, Lam PK, et al. C-reactive protein as a potential biomarker for disease progression in dengue: a multi-country observational study. *BMC Med* 2020; **18**(1): 35.

32. World Health Organization. Integrated Management of Childhood Illnesses. Geneva, Switzerland; 2014.

33. World Health Organization. IMAI District Clinician Manual: Hospital Care for Adolescents and Adults: guidelines for the management of common illnesses with limited resources. Geneva, 2011.

34. Keitel K, Kilowoko M, Kyungu E, Genton B, D'Acremont V. Performance of prediction rules and guidelines in detecting serious bacterial infections among Tanzanian febrile children. *BMC Infect Dis* 2019; **19**(1): 769.

35. Lange S, Mwisongo A, Maestad O. Why don't clinicians adhere more consistently to guidelines for the Integrated Management of Childhood Illness (IMCI)? *Soc Sci Med* 2014; **104**: 56-63.

36. Unitaid and the Foundation for Innovative New Diagnostics (FIND). Biomarkers for acute febrile illness at the point-of-care in low-resource settings. 2021. https://www.finddx.org/wp-content/uploads/2021/05/Meeting-report_Biomarkers-for-acute-febrile-illness-at-the-point-of-care-in-low-resource-settings.pdf (accessed 27 May 2021).

37. Leligdowicz A, Richard-Greenblatt M, Wright J, Crowley VM, Kain KC. Endothelial Activation: The Ang/Tie Axis in Sepsis. *Front Immunol* 2018; **9**: 838.

38. Donadello K, Scolletta S, Covajes C, Vincent JL. suPAR as a prognostic biomarker in sepsis. *BMC Med* 2012; **10**(2).

39. Richard-Greenblatt M, Boillat-Blanco N, Zhong K, et al. Prognostic Accuracy of Soluble Triggering Receptor Expressed on Myeloid Cells (sTREM-1)-based Algorithms in Febrile Adults Presenting to Tanzanian Outpatient Clinics. *Clin Infect Dis* 2020; **70**(7): 1304-12.

40. Wright SW, Lovelace-Macon L, Hantrakun V, et al. sTREM-1 predicts mortality in hospitalized patients with infection in a tropical, middle-income country. *BMC Med* 2020; **18**(1): 159.

41. Conroy AL, Hawkes M, McDonald CR, et al. Host Biomarkers Are Associated With Response to Therapy and Long-Term Mortality in Pediatric Severe Malaria. *Open Forum Infect Dis* 2016; **3**(3): ofw134.

42. Kruk ME, Gage AD, Joseph NT, Danaei G, Garcia-Saiso S, Salomon JA. Mortality due to low-quality health systems in the universal health coverage era: a systematic analysis of amenable deaths in 137 countries. *Lancet* 2018; **392**(10160): 2203-12.

43. Debarre A. Hard to Reach: Providing Healthcare in Armed Conflict: International Peace Institute, 2018.

44. Keitel K, d'Acremont V. Electronic clinical decision algorithms for the integrated primary care management of febrile children in low-resource settings: review of existing tools. *Clin Microbiol Infect* 2018; **24**: 845-55.

45. Achan J, Tibenderana J, Kyabayinze D, et al. Case management of severe malaria--a forgotten practice: experiences from health facilities in Uganda. *PLoS One* 2011; **6**(3): e17053.

46. Hercik C, Cosmas L, Mogeni OD, et al. Health Beliefs and Patient Perspectives of Febrile Illness in Kilombero, Tanzania. *Am J Trop Med Hyg* 2019; **101**(1): 263-70.

47. World Health Organization. Managing possible serious bacterial infection in young infants when referral is not feasible. Geneva, 2015.

48. Roddy P, Dalrymple U, Jensen TO, et al. Quantifying the incidence of severe-febrile-illness hospital admissions in sub-Saharan Africa. *PLoS One* 2019; **14**(7): e0220371.

49. Argent AC. Considerations for Assessing the Appropriateness of High-Cost Pediatric Care in Low-Income Regions. *Front Pediatr* 2018; **6**: 68.

50. Argent AC, Ahrens J, Morrow BM, et al. Pediatric intensive care in South Africa: an account of making optimum use of limited resources at the Red Cross War Memorial Children's Hospital*. *Pediatr Crit Care Med* 2014; **15**(1): 7-14.

51. Prescott HC, Osterholzer JJ, Langa KM, Angus DC, Iwashyna TJ. Late mortality after sepsis: propensity matched cohort study. *BMJ* 2016; **353**: i2375.

52. Rahmel T, Schmitz S, Nowak H, et al. Long-term mortality and outcome in hospital survivors of septic shock, sepsis, and severe infections: The importance of aftercare. *PLoS One* 2020; **15**(2): e0228952.

53. Nemetchek B, English L, Kissoon N, et al. Paediatric postdischarge mortality in developing countries: a systematic review. *BMJ Open* 2018; **8**(12): e023445.

54. Akech S, Kwambai T, Wiens MO, Chandna A, Berkley JA, Snow RW. Tackling post-discharge mortality in children living in LMICs to reduce child deaths. *Lancet Child Adolesc Health* 2023; **7**(3): 149-51.

55. Madrid L, Casellas A, Sacoor C, et al. Postdischarge Mortality Prediction in Sub-Saharan Africa. *Pediatrics* 2019; **143**(1): e20180606.

56. Smart Discharges. https://www.bcchr.ca/smart-discharges (accessed 2 December 2020).

57. Prendergast AJ, Walson JL. Seeking interventions to reduce post-discharge mortality among children in sub-Saharan Africa. *Lancet Glob Health* 2019; **7**(10): e1306-e7.

58. Haniffa R, Isaam I, De Silva AP, Dondorp AM, De Keizer NF. Performance of critical care prognostic scoring systems in low and middle-income countries: a systematic review. *Crit Care* 2018; **22**(1): 18.

59. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making* 2006; **26**(6): 565-74.

60. Bouwmeester W, Zuithoff NP, Mallett S, et al. Reporting and methods in clinical prediction research: a systematic review. *PLoS Med* 2012; **9**(5): 1-12.

61. Smith S, Kennedy BJ, Dermedgoglou A, et al. A simple score to predict severe leptospirosis. *PLoS Negl Trop Dis* 2019; **13**(2): e0007205.

62. Mohd Roslani AD, Tay ST, Puthucheary SD, Rukumani DV, Sam IC. Short report: Predictors of severe disease in melioidosis patients in Kuala Lumpur, Malaysia. *Am J Trop Med Hyg* 2014; **91**(6): 1176-8.

63. Kim HL, Park HR, Kim CM, Cha YJ, Yun NR, Kim DM. Indicators of severe prognosis of scrub typhus: prognostic factors of scrub typhus severity. *BMC Infect Dis* 2019; **19**(1): 283.

64. de Hond AAH, Steyerberg EW, van Calster B. Interpreting area under the receiver operating characteristic curve. *Lancet Digit Health* 2022; **4**(12): e853-e5.

65. Fackler JC, Rehman M, Winslow RL. Please Welcome the New Team Member: The Algorithm. *Pediatr Crit Care Med* 2019; **20**(12): 1200-1.

66. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med* 2015; **162**(1): 55-63.

67. Wolff RF, Moons KGM, Riley RD, et al. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. *Ann Intern Med* 2019; **170**(1): 51-8.

68. Buntinx F, Mant D, Van den Bruel A, Donner-Banzhof N, Dinant GJ. Dealing with low-incidence serious diseases in general practice. *Br J Gen Pract* 2011; **61**(582): 43-6.

69. Molyneux EM, Ahmad S, Robertson A. Improved triage and emergency care for children reduces inpatient mortality in a resource-constrained setting. *Bull World Health Organ* 2006; **84**(4): 314-9.

70. Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015; **162**(1): W1-73.

71. Royston P, Moons KG, Altman DG, Vergouwe Y. Prognosis and prognostic research: Developing a prognostic model. *BMJ* 2009; **338**(b604).

72. Van den Bruel A, Haj-Hassan T, Thompson M, Buntinx F, Mant D. Diagnostic value of clinical features at presentation to identify serious infection in children in developed countries: a systematic review. *Lancet* 2010; **375**(9717): 834-45.

73. Van den Bruel A, Thompson MJ, Haj-Hassan T, et al. Diagnostic value of laboratory tests in identifying serious infections in febrile children: systematic review. *BMJ* 2011; **342**: d3082.

74. Cheng MP, Stenstrom R, Paquette K, et al. Blood Culture Results Before and After Antimicrobial Administration in Patients With Severe Manifestations of Sepsis: A Diagnostic Study. *Ann Intern Med* 2019; **171**(8): 547-54.

75. Parry CM, Hien TT, Dougan G, White NJ, Farrar JJ. Typhoid Fever. *N Engl J Med* 2002; **347**: 1770-82.

76. D'Acremont V, Kilowoko M, Kyungu E, et al. Beyond malaria--causes of fever in outpatient Tanzanian children. *N Engl J Med* 2014; **370**(9): 809-17.

77. Wahl B, O'Brien KL, Greenbaum A, et al. Burden of Streptococcus pneumoniae and Haemophilus influenzae type b disease in children in the era of conjugate vaccines: global, regional, and national estimates for 2000-15. *Lancet Glob Health* 2018; **6**(7): e744-e57.

78. Moons KG, de Groot JA, Bouwmeester W, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. *PLoS Med* 2014; **11**(10): e1001744.

79. Riley RD, Moons KGM, Snell KIE, et al. A guide to systematic review and meta-analysis of prognostic factor studies. *BMJ* 2019; **364**: k4597.

80. Hayden JA, van der Windt DA, Cartwright JL, Coˆte´ P, Bombardier C. Assessing Bias in Studies of Prognostic Factors. *Ann Intern Med* 2013; **158**: 280-6.

81. Moons KGM, Wolff RF, Riley RD, et al. PROBAST: A Tool to Assess Risk of Bias and Applicability of Prediction Model Studies: Explanation and Elaboration. *Ann Intern Med* 2019; **170**(1): W1-W33.

82. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. 2009; **151**: 264-9.

83. Schlapbach LJ, Kissoon N. Defining Pediatric Sepsis. *JAMA Pediatr* 2018; **172**(4): 312-4.

84. Geersing GJ, Bouwmeester W, Zuithoff P, Spijker R, Leeflang M, Moons KG. Search filters for finding prognostic and diagnostic prediction studies in Medline to enhance systematic reviews. *PLoS One* 2012; **7**(2): e32844.

85. Haynes RB, McKibbon KA, Wilczynski NL, Walter SD, Werre SR. Optimal search strategies for retrieving scientifically strong studies of treatment from Medline: analytical survey. *BMJ* 2005; **330**(7501): 1179.

86. Ingui BJ, Rogers MAM. Searching for Clinical Prediction Rules in MEDLINE. *J Am Med Inform Assoc* 2001; **8**(4): 391-7.

87. Jaeschke R, Guyatt GH, Sackett DL, Evidence-Based Medicine Working Group. Users' Guides to the Medical Literature III. How to Use an Article About a Diagnostic Test B. What Are the Results and Will They Help Me in Caring for My Patients? *JAMA* 1994; **271**(9): 703-7.

88. Fischer JE, Bachmann LM, Jaeschke R. A readers' guide to the interpretation of diagnostic test properties: clinical example of sepsis. *Intensive Care Med* 2003; **29**(7): 1043-51.

89. R Core Team. R: A language and environment for statistical computing. VIenna, Austria: R Foundation for Statistical Computing; 2020.

90. Debray TP, Damen JA, Snell KI, et al. A guide to systematic review and meta-analysis of prediction model performance. *BMJ* 2017; **356**: i6460.

91. Aramburo A, Todd J, George EC, et al. Lactate clearance as a prognostic marker of mortality in severely ill febrile children in East Africa. *BMC Med* 2018; **16**(1): 37.

92. Conroy AL, Hawkes M, Hayford K, et al. Prospective validation of pediatric disease severity scores to predict mortality in Ugandan children presenting with malaria and non-malaria febrile illness. *Crit Care* 2015; **19**: 47.

93. Lowlaavar N, Larson CP, Kumbakumba E, et al. Pediatric in-Hospital Death from Infectious Disease in Uganda: Derivation of Clinical Prediction Models. *PLoS One* 2016; **11**(3): e0150683.

94. Costa de Santana M, Amoedo CDM, Nascimento-Carvalho CM. Clinical and epidemiological characteristics of children admitted with fever in emergency department with or without sepsis. *J Infect Dev Ctries* 2017; **11**(8): 597-603.

95. Elshout G, Kool M, Bohnen AM, Koes BW, Moll HA, Berger MY. Predicting prolonged duration of fever in children: a cohort study in primary care. *Br J Gen Pract* 2015; **65**(638): e578-84.

96. Freyne B, Divilley R, Kissoon-Harrison G, O'Neill MB. Field testing the utility of procalcitonin and the acute infantile observation score in febrile infants 6 to 36 months old presenting to the pediatric emergency department with no obvious focus of infection. *Clin Pediatr (Phila)* 2013; **52**(6): 503-6.

97. Kwizera A, Kissoon N, Musa N, et al. A Machine Learning-Based Triage Tool for Children With Acute Infection in a Low Resource Setting. *Pediatr Crit Care Med* 2019; **20**(12): e524-e30.

98. George EC, Walker AS, Kiguli S, et al. Predicting mortality in sick African children: the FEAST Paediatric Emergency Triage (PET) Score. *BMC Med* 2015; **13**: 174.

99. Mtove G, Nadjm B, Hendriksen IC, et al. Point-of-care measurement of blood lactate in children admitted with febrile illness to an African District Hospital. *Clin Infect Dis* 2011; **53**(6): 548-54.

100. Mwandama D, Mwale C, Bauleni A, et al. Clinical outcomes among febrile children aged 2 to 59 months with negative malaria rapid diagnostic test results in Mchinji District, Malawi. *Malawi Med J* 2016; **28**(4): 150-3.

101. Nadjm B, Mtove G, Amos B, et al. Blood glucose as a predictor of mortality in children admitted to the hospital with febrile illness in Tanzania. *Am J Trop Med Hyg* 2013; **89**(2): 232-7.

102. Scott HF, Brou L, Deakyne SJ, Kempe A, Fairclough DL, Bajaj L. Association Between Early Lactate Levels and 30-Day Mortality in Clinically Suspected Sepsis in Children. *JAMA Pediatr* 2017; **171**(3): 249-55.

103. Scott HF, Donoghue AJ, Gaieski DF, Marchese RF, Mistry RD. Effectiveness of physical exam signs for early detection of critical illness in pediatric systemic inflammatory response syndrome. *BMC Em Med* 2014; **14**(24).

104. Scott HF, Donoghue AJ, Gaieski DF, Marchese RF, Mistry RD. The utility of early lactate testing in undifferentiated pediatric systemic inflammatory response syndrome. *Acad Emerg Med* 2012; **19**(11): 1276-80.

105. Southeast Asia Infectious Disease Clinical Research Network. Causes and outcomes of sepsis in southeast Asia: a multinational multicentre cross-sectional study. *Lancet Glob Health* 2017; **5**(2): e157-e67.

106. van Nassau SC, van Beek RH, Driessen GJ, Hazelzet JA, van Wering HM, Boeddha NP. Translating Sepsis-3 Criteria in Children: Prognostic Accuracy of Age-Adjusted Quick SOFA Score in Children Visiting the Emergency Department With Suspected Bacterial Infection. *Front Pediatr* 2018; **6**: 266.

107. Walia S, M HM, Kumble A, Mahalingam S, Baliga BS, Nanjegowda CK. Yale Observation Scale as a Predictor of Bacteremia and Final Outcome in 3-36 Months Old Febrile Children Admitted in Tertiary Health Centres: A Hospital-Based Cross-Sectional Study. *Asian J Pharm Clin Res* 2016; **9**(9): 219-22.

108. Scott HF, Colborn KL, Sevick CJ, et al. Development and Validation of a Predictive Model of the Risk of Pediatric Septic Shock Using Data Known at the Time of Hospital Arrival. *J Pediatr* 2020; **217**: 145-51 e6.

109. Goldstein B, Giroir B, Randolph A, International Consensus Conference on Pediatric S. International pediatric sepsis consensus conference: definitions for sepsis and organ dysfunction in pediatrics. *Pediatr Crit Care Med* 2005; **6**(1): 2-8.

110. von Seidlein L, Olaosebikan R, Hendriksen IC, et al. Predicting the clinical outcome of severe falciparum malaria in african children: findings from a large randomized trial. *Clin Infect Dis* 2012; **54**(8): 1080-90.

111. Helbok R, Kendjo E, Issifou S, et al. The Lambarene Organ Dysfunction Score (LODS) is a simple clinical predictor of fatal malaria in African children. *J Infect Dis* 2009; **200**(12): 1834-41.

112. Berkley JA, Ross A, Wangi I, et al. Prognostic indicators of early and late death in children admitted to district hospital in Kenya: cohort study. *BMJ* 2003; **326**(361).

113. Parshuram CS, Hutchison J, Middaugh K. Development and initial validation of the Bedside Paediatric Early Warning System score. *Crit Care* 2009; **13**(4): R135.

114. Pollack M, Patel K, Ruttimann UE. PRISM III: An updated Pediatric Risk of Mortality score. *Crit Care Med* 1996; **24**(5): 743-52.

115. Leclerc F, Duhamel A, Deken V, Grandbastien B, Leteurtre S, Groupe Francophone de Reanimation et Urgences P. Can the Pediatric Logistic Organ Dysfunction-2 Score on Day 1 Be Used in Clinical Criteria for Sepsis in Children? *Pediatr Crit Care Med* 2017; **18**(8): 758-63.

116. Seymour CW, Liu VX, Iwashyna TJ, et al. Assessment of Clinical Criteria for Sepsis: For the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA* 2016; **315**(8): 762-74.

117. Kumar N, Thomas N, Singhal D, Puliyel JM, Sreenivas V. Triage score for severity of illness. *Indian Pediatr* 2003; **40**(3): 204-10.

118. McCarthy PL, Sharpe MR, Spiesel SZ, et al. Observation Scales to Identify Serious Illness in Febrile Children. *Pediatrics* 1982; **70**(5): 802-9.

119. Kinasewitz GT, Yan SB, Basson B, et al. Universal changes in biomarkers of coagulation and inflammation occur in patients with severe sepsis, regardless of causative micro-organism. *Crit Care* 2004; **8**(2): R82-90.

120. Tan R, Kagoro F, Levine GA, et al. Clinical Outcome of Febrile Tanzanian Children with Severe Malnutrition Using Anthropometry in Comparison to Clinical Signs. *Am J Trop Med Hyg* 2020; **102**(2): 427-35.

121. Keitel K, Samaka J, Masimba J, et al. Safety and Efficacy of C-reactive Protein-guided Antibiotic Use to Treat Acute Respiratory Infections in Tanzanian Children: A Planned Subgroup Analysis of a Randomized Controlled Noninferiority Trial Evaluating a Novel Electronic Clinical Decision Algorithm (ePOCT). *Clin Infect Dis* 2019; **69**(11): 1926-34.

122. Snavely ME, Maze MJ, Muiruri C, et al. Sociocultural and health system factors associated with mortality among febrile inpatients in Tanzania: a prospective social biopsy cohort study. *BMJ Glob Health* 2018; **3**(1): e000507.

123. Hansoti B, Jenson A, Keefe D, et al. Reliability and validity of pediatric triage tools evaluated in Low resource settings: a systematic review. *BMC Pediatr* 2017; **17**(1): 37.

124. Fitzgerald E, Mlotha-Mitole R, Ciccone EJ, et al. A pediatric death audit in a large referral hospital in Malawi. *BMC Pediatr* 2018; **18**(1): 75.

125. Fung JST, Akech S, Kissoon N, Wiens MO, English M, Ansermino JM. 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.

126. World Health Organization. Integrated Community Case Management. Geneva, Switzerland; 2012.

127. Li E, Mawji A, Akech S, et al. Pediatric Sepsis Predictors Standardization Working Group Guidelines for the standardized collection of predictor variables in studies for pediatric sepsis. Scholars Portal Dataverse, 2020.

128. Wooldridge G, Murthy S, Kissoon N. Core outcome set in paediatric sepsis in low- and middle-income countries: a study protocol. *BMJ Open* 2020; **10**(4): e034960.

129. Slatnick LR, Thornhill D, Deakyne Davies SJ, et al. Disseminated Intravascular Coagulation Is an Independent Predictor of Adverse Outcomes in Children in the Emergency Department with Suspected Sepsis. *J Pediatr* 2020; **225**: 198-206 e2.

130. Romaine ST, Sefton G, Lim E, et al. Performance of seven different paediatric early warning scores to predict critical care admission in febrile children presenting to the emergency department: a retrospective cohort study. *BMJ Open* 2021; **11**(5): e044091.

131. Lenihan RAF, Ang J, Pallmann P, et al. Mid-Regional Pro-Adrenomedullin in Combination With Pediatric Early Warning Scores for Risk Stratification of Febrile Children Presenting to the Emergency Department: Secondary Analysis of a Nonprespecified United Kingdom Cohort Study. *Pediatr Crit Care Med* 2022; **23**(12): 980-9.

132. Hagedoorn NN, Zachariasse JM, Borensztajn D, et al. Shock Index in the early assessment of febrile children at the emergency department: a prospective multicentre study. *Arch Dis Child* 2022; **107**(2): 116-22.

133. Romaine S.T, Potter J, Khanijau A, et al. Accuracy of a Modified qSOFA Score for Predicting Critical Care Admission in Febrile Children. *Pediatrics* 2020; **146**(4): e20200782.

134. Leligdowicz A, Conroy AL, Hawkes M, et al. Risk-stratification of febrile African children at risk of sepsis using sTREM-1 as basis for a rapid triage test. *Nat Commun* 2021; **12**(1): 6832.

135. Moxon CA, Chisala NV, Mzikamanda R, et al. Laboratory evidence of disseminated intravascular coagulation is associated with a fatal outcome in children with cerebral malaria despite an absence of clinically evident thrombosis or bleeding. *J Thromb Haemost* 2015; **13**(9): 1653-64.

136. Bigio J, MacLean E, Vasquez NA, et al. Most common reasons for primary care visits in low- and middle-income countries: A systematic review. *PLOS Glob Pub Health* 2022; **2**(5): e0000196.

137. Finley CR, Chan DS, Garrison S, et al. What are the most common conditions in primary care? *Can Fam Phys* 2018; **64**(11): 832-40.

138. Izudi J, Anyigu S, Ndungutse D. Adherence to Integrated Management of Childhood Illnesses Guideline in Treating South Sudanese Children with Cough or Difficulty in Breathing. *Int J Pediatr* 2017: 5173416.

139. Lodha R, Kabra SK, Pandey RM. Antibiotics for community-acquired pneumonia in children. *Cochrane Database Syst Rev* 2013; (6): CD004874.

140. Wilkes C, Graham H, Walker P, Duke T, group ARIR. Which children with chest-indrawing pneumonia can be safely treated at home, and under what conditions is it safe to do so? A systematic review of evidence from low- and middle-income countries. *J Glob Health* 2022; **12**: 10008.

141. Zhang S, Sammon PM, King I, et al. Cost of management of severe pneumonia in young children: systematic analysis. *J Glob Health* 2016; **6**(1): 010408.

142. McAllister DA, Liu L, Shi T, et al. Global, regional, and national estimates of pneumonia morbidity and mortality in children younger than 5 years between 2000 and 2015: a systematic analysis. *Lancet Glob Health* 2019; **7**(1): e47-e57.

143. Marangu D, Zar HJ. Childhood pneumonia in low-and-middle-income countries: An update. *Paediatr Respir Rev* 2019; **32**: 3-9.

144. Zhang R, Rai U, Ibrahim NBM, et al. Dysregulation of angiopoietin-Tie-2 axis in ugandan children hospitalized with pneumonia. *Cytokine* 2020; **133**: 155175.

145. Jullien S, Richard-Greenblatt M, Ngai M, et al. Performance of host-response biomarkers to risk-stratify children with pneumonia in Bhutan. *J Infect* 2022; **85**(6): 634-43.

146. McDonald CR, Leligdowicz A, Conroy AL, et al. Immune and endothelial activation markers and risk stratification of childhood pneumonia in Uganda: A secondary analysis of a prospective cohort study. *PLoS Med* 2022; **19**(7): e1004057.

147. Balanza N, Erice C, Ngai M, et al. Prognostic accuracy of biomarkers of immune and endothelial activation in Mozambican children hospitalized with pneumonia. *PLOS Glob Pub Health* 2023; **3**(2): e0001553.

148. Gutbier B, Neuhauss AK, Reppe K, et al. Prognostic and Pathogenic Role of Angiopoietin-1 and -2 in Pneumonia. *Am J Respir Crit Care Med* 2018; **198**(2): 220-31.

149. Balanza N, Erice C, Ngai M, Varo R, Kain KC, Bassat Q. Host-Based Prognostic Biomarkers to Improve Risk Stratification and Outcome of Febrile Children in Low- and Middle-Income Countries. *Front Pediatr* 2020; **8**: 552083.

150. Charakida M, Donald AE, Terese M, et al. Endothelial dysfunction in childhood infection. *Circulation* 2005; **111**(13): 1660-5.

151. Turner C, Turner P, Carrara V, et al. High rates of pneumonia in children under two years of age in a South East Asian refugee population. *PLoS One* 2013; **8**(1): e54026.

152. World Health Organization. Integrated Management of Childhood Illnesses. Geneva, Switzerland, 2005.

153. Myatt M, Guevarra E. zscorer: Child Anthropometry z-Score Calculator. R package version 0.3.1. 2019.

154. 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).

155. 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**(8): 743-8.

156. Olson D, Davis NL, Milazi R, et al. Development of a severity of illness scoring system (inpatient triage, assessment and treatment) for resource-constrained hospitals in developing countries. *Trop Med Int Health* 2013; **18**(7): 871-8.

157. 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**(12): e0168126.

158. Reed C, Madhi SA, Klugman KP, et al. Development of the Respiratory Index of Severity in Children (RISC) score among young children with respiratory infections in South Africa. *PLoS One* 2012; **7**(1): e27793.

159. Egdell P, Finlay L, Pedley DK. The PAWS score: validation of an early warning scoring system for the initial assessment of children in the emergency department. *Emerg Med J* 2008; **25**(11): 745-9.

160. Emukule GO, McMorrow M, Ulloa C, et al. Predicting mortality among hospitalized children with respiratory illness in Western Kenya, 2009-2012. *PLoS One* 2014; **9**(3): e92968.

161. Schlapbach LJ, Straney L, Bellomo R, MacLaren G, Pilcher D. Prognostic accuracy of age-adapted SOFA, SIRS, PELOD-2, and qSOFA for in-hospital mortality among children with suspected infection admitted to the intensive care unit. *Intensive Care Med* 2018; **44**(2): 179-88.

162. Brierley J, Carcillo JA, Choong K, et al. Clinical practice parameters for hemodynamic support of pediatric and neonatal septic shock: 2007 update from the American College of Critical Care Medicine. *Crit Care Med* 2009; **37**(2): 666-88.

163. Beane A, Silva AP, Munasinghe S, et al. Comparison of Quick Sequential Organ Failure Assessment and Modified Systemic Inflammatory Response Syndrome Criteria in a Lower Middle Income Setting. *J Acute Med* 2017; **7**(4): 141-8.

164. Mankhambo LA, Banda DL, Group IPDS, et al. The role of angiogenic factors in predicting clinical outcome in severe bacterial infection in Malawian children. *Crit Care* 2010; **14**(3): R91.

165. Parikh SM. Dysregulation of the angiopoietin-Tie-2 axis in sepsis and ARDS. *Virulence* 2013; **4**(6): 517-24.

166. Fuchs A, Gotta V, Decker ML, et al. Cytokine kinetic profiles in children with acute lower respiratory tract infection: a post hoc descriptive analysis from a randomized control trial. *Clin Microbiol Infect* 2018; **24**(12): 1341 e1- e7.

167. Florin TA, Ambroggio L, Brokamp C, et al. Biomarkers and Disease Severity in Children With Community-Acquired Pneumonia. *Pediatrics* 2020; **145**(6): e20193728.

168. Carrol ED, Thomson APJ, Mobbs KJ. The Role of RANTES in Meningococcal Disease. *J Infect Dis* 2000; **182**(1): 363-6.

169. Zhao Y, Qin L, Zhang P, et al. Longitudinal COVID-19 profiling associates IL-1RA and IL-10 with disease severity and RANTES with mild disease. *JCI Insight* 2020; **5**(13): e139834.

170. Singla M, Kar M, Sethi T, et al. Immune Response to Dengue Virus Infection in Pediatric Patients in New Delhi, India--Association of Viremia, Inflammatory Mediators and Monocytes with Disease Severity. *PLoS Negl Trop Dis* 2016; **10**(3): e0004497.

171. Van Singer M, Brahier T, Ngai M, et al. COVID-19 risk stratification algorithms based on sTREM-1 and IL-6 in emergency department. *J Allergy Clin Immunol* 2021; **147**(1): 99-106 e4.

172. Chen HL, Hung CH, Tseng HI, Yang RC. Soluble form of triggering receptor expressed on myeloid cells-1 (sTREM-1) as a diagnostic marker of serious bacterial infection in febrile infants less than three months of age. *Jpn J Infect Dis* 2008; **61**(1): 31-5.

173. Carrol ED, Newland P, Riordan FAI, Thomson APJ, Curtis N, Hart CA. Procalcitonin as a diagnostic marker of meningococcal disease in children presenting with fever and a rash. *Arch Dis Child* 2002; **86**: 282-5.

174. Pourakbari B, Mamishi S, Zafari J, et al. Evaluation of procalcitonin and neopterin level in serum of patients with acute bacterial infection. *Braz J Infect Dis* 2010; **14**(3): 252-5.

175. Leligdowicz A, Conroy AL, Hawkes M, et al. Validation of two multiplex platforms to quantify circulating markers of inflammation and endothelial injury in severe infection. *PLoS One* 2017; **12**(4): e0175130.

176. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software* 2011; **45**(3): 1-67.

177. Harrell FE, Jr. rms: Regression Modeling Strategies. R package version 6.2-0. 2021.

178. UK National Institute for Health and Care Excellence. Algorithm for managing suspected sepsis in children aged under 5 years outside an acute hospital setting. United Kingdom; 2017.

179. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software* 2010; **33**(1): 1-22.

180. Sjoberg DD. dcurves: Decision Curve Analysis for Model Evaluation. R package version 0.3.0. 2022.

181. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics* 1988; **44**(3): 837-45.

182. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 2011; **12**(77).

183. Fox J. polycor: Polychoric and Polyserial Correlations. R package version 0.8-1. 2022.

184. Therneau T, Atkinson B. rpart: Recursive Partitioning and Regression Trees. 2019.

185. Vergouwe Y, Steyerberg EW, Eijkemans MJ, Habbema JD. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. *J Clin Epidemiol* 2005; **58**(5): 475-83.

186. Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model. *BMJ* 2020; **368**: m441.

187. Ensor J, Martin EC, Riley RD. pmsampsize: Calculates the Minimum Sample Size Required for Developing a Multivariable Prediction Model. R package version 1.1.1. R; 2021.

188. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for Reporting Observational Studies. *PLoS Med* 2007; **4**(10): e296.

189. Bonafide CP, Brady PW, Keren R, Conway PH, Marsolo K, Daymont C. Development of Heart and Respiratory Rate Percentile Curves for Hospitalized Children. *Pediatrics* 2013; **131**(4): e1150-7.

190. World Health Organization. Interagency List of Priority Medical Devices for Essential Interventions for Reproductive, Maternal, Newborn and Child Health. Geneva, Switzerland, 2016.

191. Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *BMJ* 2016; **352**: i6.

192. Wright SW, Hantrakun V, Rudd KE, et al. Enhanced bedside mortality prediction combining point-of-care lactate and the quick Sequential Organ Failure Assessment (qSOFA) score in patients hospitalised with suspected infection in southeast Asia: a cohort study. *Lancet Glob Health* 2022; **10**(9): e1281-e8.

193. Chandna A, Richard-Greenblatt M, Tustin R, et al. Practical Methods to Permit the Analysis of Host Biomarkers in Resource-Limited Settings. *Am J Trop Med Hyg* 2022; **106**(6): 1765-9.

194. Eun S, Kim H, Kim HY, et al. Age-adjusted quick Sequential Organ Failure Assessment score for predicting mortality and disease severity in children with infection: a systematic review and meta-analysis. *Sci Rep* 2021; **11**(1): 21699.

195. Spruijt B, Vergouwe Y, Nijman RG, Thompson M, Oostenbrink R. Vital signs should be maintained as continuous variables when predicting bacterial infections in febrile children. *J Clin Epidemiol* 2013; **66**(4): 453-7.

196. Hakanpaa L, Sipila T, Leppanen VM, et al. Endothelial destabilization by angiopoietin-2 via integrin beta1 activation. *Nat Commun* 2015; **6**: 5962.

197. Chandna A, Mahajan R, Gautam P, et al. Host biomarkers reflect prognosis in patients presenting with moderate Covid-19 – a prospective cohort study *Open Forum Infect Dis* 2022; **6**(9): ofac526.

198. Steyerberg EW. Clinical Prediction Models: Springer International Publishing; 2019.

199. Harrell FE, Jr. Regression Modeling Strategies: Springer International Publishing; 2006.

200. Chandna A, Osborn J, Bassat Q, et al. Anticipating the future: prognostic tools as a complementary strategy to improve care for patients with febrile illnesses in resource-limited settings. *BMJ Glob Health* 2021; **6**(7): e006057.

201. Chandna A, Aderie EM, Ahmad R, et al. Prediction of disease severity in young children presenting with acute febrile illness in resource-limited settings: a protocol for a prospective observational study. *BMJ Open* 2021; **11**(1): e045826.

202. 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.

203. 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**(12): 1001-2.

204. 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**(12): 893-902.

205. Unitaid. The ACT-A Oxygen Emergency Taskforce…one year on. 2022. https://unitaid.org/news-blog/the-act-a-oxygen-emergency-taskforceone-year-on (accessed 15 May 2023).

206. Scott JA, Wonodi C, Moisi JC, et al. The definition of pneumonia, the assessment of severity, and clinical standardization in the Pneumonia Etiology Research for Child Health study. *Clin Infect Dis* 2012; **54 Suppl 2**(Suppl 2): S109-16.

207. Erdman LK, Dhabangi A, Musoke C, et al. Combinations of host biomarkers predict mortality among Ugandan children with severe malaria: a retrospective case-control study. *PLoS One* 2011; **6**(2): e17440.

208. Turner EL, Nielsen KR, Jamal SM, von Saint Andre-von Arnim A, Musa NL. A Review of Pediatric Critical Care in Resource-Limited Settings: A Look at Past, Present, and Future Directions. *Front Pediatr* 2016; **4**: 5.

209. Losonczy LI, Papali A, Kivlehan S, et al. White Paper on Early Critical Care Services in Low Resource Settings. *Ann Glob Health* 2021; **87**(1): 105.

210. Slusher TM, Kiragu AW, Day LT, et al. Pediatric Critical Care in Resource-Limited Settings-Overview and Lessons Learned. *Front Pediatr* 2018; **6**: 49.

211. Duke T, Wandi F, Jonathan M, et al. Improved oxygen systems for childhood pneumonia: a multihospital effectiveness study in Papua New Guinea. *Lancet* 2008; **372**(9646): 1328-33.

212. Tripathi S, Kaur H, Kashyap R, Dong Y, Gajic O, Murthy S. A survey on the resources and practices in pediatric critical care of resource-rich and resource-limited countries. *J Intensive Care* 2015; **3**: 40.

213. Argent AC, Chisti MJ, Ranjit S. What's new in PICU in resource limited settings? *Intensive Care Med* 2018; **44**(4): 467-9.

214. Riviello ED, Letchford S, Achieng L, Newton MW. Critical care in resource-poor settings: lessons learned and future directions. *Crit Care Med* 2011; **39**(4): 860-7.

215. 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.

216. Marshall JC, Bosco L, Adhikari NK, et al. What is an intensive care unit? A report of the task force of the World Federation of Societies of Intensive and Critical Care Medicine. *J Crit Care* 2017; **37**: 270-6.

217. Murthy S, Leligdowicz A, Adhikari NK. Intensive care unit capacity in low-income countries: a systematic review. *PLoS One* 2015; **10**(1): e0116949.

218. Pollack MM, Holubkov R, Funai T, et al. The Pediatric Risk of Mortality Score: Update 2015. *Pediatr Crit Care Med* 2016; **17**(1): 2-9.

219. Straney L, Clements A, Parslow RC, et al. Paediatric index of mortality 3: an updated model for predicting mortality in pediatric intensive care*. *Pediatr Crit Care Med* 2013; **14**(7): 673-81.

220. Diaz JV, Riviello ED, Papali A, Adhikari NKJ, Ferreira JC. Global Critical Care: Moving Forward in Resource-Limited Settings. *Ann Glob Health* 2019; **85**(1): 3.

221. Wheeler I, Price C, Sitch A, et al. Early warning scores generated in developed healthcare settings are not sufficient at predicting early mortality in Blantyre, Malawi: a prospective cohort study. *PLoS One* 2013; **8**(3): e59830.

222. George EC. Good-quality research: a vital step in improving outcomes in paediatric intensive care units in low- and middle-income countries. *Paediatr Int Child Health* 2017; **37**(2): 79-81.

223. Seifu A, Eshetu O, Tafesse D, Hailu S. Admission pattern, treatment outcomes, and associated factors for children admitted to pediatric intensive care unit of Tikur Anbessa specialized hospital, 2021: a retrospective cross-sectional study. *BMC Anesthesiol* 2022; **22**(1): 13.

224. Rusmawatiningtyas D, Rahmawati A, Makrufardi F, et al. Factors associated with mortality of pediatric sepsis patients at the pediatric intensive care unit in a low-resource setting. *BMC Pediatr* 2021; **21**(1): 471.

225. Nyirasafari R, Corden MH, Karambizi AC, et al. Predictors of mortality in a paediatric intensive care unit in Kigali, Rwanda. *Paediatr Int Child Health* 2017; **37**(2): 109-15.

226. Teshager NW, Amare AT, Tamirat KS. Incidence and predictors of mortality among children admitted to the pediatric intensive care unit at the University of Gondar comprehensive specialised hospital, northwest Ethiopia: a prospective observational cohort study. *BMJ Open* 2020; **10**(10): e036746.

227. Frankel LR, Hsu BS, Yeh TS, et al. Criteria for Critical Care Infants and Children: PICU Admission, Discharge, and Triage Practice Statement and Levels of Care Guidance. *Pediatr Crit Care Med* 2019; **20**(9): 847-87.

228. Ogero M, Sarguta RJ, Malla L, et al. Prognostic models for predicting in-hospital paediatric mortality in resource-limited countries: a systematic review. *BMJ Open* 2020; **10**(10): e035045.

229. World Health Organization. Haemoglobin concentrations for the diagnosis of anaemia and assessment of severity. Vitamin and Mineral Nutrition Information System. Geneva, 2011.

230. National Clinical Effectiveness Committee and Department of Health. National Clinical Guideline on the Irish Paediatric Early Warning System (PEWS), 2016.

231. Rosman SL, Karangwa V, Law M, Monuteaux MC, Briscoe CD, McCall N. Provisional Validation of a Pediatric Early Warning Score for Resource-Limited Settings. *Pediatrics* 2019; **143**(5): e20183657.

232. Pollack MM, Dean JM, Butler J, et al. The ideal time interval for critical care severity-of-illness assessment. *Pediatr Crit Care Med* 2013; **14**(5): 448-53.

233. Feudtner C, Christakis DA, Connell FA. Pediatric Deaths Attributable to Complex Chronic Conditions: A Population-Based Study of Washington State, 1980–1997. *Pediatrics* 2000; **106**: 205-9.

234. de Onis M, Onyango AW, Borghi E, Siyam A, Nishida C, Siekmann J. Development of a WHO growth reference for school-aged children and adolescents. *Bull World Health Organ* 2007; **85**(9): 660-7.

235. WHO Multicentre Growth Reference Study Group. WHO Child Growth Standards based on length/height, weight and age. *Acta Paediatr* 2006: 76-85.

236. Kuczmarski RJ, Ogden CL, Guo SS, et al. 2000 CDC growth charts for the United States: Methods and development. *Vital Health Stat* 2002; **11**: 1-190.

237. Tierney N, Cook D. Expanding Tidy Data Principles to Facilitate Missing Data Exploration, Visualization and Assessment of Imputations. *Journal of Statistical Software* 2023; **105**(7): 1-31.

238. Cule E, Moritz S, Frankowski D. ridge: Ridge Regression with Automatic Selection of the Penalty Parameter. 2022.

239. Du Z, Hao Y. reportROC: An Easy Way to Report ROC Analysis. 2022.

240. Celiker MY, Pagnarith Y, Akao K, Sophearin D, Sorn S. Pediatric Palliative Care Initiative in Cambodia. *Front Public Health* 2017; **5**: 185.

241. Hodkinson P, Argent A, Wallis L, et al. Pathways to Care for Critically Ill or Injured Children: A Cohort Study from First Presentation to Healthcare Services through to Admission to Intensive Care or Death. *PLoS One* 2016; **11**(1): e0145473.

242. Wiens MO, Kumbakumba E, Larson CP, et al. Postdischarge mortality in children with acute infectious diseases: derivation of postdischarge mortality prediction models. *BMJ Open* 2015; **5**(11): e009449.

243. Chisti MJ, Graham SM, Duke T, et al. Post-discharge mortality in children with severe malnutrition and pneumonia in Bangladesh. *PLoS One* 2014; **9**(9): e107663.

244. Muenchhoff M, Goulder PJ. Sex differences in pediatric infectious diseases. *J Infect Dis* 2014; **209**: S120-6.

245. Calu Costa J, Wehrmeister FC, Barros AJ, Victora CG. Gender bias in careseeking practices in 57 low- and middle-income countries. *J Glob Health* 2017; **7**(1): 010418.

246. World Health Organization. Pneumonia of unknown cause – China. Geneva, 2020.

247. Mathieu E, Ritchie H, Rodes-Guirao L, et al. Coronavirus Pandemic (COVID-19). 2020. https://ourworldindata.org/coronavirus (accessed 22 April 2023).

248. Nakkazi E. Oxygen supplies and COVID-19 mortality in Africa. *Lancet Respir Med* 2021; **9**(4): e39.

249. Aziz S, Arabi YM, Alhazzani W, et al. Managing ICU surge during the COVID-19 crisis: rapid guidelines. *Intensive Care Med* 2020; **46**(7): 1303-25.

250. RECOVERY Collaborative Group. Tocilizumab in patients admitted to hospital with COVID-19 (RECOVERY): a randomised, controlled, open-label, platform trial. *Lancet* 2021; **397**(10285): 1637-45.

251. RECOVERY Collaborative Group, Horby P, Lim WS, et al. Dexamethasone in Hospitalized Patients with Covid-19. *N Engl J Med* 2021; **384**(8): 693-704.

252. Jara A, Undurraga EA, Gonzalez C, et al. Effectiveness of an Inactivated SARS-CoV-2 Vaccine in Chile. *N Engl J Med* 2021.

253. Voysey M, Clemens SAC, Madhi SA, et al. Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. *Lancet* 2021; **397**(10269): 99-111.

254. Government of the United Kingdom. Supporting our high streets after COVID-19. London; 2021.

255. Government of Canada. Government of Canada announces easing of border measures for fully vaccinated travellers. 2021.

256. Batista C, Hotez P, Amor YB, et al. The silent and dangerous inequity around access to COVID-19 testing: A call to action. *EClinicalMedicine* 2022; **43**: 101230.

257. Medecins Sans Frontieres. Lessons learned and recommendations to ensure access to vaccines, medicines and diagnostics in the COVID-19 pandemic response and beyond., 2021.

258. Tan MKI. COVID-19 in an inequitable world: the last, the lost and the least. *Int Health* 2021; **13**(6): 493-6.

259. Kapoor G, Hauck S, Sriram A, et al. State-wise estimates of current hospital beds, intensive care unit (ICU) beds and ventilators in India: Are we repared for a surge in COVID-19 hospitalizations? *medRxiv* 2020.

260. World Health Organization. Update on Omicron. Geneva; 2021.

261. Truelove S, Abrahim O, Altare C, et al. The potential impact of COVID-19 in refugee camps in Bangladesh and beyond: A modeling study. *PLoS Med* 2020; **17**(6): e1003144.

262. Ellis-Peterson H. 'The system has collapsed': India's descent into Covid hell. The Guardian. 2021.

263. Medecins Sans Frontieres. COVID-19 leaves Amazonas health system saturated, overloaded and struggling. 2021. https://www.msf.org/coronavirus-covid-19-collapses-health-system-manaus-brazil (accessed 8 November 2021).

264. World Health Organization. COVID-19 oxygen emergency impacting more than half a million people in low- and middle-income countries every day, as demand surges. 2021. https://www.who.int/news/item/25-02-

2021-covid-19-oxygen-emergency-impacting-more-than-half-a-million-people-in-low--and-middle-income-countries-every-day-as-demand-surges (accessed 8 November 2021).

265. World Health Organization. COVID-19 Clinical Management - Living Guidance 25 January 2021. 2021.

266. Barros LM, Pigoga JL, Chea S, et al. Pragmatic Recommendations for Identification and Triage of Patients with COVID-19 Disease in Low- and Middle-Income Countries. *Am J Trop Med Hyg* 2021; **104**: 3-11.

267. Wynants L, Van Calster B, Collins GS, et al. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. *BMJ* 2020; **369**: m1328.

268. Schultz MJ, Gebremariam TH, Park C, et al. Pragmatic Recommendations for the Use of Diagnostic Testing and Prognostic Models in Hospitalized Patients with Severe COVID-19 in Low- and Middle-Income Countries. *Am J Trop Med Hyg* 2021; **104**: 34-47.

269. Prieto-Alhambra D, Ballo E, Coma E, et al. Filling the gaps in the characterization of the clinical management of COVID-19: 30-day hospital admission and fatality rates in a cohort of 118 150 cases diagnosed in outpatient settings in Spain. *Int J Epidemiol* 2021; **49**(6): 1930-9.

270. Ji D, Zhang D, Xu J, et al. Prediction for Progression Risk in Patients With COVID-19 Pneumonia: The CALL Score. *Clin Infect Dis* 2020; **71**(6): 1393-9.

271. Wei W, Sivapalasingam S, Mellis S, Geba GP, Jalbert JJ. A Retrospective Study of COVID-19-Related Urgent Medical Visits and Hospitalizations After Outpatient COVID-19 Diagnosis in the US. *Adv Ther* 2021; **38**(6): 3185-202.

272. Chang MC, Park YK, Kim BO, Park D. Risk factors for disease progression in COVID-19 patients. *BMC Infect Dis* 2020; **20**(1): 445.

273. Guo Y, Liu Y, Lu J, et al. Development and validation of an early warning score (EWAS) for predicting clinical deterioration in patients with coronavirus disease 2019. *medRxiv* 2020.

274. Higuchi T, Nishida T, Iwahashi H, et al. Early clinical factors predicting the development of critical disease in Japanese patients with COVID-19: A single-center, retrospective, observational study. *J Med Virol* 2021; **93**(4): 2141-8.

275. Shi J, Li Y, Zhou X, et al. Lactate dehydrogenase and susceptibility to deterioration of mild COVID-19 patients: a multicenter nested case-control study. *BMC Med* 2020; **18**(1): 168.

276. Sun Q, Qiu H, Huang M, Yang Y. Lower mortality of COVID-19 by early recognition and intervention: experience from Jiangsu Province. *Ann Intensive Care* 2020; **10**(1): 33.

277. Smadja DM, Guerin CL, Chocron R, et al. Angiopoietin-2 as a marker of endothelial activation is a good predictor factor for intensive care unit admission of COVID-19 patients. *Angiogenesis* 2020; **23**(4): 611-20.

278. Rovina N, Akinosoglou K, Eugen-Olsen J, Hayek S, Reiser J, Giamarellos-Bourboulis EJ. Soluble urokinase plasminogen activator receptor (suPAR) as an early predictor of severe respiratory failure in patients with COVID-19 pneumonia. *Crit Care* 2020; **24**(1): 187.

279. Yang Y, Shen C, Li J, et al. Plasma IP-10 and MCP-3 levels are highly associated with disease severity and predict the progression of COVID-19. *J Allergy Clin Immunol* 2020; **146**(1): 119-27 e4.

280. Del Valle DM, Kim-Schulze S, Huang HH, et al. An inflammatory cytokine signature predicts COVID-19 severity and survival. *Nat Med* 2020; **26**(10): 1636-43.

281. Cheng B, Hu J, Zuo X, et al. Predictors of progression from moderate to severe coronavirus disease 2019: a retrospective cohort. *Clin Microbiol Infect* 2020; **26**(10): 1400-5.

282. Arnold DT, Attwood M, Barratt S, et al. Predicting outcomes of COVID-19 from admission biomarkers: a prospective UK cohort study. *Emerg Med J* 2021; **38**(7): 543-8.

283. Dittrich S, Tadesse BT, Moussy F, et al. Target Product Profile for a Diagnostic Assay to Differentiate between Bacterial and Non-Bacterial Infections and Reduce Antimicrobial Overuse in Resource-Limited Settings: An Expert Consensus. *PLoS One* 2016; **11**(8): e0161721.

284. Ni W, Han Y, Zhao J, et al. Serum soluble urokinase-type plasminogen activator receptor as a biological marker of bacterial infection in adults: a systematic review and meta-analysis. *Sci Rep* 2016; **6**: 39481.

285. Wittenhagen P, Kronborg G, Weis N, et al. The plasma level of soluble urokinase receptor is elevated in patients with Streptococcus pneumoniae bacteraemia and predicts mortality. *Clin Microbiol Infect* 2004; **10**(5): 409-15.

286. Rabna P, Andersen A, Wejse C, et al. Utility of the plasma level of suPAR in monitoring risk of mortality during TB treatment. *PLoS One* 2012; **7**(8): e43933.

287. Luo Q, Ning P, Zheng Y, Shang Y, Zhou B, Gao Z. Serum suPAR and syndecan-4 levels predict severity of community-acquired pneumonia: a prospective, multi-centre study. *Crit Care* 2018; **22**(1): 15.

288. Pregernig A, Muller M, Held U, Beck-Schimmer B. Prediction of mortality in adult patients with sepsis using six biomarkers: a systematic review and meta-analysis. *Ann Intensive Care* 2019; **9**(1): 125.

289. Marshall JC, Murthy S, Diaz J, et al. A minimal common outcome measure set for COVID-19 clinical research. *Lancet Infect Dis* 2020; **20**(8): e192-e7.

290. Wayne MT, Weng W, O'Malley M, et al. Variation in COVID-19 disease severity at hospital admission over time and across hospitals: A multi-institution cohort of Michigan hospitals. *Medicine (Baltimore)* 2021; **100**(37): e27265.

291. Lambden S, Laterre PF, Levy MM, Francois B. The SOFA score-development, utility and challenges of accurate assessment in clinical trials. *Crit Care* 2019; **23**(1): 374.

292. Grissom CK, Brown SM, Kuttler KG, et al. A modified sequential organ failure assessment score for critical care triage. *Disaster Med Public Health Prep* 2010; **4**(4): 277-84.

293. European Union. Technology readiness levels (TRL); Extract from Part 19 - Commission Decision C(2014)4995. 2014.

294. Luo X, Zhou W, Yan X, et al. Prognostic Value of C-Reactive Protein in Patients With Coronavirus 2019. *Clin Infect Dis* 2020; **71**(16): 2174-9.

295. Wang G, Wu C, Zhang Q, et al. C-Reactive Protein Level May Predict the Risk of COVID-19 Aggravation. *Open Forum Infect Dis* 2020; **7**(5): ofaa153.

296. Herold T, Jurinovic V, Arnreich C, et al. Elevated levels of IL-6 and CRP predict the need for mechanical ventilation in COVID-19. *J Allergy Clin Immunol* 2020; **146**(1): 128-36 e4.

297. Zhang L, Yan X, Fan Q, et al. D-dimer levels on admission to predict in-hospital mortality in patients with Covid-19. *J Thromb Haemost* 2020; **18**(6): 1324-9.

298. Thachil J, Tang N, Gando S, et al. ISTH interim guidance on recognition and management of coagulopathy in COVID-19. *J Thromb Haemost* 2020; **18**(5): 1023-6.

299. Wu C, Chen X, Cai Y, et al. Risk Factors Associated With Acute Respiratory Distress Syndrome and Death in Patients With Coronavirus Disease 2019 Pneumonia in Wuhan, China. *JAMA Intern Med* 2020.

300. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* 2020; **395**(10229): 1054-62.

301. Aziz M, Fatima R, Assaly R. Elevated interleukin-6 and severe COVID-19: A meta-analysis. *J Med Virol* 2020; **92**(11): 2283-5.

302. Liu F, Li L, Xu M, et al. Prognostic value of interleukin-6, C-reactive protein, and procalcitonin in patients with COVID-19. *J Clin Virol* 2020; **127**: 104370.

303. Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020; **395**(10223): 497-506.

304. Lippi G, Plebani M. Procalcitonin in patients with severe coronavirus disease 2019 (COVID-19): A meta-analysis. *Clin Chim Acta* 2020; **505**: 190-1.

305. Wang D, Hu B, Hu C, et al. Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. *JAMA* 2020; **323**(11): 1061-9.

306. Eugen-Olsen J, Altintas I, Tingleff J, et al. Low levels of the prognostic biomarker suPAR are predictive of mild outcome in patients with symptoms of COVID-19 - a prospective cohort study. *medRxiv* 2020.

307. Virogates A/S Denmark. Instructions for use: suPARnostic Quick Triage for aLF Reader. Birkerød, Denmark; 2020.

308. Virogates A/S Denmark. Instructions for use: Enzyme immunoassay for quantitative determination of soluble urokinase plasminogen activator receptor in human plasma and serum. Birkerød, Denmark; 2020.

309. Gupta RK, Harrison EM, Ho A, et al. Development and validation of the ISARIC 4C Deterioration model for adults hospitalised with COVID-19: a prospective cohort study. *Lancet Respir Med* 2021; **9**(4): 349-59.

310. Ambler G, Benner A. mfp: Multivariable fractional polynomials. 2022.

311. Bland M, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; **327**(8476): 307-10.

312. Datta D. blandr: a Bland-Altman Method Comparison package for R. 2017.

313. Clinical and Laboratory Standards Institute. Measurement Procedure Comparison and Bias Estimation Using Patient Samples; Approved Guideline —Third Edition (EP09-A3). Philadelphia, USA, 2013.

314. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ* 2015; **351**: h5527.

315. International Severe Acute Respiratory Infection Consoritum (ISARIC). Clinical Data Collection – The COVID-19 Case Report Forms (CRFs). 2021. https://isaric.org/research/covid-19-clinical-research-resources/covid-19-crf/.

316. Webb BJ, Peltan ID, Jensen P, et al. Clinical criteria for COVID-19-associated hyperinflammatory syndrome: a cohort study. *Lancet Rheum* 2020; **2**(12): e754-e63.

317. Kyriazopoulou E, Poulakou G, Milionis H, et al. Early treatment of COVID-19 with anakinra guided by soluble urokinase plasminogen receptor plasma levels: a double-blind, randomized controlled phase 3 trial. *Nat Med* 2021; **27**(10): 1752-60.

318. Pai M, Schumacher SG, Abimbola S. Surrogate endpoints in global health research: still searching for killer apps and silver bullets? *BMJ Glob Health* 2018; **3**(2): e000755.

319. Stauning MA, Altintas I, Kallemose T, et al. Soluble Urokinase Plasminogen Activator Receptor as a Decision Marker for Early Discharge of Patients with COVID-19 Symptoms in the Emergency Department. *J Emerg Med* 2021; **61**(3): 298-313.

320. Chalkias A, Skoulakis A, Papagiannakis N, et al. Circulating suPAR associates with severity and in-hospital progression of COVID-19. *Eur J Clin Invest* 2022; **52**(7): e13794.

321. Winnicki W, Sunder-Plassmann G, Sengolge G, et al. Diagnostic and Prognostic Value of Soluble Urokinase-type Plasminogen Activator Receptor (suPAR) in Focal Segmental Glomerulosclerosis and Impact of Detection Method. *Sci Rep* 2019; **9**(1): 13783.

322. Chew-Harris J, Appleby S, Richards AM, Troughton RW, Pemberton CJ. Analytical, biochemical and clearance considerations of soluble urokinase plasminogen activator receptor (suPAR) in healthy individuals. *Clin Biochem* 2019; **69**: 36-44.

323. Kofoed K, Schneider UV, Scheel T, Andersen O, Eugen-Olsen J. Development and validation of a multiplex add-on assay for sepsis biomarkers using xMAP technology. *Clin Chem* 2006; **52**(7): 1284-93.

324. Kumar P, Kakar A, Gogia A, Waziri N. Evaluation of soluble urokinase-type plasminogen activator receptor (suPAR) quick test for triage in the emergency department. *J Family Med Prim Care* 2019; **8**(12): 3871-5.

325. Moons KGM, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? *BMJ* 2009; **338**: b375.

326. World Health Organization. Fifteen African countries hit 10% COVID-19 vaccination goal 2021. https://www.afro.who.int/news/fifteen-african-countries-hit-10-covid-19-vaccination-goal (accessed 8 November 2021).

327. Hogan A. B, Winskill P, Watson O, et al. Report 33: Modelling the allocation and impact of a COVID-19 vaccine. Imperial College London (25-09-2020), 2020.

328. RECOVERY Collaborative Group. Casirivimab and imdevimab in patients admitted to hospital with COVID-19 (RECOVERY): a randomised, controlled, open-label, platform trial. *Lancet* 2022; **399**(10325): 665-76.

329. Sendagire C, Lipnick MS, Kizito S, et al. Feasibility of the modified sequential organ function assessment score in a resource-constrained setting: a prospective observational study. *BMC Anesthesiol* 2017; **17**(1): 12.

330. McHugh LC, Snyder K, Yager TD. The effect of uncertainty in patient classification on diagnostic performance estimations. *PLoS One* 2019; **14**(5): e0217146.

331. RECOVERY Collaborative Group. Convalescent plasma in patients admitted to hospital with COVID-19 (RECOVERY): a randomised controlled, open-label, platform trial. *The Lancet* 2021; **397**(10289): 2049-59.

332. Holtman GA, Berger MY, Burger H, et al. Development of practical recommendations for diagnostic accuracy studies in low-prevalence situations. *J Clin Epidemiol* 2019; **114**: 38-48.

333. Chandna A, Shwe Nwe Htun N, Peto TJ, et al. Defining the burden of febrile illness in rural South and Southeast Asia: an open letter to announce the launch of the Rural Febrile Illness project. *Wellcome Open Research* 2021; **6**: 64.

9 Appendix

9.1 PRISMA checklist: systematic review

Table 9.1-1: PRISMA checklist for systematic review. AUC = area under the receiver operating characteristic curve; CHARMS = CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies; CHARMS-PF = CHARMS for Prognostic Factor studies; LR = likelihood ratio; PRISMA = Preferred Reporting Items for Systematic Review and Meta-Analysis, PROBAST = Prediction model Risk Of Bias ASsessment Tool, QUIPS = QUality In Prognosis Studies; SNOMED-CT = Systematized Nomenclature of Medicine Clinical Terms.

Section/topic	#	Checklist item	Page
TITLE			
Title	1	Predictors of disease severity in febrile children: a systematic review of prognostic studies	NA
ABSTRACT			
Structured summary	2	Background: Early identification of children at risk of severe febrile illness can optimise referral, admission, and treatment decisions. Objectives: Identify prognostic clinical and laboratory factors that predict progression to severe disease in febrile children. Data sources: MEDLINE, Web of Science, and Embase between 31/05/1999 and 30/04/2020. Study eligibility criteria, participants, and interventions: Studies evaluating prognostic factors or clinical prediction models in children presenting with acute febrile illness. Study appraisal and synthesis methods: Unadjusted LRs for prognostic factors and AUCs for prediction models. Risk of bias and applicability assessed using PROBAST and QUIPS tools. Results: 5,949 identified studies, 18 included evaluating 200 prognostic factors and 33 prediction models in 24,530 children. Hypoxia, altered consciousness, and markers of acidosis and poor peripheral perfusion most common prognostic factors. Limitations: Heterogeneity between studies precluded meta-analysis. Applicability and risk of bias concerns identified for many studies. Conclusions and implications of key findings: Prognostic factors identified from range of geographic contexts. Few studies address	NA

Section/topic	#	Checklist item	Page
		this question. Future studies should be multi-centre and include outpatients to explore generalisability and develop tools to guide referral and admission in children presenting from the community. Systematic review registration number: CRD42019140542	
INTRODUCTION			
Rationale	3	Identification of low-incidence serious disease is challenging. Prognostic factors are increasingly integrated into clinical decision-support algorithms but it is not clear which predictors should be considered for inclusion. Most existing reviews have focussed on diagnosis rather than prognosis and used imperfect outcome measures, such as serious bacterial infection.	30-31
Objectives	4	To identify clinical and laboratory factors that predict progression to severe disease in children presenting from the community with acute febrile illness	31
METHODS			
Protocol and registration	5	PROSPERO registration: CRD42019140542	31
Eligibility criteria	6	P = children aged between 28 days and 19 years with acute febrile illness I = Clinical and laboratory prognostic factors – measured alone or as part of a clinical prediction model at the time children presented to care, as the aim was to identify *prognostic* variables C = No comparators were considered due to expected heterogeneity between studies which renders these comparisons prone to bias O = Any prospectively defined objective measure of disease severity ascertained within the 28 days following measurement of the predictors or during hospitalisation S = Children presenting from the community to primary care or hospital outpatient or emergency departments Report characteristics: No language restriction, 31/05/99 to 30/04/20.	31-32
Information sources	7	MEDLINE, Embase, Web of Science Date last searched 05/05/20 Study authors were contacted if data to calculate LRs or AUCs were not available in the manuscript	33-35
Search	8	Please see Table 3.2-1 for detailed search terms for the three databases	34-35
Study selection	9	Title, abstract, and full-text screening were performed independently by two reviewers. Agreement was checked after the first 20 and 250 articles. Discrepancies were resolved by discussion or independent assessment by a third reviewer.	

Section/topic	#	Checklist item	Page
Data collection process	10	Data extraction sheets were developed based on the CHARMS and CHARMS-PF checklists. Data were extracted independently by one reviewer and checked by another. Discrepancies were discussed and resolved between the two reviewers. Seven authors responded to requests for clarifications and six provided additional data not available in the published manuscript.	36
Data items	11	P = Eligibility criteria for each study including age, definition of febrile illness (for example, temperature, duration), and number of participants (sample size). I = All prognostic factors evaluated. Predictors in different studies were harmonised using the SNOMED-CT definitions. The timing of predictor measurement was confirmed to be before outcome ascertainment (prognostic studies). C = No comparators were considered due to expected heterogeneity between studies which renders these comparisons prone to bias O = Definition and type of outcome (for example, 'hard': death, organ dysfunction, organ support, or PICU admission; vs. 'soft': persistence of symptoms or length of stay). The number of participants meeting the outcome and the outcome prevalence. The timing of outcome ascertainment was confirmed to be after predictor measurement (prognostic studies). S = Country, year, level of the health system (primary, secondary, tertiary), and location of recruitment (community, ED/OPD, IPD)	36
Risk of bias in individual studies	12	Risk of bias and applicability of studies were assessed using the QUIPS tool for prognostic factor studies, and PROBAST for studies developing, validating, or updating prediction models. Each study was independently assessed using QUIPS or PROBAST by two reviewers, as well as an independent senior reviewer. The information was used to inform the narrative data synthesis.	37-38
Summary measures	13	Unadjusted LRs for prognostic factors. AUCs for clinical prediction models.	36-37
Synthesis of results	14	Due to expected heterogeneity between studies (as a result of variations in case-mix and baseline risk), few common predictors for comparison, and absence of well-defined subgroups, no formal meta-analysis nor comparison of variability and bias between studies was planned, as these comparisons are recognised as being prone to bias.	37
Risk of bias across studies	22	Due to expected heterogeneity between studies (as a result of variations in case-mix and baseline risk), few common predictors for comparison, and absence of well-defined subgroups, no formal meta-analysis nor comparison of variability and bias between studies was planned, as these comparisons are recognised as being prone to bias.	NA
Additional analysis	23	To contextualise the results, we used the outcome prevalence of individual studies to calculate the pre-test probability, and display positive and negative post-test probabilities on dumbbell plots. This was pre-specified.	37
RESULTS			
Study selection	17	5,947 studies screened, 16 included in the review. Details for exclusion are presented in the flow diagram in Figure 1.	38

Section/topic	#	Checklist item	Page
Study characteristics	18	PICOS characteristics for each study (including the sample size) are tabulated in Table 1 with relevant citations.	39-44
Risk of bias within studies	19	Risk of bias and applicability assessments are summarised for both prognostic factor and clinical prediction model studies in the text and in Fig 6. The supporting information – S5 Appendix – provides granular detail on the different domains in the risk of bias and applicability assessments for each study.	55-59
Results of individual studies	20	A narrative summary of the results of individual studies is provided, categorised according to prognostic factors and clinical prediction models. A forest plot with confidence intervals summarises the AUROCs for clinical prediction models, unadjusted likelihood ratios with confidence intervals (in association with dumbbell plots that display pre- and post-test probabilities – see item 23) are provided for prognostic factors.	45-54
Synthesis of results	21	A narrative synthesis is provided for both prognostic factors and clinical prediction models. No formal meta-analysis nor comparison of variability and bias between studies was planned.	45-54
Risk of bias across studies	22	Due to expected heterogeneity between studies (as a result of variations in case-mix and baseline risk), few common predictors for comparison, and absence of well-defined subgroups, no formal meta-analysis nor comparison of variability and bias between studies was planned, as these comparisons are recognised as being prone to bias.	NA
DISCUSSION			
Summary of evidence	24	Most prognostic factors and clinical prediction models identified in this review indicate children with relatively advanced illnesses – reflecting the fact that most studies included only hospitalised children and focused on predicting mortality. Performance varied across a wide range of geographic contexts. Whilst clinicians can use these predictors to identify children at risk of progressing to serious illness, they must be cognisant not to over interpret individual predictors. Researchers must derive and validate clinical prediction models using appropriate methodology, and populations and outcomes relevant to the clinical problem.	60-63
Limitations	25	Study level: Heterogeneity of studies precluded formal meta-analysis. Applicability of studies was low as they often only included hospitalised children. Risk of bias was high, predominantly due to inappropriate analytical methods to identify prognostic factors or derive prediction models. Review level: Difficult to identify studies that recruited children at the first point-of-access to a healthcare system, as care-seeking for febrile children, particularly in LMICs, is heterogeneous.	63-64
Conclusions	26	Performance of prognostic factors varies widely across settings and clinicians must be cognisant not to over interpret individual predictors. Clinical prediction models and treatment algorithms can help identify children at risk of severe febrile illness but deriving these will require multiple, large, collaborative, research initiatives across different settings, which collect harmonised data on predictors and outcomes, and include unselected children presenting from the community.	65

Section/topic	#	Checklist item	Page
FUNDING			
Funding	27	AC is supported by a Wellcome Trust Doctoral Training Fellowship. RT is supported by the Botnar Foundation. The Cambodia Oxford Medical Research Unit and Mahidol-Oxford Tropical Medicine Research Unit are part of the Wellcome Trust Thailand Africa and Asia Programme, which receives core-funding from the UK Wellcome Trust (106698Z/14/Z). The project also received seed funding from the Tropical Health Education Trust in the UK. The funders had no role in study design, data collection, data analysis, data interpretation, or writing of the report.	6

9.2 Technical Advisory Panel

Table 9.2-1: Technical Advisory Panel of domain experts assembled for the systematic review.

Technical Advisory Panel member	Affiliation	Key authors proposed by Technical Advisory Panel
Dr. Jalemba Aluvaala	Paediatrics and Child Health, University of Nairobi, Nairobi, Kenya; KEMRI-Wellcome Trust Research Programme, Nairobi, Kenya	Ambrose Agweyu Andre Siqueira Anna Seale Anthony Scott Christopher C Moore Climent Casals-Pascual Elizabeth Molyneux Henriette Moll Kathryn Maitland Jay Berkeley Elizabeth Molyneux Quique Bassat Kristina E Rudd Martin Otyek Opio Michaëla A M Huson Mike English Mike Levin Ruud Nijman Samuel Akech Tim Baker Trevor Duke
Professor Quique Bassat	Centro de Investigação em Saúde de Manhiça, Maputo, Mozambique; ISGlobal, Hospital Clínic-Universitat de Barcelona, Barcelona, Spain; Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain	
Dr. David Bell	FIND, Campus Biotech, Building B, Level 0, Chemin des Mines 9, 1202, Geneva, Switzerland	
Professor John Crump	Division of Infectious Diseases and International Health, Duke University Medical Center, Durham, North Carolina; Centre for International Health, University of Otago, Dunedin, New Zealand	
Professor W. Conrad Liles	Department of Medicine, University of Washington, Seattle, Washington, USA	
Dr. Rianne Oostenbrink	Department of General Paediatrics, Erasmus Medical Center Sophia Children's Hospital, Rotterdam, The Netherlands	
Dr. Shunmay Yeung	Clinical Research Department, London School of Hygiene and Tropical Medicine, London, UK; Department of Paediatrics, Imperial College Healthcare NHS Trust, London, UK	

9.3 Data extraction sheet

Table 9.3-1. Data extraction sheet for the systematic review. Template was modified from the CHARMS and CHARMS-PF checklists. CHARMS = CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies; CHARMS-PF = CHARMS for Prognostic Factor studies; DOI = digital object identifier; PF = prognostic factor.

Domain	Item	General	Applicability	Risk of Bias	Extraction
Study	Study label	YES	NO	NO	
	Year of publication	YES	NO	NO	
	Journal of publication	YES	NO	NO	
	DOI	YES	NO	NO	
Source of data	Study design	YES	YES	YES	
	Target population	NO	YES	NO	
Participants	Single center or multi-center	YES	YES	YES	
	Number of centres recruiting	YES	NO	NO	
	Type of centres recruiting	YES	YES	YES	
	Location of study	YES	NO	NO	
	Recruitment method	YES	YES	YES	
	Recruitment setting	YES	YES	YES	
	Age range	YES	YES	YES	
	Fever definition and duration	YES	YES	YES	
	Inclusion criteria	YES	YES	YES	
	Exclusion criteria	YES	YES	YES	
	Participant description	YES	NO	NO	
	Study dates	YES	YES	NO	
Outcomes to be predicted	Prognostic outcome and definition	YES	YES	YES	
	Method of measurement of outcome	NO	NO	YES	
	Same outcome definition for all participants	NO	NO	YES	
	Same measurement of outcome for all participants	NO	NO	YES	
	Type of outcome (single or combined endpoints)	YES	YES	NO	
	Outcomes assessed without knowledge of the candidate predictor (blinded)	NO	NO	YES	
	Were candidate prognostic factors part of the outcome (e.g., when using a panel or consensus outcome measurement)	NO	NO	YES	
	Time of outcome occurrence	YES	YES	NO	
Prognostic factors	Demographic prognostic factors	YES	NO	NO	
	Anthropometric prognostic factors	YES	NO	NO	
	Socioeconomic prognostic factors	YES	NO	NO	
	Historical prognostic factors	YES	NO	NO	
	Clinical symptoms prognostic factors	YES	NO	NO	
	Clinical signs prognostic factors	YES	NO	NO	

Domain	Item	General	Applicability	Risk of Bias	Extraction
	Vital signs prognostic factors	YES	NO	NO	
	Laboratory measures prognostic factors	YES	NO	NO	
	Score prognostic factors with definition and weights	YES	NO	NO	
	Method for measurement	NO	NO	YES	
	Method of measurement of PFs is the same for all study participants	NO	NO	YES	
	Setting of measurement of PF	YES	YES	YES	
	Setting of measurement of PF is the same for all study participants	NO	NO	YES	
	Timing of prognostic factor measurement	NO	YES	YES	
	Prognostic factor assessed blinded for outcome	NO	NO	YES	
	Handling of prognostic factor in the analysis (continuous, linear, categorised, non-linear transformations)	NO	NO	YES	
Sample size	Number of participants	YES	NO	NO	
	Number of refusals	NO	NO	YES	
	Number of outcomes/events	YES	NO	NO	
	For model studies: Number of outcomes/events in relation to the number of candidate prognostic factors	NO	NO	YES	
Missing data	Proportion of data on PF available	NO	NO	YES	
	Number of participants with missing data for each outcome	NO	NO	YES	
	Method used for missing data	NO	NO	YES	
Analysis	Modelling method (linear, logistic, cox, parametric survival, competing risks)	YES	NO	YES	
	How modelling assumptions were checked	NO	NO	YES	
	Method for selection of PF for INCLUSION in multivariable modelling	NO	NO	YES	
	Method for selection of PF DURING multivariable modelling	NO	NO	YES	
	Inclusion of additional PF (not measured at admission or not included in above categories) for multivariable modelling	NO	YES	YES	
	Criteria used for any selection or exclusion of PF DURING multivariable modelling	NO	NO	YES	
	Method of handling each continuous PF (dichotomisation, categorisation, linear, non-linear), including values of any cut-points used and their justification	NO	NO	YES	
Results	Unadjusted effect estimates for each PF	YES	NO	NO	
	Adjusted effect estimates for each PF	YES	NO	NO	
Interpretation and discussion	Interpretation of presented results	YES	YES	YES	
	Comparison with other studies	YES	YES	NO	
	Discussion of generalisability	YES	YES	NO	
	Strengths	YES	YES	YES	
	Limitations	YES	YES	YES	

9.4 Prognostic factors that did not meet criteria for clinical relevance

9.4.1 Hard outcomes

Table 9.4-1: Prognostic factors judged to be of limited value to identify children at risk of progressing to severe febrile illness against 'hard' outcomes. Prognostic factors were judged to be of limited value if neither the PLR was ≥ 5.0 nor the NLR was ≤ 0.2. *CRS scored out of four variables including mental status, capillary refill time, peripheral pulse character, and presence of cold or mottled extremities;[103] †children with sepsis were enrolled based on modified Goldstein criteria (Table 3.3-2). Severe sepsis was defined based on Goldstein criteria for severe sepsis.[109] AVPU = Alert Voice Pain Unresponsive scale; CI = confidence interval; CRS = Clinical Recognition Signs; ED = emergency department; MUAC = mid-upper arm circumference; NLR = negative likelihood ratio; OPD = outpatient department; PLR = positive likelihood ratio; Prev = outcome prevalence (%); WAZ = weight-for-age z-score.

Study	Cohort	Outcome	Prev.	Prognostic factor	Definition / Cut-off	PLR	95% CI	NLR	95% CI
Demographic									
Kwizera 2019	Hospitalised	In-hospital mortality	1.5	Age	< 12m	0.97	(0.27 to 3.52)	1.01	(0.81 to 1.25)
Kwizera 2019	Hospitalised	In-hospital mortality	1.5	Age	1y to < 5y	0.69	(0.34 to 1.40)	1.33	(0.90 to 1.98)
Kwizera 2019	Hospitalised	In-hospital mortality	1.5	Age	5y to < 12y	1.41	(0.69 to 2.87)	0.86	(0.58 to 1.27)
Kwizera 2019	Hospitalised	In-hospital mortality	1.5	Age	12y to < 18y	1.76	(0.48 to 6.46)	0.93	(0.75 to 1.16)
Scott 2017	Hospital OPD/ED	30-day mortality	1.9	Age	13y to 17y	0.57	(0.20 to 1.67)	1.11	(0.96 to 1.29)
Scott 2017	Hospital OPD/ED	30-day mortality	1.9	Age	6y to < 13y	1.21	(0.71 to 2.06)	0.91	(0.68 to 1.22)
Scott 2017	Hospital OPD/ED	30-day mortality	1.9	Age	12m to < 6y	1.15	(0.76 to 1.73)	0.90	(0.61 to 1.31)
Scott 2017	Hospital OPD/ED	30-day mortality	1.9	Age	< 12m	0.53	(0.08 to 3.66)	1.04	(0.96 to 1.13)
Conroy 2015	Hospitalised	In-hospital mortality	4.7	Age	< 12m	1.16	(0.88 to 1.52)	0.93	(0.80 to 1.08)
Mtove 2011	Hospitalised	In-hospital mortality	5.0	Age	< 12m	1.47	(1.22 to 1.78)	0.81	(0.71 to 0.92)
Nadjm 2013	Hospitalised	In-hospital mortality	5.1	Age	< 12m	1.42	(1.19 to 1.70)	0.81	(0.71 to 0.93)
Nadjm 2013	Hospitalised	In-hospital mortality	5.1	Age	< 24m	1.20	(1.10 to 1.30)	0.63	(0.47 to 0.83)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Age	< 12m	1.47	(1.20 to 1.79)	0.90	(0.83 to 0.96)
Kwizera 2019	Hospitalised	In-hospital mortality	1.5	Sex	Female	0.87	(0.47 to 1.60)	1.12	(0.88 to 1.06)
Scott 2017	Hospital OPD/ED	30-day mortality	1.9	Sex	Female	1.05	(0.67 to 1.64)	0.97	(0.68 to 1.37)
Conroy 2015	Hospitalised	In-hospital mortality	4.7	Sex	Female	0.97	(0.77 to 1.22)	1.02	(0.86 to 1.23)
Lowlaavar 2016	Hospitalised	In-hospital mortality	5.0	Sex	Female	1.02	(0.78 to 1.34)	0.98	(0.78 to 1.23)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Sex	Female	1.01	(0.89 to 1.15)	0.99	(0.89 to 1.10)

Study	Cohort	Outcome	Prev.	Prognostic factor	Definition / Cut-off	PLR	95% CI	NLR	95% CI
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Sex	Female	1.04	(0.91 to 1.17)	0.97	(0.87 to 1.08)
Anthropometric									
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Weight	< 6kg	1.57	(0.92 to 2.68)	0.98	(0.96 to 1.01)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Weight	< 8kg	1.29	(1.06 to 1.58)	0.92	(0.86 to 0.99)
Historical									
Scott 2017	Hospital OPD/ED	30-day mortality	1.9	Medical history	Non-oncological comorbidity	1.06	(0.68 to 1.66)	0.96	(0.67 to 1.36)
Scott 2017	Hospital OPD/ED	30-day mortality	1.9	Medical history	Oncological comorbidity	1.73	(1.17 to 2.54)	0.69	(0.46 to 1.03)
Scott 2014	Hospital OPD/ED	24-hour organ dysfunction	5.4	Medical history	Immunosuppressed	4.64	(1.79 to 12.00)	0.74	(0.52 to 1.07)
Clinical symptoms									
Conroy 2015	Hospitalised	In-hospital mortality	4.7	Convulsions	Caretaker history	1.90	(1.41 to 2.54)	0.81	(0.70 to 0.93)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Convulsions	Caretaker history	1.39	(1.05 to 1.84)	0.95	(0.90 to 1.00)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Fever	Caretaker history	1.00	(0.99 to 1.01)	1.42	(0.32 to 6.25)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Urine looks dark	Caretaker history	1.04	(0.76 to 1.41)	0.99	(0.95 to 1.04)
Clinical signs									
van Nassau 2018	Hospitalised	PICU transfer and/or in-hospital mortality	2.7	Abnormal temperature	> 38.5°C or < 36°C	1.43	(0.93 to 2.20)	0.77	(0.50 to 1.17)
Conroy 2015	Hospitalised	In-hospital mortality	4.7	Hyperthermia	Temperature > 38°C	0.59	(0.42 to 0.82)	1.34	(1.18 to 1.51)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Hyperthermia	Axillary temperature > 37°C	0.72	(0.66 to 0.80)	2.24	(1.92 to 2.62)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Hyperthermia	Axillary temperature > 39°C	0.60	(0.45 to 0.79)	1.13	(1.07 to 1.19)
Conroy 2015	Hospitalised	In-hospital mortality	4.7	Hypothermia	Temperature < 36°C	3.16	(1.73 to 5.77)	0.92	(0.86 to 0.99)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Hypothermia	Axillary temperature < 36°C	3.47	(2.58 to 4.67)	0.87	(0.83 to 0.92)
van Nassau 2018	Hospitalised	PICU transfer and/or in-hospital mortality	2.7	Heart rate	Age-adjusted	1.98	(1.33 to 2.94)	0.63	(0.40 to 0.99)
Conroy 2015	Hospitalised	In-hospital mortality	4.7	Heart rate	Age-adjusted	0.92	(0.77 to 1.09)	1.15	(0.90 to 1.46)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Heart rate	≥ 200 bpm	4.61	(1.99 to 10.67)	0.98	(0.96 to 1.00)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Heart rate	Age-adjusted	0.74	(0.66 to 0.82)	1.69	(1.47 to 1.93)
Conroy 2015	Hospitalised	In-hospital mortality	4.7	Capillary refill time	≥ 3 secs	4.67	(3.00 to 7.28)	0.83	(0.75 to 0.92)
Scott 2014	Hospital OPD/ED	24-hour organ dysfunction	5.4	Capillary refill time	Flash or > 2 secs	0.50	(0.07 to 3.35)	1.09	(0.92 to 1.29)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Capillary refill time	> 2 secs	1.28	(1.21 to 1.35)	0.48	(0.37 to 0.62)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Capillary refill time	< 2 secs	0.50	(0.39 to 0.65)	1.26	(1.19 to 1.33)

Study	Cohort	Outcome	Prev.	Prognostic factor	Definition / Cut-off	PLR	95% CI	NLR	95% CI
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Capillary refill time	≥ 3 secs	1.98	(1.73 to 2.27)	0.69	(0.62 to 0.77)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Capillary refill time	2-3 secs	0.84	(0.72 to 0.99)	1.11	(1.02 to 1.22)
Scott 2014	Hospital OPD/ED	24-hour organ dysfunction	5.4	Poor peripheral perfusion	Cold extremity	4.35	(0.52 to 36.17)	0.94	(0.80 to 1.10)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Poor peripheral perfusion	Limb-core temp. gradient	1.34	(1.25 to 1.44)	0.54	(0.44 to 0.67)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Poor peripheral perfusion	Limb-core temp. gradient	1.32	(1.23 to 1.42)	0.57	(0.47 to 0.70)
van Nassau 2018	Hospitalised	PICU transfer and/or in-hospital mortality	2.7	Respiratory rate	Age-adjusted	1.11	(0.93 to 1.33)	0.62	(0.22 to 1.78)
Conroy 2015	Hospitalised	In-hospital mortality	4.7	Respiratory rate	Age-adjusted	2.05	(1.66 to 2.53)	0.66	(0.54 to 0.80)
Conroy 2015	Hospitalised	In-hospital mortality	4.7	Respiratory distress	Subcostal recession	3.76	(3.18 to 4.45)	0.41	(0.31 to 0.54)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Respiratory distress	Chest wall retraction	1.15	(1.07 to 1.23)	0.70	(0.56 to 0.86)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Respiratory distress	Increased work of breathing or deep breathing	1.13	(1.09 to 1.17)	0.43	(0.29 to 0.63)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Respiratory distress	Increased work of breathing or deep breathing	1.12	(1.08 to 1.16)	0.48	(0.34 to 0.69)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Respiratory crackles	Physician assessment	1.82	(1.54 to 2.14)	0.79	(0.72 to 0.86)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Respiratory crackles	Physician assessment	1.81	(1.53 to 2.13)	0.80	(0.73 to 0.87)
Kwizera 2019	Hospitalised	In-hospital mortality	1.5	Focus of infection	Meningeal	2.71	(0.17 – 43.95)	0.98	(0.89 – 1.08)
Kwizera 2019	Hospitalised	In-hospital mortality	1.5	Focus of infection	Respiratory	0.50	(0.14 – 1.81)	1.20	(0.97 – 1.49)
Kwizera 2019	Hospitalised	In-hospital mortality	1.5	Focus of infection	Gastrointestinal	0.56	(0.08 – 3.71)	1.07	(0.92 – 1.23)
Kwizera 2019	Hospitalised	In-hospital mortality	1.5	Focus of infection	Urinary	1.69	(0.11to26.71)	0.99	(0.90 – 1.08)
Kwizera 2019	Hospitalised	In-hospital mortality	1.5	Focus of infection	Skin and/or soft-tissue	3.28	(0.20 to53.88)	0.98	(0.89 – 1.07)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Cough	Physician assessment	1.00	(0.93 to 1.08)	1.00	(0.83 to 1.21)
Scott 2014	Hospital OPD/ED	24-hour organ dysfunction	5.4	Agitation	Physician assessment	4.17	(2.08 to 8.35)	0.61	(0.37 to 1.00)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Seizures	Physician assessment	1.24	(0.96 to 1.61)	0.96	(0.91 to 1.01)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Normal consciousness	AVPU = A	0.29	(0.19 to 0.44)	1.23	(1.18 to 1.28)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Vomiting	Physician assessment	1.14	(1.03 to 1.27)	0.85	(0.74 to 0.98)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Dehydration	Decreased skin turgor	2.83	(2.07 to 3.89)	0.90	(0.86 to 0.95)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Dehydration	Sunken eyes or reduced skin turgor	2.52	(1.89 to 3.36)	0.89	(0.85 to 0.94)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Jaundice	Physician assessment	1.39	(1.21 to 1.60)	0.82	(0.75 to 0.91)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Severe pallor	Physician assessment	1.47	(1.35 to 1.59)	0.56	(0.47 to 0.67)

Study	Cohort	Outcome	Prev.	Prognostic factor	Definition / Cut-off	PLR	95% CI	NLR	95% CI
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Severe pallor	Physician assessment	1.53	(1.37 to 1.71)	0.70	(0.62 to 0.80)
Laboratory									
Nadjm 2013	Hospitalised	In-hospital mortality	5.1	Glucose	> 5 mmol/l	0.56	(0.47 to 0.67)	2.33	(2.02 to 2.68)
Nadjm 2013	Hospitalised	In-hospital mortality	5.1	Glucose	2.5-5 mmol/l	1.39	(1.11 to 1.75)	0.88	(0.80 to 0.98)
Mtove 2011	Hospitalised	In-hospital mortality	5.0	Haemoglobin	< 4 g/dl	1.98	(1.53 to 2.56)	0.83	(0.76 to 0.92)
Nadjm 2013	Hospitalised	In-hospital mortality	5.1	Haemoglobin	< 5 g/dl	1.93	(1.51 to 2.46)	0.83	(0.75 to 0.92)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Haemoglobin	< 5 g/dl	1.43	(1.24 to 1.64)	0.81	(0.73 to 0.89)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Haemoglobin	5-7 g/dl	0.90	(0.68 to 1.19)	1.02	(0.97 to 1.07)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Haemoglobin	7-10 g/dl	0.97	(0.79 to 1.18)	1.01	(0.94 to 1.09)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Haemoglobin	≥ 10 g/dl	0.56	(0.42 to 0.75)	1.14	(1.09 to 1.20)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Lactate	< 2.5 mmol/l	0.26	(0.17 to 0.37)	1.35	(1.29 to 1.40)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Lactate	2.5-5 mmol/l	0.45	(0.34 to 0.58)	1.28	(1.21 to 1.35)
van Nassau 2018	Hospitalised	PICU transfer and/or in-hospital mortality	2.7	Leukocyte count	High or low (age-adjusted)	0.97	(0.64 to 1.48)	1.03	(0.67 to 1.58)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	pH	< 7.2	4.85	(3.79 to 6.21)	0.70	(0.63 to 0.77)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	pH	< 7.2	4.43	(3.45 to 5.68)	0.72	(0.65 to 0.79)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Urea	> 20 mg/dl	2.50	(2.08 to 3.00)	0.67	(0.58 to 0.76)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Urea	> 20 mg/dl	2.37	(1.97 to 2.84)	0.69	(0.61 to 0.78)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Base deficit	> -8 mmol/l	1.74	(1.61 to 1.88)	0.31	(0.23 to 0.43)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Bicarbonate	< 15 mmol/l	2.25	(2.03 to 2.51)	0.40	(0.31 to 0.50)
Composite scores									
Scott 2014	Hospital OPD/ED	24-hour organ dysfunction	5.4	CRS[*]	≥ 1	2.02	(1.26 to 3.23)	0.55	(0.28 to 1.11)
Scott 2014	Hospital OPD/ED	24-hour organ dysfunction	5.4	CRS[*]	≥ 2	4.35	(1.40 to 13.52)	0.81	(0.60 to 1.10)
SEAIDCRN 2017	Hospitalised	28-day mortality	1.9	Severe sepsis[†]	Goldstein criteria	3.08	(2.28 to 4.16)	0.29	(0.11 to 0.79)

9.4.2 Soft outcomes

Table 9.4-2: Prognostic factors judged to be of limited value to identify children at risk of progressing to severe febrile illness against 'soft' outcomes. Prognostic factors were judged to be of limited value if neither the PLR was ≥ 5.0 nor the NLR was ≤ 0.2. *AIOS calculated as described for Yale Observation Score in Table 3.3-3. AIOS = Acute Infantile Observation Score; CI = confidence interval; ITN = insecticide-treated bednet; NLR = negative likelihood ratio; PLR = positive likelihood ratio; Prev. = outcome prevalence (%).

Study	Cohort	Outcome	Prev.	Prognostic factor	Definition / Cut-off	PLR	95% CI	NLR	95% CI
Demographic									
Mwandama 2016	Primary care	Persistent symptoms at D7	10.4	Sex	Female	0.87	(0.53 to 1.42)	1.16	(0.73 to 1.83)
Socioeconomic									
Mwandama 2016	Primary care	Persistent symptoms at D7	10.4	Household socioeconomic status	Highest wealth quintile	1.24	(0.89 to 1.73)	0.71	(0.36 to 1.40)
Mwandama 2016	Primary care	Persistent symptoms at D7	10.4	Household socioeconomic status	Slept under ITN	0.73	(0.49 to 1.07)	2.08	(1.13 to 3.82)
Mwandama 2016	Primary care	Persistent symptoms at D7	10.4	Parental education	None	0.71	(0.10 to 5.20)	1.02	(0.91 to 1.15)
Mwandama 2016	Primary care	Persistent symptoms at D7	10.4	Parental education	Primary	1.05	(0.82 to 1.34)	0.86	(0.34 to 2.13)
Mwandama 2016	Primary care	Persistent symptoms at D7	10.4	Parental education	Secondary	0.92	(0.31 to 2.74)	1.02	(0.83 to 1.25)
Clinical symptoms									
Mwandama 2016	Primary care	Persistent symptoms at D7	10.4	URTI/cold presentation	Caretaker history	1.27	(0.81 to 2.00)	0.79	(0.46 to 1.35)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Sore throat	Caretaker history	2.21	(1.39 to 3.51)	0.82	(0.70 to 0.96)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Rhinorrhoea	Caretaker history	1.19	(0.84 to 1.67)	0.91	(0.74 to 1.12)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Cough	Caretaker history	1.20	(0.86 to 1.67)	0.90	(0.73 to 1.11)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Moaning respiration	Caretaker history	1.27	(1.01 to 1.60)	0.77	(0.56 to 1.05)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Abdominal pain	Caretaker history	1.45	(0.87 to 2.42)	0.92	(0.80 to 1.05)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Diarrhoea > 2/day	Caretaker history	1.45	(0.98 to 2.15)	0.87	(0.72 to 1.04)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Vomiting	Caretaker history	1.31	(0.95 to 1.81)	0.85	(0.68 to 1.07)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Febrile convulsions	Caretaker history	2.36	(1.04 to 5.35)	0.93	(0.85 to 1.02)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Drowsy or difficult to wake	Caretaker history	0.85	(0.63 to 1.15)	1.15	(0.91 to 1.46)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Restlessness or confusion	Caretaker history	1.09	(0.74 to 1.62)	0.96	(0.80 to 1.15)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Feeling irritable	Caretaker history	1.31	(0.97 to 1.78)	0.84	(0.66 to 1.06)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Inconsolable crying	Caretaker history	0.99	(0.75 to 1.32)	1.00	(0.79 to 1.28)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Crying during diaper change	Caretaker history	0.95	(0.63 to 1.45)	1.02	(0.86 to 1.21)

Study	Cohort	Outcome	Prev.	Prognostic factor	Definition / Cut-off	PLR	95% CI	NLR	95% CI
Elshout 2015	Primary care	Persistent fever at D3	13.1	Crying when picked up	Caretaker history	0.86	(0.55 to 1.35)	1.06	(0.90 to 1.24)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Different illness than usual	Caretaker history	1.16	(0.94 to 1.44)	0.81	(0.59 to 1.13)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Parental concern	Caretaker history	1.51	(0.92 to 2.47)	0.90	(0.78 to 1.05)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Drinking less than half usual	Caretaker history	1.07	(0.76 to 1.51)	0.96	(0.78 to 1.18)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Skin rash	Caretaker history	0.92	(0.54 to 1.59)	1.02	(0.90 to 1.16)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Pale, grey or spotted skin	Caretaker history	0.91	(0.69 to 1.21)	1.09	(0.85 to 1.39)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Normal play behaviour	Caretaker history	1.09	(0.89 to 1.34)	0.87	(0.62 to 1.24)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Normal reaction to parents	Caretaker history	0.70	(0.26 to 1.89)	1.03	(0.96 to 1.11)
Clinical signs									
Mwandama 2016	Primary care	Persistent symptoms at D7	10.4	Hyperthermia	Axillary temperature ≥ 37.5°C	1.64	(1.01 to 2.66)	0.67	(0.39 to 1.15)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Hyperthermia	Rectal temperature ≥ 38°C	1.47	(1.08 to 2.01)	0.80	(0.63 to 1.00)
van Nassau 2018	Hospitalised	Length of stay >/= 7d	22.2	Abnormal temperature	> 38.5°C or < 36°C	0.81	(0.63 to 1.04)	1.11	(0.99 to 1.24)
van Nassau 2018	Hospitalised	Length of stay >/= 7d	22.2	Heart rate	Age-adjusted	1.67	(1.18 to 2.37)	0.88	(0.80 to 0.97)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Capillary refill time	> 2 secs	0.98	(0.35 to 2.71)	1.00	(0.93 to 1.07)
van Nassau 2018	Hospitalised	Length of stay >/= 7d	22.2	Respiratory rate	Age-adjusted	0.99	(0.89 to 1.10)	1.04	(0.73 to 1.47)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Respiratory distress (dyspnoea)	Physician assessment	1.06	(0.71 to 1.58)	0.98	(0.82 to 1.16)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Pharyngitis	Sign of throat infection	1.64	(1.25 to 2.16)	0.70	(0.54 to 0.91)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Rhinorrhoea	Physician assessment	0.91	(0.70 to 1.18)	1.11	(0.85 to 1.44)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Cough	Physician assessment	1.28	(0.95 to 1.72)	0.84	(0.66 to 1.07)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Palpable lymph nodes	Physician assessment	1.39	(1.09 to 1.77)	0.73	(0.54 to 0.98)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Meningism	Able to put chin to chest	0.34	(0.05 to 2.47)	1.03	(0.99 to 1.07)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Ill appearance	Physician assessment	1.32	(0.61 to 2.85)	0.97	(0.89 to 1.06)
Laboratory									
van Nassau 2018	Hospitalised	Length of stay >/= 7d	22.2	Leukocyte count	High or low (age-adjusted)	1.15	(0.96 to 1.36)	0.86	(0.70 to 1.06)
Freyne 2013	Hospitalised	Length of stay > 96h	26.1	Leukocyte count	> 15,000 cells/mm	0.97	(0.44 to 2.15)	1.02	(0.57 to 1.82)
Freyne 2013	Hospitalised	Length of stay > 96h	26.1	Procalcitonin	> 1.0 ng/l	1.00	(0.31 to 3.23)	1.00	(0.68 to 1.48)
Freyne 2013	Hospitalised	Length of stay > 96h	26.1	C-reactive protein	> 20 mg/dl	1.27	(0.61 to 2.64)	0.82	(0.43 to 1.56)
Composite scores									
Freyne 2013	Hospitalised	Length of stay > 96h	26.1	AIOS*	> 10	1.00	(0.51 to 1.97)	1.00	(0.51 to 1.97)

9.5 Case report form for acute illness episode in original birth cohort study

Acute Respiratory Infection Study - Acute Infection

ID code
Date of Birth (dd/mm/year) Age at Examination (months)
Sex Weight (kg)
Head Circumference (cm) Length (cm)

If the child is severely unwell use the Emergency Protocol for the Seriously Ill Child

History

Presenting complaints

1. Cough **YES/NO**
   If yes duration _____days
2. Fever **YES/NO**
   If yes duration _____days
3. Runny nose **YES/NO**
   If yes duration _____days
4. Noisy breathing **YES/NO**
   Describe _____
5. Earache **YES/NO**
   If yes duration _____days
6. Sore throat **YES/NO**
   If yes duration _____days
7. Vomiting **YES/NO**
   If yes duration _____days
8. Diarrhoea **YES/NO** If yes
   Duration _____days
   Frequency of stools _____/day
   Is there blood in the stool Yes/No
9. Rash **YES/NO**
   If yes duration _____days
   Describe the rash _____

10. When did your infant last pass urine _____(hours ago)

11. How are you feeding your infant (circle)?
    Breast feeding Formula milk
    Mixed feeding Other, please give details _____

12. Do you have any feeding concerns (circle)? **YES/NO**
    Details_____

13. How long is the infant able to feed for (minutes?) _____

14. Has your baby taken any antibiotics since last seen (circle)? **YES/NO**

Episode	1	2	3	4	5
What antibiotic?					
Started					
For how long (days)?					

15. Has your baby taken any other medication since last seen (circle)?**YES/NO**

Episode	1	2	3	4	5
What medication?					
Started					
For how long (days)?					

16. Does your infant have any heart problems (circle)?
    YES/NO Details_____
17. Has your infant had breathing problems in the past (circle)?
    YES/NO Details_____
18. Does your infant have any abnormalities of ear, nose or throat (circle)? **YES/NO**
    Details_____
19. Does your infant have any abnormalities of their central nervous system such as seizures or cerebral palsy (circle)?
    YES/NO Details_____
20. Does your infant have any abnormalities of their kidneys such as many UTIs (circle)?
    YES/NO Details_____
21. Have there been any changes to the household structure since last visit? **YES/NO**
    Details_____
22. Are any family members unwell at the current time (circle)?
    YES/NO Details_____

Examination

Observations

1. Is the infant (circle):

Alert *Miserable* *Inconsolable*
Responding only to voice *Responding only to pain* *Unresponsive*

2. What is the colour of the infant (circle)?

Pale *Mottled* *Cyanosed* *Normal* *Jaundiced*

3. What is the infants pulse rate (per minute) _____
4. Respiratory rate (count over 60s) _____
5. O_2 Saturations _____%
6. What is the infant's temperature? _____^{0}C
7. Is there a rash present (circle)? **YES/NO**
   Describe _____
8. Is the infant able to suck and sustain a feed (circle)? **YES/NO**
9. Is the infant dehydrated (circle)? **YES/NO**
   If yes what % _____

10. Is the infant's cry normal (circle)? **YES/NO**
    If no describe _____

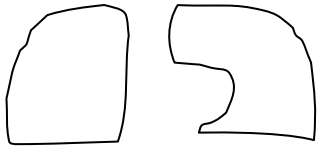
Cardiovascular Examination

1. Central capillary refill time (seconds) _____
2. Heart sounds (circle)
   a. Normal
   b. Murmur present, describe _____
   c. Gallop rhythm present (circle)? **YES/NO**

Respiratory Examination

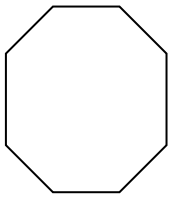
1. Any of the following present (circle)?

Grunting	*Intercostal recession*	*Subcostal recession*
Tracheal tug	*Nasal flaring*	*Head bobbing*
Stridor	*Sternal recession*	*Chest wall deformity*



2. Auscultation of the chest (circle)

Normal Abnormal, describe, quality and location on diagram



Abdominal Examination

1. On palpation of the abdomen: any mass present (circle)? **YES/NO**
   is the liver palpable (circle)? **YES/NO**
   is the spleen palpable (circle)? **YES/NO**
   If yes demonstrate on diagram with measurements (cm)

Neurological Examination

1. Does the infant have normal muscle tone (circle)? **YES/NO**
   Describe if no _____
2. Does the infant have normal posture (circle)? **YES/NO**
   Describe if no _____
3. Is the infant moving all limbs normally (circle)? **YES/NO**
   Describe if no _____
4. Are the infant's pupils equal and reactive? (circle)? **YES/NO**
   Describe if no _____
5. Describe the anterior fontanelle
   a. Normal
   b. Raised
   c. Depressed

ENT Examination

1. Does the infant have any structural abnormality of ears, nose or palate (circle)? **YES/NO**
   Describe if yes_____
2. Throat **Normal/Abnormal**
   If abnormal describe _____
3. Right ear **Normal/Abnormal**
   If abnormal describe _____
4. Left ear **Normal/Abnormal**
   If abnormal describe _____

5. Nose **Normal/Abnormal**
   If abnormal describe _____

Bone and Joint

1. Does the infant have any hot or swollen joints (circle)? **YES/NO**
   Describe if yes _____

Skin

1. Does the infant have any areas of cellulitis (circle)? **YES/NO**
   Describe if yes _____

Diagnosis

Circle one
Very severe pneumonia
Severe pneumonia
Pneumonia
Cardiac failure
Bronchiolitis
Acute Asthma
Croup
Foreign body aspiration
Other, details _____

Management

Oxygen **YES/NO**
Details (amount and delivery) _____
IV bolus **YES/NO**
Details _____
IV maintenance fluids **YES/NO**
Details _____

NGT fluids **YES/NO**
Details _____
Antibiotics **YES/NO**
Details including which antibiotic, route, and dose _____

Investigations

CBC **YES/NO**
Blood Culture **YES/NO**
Nasopharyngeal aspirate **YES/NO**
Nasopharyngeal swab **YES/NO**
Chest x-ray **YES/NO**
Lumbar puncture **YES/NO**
Lung aspirate **YES/NO**
Others **YES/NO**
Details _____

Form completed by:
Name _____
Signature _____

9.6 Standard Operating Procedure: Maela study – biomarker assays

MAELA-BIO

Standard Operating Procedure

Title:	Biomarker analysis using the Ella system
Version:	3.0
Date:	18 February 2022

1. Abbreviations

SOP Standard operating procedure

2. Background

Maela-BIO is a secondary analysis of a prospective longitudinal observational study (the Maela ARI study). The study aims to identify clinical and biochemical prognostic markers in young children with acute respiratory infections. The end goal is to externally validate existing clinical risk scores and to determine the added value of candidate biomarkers for risk stratification of children with pneumonia.

3. Purpose and Scope

The purpose of this document is to describe the processing of serum collected from the children with acute respiratory infections for biomarker analysis. Serum biomarker concentrations will be measured using the Ella platform (Simple Plex, ProteinSimple™).

4. Requirements

- Ella Simple Plex System (ProteinSimple™)
- Ella cartridge kits (contain wash buffer A, sample diluent, and cartridge)
- Microcentrifuge
- Class II Biosafety Cabinet
- P1000 pipette
- P200 pipette
- P1000 sterile pipette tips
- P200 sterile pipette tips
- 1.5 mL sterile snap cap microcentrifuge tubes (~1500 tubes)
- Microcentrifuge tube racks (x3)
- Gloves
- Lab coat
- Biohazard Bags
- 0.3% Chlorine
- Sharpie
- Computer with Microsoft Excel to prepare biomarker result database

5. Procedure

The recommendations in this SOP are intended for staff already familiar with processing blood specimens and performing laboratory methods in biochemistry (particularly enzyme-linked immunosorbent assays) who have been authorised to do so by the Principal Investigator(s).

Gloves and laboratory coat should be worn at all times when handling specimens or cartridges.

1. Storage and Thawing Serum

1.1 Store collected serum samples at -80°C until processing.
1.2 All Ella cartridge kits should be stored at 4°C until use.
1.3 The day before processing serum specimens, remove 10 samples from -80°C storage and allow to thaw at 4°C overnight. Move the specimens from -80°C as late as possible the day prior to running on the Ella.

2. Processing Serum Samples

Starting the Ella

2.1 Turn on ELLA (power switch is located on the back of the instrument)

Figure 1. Ella power switch.

2.2 Turn on Ella's computer and monitor.
2.3 Login to the computer.
2.4 Wait until the Ella's status light is green as this indicates the system is ready (Figure 2 and Table 1).

Figure 2. Ella ready status light.

Table 1. Status of Ella lights

Description	Status
Green	Ella is powered and ready for use
Blue	Ella is running a cartridge
Blinking red	An error has occurred (Ella may optionally sound an audible alarm)

2.5 Start the Simple Plex Runner software. The software can be accessed by either:
i) Simple Plex Runner application on the computer desktop
ii) Windows **Start** menu, select **Programs** > **Simple Plex** > **Simple Plex Runner**

Cartridge Preparation

DO NOT PREPARE CARTRIDGES IN ADVANCE. THESE SHOULD BE PREPARED BEFORE THE START OF THE RUN. IF SERUM SAMPLES ARE NOT BEING USED, ENSURE THEY ARE STORED AT 4°C AT ALL TIMES TO LIMIT BIOMARKER DEGRADATION.

Each time a kit is run, a new kit and cartridge ID will need to be scanned into the system.

2.6 Select Kit ID on the left-hand panel of the Single Plex Runner application (Figure 3).

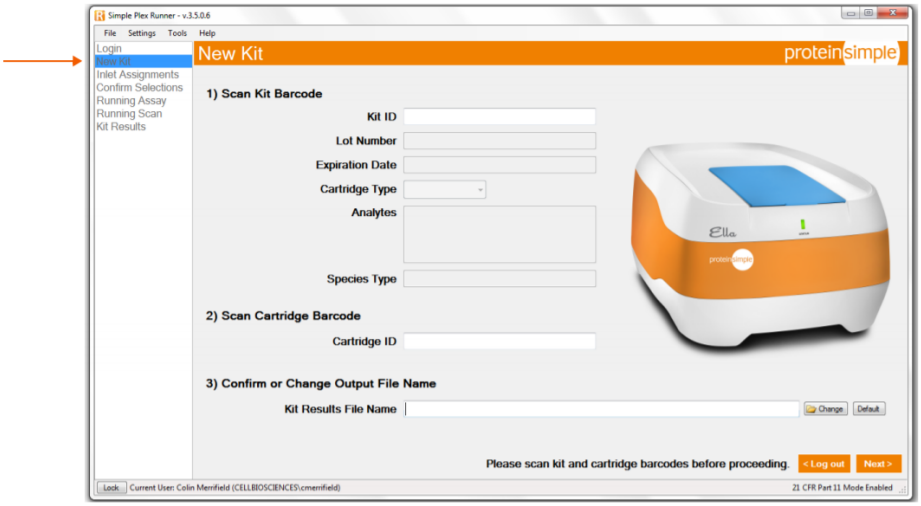


Figure 3. Screen for programming application for a new kit.

2.7 Remove a single Ella kit from its 4°C storage.

2.8 Scan the kit barcode on the packaging of the cartridge. See Figure 6C for an example of the package barcode.

2.9 Remove the cartridge from the vacuum bag. **Do not remove the plastic lining at the bottom of the cartridge until the cartridge is ready to be loaded**.

2.10 Scan the cartridge barcode for the cartridge ID and return the cartridge to the bag until it is ready to be loaded. This prevents dust and debris from getting into the cartridge wells.

2.11 Select the file where cartridge results will be saved.

2.12 Click the orange icon “Next”.

2.13 The inlet assignment application will now be available (Figure 4).

Assign:

Sample type = unknown (sample with an unknown concentration)

Sample name = specimen number (specnum) from the ARI serum sample list

Dilution factor = e.g., 50,000 for Cartridge E with CRP

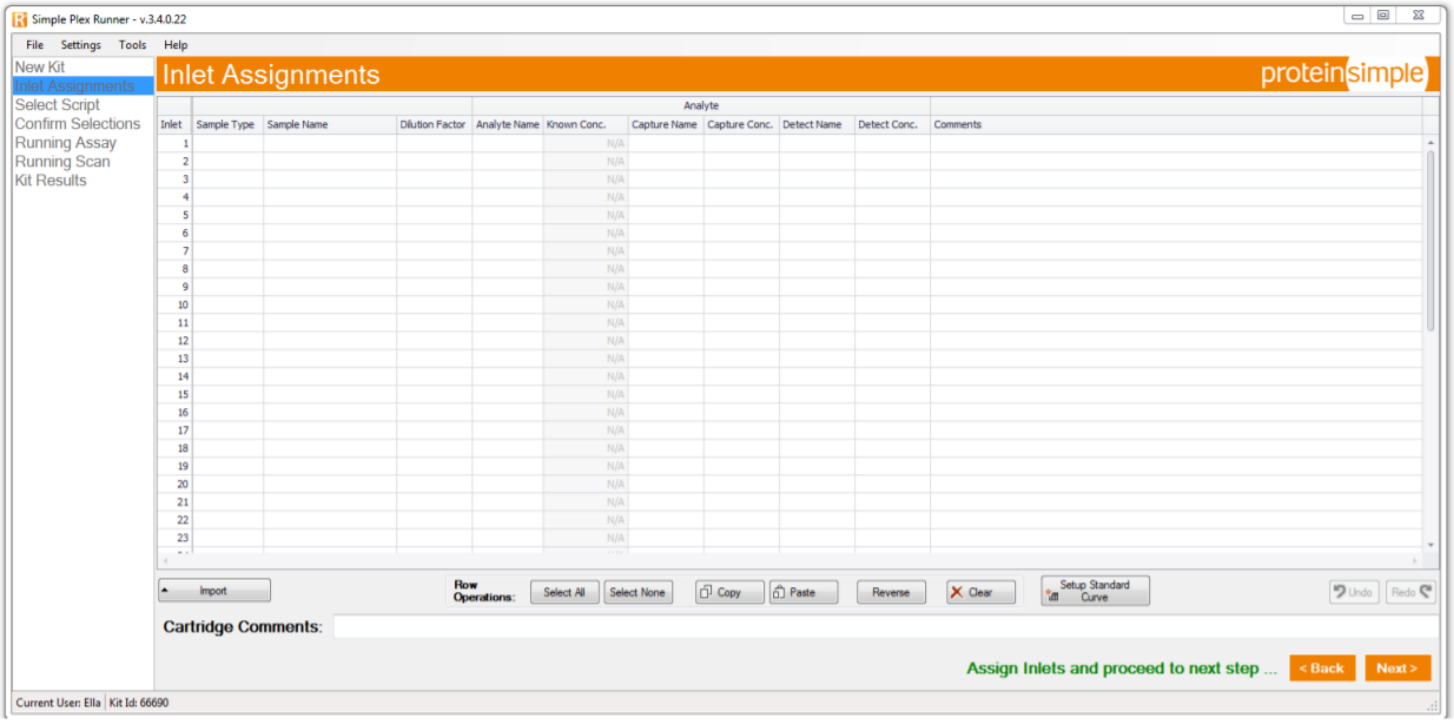


Figure 4. Inlet assignment screen.

2.14 Select the orange button “Next”. Leave the program in this Confirm Selections Screen (Figure 5). Do not click Start.

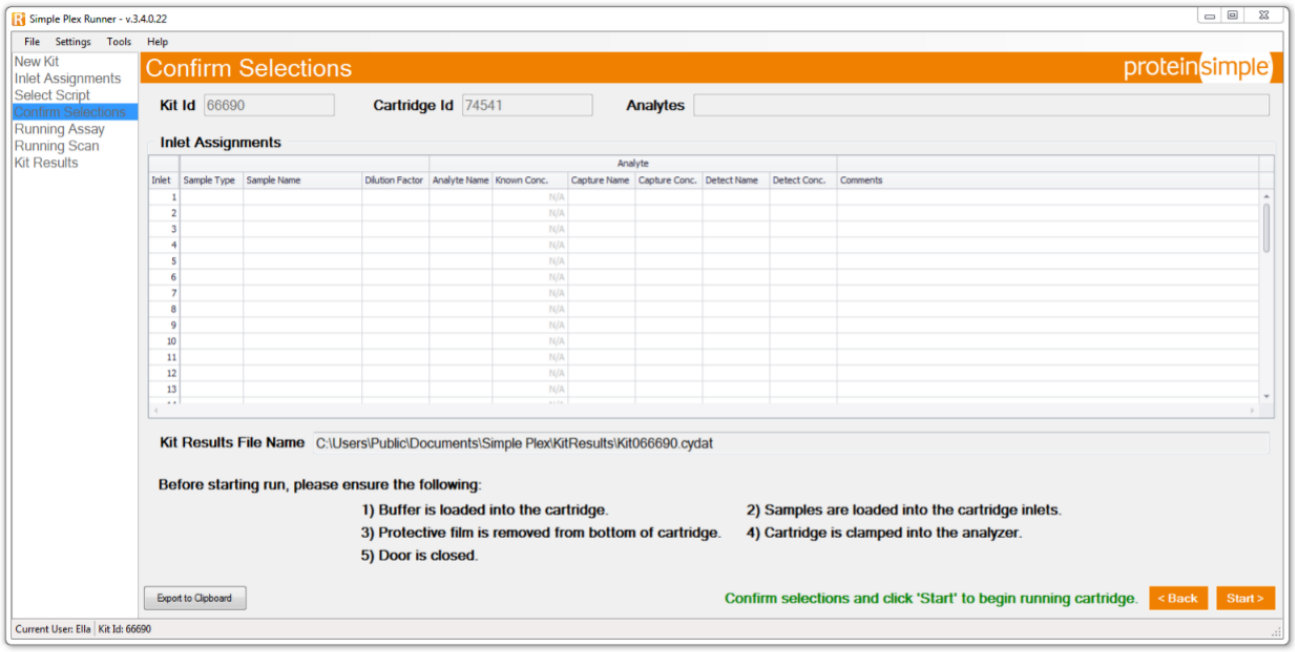


Figure 5. Confirm selections screen.

2.15 The plate is now ready to have samples and wash buffer loaded.

Sample Preparation

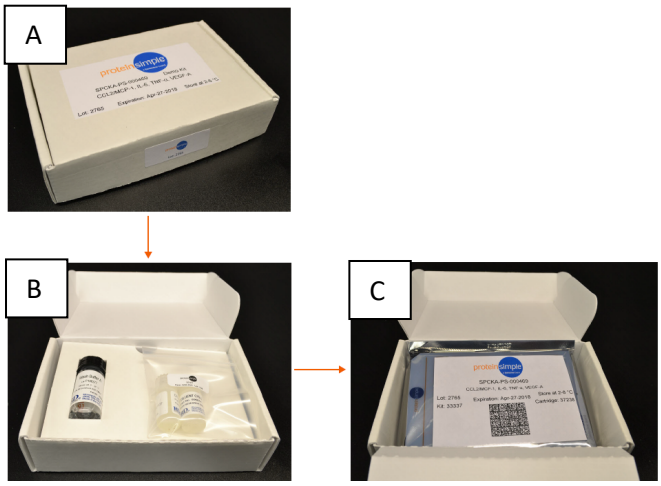


Figure 6. Ella simple cartridge kit. (A) Unopened Ella simple cartridge kit, (B) Wash buffer A (left) and sample diluent (right), (C) Barcoded cartridge found underneath wash buffer and sample diluent.

2.16 A total of 5 plates should be processed each day to avoid multiple freeze thaws of serum (see below). Remove the cartridge from its packaging but **DO NOT throw packaging away until the plate has completed its run**.

A) 1 plate for VEGFR1 (1 x 32 samples) – Dilution 1:2
B) 1 plate for Ang-2, IP-10, IL-10, IL-1ra, IL-6, IL-8 (6 x 32 samples) – Dilution 1:10
C) 1 plate for Procalcitonin, TREM-1 (2 x 32 samples) – Dilution 1:10
D) 1 plate for Ang-1, CHI3L1, TNFR1 (3 x 32 samples) – Dilution 1:50
E) 1 plate for CRP (1 x 32 samples) – Dilution 1:50,000

2.17 Prepare serial dilutions of the serum in the 96-well plate (Figure 7 and 8). Between each dilution tube, ensure the pipette tip is changed and the tube is vortexed before transferring to the next tube in the series. ** Each dilution is calculated to have enough volume for all cartridges (only one dilution series needed). After loading Cartridge A, immediately put the serum and serum dilution series at 4-8°C until the next cartridge will be loaded. Only place serum back in freezer at the end of the day to avoid any freeze thaws if additional dilutions need to be prepared.

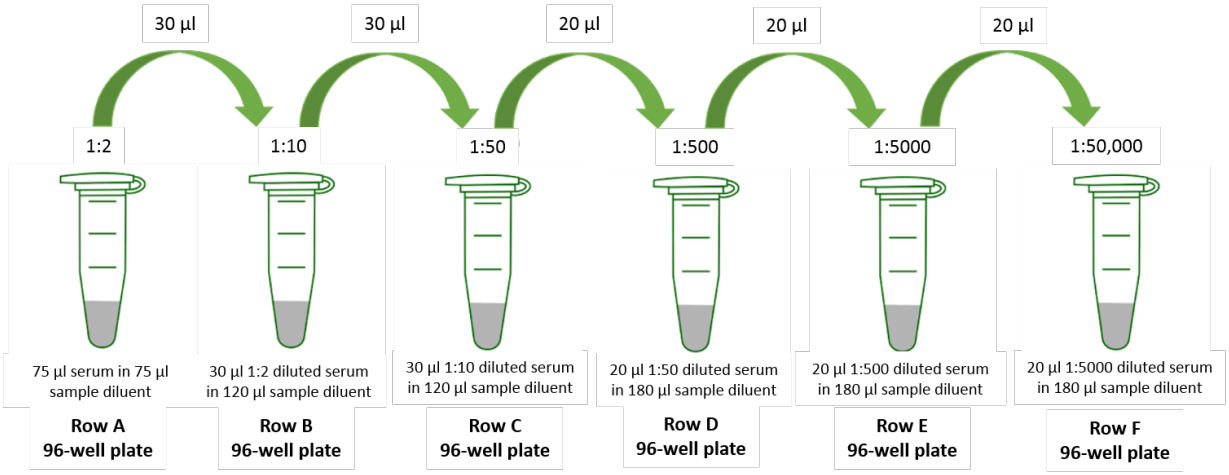
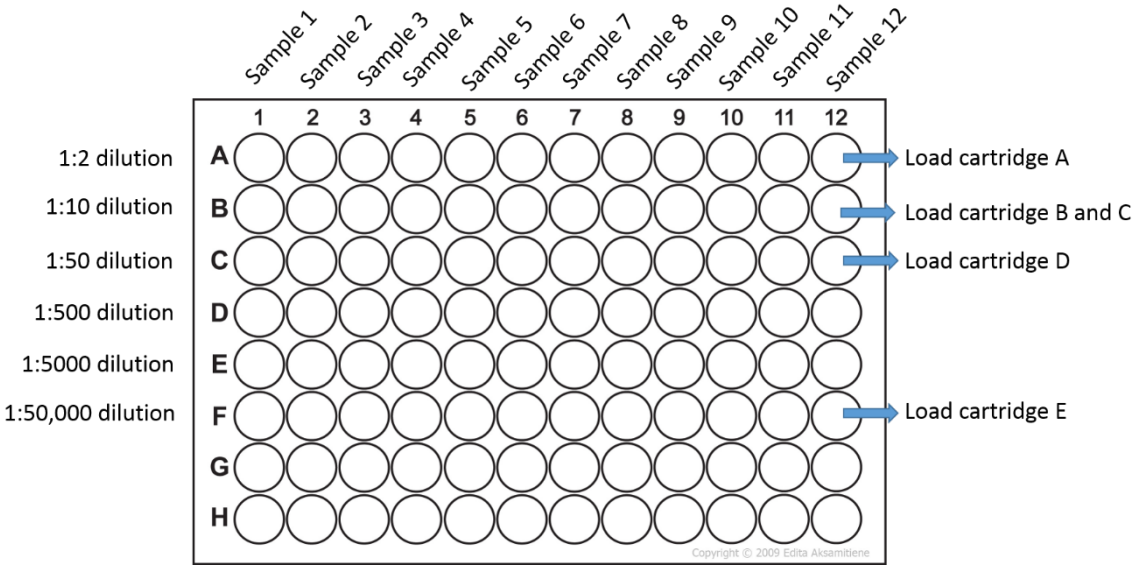
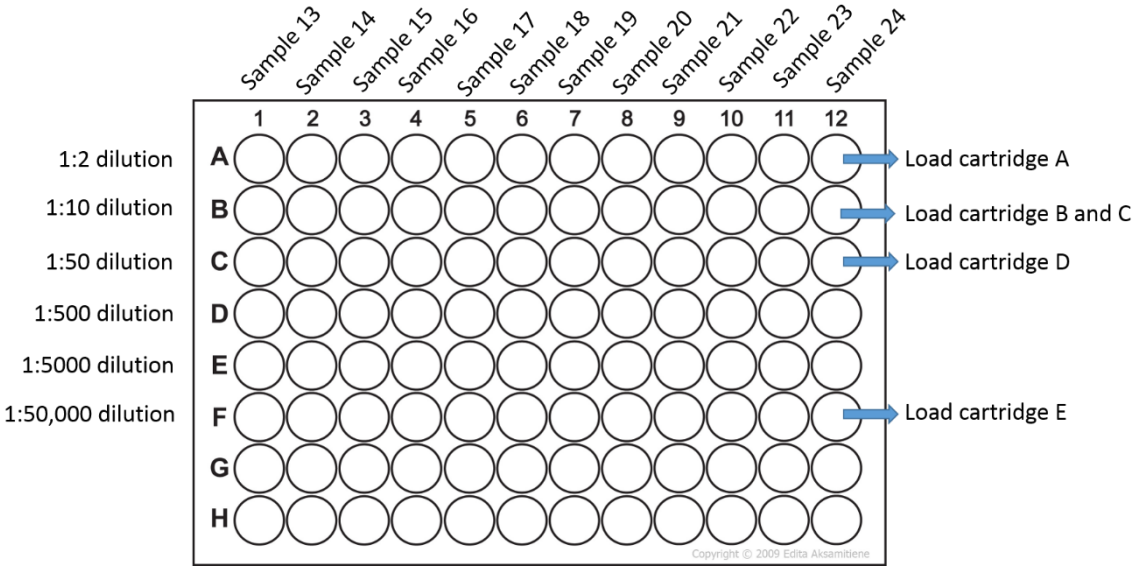


Figure 7. Dilution series for biomarker dilution optimization studies.

A


Sample 1
Sample 2
Sample 3
Sample 4
Sample 5
Sample 6
Sample 7
Sample 8
Sample 9
Sample 10
Sample 11
Sample 12
1 2 3 4 5 6 7 8 9 10 11 12
A
B
C
D
E
F
G
H
1:2 dilution
1:10 dilution
1:50 dilution
1:500 dilution
1:5000 dilution
1:50,000 dilution
Load cartridge A
Load cartridge B and C
Load cartridge D
Load cartridge E
Copyright © 2009 Edita Aksamitiene

B


Sample 13
Sample 14
Sample 15
Sample 16
Sample 17
Sample 18
Sample 19
Sample 20
Sample 21
Sample 22
Sample 23
Sample 24
1 2 3 4 5 6 7 8 9 10 11 12
A
B
C
D
E
F
G
H
1:2 dilution
1:10 dilution
1:50 dilution
1:500 dilution
1:5000 dilution
1:50,000 dilution
Load cartridge A
Load cartridge B and C
Load cartridge D
Load cartridge E
Copyright © 2009 Edita Aksamitiene

C

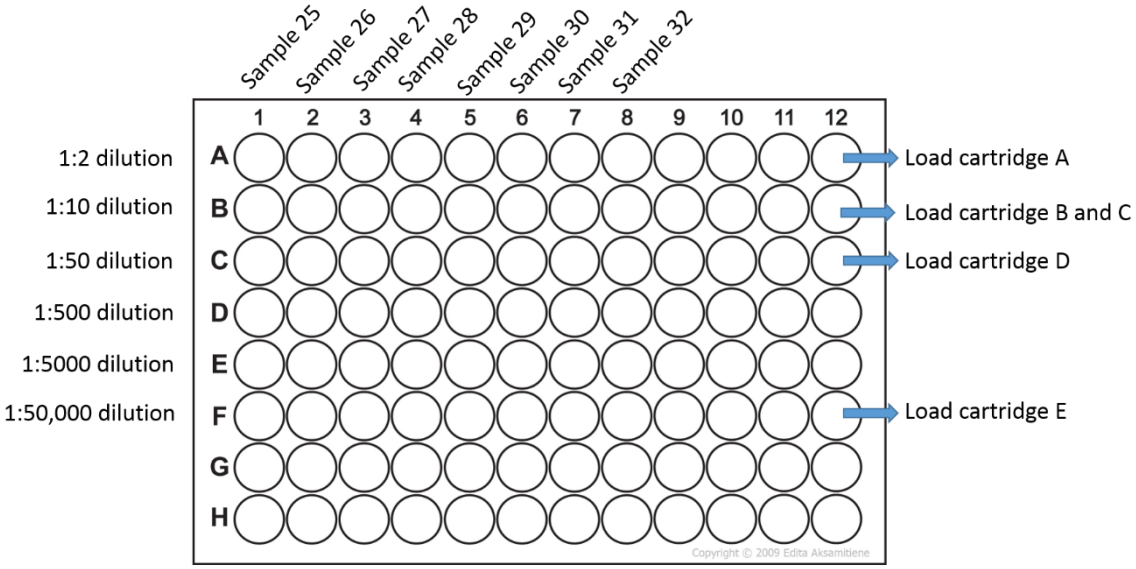


Figure 8. 96-well plate set-up for dilution series. To run 32 samples, a total of 3 96-well plates will be required for set up. Panel A, samples 1-12; Panel B, samples 13-24; Panel C, samples 25-32.

2.18 For each sample, combine 75 µl of serum to 75 µl of sample diluent (Figure 7 and Figure 8 row A). Using a multi-channel pipette, the remaining dilution series can be completed (follow Figure 7 dilution series in a 96-well plate), changing the tips between each dilution.

2.19 Load 50 µl diluted serum as follows onto the appropriate Ella cartridge:

- Cartridge A (VEGFR1/flt-1): load 1:2 dilution (row A in 96-well plate in figure 8)
- Cartridge B (Ang-2, IP-10, IL-10, IL-1ra, IL-6, IL-8) – load dilution 1:10 (row B in 96-well plate in Figure 8)
- Cartridge C (Procalcitonin, TREM-1) – load dilution 1:10 (row B in 96-well plate in Figure 8)
- Cartridge D (Ang-1, CHI3L1, TNFR1) – load dilution 1:50 (row C in 96-well plate in Figure 8)
- Cartridge E (CRP) – load dilution 1:50,000 (row F in 96-well plate in Figure 8)

2.20 Pipette 1 mL of wash buffer A into the oval-shaped wells on the cartridges (16 wash wells on 32 sample cartridges).

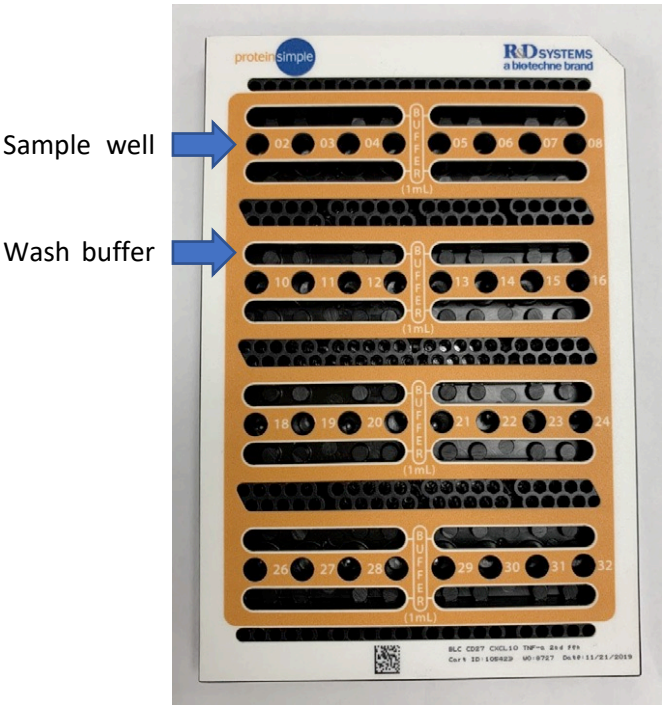


Figure 9. Cartridge layout of 32 sample plate (multi-analyte). Arrows denote the patient sample wells (circular) and wash buffer wells (oval/cylindrical)

2.21 Carefully carry the cartridge over to the instrument.

Loading the Cartridge on to the Ella System

2.22 Double check that the status light is still green on the front of the Ella instrument. If green, open the Ella door (blue lid on instrument represented in Figure 9 and 10).

2.23 Gently peel off the protective plastic lining at the bottom of the cartridge. Make sure the exposed bottom does not touch anything.

2.24 Lift the cartridge clamp and place the cartridge inside the instrument.

2.25 Slowly lower the cartridge clamp. Close the door. See figure 10 for the loaded cartridge and closed clamp.

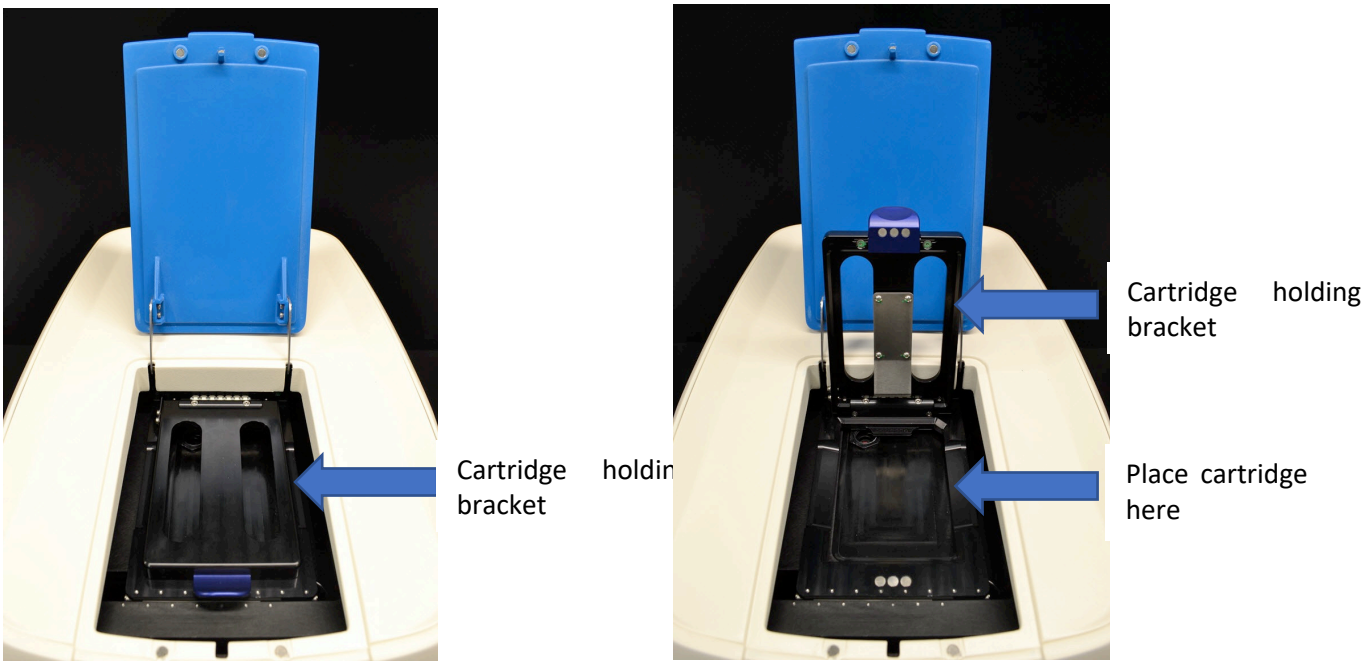


Figure 9. Ella door open displaying the cartridge clamp closed (A) and open (B).

Figure 10. Closed cartridge clamp.

2.26 Press start on the Confirm Selections screen (Figure 5).
2.27 Confirm you would like to start the run by pressing Start again in the Confirm dialogue box.
2.28 A dialogue box will now appear showing the status of the run (Figure 11). The approximate run time is 75 minutes.
2.29 Once the run is complete, the cartridge can be discarded in biohazardous waste for sharps.

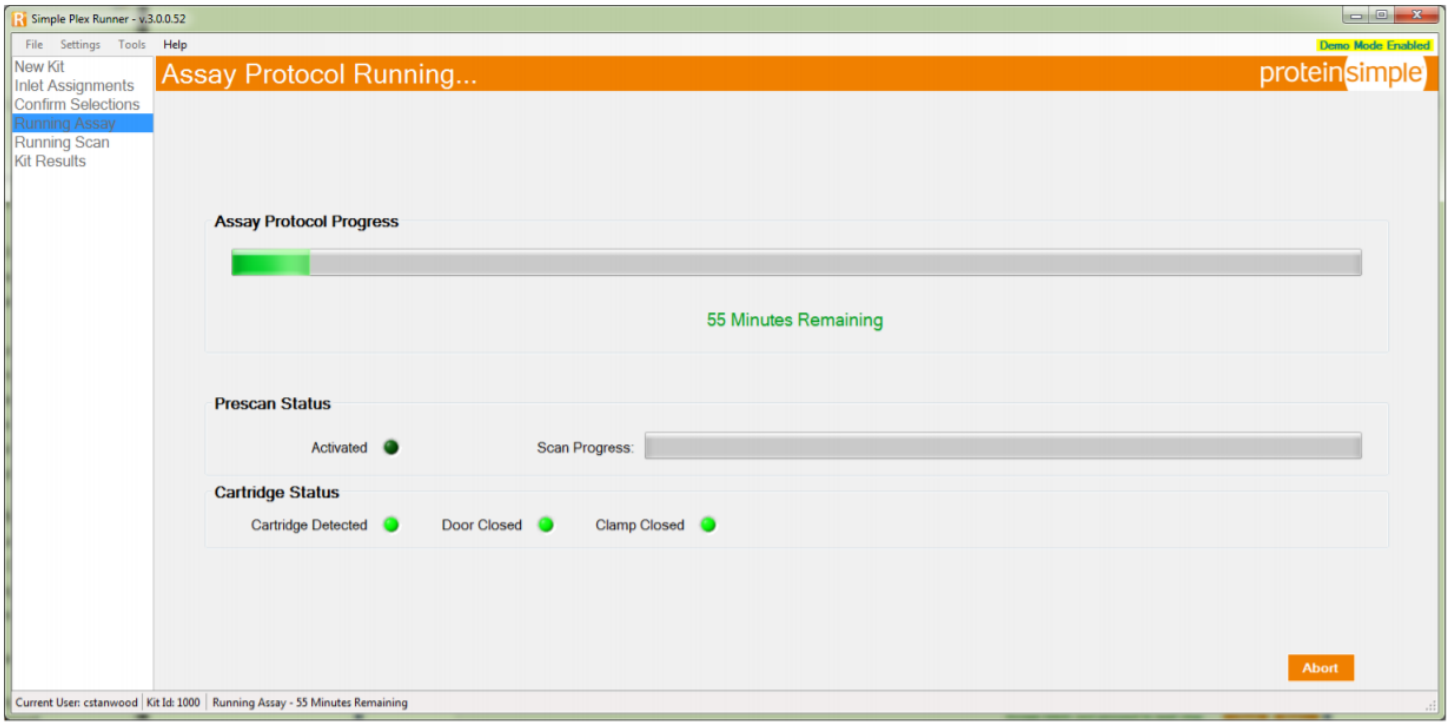


Figure 11. Running assay screen.

3. Viewing and Exporting Results

3.1 The run results will be initially displayed in a dashboard style format (Figure 12). Concentrations of each sample inlet will be displayed.

3.2 The export button allows export of results to a clipboard as a tab-delimited spreadsheet compatible with string. Save the data on a USB key and transfer to a secondary computer.

3.3 Compile all biomarker data for each run into an excel spreadsheet to generate a database for all samples.

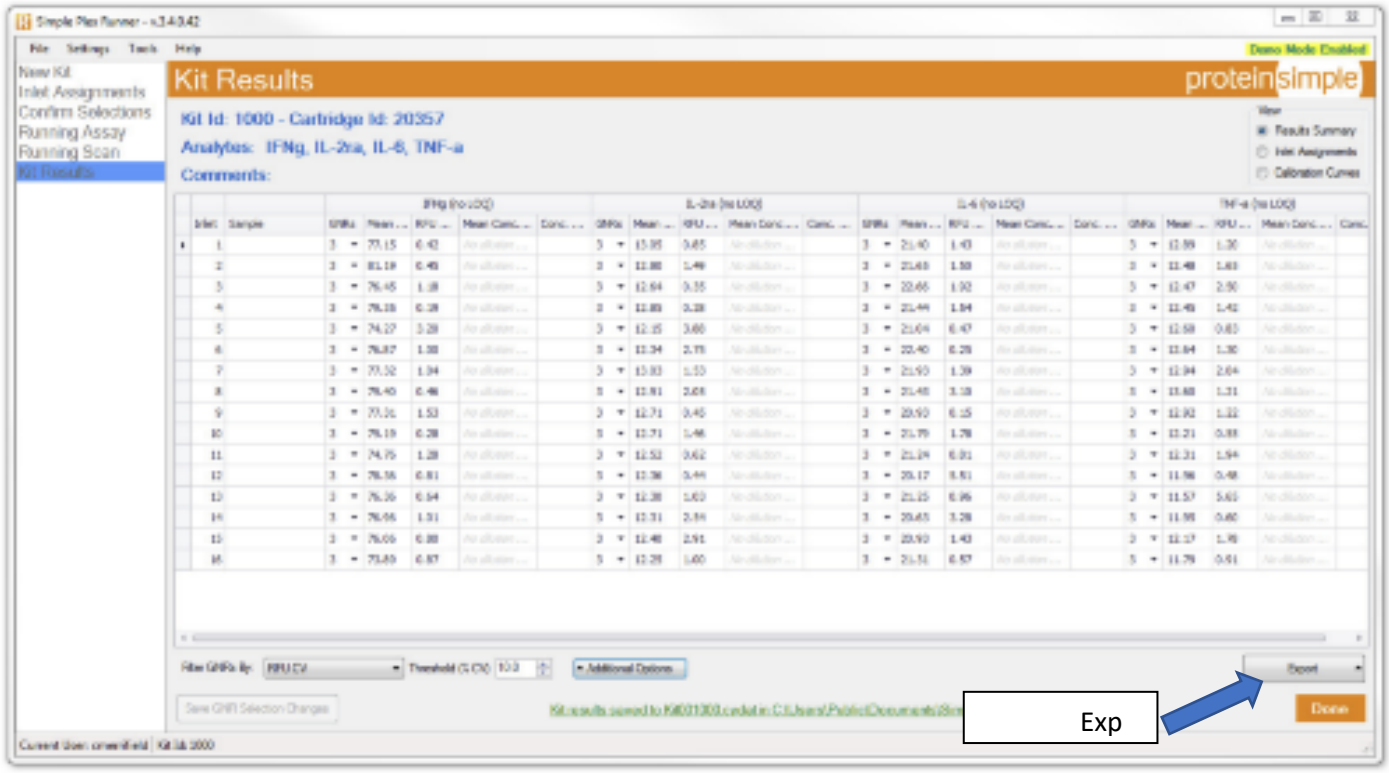


Figure 12. Cartridge results screen and export data option.

4. Instrument shutdown

4.1 To shut down Ella, close the Simple Plex Runner application by doing one of the following:
   i) Click File > Exit on the menu
   ii) Click Close (X) in the upper-right corner of the Simple Plex Runner application

4.2 Shut down Ella's computer

4.3 Turn off Ella by pressing the power switch on the back panel (Figure 1).

5. Biohazard Waste Management

5.1 Transfer used cartridges to a clearly labelled biohazard waste container that will undergo BSL-2 waste processing (e.g., autoclaving) prior to disposal. As the cartridges have the potential to tear biohazard bags, these bags should be handled similarly as pipette tips or other "sharp" biohazard waste.

5.2 Send for safe disposal to the validated biowaste destruction company.

6. Update history

Version	Date	Summary of changes
1.0	31 October 2021	Written and approved by MRG and AC
2.0	08 November 2021	Serum volumes updated
3.0	18 February 2022	Sample volumes adjusted based on optimisation results

9.7 TRIPOD checklist: Maela study – ARI cohort

Table 9.7-1: TRIPOD checklist for derivation and validation of the severity scores and clinical prediction models in the ARI cohort. ARI = acute respiratory infection; CI = confidence interval; D = derivation; TRIPOD = Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis; V = validation.

Section	Item	D / V	Checklist item	Page
Title and abstract				
Title	1	D, V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	NA
Abstract	2	D, V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	NA
Introduction				
Background and objectives	3a	D, V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	69-70
	3b	D, V	Specify the objectives, including whether the study describes the development or validation of the model or both.	70
Methods				
Source of data	4a	D, V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	71
	4b	D, V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	71
Participants	5a	D, V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	71
	5b	D, V	Describe eligibility criteria for participants.	71
	5c	D, V	Give details of treatments received, if relevant.	NA
Outcome	6a	D, V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	72-73
	6b	D, V	Report any actions to blind assessment of the outcome to be predicted.	72
Predictors	7a	D, V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	71-72
	7b	D, V	Report any actions to blind assessment of predictors for the outcome and other predictors.	NA
Sample size	8	D, V	Explain how the study size was arrived at.	86
Missing data	9	D, V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	81-83
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	84
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	84
	10c	V	For validation, describe how the predictions were calculated.	NA

Section	Item	D / V	Checklist item	Page
	10d	D, V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	84
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	84
Risk groups	11	D, V	Provide details on how risk groups were created, if done.	NA
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	76
Results				
Participants	13a	D, V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	87-88
	13b	D, V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	89-92
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	NA
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	89-92
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	99
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	105
	15b	D	Explain how to the use the prediction model.	106-109
Model performance	16	D, V	Report performance measures (with CIs) for the prediction model.	93-102
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	103-104
Discussion				
Limitations	18	D, V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	145-147
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	140-142
	19b	D, V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	147
Implications	20	D, V	Discuss the potential clinical use of the model and implications for future research.	147
Other information				
Supplementary information	21	D, V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	NA
Funding	22	D, V	Give the source of funding and the role of the funders for the present study.	6

9.8 STROBE checklist: Maela study – pneumonia cohort

Table 9.8-1: STROBE checklist for host biomarker analyses in the pneumonia cohort. STROBE = STrengthening the Reporting of OBservational Studies in Epidemiology.

Section	Item	Recommendation	Page
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	NA
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	NA
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	69-70
Objectives	3	State specific objectives, including any prespecified hypotheses	70
Methods			
Study design	4	Present key elements of study design early in the paper	71
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	71
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	71
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	NA
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	71-73
Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	71-72; 79
Bias	9	Describe any efforts to address potential sources of bias	NA
Study size	10	Explain how the study size was arrived at	162
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	NA
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	85-86
		(*b*) Describe any methods used to examine subgroups and interactions	85
		(*c*) Explain how missing data were addressed	83
		(*d*) If applicable, explain how loss to follow-up was addressed	87
		(*e*) Describe any sensitivity analyses	85
Results			
Participants	13	(a) Report numbers of individuals at each stage of study—e.g., numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	87
		(b) Give reasons for non-participation at each stage	88
		(c) Consider use of a flow diagram	88
Descriptive data	14	(a) Give characteristics of study participants (e.g., demographic, clinical, social) and information on exposures and potential confounders	113-117
		(b) Indicate number of participants with missing data for each variable of interest	83, 87
		(c) Summarise follow-up time (e.g., average and total amount)	NA

Section	Item	Recommendation	Page
Outcome data	15	Report numbers of outcome events or summary measures over time	114
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included	120-121
		(*b*) Report category boundaries when continuous variables were categorized	NA
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	NA
Other analyses	17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	122-139
Discussion			
Key results	18	Summarise key results with reference to study objectives	140; 142-144
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	146-147
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	147
Generalisability	21	Discuss the generalisability (external validity) of the study results	147
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	6

9.9 Case report form: PICU study

Variables Influencing Survival After Admission to Paediatric Intensive Care in a Resource-Limited Setting

Angkor Hospital for Children Retrospective PICU Study

SCREENING	
Inclusion criteria confirmed (all should be "Yes")	○
Aged > 28 days [day of birth = Day 1]	
Admitted to PICU between 1 January 2018 and 31 December 2019	
Exclusion criteria (all should be "No")	○
Elective admission (e.g., planned post-operative admission)	
Non-clinical reason for admission (e.g., admission due to lack of bed space on inpatient ward)	

Study ID number	|_R_|_P_|_S_|-|___|___|___|___|
AHC Hospital ID number	|___|___|___|___|-|___|___|___|___|___|___|

Date |___|___|-|___|___|___|-|___|___|___|___|	**Eligibility confirmed by** |___|___|

DEMOGRAPHICS					
Date of Birth	|___|___|-|___|___|___|-|___|___|___|___|	**Sex**	Male ○ Female ○ Unknown ○		
Province	_____	**District**	_____	**Commune**	_____

PERINATAL HISTORY [**NR** = Not recorded; **wk** = weeks; **mo** = months; **PT** = preterm; **FT** = full term; **LBW** = low birth weight]	
Gestational age	|___|___| wk |___| mo **If not documented:** PT ○ FT ○ NR ○
Birth weight (BW)	|___|.|___| kg **If not documented:** LBW ○ Normal BW ○ NR ○

MEDICAL HISTORY [**NR** = Not recorded]		
Comorbidity	Yes ○	1. _____
	No ○ **If yes, what:**	2. _____
	NR ○	3. _____

HEALTH JOURNEY PRIOR TO AHC [**NR** = Not recorded; **DOA =** Date of Admission]	
Date of AHC admission	|___|___|-|___|___|___|-|___|___|___|___|
Duration of illness prior to AHC admission [incl. DOA]	|___|___| days NR ○
Number of care providers prior to AHC admission [select all that apply]	Traditional ○ Pharmacy ○ Private clinic ○ KBH ○ |___|___| **Type:** HC / HP ○ Govt. hosp. ○ AHC ○ SC ○ None / NR ○ Other ○ _____
Admission overnight at another facility prior to AHC	Yes ○ No / NR ○
Transfer from another facility to AHC	Yes ○ No / NR ○ **If Yes:** SC ○ Other ○ Self-referral ○
Route of admission to PICU	ER ○ IPD ○ Surgical unit ○ Operating theatre ○ LAU ○ Other ○

PICU ADMISSION [observations at time PICU admission decision made]						Cardiac arrest at time of admission ○
Date |___|___|-|___|___|___|-|___|___|___|___| **Time** |___|___|:|___|___|						
Temperature	|___|___|.|___| °C	**Heart rate**	|___|___|___| bpm	**Respiratory rate**	|___|___| bpm	
Blood Pressure	|___|___|___| / |___|___|___| mmHg			**SpO_2**	|___|___|___| %	
FiO_2	|___|___| LPM	|___|___| %	|___| cmH_20	RA ○	NC ○	FM ○ CPAP ○ Ventilated ○
Alert [MD to complete]	Yes ○	No ○	Unknown ○	**GCS**	|___|___|	**Blood glucose** |___|___|___| mg/dL Hi ○ Low ○
Weight	|___|___|.|___| kg	**Height / Length**	|___|___|___|.|___| cm			

VITAL SYSTEMS [**Y** = Yes; **N** = No; **NR** = Not recorded; **ND** = Not detected]									
	Y	**N**	**NR**		**Y**	**N**	**NR**		**Y**	**N / NR**
Lung crackles	○	○	○	**Stridor**	○	○	○	**Active seizures**	○	○
Resp. distress	○	○	○	**Cool extremities**	○	○	○	**History of seizure**	○	○
								History of fever	○	○
Capillary refill time	|___| seconds			Normal / < 2 sec ○	Rapid ○	> 2 sec ○		**Pulse**	Strong ○	Fair ○ Weak ○ ND ○

CLINICIAN ASSESSMENT	
Diagnosis on admission to PICU [select all that apply]	Bronchiolitis ○ Pneumonia ○ Gastroenteritis ○ Meningo-encephalitis ○ Sepsis ○ SSTI ○ Melioidosis ○ Dengue ○ Asthma / RAD ○ Heart failure ○ Poisoning ○ Trauma / accident ○ Malnutrition ○ HCAI ○ Other _____
Reason for PICU admission [to be completed by MD, select all that apply]	Respiratory distress / support ○ Shock / circulatory support ○ Seizing / post-ictal ○ Impaired consciousness ○ Trauma / accident ○ Unclear ○ Other _____

LABORATORY RESULTS [closest result to PICU admission]			
Complete blood count	**Date** |___|___|-|___|___|___|-|___|___|___|___|		**Time** |___|___|:|___|___|
	WBC |___|___|.|___| x10^9/L	**Lymphocyte** |___|___|.|___| x10^9/L	
	Hb |___|___|___| g/L	**Neutrophil** |___|___|.|___| x10^9/L	
	Platelets |___|___|___| x10^9/L		
Coagulation studies	**Date** |___|___|-|___|___|___|-|___|___|___|___|		**Time** |___|___|:|___|___|
	INR |___|___|.|___|	**APTT** |___|___|.|___| seconds	
Biochemistry	**Date** |___|___|-|___|___|___|-|___|___|___|___|		**Time** |___|___|:|___|___|
	Urea |___|___|.|___| mmol/L	**Total bilirubin** |___|___|___| µmol/L	
	Creatinine |___|___|___| µmol/L ○ <30	**CRP** |___|___|___| mg/L	

FIRST 24H TREND **Date** |___|___|-|___|___|___|-|___|___|___|___| **Time** |___|___|:|___|___|					
TIME	**0h**	**+6h** -1h □ +1h □	**+12h** -1h □ +1h □	**+18h** -1h □ +1h □	**+24h** -1h □ +1h □
Temperature	|___|___|.|___|	|___|___|.|___|	|___|___|.|___|	|___|___|.|___|	|___|___|.|___|
Heart rate	|___|___|___|	|___|___|___|	|___|___|___|	|___|___|___|	|___|___|___|
Blood pressure	|___|___|___| / |___|___|___|	|___|___|___| / |___|___|___|	|___|___|___| / |___|___|___|	|___|___|___| / |___|___|___|	|___|___|___| / |___|___|___|
Respiratory rate	|___|___|	|___|___|	|___|___|	|___|___|	|___|___|
SpO_2	|___|___|___|	|___|___|___|	|___|___|___|	|___|___|___|	|___|___|___|
FiO_2 **LPM**	|___|___|	|___|___|	|___|___|	|___|___|	|___|___|
cmH_20	|___|___|	|___|___|	|___|___|	|___|___|	|___|___|
%	|___|___|___|	|___|___|___|	|___|___|___|	|___|___|___|	|___|___|___|
Mode	RA ○ NC ○ FM ○ CPAP ○ Ventilated ○	RA ○ NC ○ FM ○ CPAP ○ Ventilated ○	RA ○ NC ○ FM ○ CPAP ○ Ventilated ○	RA ○ NC ○ FM ○ CPAP ○ Ventilated ○	RA ○ NC ○ FM ○ CPAP ○ Ventilated ○
Hi temp last 6h	|___|___|.|___|	|___|___|.|___|	|___|___|.|___|	|___|___|.|___|	|___|___|.|___|
Low temp last 6h	|___|___|.|___|	|___|___|.|___|	|___|___|.|___|	|___|___|.|___|	|___|___|.|___|
No. PU last 6h	|___|___| FC ○	|___|___| FC ○	|___|___| FC ○	|___|___| FC ○	|___|___| FC ○
Crystalloid bolus (ml/kg last 6h)	|___|___|	|___|___|	|___|___|	|___|___|	|___|___|

Colloid bolus (ml/kg last 6h)	|___|___|	|___|___|	|___|___|	|___|___|	|___|___|
RBC last 6h	Yes ○ No ○	Yes ○ No ○	Yes ○ No ○	Yes ○ No ○	Yes ○ No ○

PICU STAY [continue on separate sheet if required]

ORGAN SUPPORT		**DATE**	**TIME**
Ventilation	START (1)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
	STOP (1)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
	START (2)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
	STOP (2)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
CPAP	START (1)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
	STOP (1)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
	START (2)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
	STOP (2)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
Inotropes	START (1)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
	STOP (1)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
	START (2)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
	STOP (2)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
Peritoneal dialysis	START (1)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
	STOP (1)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
Surgery required		Yes ○ No ○	

To be considered a 'new episode', cessation of therapy must be for:

- Ventilation: ≥ 48 hours
- CPAP: ≥ 24 hours
- Inotropes: ≥ 24 hours

PICU DISCHARGE	**Date** |___|___|-|___|___|___|-|___|___|___|___|
Alive	Yes ○ No ○ **If Yes: Discharged home to die** Yes ○ No ○
Diagnosis [select all that apply]	Bronchiolitis ○ Pneumonia ○ Gastroenteritis ○ Meningo-encephalitis ○ Sepsis ○ SSTI ○ Melioidosis ○ Dengue ○ Asthma / RAD ○ Heart failure ○ Poisoning ○ Trauma ○ Malnutrition ○ HCAI ○ Other _____
Destination	AHC inpatient ward ○ Home ○ Transfer ○ LAMA ○

DEATH OR DISCHARGED HOME TO DIE DURING PICU STAY [to be completed by MD]	
New illness acquired during PICU stay	Yes ○ No ○ **If Yes:** End date of admission illness |___|___|-|___|___|___|-|___|___|___|___| Diagnosis of admission illness _____ Reason to remain in PICU _____ Date of onset of final illness |___|___|-|___|___|___|-|___|___|___|___|

HOSPITAL DISCHARGE	Discharged from AHC directly from PICU ○
Date	|___|___|-|___|___|___|-|___|___|___|___|
Alive	Yes ○ No ○ **If Yes: Discharged home to die** Yes ○ No ○
Diagnosis [select all that apply]	Bronchiolitis ○ Pneumonia ○ Gastroenteritis ○ Meningo-encephalitis ○ Sepsis ○ SSTI ○ Melioidosis ○ Dengue ○ Asthma / RAD ○ Heart failure ○ Poisoning ○ Trauma ○ Malnutrition ○ HCAI ○ Other _____
Destination	Home ○ Transfer ○ LAMA ○

12-MONTH FOLLOW-UP		
Mode of follow-up	Notes ○ Telephone consent ○ Telephone refusal ○ Uncontactable ○	
Alive	Yes ○ No ○	**If No:** **Date of death** |___|___|___|-|___|___|___|___| **Place of death** Community ○ Health facility ○

PICU STAY CONTINUATION SHEET			
ORGAN SUPPORT		**DATE**	**TIME**
Ventilation	START (3)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
	STOP (3)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
	START (4)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
	STOP (4)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
	START (5)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
	STOP (5)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
CPAP	START (3)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
	STOP (3)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
	START (4)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
	STOP (4)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
	START (5)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
	STOP (5)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
Inotropes	START (3)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
	STOP (3)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
	START (4)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
	STOP (4)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
	START (5)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
	STOP (5)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
Peritoneal dialysis	START (2)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
	STOP (2)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
	START (3)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|
	STOP (3)	|___|___|-|___|___|___|-|___|___|___|___|	|___|___|:|___|___|

To be considered a 'new episode', cessation of therapy must be for:

- Ventilation: ≥ 48 hours
- CPAP: ≥ 24 hours
- Inotropes: ≥ 24 hours

9.10 Admission proforma at Angkor Hospital for Children

Angkor Hospital for Children Admission Form

Patient Label with Address

ICU 2022

Date: ___/___/___ **Time**_____

Doctor: _____

Chief Complaint.....

History .....
.....
.....

Past Medical History / Birth History .....
.....
.....

Current Medications.....
.....

Past Medications.....
.....

Allergies .....

Yellow Card Reviewed: ☐ Yes, ☐ No **Growth Chart Plotted:** ☐ Yes, ☐ No

Immunizations:

(Birth)	(6weeks)	(10weeks)	(14weeks)	(6months)	(9months)	(18months)	
☐ BCG ☐ $HepB_0$	☐ DPT- HepB Hib_1 ☐ PCV_1 ☐ $Polio_1$	☐ DPT- HepB Hib_2 ☐ PCV_2 ☐ $Polio_2$	☐ DPT- HepB Hib_3 ☐ PCV_3 ☐ $Polio_3$ ☐ IPV	☐ MR_0	☐ MR_1 ☐ JE	☐ MR_2	☐ Unknown ☐ Other:.... ....

Development History

Gross Motor head control-rolls-sits-crawls-walks-runs-stairs-jumps

Fine Motor eyes fix-reaches-transfers-unilateral reach-pincer-throws-dresses

Speech/Hearing alert to sound –coos-laughs-babbles-1 word-2 words-4 to 6 words-many words-asked questions

Social looks at face-social smile-recognizes parent-explores-uses spoons & cup-plays-continent urine and stool

Based on above is development appropriate for age? ☐ Yes, ☐ No.

Further Development History.....

Social History .....

.....

.....

Family History ○ – □ ○ - female □ - male

Examination

Appearance .....

Temperature ..... ˚C HR..... RR..... BP...../..... O_2 Sat..... RA or ..... L/mn O_2

Wt ..... kg Height ..... cm Wt/Ht Z-score: .....SD

HEENT

Head Fontanelle.....Lymphadenopathy.....

Eyes R.....L.....Fundus.....

Ears R.....L.....

Nose

Throat

Other .....

Cardiovascular System

Pulse (circle) strong fair weak cannot detect

Heart Sounds .....

Capillary Refill Time .....

Other (JVP, precordium) .....

.....

Respiratory System

Auscultation .....

Grunting/flaring .....

Indrawing .....

Other (percussion, tracheal position) .....

Gastro-Intestinal System

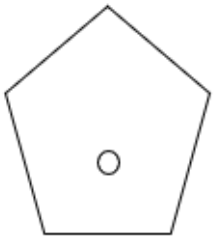
Inspection .....

Palpation .....

Organs .....

Bowel Sounds .....

Masses .....



Other (genitalia, rectum, hernia, ascites) .....

Nervous System

Mental Status ..... Kernig's .....
Neck Stiffness ..... GCS ..... (Use Chart Below)
Cranial Nerves .....

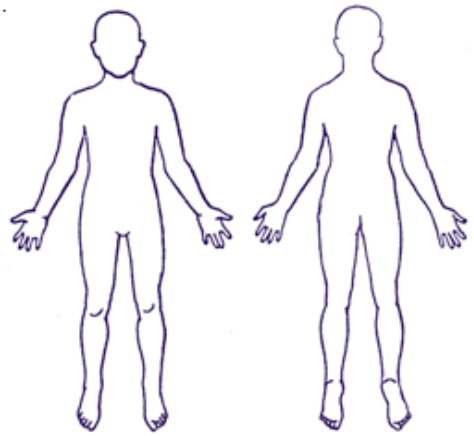
	1	2	3	4	5	6
EYES	Does not open	Opens in response to pain	Opens in response to voice	Opens spontaneously	X	X
VERBAL	Makes no sounds	Incomprehensible	Says inappropriate words	Confused and disoriented	Oriented and converses normally	X
MOTOR	No movements	Extension to pain	Abnormal flexion to pain	Flexion/Withdrawal to pain	Localizes painful stimuli	Obeys commands

Peripheral Nerves **R** arm **L** arm **R** leg **L** leg

Tone.....
Power.....
Reflexes.....
Sensation.....

Clonus .....
Babinski .....

Development .....
.....



Skin & Extremities

Rash .....
Edema .....
Wounds .....
Other .....
.....

Other.....
.....
.....

9.11 PICU admission proforma at Angkor Hospital for Children

មន្ទីរពេទ្យកុមារអង្គរ

Phreah Sangreach Tep Vong Road & Umchhay Street, Mondul 1, Sangkat Svay Dangkum P.O. Box 50, Siem Reap, Kingdom of Cambodia
Tel: (855) 63 963 409 Fax: (855) 63 760 452 Email: admin@angkorhospital.org http://www.angkorhospital.org

EMERGENCY / INTENSIVE CARE UNIT
Night time Pediatric Nursing Assessment

Yesterday's weight……….Kg Bed Nº:………..
Today's Weight…………..Kg IBW… ………Kg
Days hospitalized……….Days in ICU…..........

Diagnosis: ____
History : ____
____
____
____
____
____
____

Ventilation: Yes☐ No☐
ETT size………Tape……….
Setting: Mode…….Rate…….
PIP/ PEEP……….. Ti………
FiO2…….. VT…….Tpl……..
VMAX ______PS__________.
Plans …. ……………………
………………………………
………………………………
………………………………

ENVIRONMENTAL / SAFETY

☐ All alarms on, audible, and functioning ☐ Side rails up
☐ Pre-calculated drug sheet for patient weight
☐ Emergency equipments are at the bedside
☐ Hypothermic ☐ Hyperthermic Temp: ……….ºC

INTEGUMENTARY / MUSCULOSKELETAL Normal ☐

☐Dry ☐Rash ☐Lesion ☐Breakdown ☐Petechiae ☐Thrush
☐Edema/swelling ☐Other problems:………………………………
Location: ………………………………………………………………….

NEUROLOGICAL / FLACC PAIN SCALE Normal ☐

Anterior Fontanel ☐ Bulging ☐ Sunken
Movement of Extremities ☐ Weak ☐ Absent ☐ Limit ROM
☐ Developmental Delay / CP ☐ Sedated ☐ Paralyzed
GCS Scale: ………../15
Pupils reaction ☐ Slow ☐ Fixed ☐ Dilated
☐ Unequal R:….….mm L:………mm
Muscle tone ☐ Hypertonic ☐ Hypotonic ☐ Quiet
☐ Flaccid ☐ Unresponsive
☐other…………………………………………………………………

Category	Score 0	Score 1	Score 2
Face	No particular expression or smile	Occasional grimace or frown, withdraw, disinterested	Frequent to constant frown, clenched jaw, quivering chin
Legs	Normal position or Relaxed	Uneasy Restless Tense	Kicking Or Legs drawn up
Activity	Lying quietly Normal position Moves easily	Squirming Shifting back/forth Tens	Arched Rigid Or Jerking
Cry	No Cry Awake or asleep	Moans or Whimpers Occasional complaint	Crying Steadily Screams or Sobs Frequent Complaint
Consol ability	Content Relaxed	Reassured by occasional touching, hugging or ' talking to' Distractible	Difficult to console or comfort

COMFORT / SEDATION / IV SITE Normal ☐

☐ Needs pain killer ☐ Needs sedation ☐ Needs paralysis
Lines: ☐ Erythema ☐ Skin damage ☐ Swelling
☐ Central ☐ Arterial ☐ IO
Location: ……………………Others………………………………

FLUIDS / NUTRITION Normal ☐

☐ Abnormal
Patient is receiving total fluid. …………………………cc/kg/day
☐ Fluid restricted: ………………………………………………….
☐ Extra fluid:………………………………………………………
☐ Abnormal Serum Electrolyte
Na: ……….. K: ……….. Ca: ……….. mmol/L ☐ Dr. Notified
☐ Fluid bolus / Blood products required:..………………………..

OTHER ASSESSMENTS Normal ☐

☐ Food support ☐ Abondoned child ☐ No parents visit
………………………………………………………………………

RESPIRATORY SYSTEM Normal ☐

☐ Cough ☐ Dyspnea ☐ Slow Breathing
☐ Apnea ☐ Nasal Flaring ☐ Tachypnea
☐ Stridor ☐ Grunting ☐ Gasping
☐ Insufficient Respiratory Effort ☐ Thorax asymmetrical
O2supply ☐ Nasal Cannula ☐ NCPAP ……..CmH2O
☐ Face Mask ☐ Face Mask with bag
☐ Flow: ……………….LPM
Retraction ☐ Suprasternal ☐ Supraclavicular
☐ Intercostal ☐ Subcostal
☐ Substernal
Lungs' sound ☐ Crackle ☐ Left ☐ Right
☐ Wheezing ☐ Left ☐ Right
☐ Decreased ☐ Left ☐ Right
☐ Other sound: ………………………….........
Ventilator: ETT Size: …………mm Depth: …………...cm
ET Suction size: …..….Fr Depth: …………...cm
Mode: ……………………PIP/PEEP: ……………cmH2O
Rate: …………. FiO2: ………..% TV: ………………
IT: ………..sec Other settings: ……………….……….
Chest tube: Size: ……..Fr ☐ Right ☐ Left ☐ Suction
☐ Gravity ☐ Chest fluid color: ……………………
Others: ……………………………………………………………….

CARDIOVASCULAR SYSTEM Normal ☐

☐ Tachycardia ☐ Bradycardia ☐ Murmur ☐ CRT>2sec
☐ Hypertensive ☐ Hypotensive
☐ Active precordium ☐ BP not detectable
☐ Cool extremities Last Hct:…………% On …………………..
☐ Sweating ☐ JVP
Color: ☐ Mottled ☐ Pallor ☐ Jaundice
Pulse: ☐ Weak ☐ Unequal ☐ Not palpable
Cyanosis: ☐ Peripheral ☐ Central
Others: ………… ………………………………………………….

GASTROINTESTINAL / GENITOURINARY Normal ☐

☐ Vomiting ☐ Constipation ☐ AG…...cm
☐ Poor appetite ☐ Colostomy ☐ NPO
☐ Imperforate Anus
Abdomen ☐ Pain ☐ Firm ☐ Distended
☐ Hepatomegaly ☐ Spleen palpable
Bowel sound ☐ Hypoactive ☐ Hyperactive ☐ Absent
Stool ☐ Watery ☐ Mucusy ☐ Bloody
☐ Loose ☐ Black
Feeding Tube ☐ Oral ☐ Nasal size: …..Fr, Day……..
Urine ☐ Urine<1ml/kg/h ☐ Urine>5ml/kg/h ☐ Anuria
☐ Cloudy ☐ Dark Yellow ☐ Hematuria
Foley catheter Size: ……………Fr, Day…………
Others: ………………………………………………………………

ASSESSMENT COMPLETED BY

Initial :.. ………..**Signature**:………………… ….**Time:**………………

Date:..... **Bed Number:**..... **Name:**..... **Age**:.....

Hours			7	8	9	10	11	12	13	14	15
Pain / Sedation Assess **NN** = No Need **PN** = Pain Need **SN** = Sedation Need GTT= Continuous Drip **Color** **Abd. Description** **PK** = Pink **S** = Soft **W** = Pale **T** = Tense **C** = Cyanotic **D** = Distended **M** = Mottled	**Temperature**										
	Patient Position										
	Heart Rate										
	NBP / ABP										
	Color / MBP										
	CRT/ CVP										
	Abd. Description / AG										
	Pain Scale Scores (1-10)										
	Pain / Sedation										
R **Amount** **Tolerated** E **S** = Small **W** = Well S **M** = Medium **F** = Fair P **L** = Large **P** = Poor I **1** = Thick **2** = Thin R	**Set RR**	**Total RR**									
	SpO_2 %	**FIO_2 %**									
	Mode										
	P_{PEAK} / P_{MEAN} (CmH_2O)										
	PEEP (CmH_2O)										
	I : E ratio		:	:	:	:	:	:	:	:	:
A **Description** T C = Clear **Y**= Yellow O **W**= White **B** = Brown R **R** = Bloody **G** = Green	**I Time / T_{PL}**										
	□TV □V_{TE} □Insp TV (ml)										
	□Flow □V_{ETOT} (LPM)										
Y **Numeric pain Scale** **0** =No pain, **10**= worst pain	**Oral Care/Pt cir drain/O_2 check**										
	Suction	Amount									
	Tolerate	Describe									
P_{PEAK}: Peak Insp Pressure **P_{MEAN}**: Mean Airway Pressure **T_{PL}**: Plateau Time **V_{ETOT}**: Exhaled Minute Volume **V_{TE}**: Exhaled Tidal Volume	**NG Measurement Method** 1.Aspirate fluid from stomach 2.Listening by stetoscope 3.pH check (Normal ≤5)										
GTTS:	**mcg/mg | kg | hr/min**										
GTTS:	**mcg/mg | kg | hr/min**										
GTTS:	**mcg/mg | kg | hr/min**										
GTTS:	**mcg/mg | kg | hr/min**										
F **Method** PO=By mouth E Bt=Bottle BF=Breastfed E OG/NG= Oral/Nasal Gastric D C: continuous B: Bolus	**NG/OG/PO**	**Residual**									
	Amount In	**Total**									
	Food type/Feed type		I	I	I	I	I	I	I	I	I
	Post feeding position										
	NG Measurement method		I I	I I	I I	I I	I I	I I	I I	I I	I I
IV Line Assessment **S**= Soft **P** = Patent L= Locked **I** = Infiltrated	**NG mark / Co-sign by TL/SN**		I	I	I	I	I	I	I	I	I
	1	3	I	I	I	I	I	I	I	I	I
	2	4	I	I	I	I	I	I	I	I	I
I	Amount In	Total	I	I	I	I	I	I	I	I	I
N	Amount In	Total	I	I	I	I	I	I	I	I	I
T	Amount In	Total	I	I	I	I	I	I	I	I	I
A	Amount In	Total	I	I	I	I	I	I	I	I	I
K	Amount In	Total	I	I	I	I	I	I	I	I	I
E	Amount In	Total	I	I	I	I	I	I	I	I	I
Blood Product1	Amount In	Total	I	I	I	I	I	I	I	I	I
Blood Product 2	Amount In	Total	I	I	I	I	I	I	I	I	I
O **Stool** U **Description** T s = soft P w = watery U r = bloody T l = loose m= mucus bl= black	**Right CT**	Amount Out	Total	I	I	I	I	I	I	I	I	I
	Left CT	Amount Out	Total	I	I	I	I	I	I	I	I	I
	OG / NG	Amount Out	Total	I	I	I	I	I	I	I	I	I
		Amount Out	Total	I	I	I	I	I	I	I	I	I
	Urine(ml)	Amount Out	Total	I	I	I	I	I	I	I	I	I
	Urine in cc/kg/h											
	Stool S M L	Description		I	I	I	I	I	I	I	I	I
Intake Output Balance											

9.12 TRIPOD checklist: PICU study

Table 9.12-1: TRIPOD checklist for derivation and validation of the existing severity scores and derivation of the new clinical prediction model. CI = confidence interval; D = derivation; TRIPOD = Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis; V = validation.

Section	Item	D / V	Checklist item	Page
Title and abstract				
Title	1	D, V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	NA
Abstract	2	D, V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	NA
Introduction				
Background and objectives	3a	D, V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	148-149
	3b	D, V	Specify the objectives, including whether the study describes the development or validation of the model or both.	149
Methods				
Source of data	4a	D, V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	149-150
	4b	D, V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	149
Participants	5a	D, V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	149-150
	5b	D, V	Describe eligibility criteria for participants.	149
	5c	D, V	Give details of treatments received, if relevant.	149-150
Outcome	6a	D, V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	150-151
	6b	D, V	Report any actions to blind assessment of the outcome to be predicted.	150
Predictors	7a	D, V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	150
	7b	D, V	Report any actions to blind assessment of predictors for the outcome and other predictors.	150
Sample size	8	D, V	Explain how the study size was arrived at.	162
Missing data	9	D, V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	159-160
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	161
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	161

Section	Item	D / V	Checklist item	Page
	10c	V	For validation, describe how the predictions were calculated.	NA
	10d	D, V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	161
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	NA
Risk groups	11	D, V	Provide details on how risk groups were created, if done.	NA
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	NA
Results				
Participants	13a	D, V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	163
	13b	D, V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	163-174
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	NA
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	163-167
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	181
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	182
	15b	D	Explain how to the use the prediction model.	183-186
Model performance	16	D, V	Report performance measures (with CIs) for the prediction model.	183-185
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	NA
Discussion				
Limitations	18	D, V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	191-192
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	NA
	19b	D, V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	188-191
Implications	20	D, V	Discuss the potential clinical use of the model and implications for future research.	192
Other information				
Supplementary information	21	D, V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	NA
Funding	22	D, V	Give the source of funding and the role of the funders for the present study.	6

9.13 Case report form: PRIORITISE study

PRIORITISE

PRognost**I**cation of **O**xygen **R**equ**I**remen**T** **I**n non-severe **S**ARS-CoV-2 inf**E**ction

ELIGIBILITY CHECK		
Inclusion criteria (all should be "Yes")	**YES**	**NO**
Aged ≥ 18 years, and willing and able to give informed consent and comply with study procedures	○	○
RT-PCR or antigen test positive for SARS-CoV-2 during current illness[1]	○	○
Systemic manifestation of SARS-CoV-2 infection defined as: Breathing difficulty **OR** History of fever during current illness **AND** chest pain OR abdominal pain OR loose stool OR severe myalgia	○	○
Exclusion criteria (all should be "No")	**YES**	**NO**
Requires supplemental oxygen[2] or mechanical ventilation (invasive / non-invasive) at presentation	○	○
Laboratory confirmed SARS-CoV-2 infection (virological or serological) during a previous illness episode	○	○

Participant's eligibility confirmed?	Yes ○ No ○
Participant study ID	|___|___|___|-|___|___|___|___|
Enrolled by (initials)	|___|___|

NOTES

○ = Single selection permitted

☐ = Multiple selections permitted

[1] Patients presenting without virologically-confirmed SARS-CoV-2 but meeting all other eligibility criteria will be recruited. They will be removed from the study if they are subsequently confirmed to be negative for SARS-CoV-2 via RT-PCR.

[2] $SpO_2 \leq 93\%$ OR respiratory rate > 30 breaths per minute OR clinical decision to give supplemental oxygen.

ENROLMENT

DATE: |___|___|-|___|___|-|___|___|___|___|

PARTICIPANT BACKGROUND

Age	|___|___| years	**Sex**	Male ○ Female ○	**If Female: Pregnant**	Yes ○ No ○
Location of residence	_____				

VITAL SIGNS

Time of measurement |___|___|:|___|___|

Respiratory rate	|___|___| bpm	**Oxygen saturation (room air)**	|___|___|___| %
Heart rate	|___|___|___| bpm	**Blood Pressure**	|___|___|___| / |___|___|___| mmHg
Axillary temperature	|___|___|___|.|___| °F	**Mental status**	Alert ○ Voice ○ Pain ○ Unresponsive ○
Weight	|___|___|___|.|___| kg	**Height**	|___|___|___| cm

PRESENTING CLINICAL SYNDROME

	Yes	No		Yes	No
History of fever	○	○	**Headache**	○	○
Cough	○	○	**Altered consciousness**	○	○
Sore throat	○	○	**Seizures**	○	○
Runny nose	○	○	**Abdominal pain**	○	○
Ear pain	○	○	**Nausea / vomiting**	○	○
Wheezing	○	○	**Diarrhoea**	○	○
Chest pain	○	○	**Conjunctivitis**	○	○
Muscle aches	○	○	**Skin rash**	○	○
Joint pain (arthralgia)	○	○	**Skin ulcers**	○	○
Fatigue / malaise	○	○	**Lymphadenopathy**	○	○
Shortness of breath	○	○	**Loss of smell**	○	○
Lower chest indrawing	○	○	**Loss of taste**	○	○
Anorexia / loss of appetite	○	○	**Other**	_____	
Onset of first symptom	|___|___| days		**Impact on daily activities**	Yes ○	No ○

CURRENT / RECENT MEDICATIONS	
New medications taken in last 14 days	Steroids ☐ Azithromycin ☐ Remdesivir ☐ Lopinavir / Ritonavir (Kaletra) ☐ Hydroxychloroquine / Chloroquine ☐ None ☐ Other ☐ _____ **If Other, please specify:** Intravenous ☐ Intramuscular ☐ Oral ☐

PAST MEDICAL HISTORY					
Current smoker	Yes ○ No ○ **If No, please specify**: Former Smoker ○ Never smoked ○				
Known comorbidities	Yes ○ No ○				
If Yes:	**Yes**	**No**		**Yes**	**No**
Cardiovascular disease	○	○	**Chronic kidney disease**	○	○
Diabetes	○	○	**Chronic neurological disorder**	○	○
Hypertension	○	○	**HIV**	○	○
Malignant neoplasm	○	○	**Other immunosuppression**	○	○
Chronic lung disease	○	○	**Liver disease**	○	○
Asthma	○	○	**Active TB**	○	○
Other	_____				

ENROLMENT SAMPLES	
Venous blood sample	Collected ○ Not collected ○ **Time of collection** | ___ | ___ | : | ___ | ___ |
Respiratory swab	Collected ○ Not collected ○

HEALTH WORKER DECISION			
Admission decision	Admit ○ Not admit ○	**If admitted, reason:**	Clinical ○ Public health / isolation ○

ENROLMENT SAMPLES

Date of full blood count collection	|___|___|-|___|___|-|___|___|___|___|		
Leukocyte count	|___|___|.|___| x10^3/µL	**Lymphocyte count**	|___|___|.|___| x10^3/µL
Neutrophil count	|___|___|.|___| x10^3/µL	**Platelet count**	|___|___|___| x10^3/µL
Date of respiratory swab collection	|___|___|-|___|___|-|___|___|___|___|		
Respiratory swab (SARS-CoV-2 RT-PCR)		Positive ○ Negative ○ **If positive:** C_t value |___|___|.|___|	

<table>
<tr><th>DAILY FU</th><th colspan="3">DAY 0</th><th colspan="3">DAY 1</th><th colspan="3">DAY 2</th></tr>
<tr><td>Date of follow-up</td><td colspan="3">|__|__| /|__|__| / 20 |__|__|</td><td colspan="3">|__|__| /|__|__| / 20 |__|__|</td><td colspan="3">|__|__| /|__|__| / 20 |__|__|</td></tr>
<tr><td>Discharged</td><td colspan="3">☐YES ☐NO</td><td colspan="3">☐YES ☐NO</td><td colspan="3">☐YES ☐NO</td></tr>
<tr><td>Alive</td><td colspan="3">☐YES ☐NO</td><td colspan="3">☐YES ☐NO</td><td colspan="3">☐YES ☐NO</td></tr>
<tr><td>Ventilated</td><td colspan="3">☐YES ☐NO</td><td colspan="3">☐YES ☐NO</td><td colspan="3">☐YES ☐NO</td></tr>
<tr><td>Sx resolved</td><td colspan="3">☐YES ☐NO</td><td colspan="3">☐YES ☐NO</td><td colspan="3">☐YES ☐NO</td></tr>
<tr><td>Supplemental O_2</td><td colspan="3">☐YES ☐NO</td><td colspan="3">☐YES ☐NO</td><td colspan="3">☐YES ☐NO</td></tr>
<tr><td>*First RR > 30</td><td colspan="3">|__|__| bpm Time: |__|__|:|__|__|</td><td colspan="3">|__|__| bpm Time: |__|__|:|__|__|</td><td colspan="3">|__|__| bpm Time: |__|__|:|__|__|</td></tr>
<tr><td>*Highest RR >30</td><td colspan="3">|__|__| bpm Time: |__|__|:|__|__|</td><td colspan="3">|__|__| bpm Time: |__|__|:|__|__|</td><td colspan="3">|__|__| bpm Time: |__|__|:|__|__|</td></tr>
<tr><td>*First $SpO_2 \leq 93\%$</td><td colspan="3">|__|__| % Time: |__|__|:|__|__|</td><td colspan="3">|__|__| % Time: |__|__|:|__|__|</td><td colspan="3">|__|__| % Time: |__|__|:|__|__|</td></tr>
<tr><td>*Lowest $SpO_2 \leq 93\%$</td><td colspan="3">|__|__| % Time: |__|__|:|__|__|</td><td colspan="3">|__|__| % Time: |__|__|:|__|__|</td><td colspan="3">|__|__| % Time: |__|__|:|__|__|</td></tr>
<tr><td rowspan="2">Current medication</td><td colspan="9">0. None 1.Steroids 2. Azithromycin 3. Remdesivir 4. Lopinavir/Ritonavir 5. Hydroxychloroquine/Chloroquine 6. Convalescent Plasma 7. Other</td></tr>
<tr><td colspan="3">☐0 ☐1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7
Others________________________________</td><td colspan="3">☐0 ☐1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7
Others________________________________</td><td colspan="3">☐0 ☐1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7
Others________________________________</td></tr>
<tr><td rowspan="2">**O_2 delivery</td><td colspan="9">1.NC 2. FM 3. Venturi 4. HFNO/NIV 5. Ventilated 6. None</td></tr>
<tr><td>☐1 ☐ 2
☐3 ☐ 4
☐5 ☐ 6</td><td>☐1 ☐ 2
☐3 ☐ 4
☐5 ☐ 6</td><td>☐1 ☐ 2
☐3 ☐ 4
☐5 ☐ 6</td><td>☐1 ☐ 2
☐3 ☐ 4
☐5 ☐ 6</td><td>☐1 ☐ 2
☐3 ☐ 4
☐5 ☐ 6</td><td>☐1 ☐ 2
☐3 ☐ 4
☐5 ☐ 6</td><td>☐1 ☐ 2
☐3 ☐ 4
☐5 ☐ 6</td><td>☐1 ☐ 2
☐3 ☐ 4
☐5 ☐ 6</td><td>☐1 ☐ 2
☐3 ☐ 4
☐5 ☐ 6</td></tr>
<tr><td>**Flow rate (L/min)</td><td>|__|__|.|__|</td><td>|__|__|.|__|</td><td>|__|__|.|__|</td><td>|__|__|.|__|</td><td>|__|__|.|__|</td><td>|__|__|.|__|</td><td>|__|__|.|__|</td><td>|__|__|.|__|</td><td>|__|__|.|__|</td></tr>
<tr><td>**FiO_2 (%)</td><td>|__|__|__|</td><td>|__|__|__|</td><td>|__|__|__|</td><td>|__|__|__|</td><td>|__|__|__|</td><td>|__|__|__|</td><td>|__|__|__|</td><td>|__|__|__|</td><td>|__|__|__|</td></tr>
<tr><td>**Lowest SpO_2 (%)</td><td>|__|__|__|</td><td>|__|__|__|</td><td>|__|__|__|</td><td>|__|__|__|</td><td>|__|__|__|</td><td>|__|__|__|</td><td>|__|__|__|</td><td>|__|__|__|</td><td>|__|__|__|</td></tr>
<tr><td>**Time</td><td>|_|_|:|_|_|</td><td>|_|_|:|_|_|</td><td>|_|_|:|_|_|</td><td>|_|_|:|_|_|</td><td>|_|_|:|_|_|</td><td>|_|_|:|_|_|</td><td>|_|_|:|_|_|</td><td>|_|_|:|_|_|</td><td>|_|_|:|_|_|</td></tr>
</table>

*If on supplemental O_2, document values prior to initiation; ** Only complete this section if participant has not already met the endpoint (RR > 30 or $SpO_2 \leq 93\%$)

[1]D7 FOLLOW-UP	Date of enrolment |___|___|-|___|___|-|___|___|___|___|	
Date of follow-up |___|___|-|___|___|-|___|___|___|___|		In person ○ Telephone ○
Able to complete D7 follow-up	Yes ○ No ○	**If no, specify reason:** Uncontactable ○ Refusal ○

Alive	Yes ○ No ○	**If no, date of death** |___|___|-|___|___|-|___|___|___|___|
Symptoms resolved	Yes ○ No ○	**If yes, how many days ago** |___| **If no, symptoms worsening** [2] Yes ○ No ○
Admitted (not study site)	Yes ○ No ○	**If yes, name of facility** _____ **If yes, access to medical records** [3] Yes ○ No ○
Received oxygen	Yes ○ No ○	**If yes, supplemental oxygen route** Nasal cannula ○ FM / Venturi ○ HFNO / NIV ○ Ventilated ○ Unknown ○ **If yes, date initiated** |___|___|-|___|___|-|___|___|___|___| **If yes, location received** _____
Medications taken since enrolment (not at study site)	Steroids ☐ Hydroxychloroquine / Chloroquine ☐ Azithromycin ☐ Remdesivir ☐ Lopinavir / Ritonavir ☐ Convalescent plasma ☐ None ☐ Other ☐ _____ **If other, please specify:** Intravenous ☐ Intramuscular ☐ Oral ☐	

NOTES

1. This form should be completed for all participants not admitted at the study site on D7
2. If symptoms worsening invite participant to re-attend study site and complete study site recall form
3. If participant admitted at another health facility and has access to their medical records, arrange to review them and complete an off-site admission form

[1]D14 FOLLOW-UP	Date of enrolment |___|___|-|___|___|-|___|___|___|___|
Date of follow-up |___|___|-|___|___|-|___|___|___|___|	In person ○ Telephone ○
Able to complete D14 follow-up	Yes ○ No ○ **If no, specify reason:** Uncontactable ○ Refusal ○

Alive	Yes ○ No ○	**If no, date of death** |___|___|-|___|___|-|___|___|___|___|
Symptoms resolved	Yes ○ [2] No ○	**If yes, how many days ago** |___|
Admitted (not study site)	Yes ○ No ○	**If yes, name of facility** _____ **If yes, access to medical records** [3] Yes ○ No ○
Received oxygen	Yes ○ No ○	**If yes, supplemental oxygen route** Nasal cannula ○ FM / Venturi ○ HFNO / NIV ○ Ventilated ○ Unknown ○ **If yes, date initiated** |___|___|-|___|___|-|___|___|___|___| **If yes, location received** _____
Medications taken since D7 (not at study site)	Steroids ☐ Hydroxychloroquine / Chloroquine ☐ Azithromycin ☐ Remdesivir ☐ Lopinavir / Ritonavir ☐ Convalescent plasma ☐ None ☐ Other ☐ _____ **If other, please specify:** Intravenous ☐ Intramuscular ☐ Oral ☐	

CRF REVIEWED BY SITE PI / SUPERVISOR	Yes ○ No ○

NOTES

1. This form should be completed for all participants not admitted at the study site on D14
2. If symptoms still present invite participant to re-attend study site and complete study site recall form
3. If participant admitted at another health facility and has access to their medical records, arrange to review them and complete an off-site admission form

STUDY SITE RECALL D7	**Date of enrolment** |___|___|-|___|___|-|___|___|___|___|			
Date of recall |___|___|-|___|___|-|___|___|___|___|				**Time** |___|___|:|___|___|
Respiratory rate > 30 bpm	Yes ○	No ○	Unknown ○	**If yes, RR** |___|___| bpm
Oxygen saturation ≤ 93%	Yes ○	No ○	Unknown ○	**If yes, SpO_2** |___|___| %

STUDY SITE RECALL D14	**Date of enrolment** |___|___|-|___|___|-|___|___|___|___|			
Date of recall |___|___|-|___|___|-|___|___|___|___|				**Time** |___|___|:|___|___|
Respiratory rate > 30 bpm	Yes ○	No ○	Unknown ○	**If yes, RR** |___|___| bpm
Oxygen saturation ≤ 93%	Yes ○	No ○	Unknown ○	**If yes, SpO_2** |___|___| %

OFF-SITE ADMISSION	**Date of enrolment** | ___ | ___ | - | ___ | ___ | - | ___ | ___ | ___ | ___ |	
Date of admission | ___ | ___ | - | ___ | ___ | - | ___ | ___ | ___ | ___ |		
Respiratory rate > 30 bpm prior to supplemental oxygen	Yes ○ No ○ Unknown ○ **If yes, date first > 30** | ___ | ___ | - | ___ | ___ | - | ___ | ___ | ___ | ___ | **If yes, first RR > 30** | ___ | ___ | bpm **Time** | ___ | ___ | : | ___ | ___ |	
Oxygen saturation ≤ 93% prior to supplemental oxygen	Yes ○ No ○ Unknown ○ **If yes, date first ≤ 93** | ___ | ___ | - | ___ | ___ | - | ___ | ___ | ___ | ___ | **If yes, first SpO_2 ≤ 93** | ___ | ___ | % **Time** | ___ | ___ | : | ___ | ___ |	
<u>If No to both of above:</u> **For each mode / FiO_2 combination**	**Date**	| ___ | ___ | - | ___ | ___ | - | ___ | ___ | ___ | ___ |
	Mode	NC ○ FM ○ Venturi ○ HFNO / NIV ○ Ventilated ○
	FiO_2	[1] **Flow rate:** | ___ | ___ | . | ___ | L / min [2] **Percentage:** | ___ | ___ | ___ | %
	Lowest SpO_2	| ___ | ___ | ___ | % **Time** | ___ | ___ | : | ___ | ___ |
<u>If No to both of above:</u> **For each mode / FiO_2 combination**	**Date**	| ___ | ___ | - | ___ | ___ | - | ___ | ___ | ___ | ___ |
	Mode	NC ○ FM ○ Venturi ○ HFNO / NIV ○ Ventilated ○
	FiO_2	[1] **Flow rate:** | ___ | ___ | . | ___ | L / min [2] **Percentage:** | ___ | ___ | ___ | %
	Lowest SpO_2	| ___ | ___ | ___ | % **Time** | ___ | ___ | : | ___ | ___ |

NOTES

1. Flow rate should be completed if participant has received O2 via nasal cannula or face mask
2. Percentage should be completed if participant has received O2 via Venturi mask, high-flow nasal oxygen, or non-invasive or mechanical ventilation

9.14 Standard Operating Procedure: PRIORITISE study – biomarker assays

PRIORITISE

Standard Operating Procedure

Title:	PRIORITISE Biomarker analysis using the Ella system
Version:	2.0
Date:	26 May 2021

1. Abbreviations

SARS-CoV-2 Severe Acute Respiratory Syndrome Coronavirus 2
SOP Standard operating procedure

2. Background

PRIORITISE is a prospective longitudinal observational study that will recruit adults with symptomatic non-severe SARS-CoV-2 infection. The study aims to identify clinical and biochemical prognostic markers in adults with laboratory-confirmed SARS-CoV-2 infection who do not require oxygen supplementation. The end goal is to be able to accurately identify patients with a low risk of progression to subsequent need for supplemental oxygen, who can thus be discharged safely away from health facilities ensuring the available resources are allocated to patients most likely to benefit.

3. Purpose and Scope

The purpose of this document is to describe the processing of plasma collected from laboratory-confirmed SARS-CoV-2 patients for biomarker analysis. Plasma biomarker concentration will be measured using the Ella platform (Simple Plex, ProteinSimple™).

4. Requirements

- Ella Simple Plex System (ProteinSimple™)
- Ella cartridge kits (contain wash buffer A, sample diluent, and cartridge)
- Microcentrifuge
- Class II Biosafety Cabinet
- P1000 pipette
- P200 pipette
- P1000 sterile pipette tips
- P200 sterile pipette tips
- 1.5 mL sterile snap cap microcentrifuge tubes (~1500 tubes)
- Microcentrifuge tube racks (x3)
- Gloves
- Lab coat
- Biohazard Bags
- O.3% Chlorine
- Sharpie
- Computer with Microsoft Excel to prepare biomarker result database

5. Procedure

The recommendations in this SOP are intended for staff already familiar with processing blood specimens and performing laboratory methods in biochemistry (particularly enzyme-linked immunosorbent assays) who have been authorised to do so by the Principal Investigator(s).

Gloves and laboratory coat should be worn at all times when handling specimens or cartridges.

6. Storage and Thawing Plasma

6.1 Store collected plasma samples at -80°C until processing.
6.2 All Ella cartridge kits should be stored at 4°C until use.
6.3 The day before processing plasma specimens, remove the samples from -80°C storage and allow to thaw at 4°C overnight. Move the specimens from -80°C as late as possible the day prior to running on the Ella.

7. Processing Plasma Samples

Starting the Ella

2.30 Turn on ELLA (power switch is located on the back of the instrument)

Figure 1. Ella power switch.

2.31 Turn on Ella's computer and monitor.
2.32 Login to the computer.

ELLA systems are delivered with a local admin account as follows:
User ID: Ella
Password: Ella

2.33 Wait until the Ella's status light is green as this indicates the system is ready (Figure 2 and Table 1).

Figure 2. Ella ready status light.

Description	Status
Green	Ella is powered and ready for use
Blue	Ella is running a cartridge
Blinking red	An error has occurred (Ella may optionally sound an audible alarm)

Table 1. Status of Ella lights

2.34 Start the Simple Plex Runner software. The software can be accessed by either:

iii) Simple Plex Runner application on the computer desktop

iv) Windows **Start** menu, select **Programs** > **Simple Plex** > **Simple Plex Runner**

Cartridge Preparation

DO NOT PREPARE CARTRIDGES IN ADVANCE. THESE SHOULD BE PREPARED BEFORE THE START OF THE RUN. IF PLASMA SAMPLES ARE NOT BEING USED, ENSURE THEY ARE STORED AT 4°C AT ALL TIMES TO LIMIT BIOMARKER DEGRADATION.

Each time a kit is run, a new kit and cartridge ID will need to be scanned into the system.

2.35 Select Kit ID on the left hand panel of the Single Plex Runner application (Figure 3).

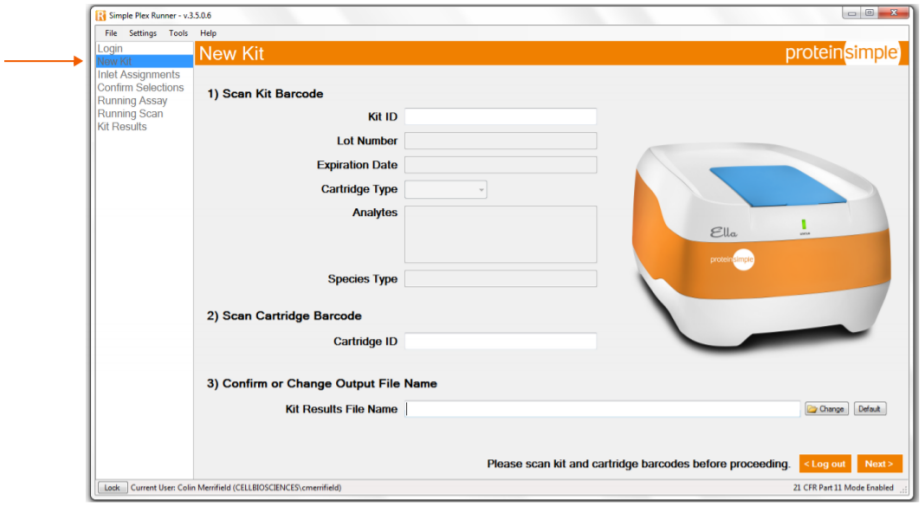


Figure 3. Screen for programming application for a new kit.

2.36 Remove a single Ella kit from its 4°C storage.

2.37 Scan the kit barcode on the packaging of the cartridge. See Figure 6C for an example of the package barcode.

2.38 Remove the cartridge from the vacuum bag. **Do not remove the plastic lining at the bottom of the cartridge until the cartridge is ready to be loaded**.

2.39 Scan the cartridge barcode for the cartridge ID and return the cartridge to the bag until it is ready to be loaded. This prevents dust and debris from getting into the cartridge wells.

2.40 Select the file where cartridge results will be saved.

2.41 Click the orange icon “Next”.

2.42 The inlet assignment application will now be available (Figure 4).

Assign:

Sample type = unknown (sample with an unknown concentration)

Sample name = sample ID from the PRIORITISE study

Dilution factor = Cartridge A = 10
Cartridge B = 5
Cartridge C = 5,000 and 500,000 (2 separate dilutions plated)

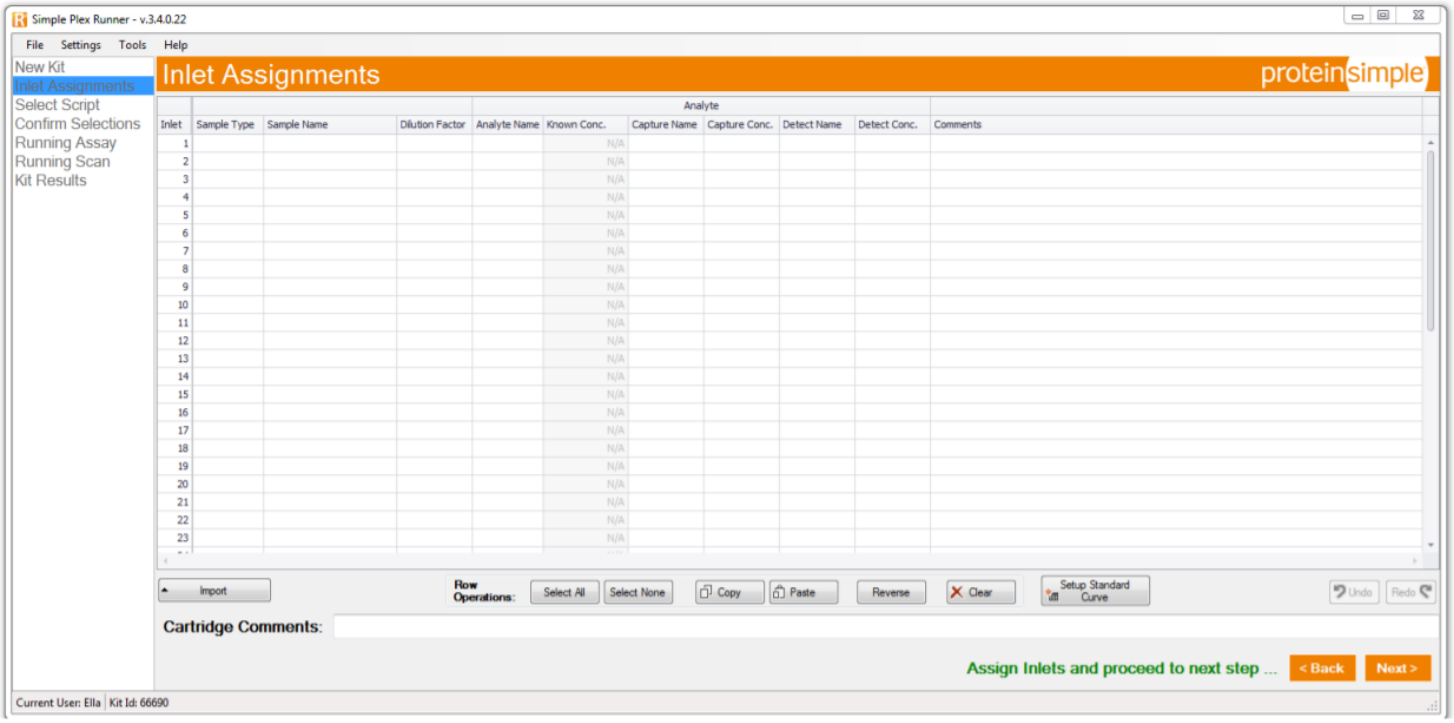


Figure 4. Inlet assignment screen.

2.43 Select the orange button “Next”. Leave the program in this Confirm Selections Screen (Figure 5). Do not click Start.

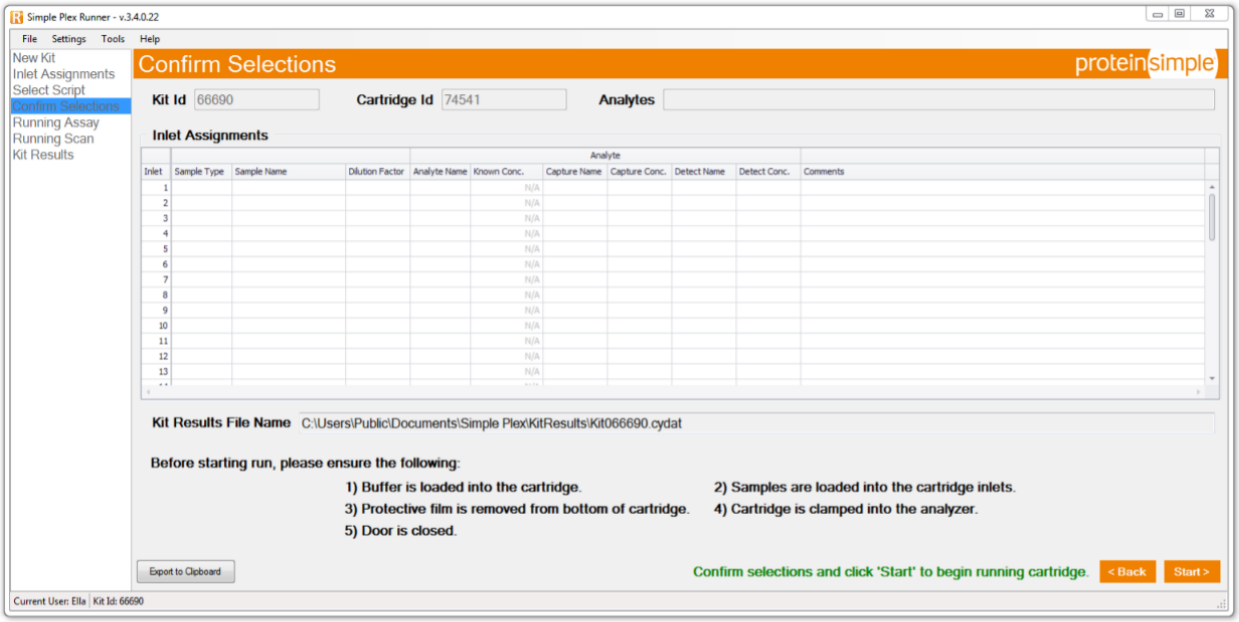


Figure 5. Confirm selections screen.

2.44 The plate is now ready to have samples and wash buffer loaded.

Sample Preparation

2.45 Aliquot 200 µl of plasma into a 1.5 mL microcentrifuge tube.

2.46 Centrifuge the 200 µl of plasma at 5,000 rpm for 5 minutes. Do not vortex once the plasma has been spun down.

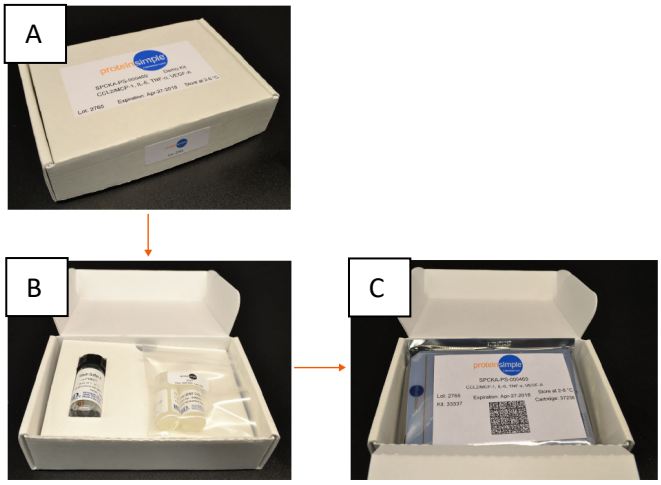


Figure 6. Ella simple cartridge kit. (A) Unopened Ella simple cartridge kit, (B) Wash buffer A (left) and sample diluent (right), (C) Barcoded cartridge found underneath wash buffer and sample diluent.

2.47 A total of 3 plates should be processed each day to avoid multiple freeze thaws of plasma (see below). Remove the cartridge from its packaging but **DO NOT throw packaging away until the plate has completed its run**.

F) 1 plate for Ang-2, D-dimer, TREM-1 (3 x 32 samples)
G) 1 plate for IP-10, IL-10, IL-1ra, IL-6, IL-8, PCT (6 x 32 samples)
H) 1 plate for CRP (1 x 72 samples)

2.48 Label each cartridge using a sharpie. Make sure this label is not covering the cartridge barcode. The cartridges should be labelled in manner that matches the plate maps so each sample can be easily traced back to the date and cartridge it was run on.
Labelling suggestion: Refer to each type of cartridge as A, B, and C and provide a number based on the day it was run. For example, for a CRP cartridge, label as C1 (day 1 of samples run), C2 (day 2 samples run), etc.

2.49 Prepare serial dilutions of the centrifuged plasma to obtain the appropriate dilution that can be accurately measured by the Ella assay standard curve. Between each dilution tube, ensure the pipette tip is changed and the tube is vortexed before transferring to the next tube in the series.

Cartridge A and B: For each sample, combine 125 µl of spun down plasma to 125 µl of sample diluent (Figure 7A, Tube #1).
** Each dilution is calculated to have enough volume for both cartridge A and cartridge B (only one dilution series needed) and to continue to the dilution series for cartridge C (from Tube #3). After loading cartridge A, immediately put the plasma and plasma dilution series at 4-8°C until the next cartridge will be loaded.

Cartridge C: For each sample, follow the dilution series provided in Figure 7B (up to 1:100,000). Tube #3 can be used from the dilution series of cartridges A and B to perform this dilution series. On cartridge C, the dilution Tubes #6 (1:5000) and #8 (1:500,000) should be plated.

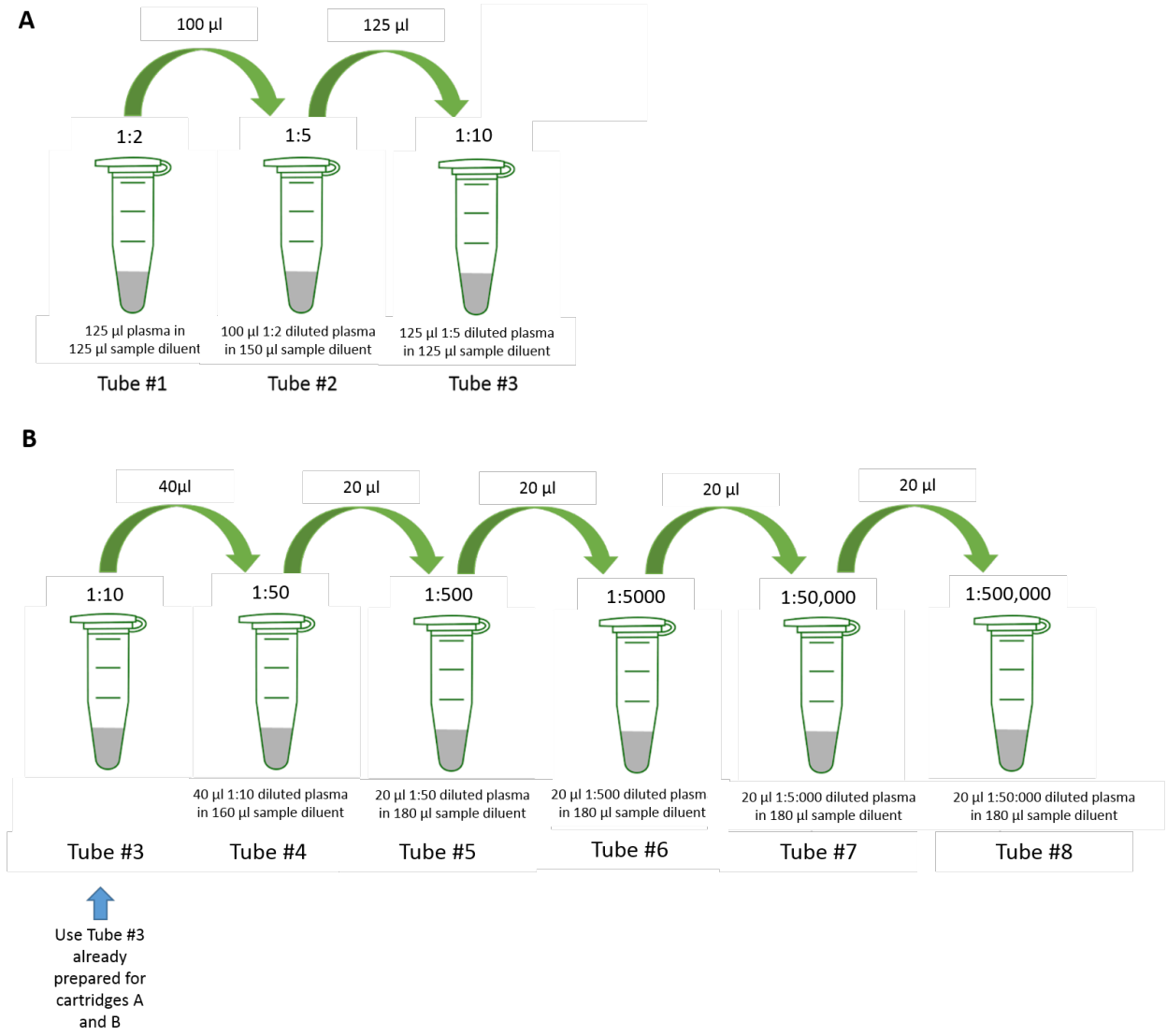


Figure 7. Dilution series for biomarker dilution optimization studies. (A) dilution series for cartridge A (Ang-2, D-dimer, TREM-1) and cartridge B (IP-10, IL-10, IL-1ra, IL-6, IL-8, PCT). (B) dilution series for cartridge C (CRP).

2.50 Pipette 50 µl of the diluted plasma sample into circular wells labelled with numbers (Figure 8). Cartridges A (Tube #3, 1:10 dilution) and B (Tube #2, 1:5 dilution) will have a single dilution plated, whereas cartridge C will have two dilutions plated (Tube #6 and Tube #7, 1:5000 and 1:500,000 dilution, respectively).

2.51 Pipette 1 mL of wash buffer A into the oval-shaped wells on the cartridges (16 wash wells on 32 sample cartridges and 10 wash wells on 72 sample cartridges).

2.52 Carefully carry the cartridge over to the instrument.

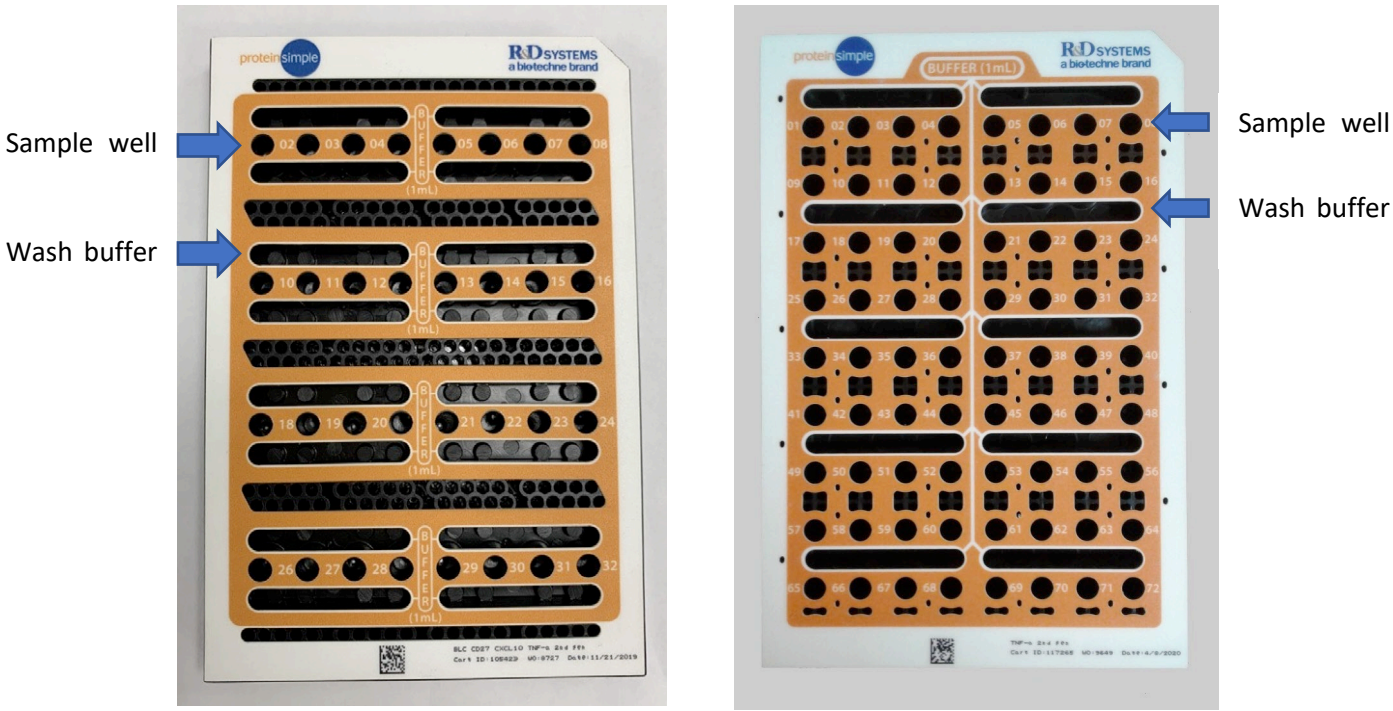


Figure 8. Cartridge layout of 32 sample plate (multi-analyte) and 72 sample plate (single analyte). Arrows denote the patient sample wells (circular) and wash buffer wells (oval/cylindrical)

Loading the Cartridge on to the Ella System

2.53 Double check that the status light is still green on the front of the Ella instrument. If green, open the Ella door (blue lid on instrument represented in Figure 9 and 10).
2.54 Gently peel off the protective plastic lining at the bottom of the cartridge. Make sure the exposed bottom does not touch anything.
2.55 Lift the cartridge clamp and place the cartridge inside the instrument.
2.56 Slowly lower the cartridge clamp. Close the door. See figure 10 for the loaded cartridge and closed clamp.

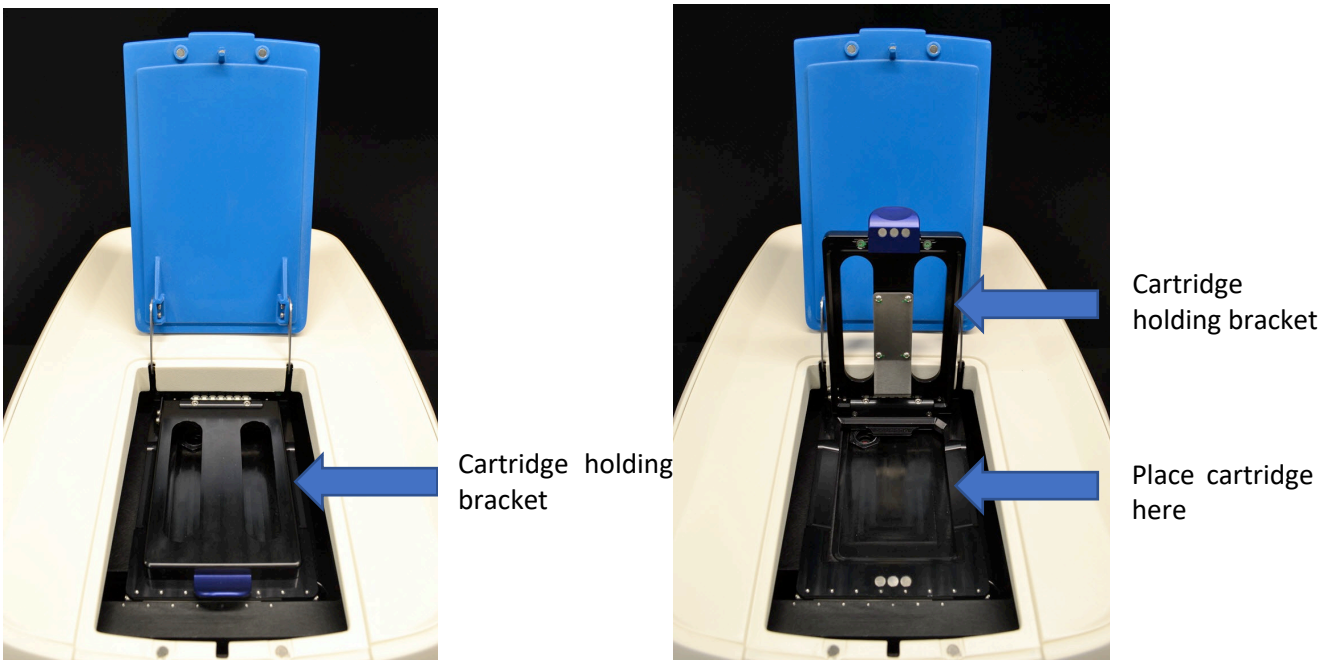


Figure 9. Ella door open displaying the cartridge clamp closed (A) and open (B).

Figure 10. Closed cartridge clamp.

2.57 Press start on the Confirm Selections screen (Figure 5).
2.58 Confirm you would like to start the run by pressing Start again in the Confirm dialogue box.
2.59 A dialogue box will now appear showing the status of the run (Figure 11). The approximate run time is 75 minutes.
2.60 Once the run is complete, the cartridge can be discarded in biohazardous waste for sharps.

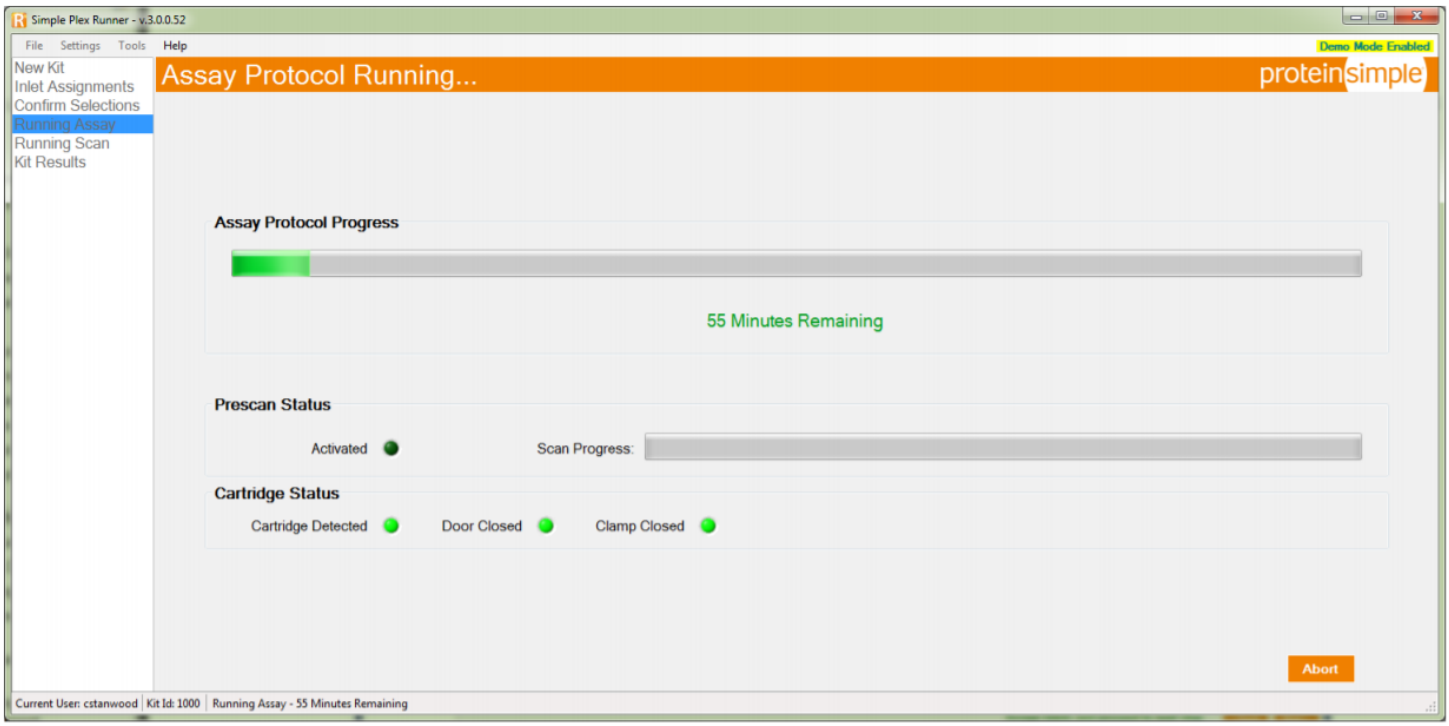


Figure 11. Running assay screen.

3 Viewing and Exporting Results

3.1 The run results will be initially displayed in a dashboard style format (Figure 12). Concentrations of each sample inlet will be displayed.

3.2 The export button allows export of results to a clipboard as a tab-delimited spreadsheet compatible with string. Save the data on a USB key and transfer to a secondary computer.

3.3 Compile all biomarker data for each run into an excel spreadsheet to generate a database for all samples.

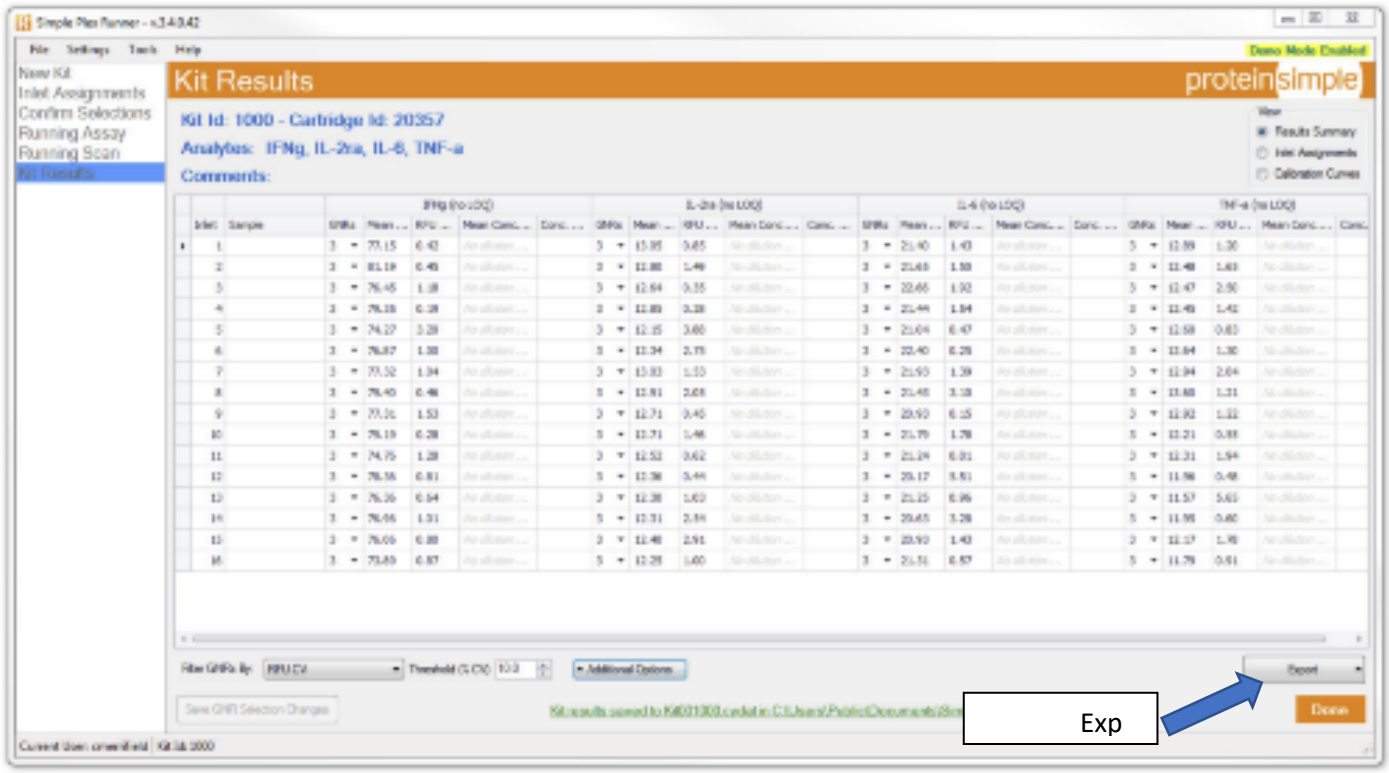


Figure 12. Cartridge results screen and export data option.

4 Instrument shutdown

4.1 To shut down Ella, close the Simple Plex Runner application by doing one of the following:
iii) Click File > Exit on the menu
iv) Click Close (X) in the upper-right corner of the Simple Plex Runner application

4.2 Shut down Ella's computer

4.3 Turn off Ella by pressing the power switch on the back panel (Figure 1).

5 Biohazard Waste Management

5.1 Transfer used cartridges to a clearly labelled biohazard waste container that will undergo BSL-2 waste processing (e.g. autoclaving) prior to disposal. As the cartridges have the potential to tear biohazard bags, these bags should be handled similarly as pipette tips or other "sharp" biohazard waste.

5.2 Send for safe disposal to the validated biowaste destruction company.

6. Update history

Version	Date	Summary of changes
1.0	15 October 2020	Written and approved by MRG and AC
2.0	26 May 2021	Sample volumes adjusted based on optimisation results

9.15 TRIPOD checklist: PRIORITISE study

Table 9.15-1: TRIPOD checklist for derivation and validation of the clinical prediction models. CI = confidence interval; D = derivation; TRIPOD = Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis; V = validation.

Section	Item	D / V	Checklist item	Page
Title and abstract				
Title	1	D, V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	NA
Abstract	2	D, V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	NA
Introduction				
Background and objectives	3a	D, V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	193-195
	3b	D, V	Specify the objectives, including whether the study describes the development or validation of the model or both.	195-196
Methods				
Source of data	4a	D, V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	196
	4b	D, V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	207
Participants	5a	D, V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	196
	5b	D, V	Describe eligibility criteria for participants.	196-197
	5c	D, V	Give details of treatments received, if relevant.	NA
Outcome	6a	D, V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	198
	6b	D, V	Report any actions to blind assessment of the outcome to be predicted.	198
Predictors	7a	D, V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	198-201; 203
	7b	D, V	Report any actions to blind assessment of predictors for the outcome and other predictors.	201
Sample size	8	D, V	Explain how the study size was arrived at.	206
Missing data	9	D, V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	203-204
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	204
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	204
	10c	V	For validation, describe how the predictions were calculated.	204
	10d	D, V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	204-205

Section	Item	D / V	Checklist item	Page
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	NA
Risk groups	11	D, V	Provide details on how risk groups were created, if done.	NA
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	NA
Results				
Participants	13a	D, V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	207-208
	13b	D, V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	209-213
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	209-213
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	209-213
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	215
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	218
	15b	D	Explain how to the use the prediction model.	220-224
Model performance	16	D, V	Report performance measures (with CIs) for the prediction model.	219-220
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	NA
Discussion				
Limitations	18	D, V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	250-252
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	245-247
	19b	D, V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	253
Implications	20	D, V	Discuss the potential clinical use of the model and implications for future research.	253
Other information				
Supplementary information	21	D, V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	NA
Funding	22	D, V	Give the source of funding and the role of the funders for the present study.	6

9.16 STARD checklist: PRIORITISE study

Table 9.16-1: STARD checklist for evaluation of the suPAR RDT. AUC = area under the receiver operating characteristic curve; RDT = rapid diagnostic test; STARD = Standards for Reporting Diagnostic accuracy studies; suPAR = soluble urokinase plasminogen activator receptor.

Section & Topic	No	Item	Page
TITLE OR ABSTRACT			
Title	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	NA
Abstract	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	NA
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	195
	4	Study objectives and hypotheses	195-196
METHODS			
Study design	5	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	196
Participants	6	Eligibility criteria	196-197
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	196
	8	Where and when potentially eligible participants were identified (setting, location and dates)	196-197; 207
	9	Whether participants formed a consecutive, random or convenience series	196
Test methods	10a	Index test, in sufficient detail to allow replication	202
	10b	Reference standard, in sufficient detail to allow replication	202
	11	Rationale for choosing the reference standard (if alternatives exist)	NA
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	NA
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	NA
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	202
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	202
Analysis	14	Methods for estimating or comparing measures of diagnostic accuracy	205
	15	How indeterminate index test or reference standard results were handled	205
	16	How missing data on the index test and reference standard were handled	208
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	NA
	18	Intended sample size and how it was determined	206
RESULTS			
Participants	19	Flow of participants, using a diagram	208

Section & Topic	No	Item	Page
	20	Baseline demographic and clinical characteristics of participants	209-213
	21a	Distribution of severity of disease in those with the target condition	209-213
	21b	Distribution of alternative diagnoses in those without the target condition	209-213
	22	Time interval and any clinical interventions between index test and reference standard	NA
Test results	23	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	225
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	226-234
	25	Any adverse events from performing the index test or the reference standard	NA
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	252
	27	Implications for practice, including the intended use and clinical role of the index test	249-250
OTHER INFORMATION			
	28	Registration number and name of registry	206-207
	29	Where the full study protocol can be accessed	206-207
	30	Sources of funding and other support; role of funders	6