

The NeoSep Severity and Recovery scores to predict mortality in hospitalized neonates and young infants with sepsis derived from the global NeoOBS observational cohort study

Neal Russell*,¹ Wolfgang Stöhr*,² Aislinn Cook,¹ James A Berkley,^{3,4,5} Bethou Adhisivam,⁶ Ramesh Agarwal,⁷ Nawshad Uddin Ahmed,⁸ Manica Balasegaram,⁹ Neema Chami,^{10,11} Adrie Bekker,¹² Davide Bilardi,¹³ Cristina G. Carvalho,¹⁴ Suman Chaurasia,¹⁵ Viviane Rinaldi Favarin Colas,¹⁶ Simon Cousens,¹⁷ Ana Carolina Dantas de Assis,¹⁴ Han Dong,¹⁸ Angela Dramowski,¹² Nguyen Trong Dung,¹⁹ Jinxing Feng,²⁰ Youri Glupczynski,²¹ Srishti Goel,²² Herman Goossens,²¹ Doan Thi Huong Hao,¹⁹ Mahmudul Hasan,⁸ Tatiana Munera Huertas,¹ Nathalie Khavessian,⁹ Angeliki Kontou,²³ Tomislav Kostyanev,²¹ Preamsak Laoyookhon,²⁴ Sorasak Lochindarat,²⁴ Maia De Luca,²⁵ Surbhi Malhotra-Kumar,²¹ Nivedita Mondal,⁶ Nitu Mundhra,²⁶ Philippa Musoke,²⁷ Marisa M. Mussi-Pinhata,¹⁴ Ruchi Nanavati,²⁶ Firdose L. Nakwa,¹⁰ Sushma Nangia,²² Alessandra Nardone,¹³ Borna Nyaoke,⁹ Christina W Obiero,^{3,28} Wang Ping,¹⁸ Kanchana Preedisripipat,²⁹ Shamim Qazi,³⁰ Lifeng Qi,³¹ Amy Riddell,¹ Lorenza Romani,²⁵ Praewpan Roysuwan,³² Robin Saggars,^{10,11} Samir Saha,⁸ Kosmas Sarafidis,²³ Valerie Tusibira,²⁷ Sithembiso Velaphi,¹⁰ Tuba Vilken,²¹ Xiaojiao Wang,³³ Yajuan Wang,³⁴ Yonghong Yang,²⁰ Sally Ellis,⁹ Julia Bielicki,¹ A Sarah Walker,² Paul T. Heath,¹ Mike Sharland,¹ *Co-First authors, listed alphabetically

¹ St George's Centre for Neonatal and Paediatric Infection, Institute of Infection & Immunity, St George's University of London

² Medical Research Council Clinical Trials Unit at University College London, London, UK

³ Clinical Research Department, KEMRI-Wellcome Trust Research Programme, Kilifi, Kenya

⁴ Centre for Tropical Medicine & Global Health, Nuffield Department of Medicine, University of Oxford, Oxford UK

⁵ The Childhood Acute Illness & Nutrition (CHAIN) Network, Nairobi, Kenya

⁶ Department of Neonatology, Jawaharlal Institute of Postgraduate Medical Education & Research (JIPMER), Pondicherry, India

⁷ Newborn Division and WHO-CC, All India Institute of Medical Sciences, New Delhi, India.

⁸ Child Health Research Foundation (CHRF), Dhaka Shishu Hospital, Dhaka, Bangladesh

⁹ Global Antibiotic Research and Development Partnership (GARDP), Geneva, Switzerland

¹⁰ Department of Paediatrics and Child Health, Chris Hani Baragwanath Academic Hospital, School of Clinical Medicine, Faculty of Health Sciences, University of Witwatersrand, Johannesburg, South Africa

¹¹ Department of Paediatrics and Child Health, Charlotte Maxeke Johannesburg Academic Hospital, Johannesburg, South Africa

¹² Department of Paediatrics and Child Health, Faculty of Medicine and Health Sciences, Stellenbosch University, South Africa

¹³ Penta Foundation, Padova, Italy

¹⁴ Department of Pediatrics, Ribeirão Preto Medical School, University of São Paulo, Brazil

¹⁵ All India Institute of Medical Sciences, Department of Paediatrics, New Delhi, India

¹⁶ Pediatric Infectious Diseases Unit, Santa Casa de São Paulo, Rua Dr. Cesário Mota Júnior 112, São Paulo, SP 01221-020, Brazil

¹⁷ Faculty of Epidemiology and Population Health, Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine

¹⁸ Beijing Obstetrics and Gynecology Hospital, Capital Medical University, Beijing, China

¹⁹ Vietnam National Children's Hospital, Hanoi, Vietnam and Department of Surgery, Vietnam National Children's Hospital, Hanoi, Vietnam

²⁰ Department of Neonatology, Shenzhen Children's Hospital, Shenzhen, China

²¹ Laboratory of Medical Microbiology, University of Antwerp, Antwerp, Belgium

²² Department of Neonatology, Lady Hardinge Medical College & associated SSK & KSC Hospitals, New Delhi

²³ Neonatology Dept, School of Medicine, Faculty of Health Sciences, Aristotle University and Hippokratia General Hospital, Thessaloniki, Greece

²⁴ Queen Sirikit National Institute of Child Health, Bangkok, Thailand

²⁵ Academic Hospital Paediatric Department, Bambino Gesù Children's Hospital, Rome, Italy

²⁶ Neonatology Department, Seth GS Medical College and King Edward Memorial Hospital, Mumbai, India.

²⁷ Department of Paediatrics and Child Health, College of Health Sciences, Makerere University and MUJHU Care

²⁸ Amsterdam UMC, University of Amsterdam, Emma Children's Hospital, Department of Global Health, Amsterdam, the Netherlands.

²⁹ Chiangrai Prachanukroh Hospital, Chiang Rai, Thailand

³⁰ World Health Organisation, Maternal, Newborn, Child and Adolescent Health Department, Geneva, Switzerland

³¹ Department of Infectious Diseases, Shenzhen Children's Hospital, Shenzhen, China

³² PHPT/IRD-MIVEGEC, Faculty of Associated Medical Sciences, Chiang Mai University, Thailand

³³ Department of Neonatology, Beijing Children's Hospital, Capital Medical University, National Centre for Children's Health, Beijing, China

³⁴ Department of Neonatology, Children's Hospital, Capital Institute of Pediatrics, Yabao Road, Chaoyang District, Beijing, China

2 **Abstract**

3 **Background**

4 Sepsis severity scores are used in clinical practice and trials to define risk groups. There are limited
5 data to derive hospital-based sepsis severity scores for neonates and young infants in high-burden
6 low- and middle-income country (LMIC) settings where trials are urgently required. We aimed to
7 create linked sepsis severity and recovery scores applicable to hospitalized neonates and young
8 infants in LMIC which could be used to inform antibiotic trials.

10 **Methods & Findings**

11 A prospective observational cohort study was conducted across 19 hospitals in 11 countries in sub-
12 Saharan Africa, Asia, Latin America and Europe. Infants aged <60 days with clinical sepsis fulfilling at
13 least two clinical or laboratory criteria (≥ 1 clinical) were enrolled. Primary outcome was 28-day
14 mortality. Two prediction models were developed for 1) 28-day mortality from factors at sepsis
15 presentation (baseline NeoSep Severity Score), and 2) daily risk of death on IV antibiotics from daily
16 updated assessments (NeoSep Recovery Score). Multivariable Cox regression models included a
17 randomly selected 85% of infants, with 15% for validation.

18
19 3204 infants were enrolled between 2018-2020. Median age was 5 days (IQR 2-15), 90.4% (n=2,895)
20 were <28 days. Median birth weight was 2500g (1400-3000g), and a median of 4 clinical (IQR 2-5)
21 and 1 laboratory (0-2) signs were present. Overall mortality was 11.3% (95%CI 10.2-12.5%; n=350). A
22 baseline NeoSep Severity Score from infants characteristics, respiratory support, and clinical signs
23 (no laboratory tests) at presentation had a C-index 0.77 (95%CI: 0.75-0.80) and 0.76 (0.69-0.82) in
24 derivation and validation samples, respectively. Mortality in the validation sample was 1.6% (3/189;
25 95%CI: 0.5-4.6%), 11.0% (27/245; 7.7-15.6%), and 27.3% (12/44; 16.3-41.8%) in low (score 0-4),
26 medium (5-8) and high (9-16) risk groups, respectively, with similar performance across subgroups.

27
28 A related NeoSep Recovery Score based on evolving post-baseline clinical signs and supportive care
29 discriminated well between infants who died or survived the following day or subsequent few days.
30 The area under the ROC curve for score on day 2 and death in the following 5 days was 0.82 (95%CI
31 0.78-0.85) and 0.85 (95%CI 0.78-0.93) in the derivation and validation data, respectively.

33 **Conclusion**

34 The baseline NeoSep Severity Score predicted 28-day mortality and could identify infants with high
35 risk of mortality for inclusion in hospital-based sepsis trials. The NeoSep Recovery Score predicts
36 day-by-day inpatient mortality and could, with further validation, help to identify poor response to
37 antibiotics.

43 **Funding**

44 This study was funded by the Global Antibiotic Research and Development Partnership (GARDP),
45 made possible by Bill & Melinda Gates Foundation; German Federal Ministry of Education and
46 Research; German Federal Ministry of Health; Government of the Principality of Monaco; the Indian
47 Council for Medical Research; Japanese Ministry of Health, Labour and Welfare; Netherlands

48 Ministry of Health, Welfare and Sport; South African Medical Research Council; UK Department of
49 Health and Social Care (UK National Institute of Health Research and the Global Antimicrobial
50 Resistance Innovation Fund – GAMRIF); UK Medical research Council; Wellcome Trust. GARDP has
51 also received core funding from the Leo Model Foundation; Luxembourg Ministry of Development
52 Cooperation and Humanitarian Aid; Luxembourg Ministry of Health; Médecins Sans Frontières; Swiss
53 Federal Office of Public Health; UK Foreign, Commonwealth & Development Office (previously the
54 UK Department for International Development).

55

56

57 **Author Summary**

58

59 **Why was this study done?**

- 60 ➤ Evidence to guide hospital-based antibiotic treatment of sepsis in neonates and young
61 infants is scarce, and clinical trials are particularly urgent in low- and middle-income (LMIC)
62 settings where antimicrobial resistance threatens to undermine existing guidelines
- 63 ➤ There is limited data to inform the design of antibiotic trials in LMIC settings, particularly to
64 define risk stratification and inclusion and escalation criteria in hospitalised neonates and
65 young infants

66

67 **What did the researchers do and find?**

- 68 ➤ To our knowledge this is the first global, prospective, hospital-based observational study of
69 clinically diagnosed neonatal sepsis across 4 continents including LMIC settings, with
70 extensive daily data collection on clinical status, antibiotic use and outcomes.
- 71 ➤ There was a high mortality among infants with sepsis in LMIC hospital settings. 4 non-
72 modifiable and 6 modifiable factors predicted mortality and were included in a NeoSep
73 Severity score which defines patterns of mortality risk at baseline
- 74 ➤ A NeoSep Recovery Score including the same modifiable factors (with the addition of
75 cyanosis) predicted mortality on the following day during the course of treatment.

76

77 **What do these findings mean?**

- 78 ➤ The NeoSep Severity Score and NeoSep Recovery score are now informing inclusion and
79 escalation criteria in the NeoSep1 antibiotic trial (ISRCTN48721236) which aims to identify
80 novel first- and second-line empiric antibiotic regimens for neonatal sepsis
- 81 ➤ The NeoSep Severity Score could be used to predict mortality at baseline in future studies of
82 targeting resources in routine care. With further validation, the NeoSep Recovery Score
83 could potentially be used to identify poor response to empiric antibiotic treatment

84

85

86

87

88

89

90 **Introduction**

91 There are an estimated 1.3 million episodes of neonatal sepsis¹ and 200,000 sepsis-attributable
92 deaths each year, largely in low- and middle-income countries (LMIC).² Increasing antimicrobial
93 resistance (AMR), particularly in hospitals, threatens to undermine the effectiveness of available
94 antibiotic treatment, with bacterial isolates increasingly resistant to World Health Organization
95 (WHO) recommended first- and second-line regimens.³⁻⁸

96 In this context, global routine antibiotic use for neonatal sepsis frequently diverges from WHO
97 guidelines.⁹ There are limited clinical data or scoring systems applicable to LMIC hospitals to stratify
98 risk and predict mortality, assess response to treatment and guide decisions on antibiotic escalation
99 and de-escalation.¹⁰ Most recent large-scale sepsis trials in neonates and young infants in LMIC have
100 been community-based, focusing on simplifying antibiotic regimens,¹¹⁻¹³ with very limited evidence
101 available to optimize treatment in higher risk hospital settings where treatment and recovery is
102 more complex.

103 Challenges to conducting antibiotic trials in neonates and young infants include lack of regulatory
104 guidance and no widely accepted case definition for neonatal sepsis.¹⁴ European Medicines Agency
105 (EMA) criteria for neonatal sepsis are comprehensive but perform variably¹⁵ and require laboratory
106 tests which are not always available in LMIC. Criteria for possible serious bacterial infection (pSBI)
107 are well adapted for LMIC but focus on community-acquired sepsis at primary care level and identify
108 large numbers of infants with mild illness.¹¹⁻¹³ There is a need for criteria to stratify risk and define
109 inclusion criteria for trials in hospital-based populations of neonates and young infants.

110 The aim of this analysis was to develop two linked clinically based scores relevant to LMIC
111 populations: 1) a sepsis severity score to predict 28-day mortality from factors known at sepsis
112 presentation and 2) a recovery score to predict the daily risk of death on treatment with IV
113 antibiotics using daily updated assessments of clinical status. Additional results of the NeoOBS
114 cohort are published separately.

115

116 **Methods**

117 **Study design and participants**

118 Hospitalized infants aged <60 days with a new episode of clinically suspected sepsis were enrolled
119 from 19 hospitals across 11 countries, in Asia (Bangladesh, China, India, Thailand, Vietnam), Africa
120 (Kenya, South Africa, Uganda), Europe (Italy, Greece) and South America (Brazil). Sites were selected
121 after a feasibility study¹⁶ to represent diverse regions and to include secondary and tertiary referral
122 hospitals, public facilities, and facilities with varying proportions of in-born and out-born neonates.
123 Access to blood cultures and high-quality microbiology was required.

124 Infants were eligible if the local physician had decided to treat the infant with antibiotics for a new
125 distinct sepsis episode meeting the inclusion criteria (supplement figure 1), which combined clinical
126 and laboratory criteria from WHO pSBI¹⁷ and EMA Criteria.¹⁵ To allow for variation in access to
127 laboratory testing, and ensure generalizability to varying LMIC hospital contexts, laboratory values
128 could be used for inclusion but were not mandatory. A minimum of 2 clinical, or 1 clinical and 1
129 laboratory, sepsis criteria were required with a locally adapted sampling frame used to recruit up to

130 200 infants per site. Infants were excluded if an alternative primary diagnosis other than sepsis was
131 suspected, or a serious non-infective comorbidity was expected to cause death within 72 hours.

132 Ethical approval was obtained from St. George's, University of London (SGUL) Research Ethics
133 Committee and sites' local, central or national ethics committees and other relevant local bodies,
134 where required. Design and reporting were guided by the STROBE-NI framework.¹⁸

135 **Procedures**

136 A blood culture taken before new antibiotics were started was the only compulsory procedure. After
137 written informed consent from parents/guardians, with a witness for those who could not
138 read/write, data were collected on demographics, maternal history, birth history, pre-existing
139 comorbidities and previous infections, as well as available clinical data from 24 hours before the
140 time of blood culture sampling. Infants were prospectively followed with daily observations
141 recorded on vital signs, clinical signs and supportive treatment.

142 Infants were followed for 28 days after enrolment in-person if still hospitalized, or by telephone
143 post-discharge. A final diagnosis was documented by clinicians, as was primary and secondary causes
144 of death, and any clinical illness or readmission occurring after discharge and within 28 days of
145 enrolment. 'Infection-related death' was independently classified by the research team based on
146 isolation of a presumed pathogen from blood or CSF culture, reporting by clinicians of infection as a
147 cause of death, or a mode of death consistent with infection in the absence of another reported
148 cause.

149 Data were collected by research and clinical staff based on clinical observation and routine source
150 documentation (e.g. medical/nursing notes, vital signs and prescription charts), and entered and
151 managed using REDCap™ electronic data capture tools¹⁹ hosted at SGUL.

152 **Statistical analysis**

153 The pre-specified primary outcome was mortality through 28 days post-enrolment. We developed
154 two prediction models and corresponding risk scores: 1) a baseline NeoSep Severity Score to predict
155 28-day mortality from factors known at sepsis presentation, and 2) a NeoSep Recovery Score to
156 predict the daily risk of death while treated with IV antibiotics from daily updated assessments of
157 clinical status. For both analyses, a randomly selected 15% of infants per site was reserved for model
158 validation and not used in model development. In the remaining 85% of infants, Cox proportional
159 hazards regression with site-level random effects were used, with time measured from the initial
160 blood culture sample, censoring at the earliest of day 28, withdrawal or last contact if lost post-
161 discharge, or, for the recovery analysis, when IV antibiotics stopped (cause-specific model). Analyses
162 used Stata version 16.1.

163 *NeoSep Severity Score*

164 Backwards elimination (exit $p=0.05$) was used to identify independent predictors of 28-day mortality
165 from a pre-defined set of factors including infants' characteristics, supportive care and clinical signs
166 that could be considered known at sepsis presentation (Table 1). A points-based risk score
167 integerised model coefficients for selected predictors, with higher scores indicating higher mortality
168 risk. Discrimination was assessed using Harrell's C-index with bootstrapped confidence intervals. The
169 NeoSep Severity Score was compared with a score based on WHO pSBI¹⁷ allocating one point for
170 every criterion.

171 *NeoSep Recovery Score*

172 To estimate the daily risk of death after initiating IV antibiotics, we initially included all clinical
173 factors in the baseline NeoSep severity score as time-updated factors, regardless of significance;
174 excluding unmodifiable infant and birth characteristics, such as gestational age, that could not
175 evolve and therefore could not reflect recovery. Forward selection (entry $p=0.05$) was then used to
176 identify additional independent time-updated clinical predictors. To avoid selecting factors
177 representing the mechanism of dying rather than being predictors of subsequently dying, clinical
178 parameters reported on a given day were used to predict death on the following day in all time-
179 updated models. A points-based risk score was then derived similarly to the baseline severity score.
180 Discrimination was assessed using time updated area under the receiver operating curve (AUROCs).
181 To examine the potential usefulness of the recovery score for informing the decision to potentially
182 switch to second line antibiotics in future empiric clinical trials, we focused on day 2 (48-72 hours)
183 post baseline, a key decisional time point during clinical management when culture results become
184 available and response to treatment is commonly evaluated (see supplement statistical methods for
185 details).

186

187 **Results**

188 3204 infants (90.4% neonates aged <28 days, $n=2895$; 42.1% female, $n=1348$) were recruited from
189 20th August 2018-29th February 2020. The median postnatal age was 5 days (IQR 1-15). 3088 (96.4%)
190 infants had been born in a hospital/facility (1550 in the enrolling facility), 1412 (44.3%) by caesarean
191 section (969 as an emergency). The median (IQR) gestational age at birth was 37 (31-39) weeks, with
192 birth weight 2500g (1400-3000g). When enrolled, 69.1% ($n=2215$) infants had been hospitalized
193 since birth, and 30.9% ($n=989$) were admitted from the community. 2759 (86.1%) were recruited in a
194 neonatal unit (supplement table 1). 41.1% ($n=1318$) of sepsis episodes were healthcare-associated
195 (occurring >48h after hospital admission). Among 309 (9.6%) infants enrolled aged ≥ 28 days, the
196 majority ($n=181$; 58.6%) were ex-premature ($n=146$; 47.4%) and/or had been admitted since the
197 neonatal period ($n=136$, 44.2%), the majority of which since birth ($n=136$, 86.8%). Other than
198 infection, the most common diagnoses/problems reported at admission were respiratory distress
199 (51.4%), prematurity (41.1%) and low birth weight (36.5% (supplement figure 2)). Supportive
200 measures such as invasive ventilation varied across sites, as described in supplement figure 3.

201 A median of 4 (IQR 2-5) clinical signs were present at baseline, the most common being respiratory
202 signs (65.8%, $n=2107$), difficulty feeding (45.7%, $n=1464$), lethargy or reduced movement (35.3%,
203 $n=1,131$), abdominal distension (24.3%, $n=777$) and evidence of shock (21.3%, $n=683$), with
204 prevalence decreasing over time (table 1 & supplement figure 4). Signs associated with meningitis
205 were relatively uncommon (irritability 10% ($n=322$), convulsions 7.4% ($n=236$), abnormal posturing
206 4.4% ($n=140$), bulging fontanelle 1.8% ($n=57$)) (supplement figure 5). Initial blood culture results
207 were available a median of 2 days (IQR 1-3) after blood draw for 3195 (99.7%) infants, among whom
208 pathogens were isolated in 564 (17.6%) (>1 pathogen in 29).

209 *Mortality*

210 Overall, 350 infants (11.3%; 95%CI 10.2-12.5%) died within 28 days of the baseline blood culture,
211 with wide variation between sites (1.0% to 27.6%, supplement figure 6). Mortality among infants
212 with a pathogen-positive baseline culture was 17.7% (99/564; 95%CI 14.7-21.1%) compared with
213 9.9% (250/2631; 8.8-11.2%) in infants without pathogens. 315/350 infants died in the initial hospital
214 (286 whilst still on IV antibiotics), and 35 after discharge (35/350, 10.0%). Reported causes of death
215 included infection in 74.9% ($n=262$), prematurity in 48.0% ($n=168$), birth asphyxia in 10.6% ($n=37$)

216 and congenital malformations in 10.3% (n=36), with infection and prematurity being the most
217 common concomitant diagnoses. After standardized review, 309 (88.3%) deaths were classified as
218 'infection-related'. Mortality at 28 days was unknown in 62 infants (5 withdrew in hospital, 57 were
219 lost post-discharge).

220 *Baseline predictors of 28-day mortality*

221 Ten clinical factors known at presentation independently predicted mortality in the final model,
222 including infant characteristics (birth weight, gestational age, duration of time in hospital, congenital
223 anomalies), level of respiratory support, and clinical signs (abnormal temperature, abdominal
224 distension, lethargy/no or reduced movement, difficulty feeding, and evidence of shock) (table 2).
225 Risk was increased with both low (<35°C) and high temperature (≥38°C) with evidence of even higher
226 risk if ≥39°C. A NeoSep Severity Score developed from these baseline predictors had a maximum 16
227 points (table 2), with C-statistics 0.77 (95%CI: 0.75-0.80) and 0.76 (0.69-0.82) in the derivation and
228 validation samples, respectively, and a good fit in the validation sample (Hosmer-Lemeshow p=0.53).
229 Defining mortality risk thresholds at 5% and 25% in the derivation sample categorized scores as low
230 (score 0-4), medium (5-8) and high (9-16) risk, with mortality in the derivation sample of 1.6%
231 (17/1071; 95%CI: 1.0-2.5%), 14.4% (203/1409; 12.7-16.3%), and 35.8% (88/246; 30.0-41.9%), and in
232 the validation sample of 1.6% (3/189; 95%CI: 0.5-4.6%), 11.0% (27/245; 7.7-15.6%), and 27.3%
233 (12/44; 16.3-41.8%)(figure 1).

234 The association between the NeoSep Severity Score and mortality was similar within multiple sub-
235 groups, for example in early or late-onset, community or healthcare- associated, high-, middle- or
236 low-income settings (figure 1), term vs. preterm infants, blood culture-positive vs. culture-negative
237 cases (supplement figure 7).

238 A score based on WHO pSBI had a C-statistic of 0.63 (95%CI 0.56-0.70) in the validation sample,
239 lower than the NeoSep Severity Score. Because the WHO pSBI do not include any infant/birth
240 characteristics, to ensure an appropriate comparison we also calculated a modified NeoSep Severity
241 Score excluding these factors which had a C-statistic of 0.72 (95%CI 0.65-0.77) (supplement figure 8).

242 *Predictors of daily risk of dying of sepsis while on intravenous antibiotics in hospital: NeoSep 243 Recovery Score*

244 Seven of the time-varying factors independently predicted mortality in the final model, including all
245 the non-modifiable signs in the NeoSep Severity Score (respiratory support and the five clinical
246 signs), plus cyanosis, which had not added independent information in the baseline model (p=0.228)
247 (table 2). A NeoSep Recovery Score developed from these time-updated predictors had a maximum
248 11 points and discriminated well between infants who died or survived the next day (figures 2&3).
249 The area under the receiver operating curve (AUROC) of the recovery score on one day, for
250 predicting death the next day, ranged between 0.8 and 0.9 over the first week post baseline in the
251 derivation sample (figure 3). As expected, it slowly decreased over subsequent days. AUROC over
252 time in the validation sample was similar although based on fewer numbers (figure 3).

253 The NeoSep Recovery Score on day 2 had an AUROC for dying in the following 5 days of 0.82 (95%CI
254 0.78-0.85) and 0.85 (95%CI 0.78-0.93) in the derivation and validation samples, respectively. A score
255 ≥4 was the most discriminative, whether this was an increase from baseline, lack of initial response
256 or an improvement from a higher score down to 4, with sensitivity and specificity of 0.74 (95%CI

257 0.64-0.82) and 0.74 (0.72-0.75) in the derivation, and 0.87 (0.60-0.98) and 0.76 (0.71-0.79) in the
258 validation samples.

259 Twenty-nine infants (derivation: n=27; validation: n=2) died between day 3 and 7 despite having had
260 a score <4 on day 2. Of these, 6 deaths were not classified as infection-related, 10 had a subsequent
261 increase in score to ≥ 4 before dying, 9 had at least 2 unmodifiable risk factors included as predictors
262 of mortality in the NeoSep Severity Score (most commonly congenital abnormalities and very low
263 birthweight), and 1 had congenital varicella; the remaining 4 had a day 2 score of 2 or more.

264 Of note, the change in score from baseline to day 2 had poorer discrimination than the absolute
265 score on day 2, with an AUROC for dying in the following 5 days of 0.68 (95%CI 0.63-0.73) and 0.62
266 (95%CI 0.48-0.77) in the derivation and validation samples, respectively. Combining change in score
267 using various cut-offs with <4/ ≥ 4 absolute score on day 2 also did not improve discrimination
268 (supplement figure 9).

269

270 Discussion

271 NeoOBS represents the largest hospital-based multi-country prospective observational cohort study
272 to include high quality daily clinical data and outcomes in neonates and infants with sepsis in
273 predominantly (>95%) LMIC settings. The cohort was largely neonatal (90%), with most of the
274 remainder (>2/3) born prematurely and/or previously admitted during the neonatal period. All
275 infants had a primary clinical diagnosis of sepsis and most deaths (88%) were infection-related.
276 Overall mortality was 11.3%, with significant variation across different settings.

277 Mortality was independently associated with a range of factors, developed into a baseline sepsis
278 severity score including unmodifiable infant characteristics and clinical characteristics at
279 presentation/antibiotic initiation. A recovery score was developed accounting for evolving clinical
280 signs after baseline. The scores use only simple clinical characteristics identifiable from routine
281 history and examination and showed good discrimination. A score of 5 or higher at baseline was
282 associated with 28-day mortality over 10%, and a recovery score 2 days after antibiotic initiation of 4
283 or higher had both sensitivity and specificity of 74% for mortality over the following 5 days. The
284 baseline and recovery scores have informed the criteria for inclusion and escalation of therapy in the
285 NeoSep1 antibiotic trial (ISRCTN48721236) for neonatal sepsis with a Personalised RANdomised
286 Controlled Trial (PRACTical).²²

287 To our knowledge, these are the first scores to specifically predict mortality risk developed from
288 inpatient neonatal and young infant populations in multiple LMIC hospitals. While several single-
289 centre studies in LMIC have identified individual risk factors for culture positivity and mortality in
290 septic neonates,²³ previous severity scores derived from hospital-based LMIC cohorts have been
291 based on smaller studies assessing general illness severity rather than sepsis specifically.²⁴⁻²⁹ High-
292 income setting based general illness severity scores are widely used,²⁴⁻²⁶ and some have been
293 assessed in small populations of septic neonates.^{30,31} The nSOFA score is a recent sepsis-specific
294 score developed on 60 very low birth weight (<1500g) infants, but relies on measurement of
295 parameters which are often unavailable in LMIC settings (e.g. use of inotropes/vasoactive drug
296 use).³²

297 The NeoSep Severity and Recovery scores include clinical signs which have been designed to closely
298 align with the WHO pSBI criteria (compared in supplement table 2) used in recent community-based
299 studies, including AFRINEST¹¹ and SATT.^{12,13} However, the NeoSep scores include additional factors

300 more relevant to hospital settings which offer further predictive value, such as clinical evidence of
301 shock, and need for oxygen and/or respiratory support. Importantly, the study population also
302 included a more heterogenous mix of both preterm (including <1500g) and term infants, with early
303 and late onset, and community and healthcare-associated sepsis, and the severity and recovery
304 scores performed similarly across all these subgroups. Convulsions were uncommon in the study
305 (7%) and were not associated with mortality despite being a pSBI criteria for critical illness,
306 potentially due to low power.

307 The hospital sites included in this cohort were diverse, which is both a strength and potential
308 limitation. Routine clinical practice was determined by sites, and there was variation in access to
309 supportive care, with some sites using invasive ventilation and inotropic support and others with
310 only oxygen and no access to vasoactive medication. Variation in intensity of clinical monitoring may
311 also have influenced ascertainment of some factors such as shock, particularly as invasive
312 monitoring was unavailable in some sites. Patient populations also varied between sites, as some
313 were tertiary neonatal units, others treated more general populations, and there was variation in
314 proportions of in-born versus out-born neonates. We attempted to account for inter-site variation
315 by using multivariable models with site-level random-effects, excluding factors with a high
316 proportion of missing values. We also examined performance of the severity score across different
317 sites in high-, middle- and low-income contexts and demonstrated reproducible risk trends across all
318 settings. Additional limitations of this study include the potential for selection bias. Prospective
319 enrolment at the time of sepsis diagnosis may bias towards particular populations for whom
320 parental consent is most practical, potentially leading to a milder phenotype than retrospective
321 studies. Furthermore, the scores were validated on a predefined reserved 15% internal sample,
322 including only 42 deaths (and relatively few infants with high-risk scores). External validation would
323 strengthen the applicability of the scores. Nevertheless, inter-site variation reflected the diverse
324 contexts across LMIC, and the data are likely to be representative of populations and subgroups that
325 could feasibly be included in prospective antibiotic trials and subsequent global guideline
326 development.

327 Overall mortality was high in the NeoOBS cohort of hospitalized neonates and young infants with
328 sepsis in predominantly LMIC settings, with significant variation across different sites, highlighting
329 the importance of standardised criteria to define severity. The NeoSep Severity and Recovery scores
330 developed in this analysis define high-risk populations with clinical sepsis at treatment initiation and
331 during recovery which can inform future research. In addition, the NeoOBS study demonstrates the
332 feasibility of developing global hospital networks for conducting future trials focusing on improving
333 outcomes through optimized prevention and treatment of sepsis in neonates and young infants.

334

335
336

Table 1: Baseline characteristics

		N (%) or median (IQR)	
Infant characteristics	WHO region of enrolment:		
		Africa	998 (31.1%)
		Americas	79 (2.5%)
		South-East Asia	1201 (37.5%)
		Europe	121 (3.8%)
		Western Pacific	805 (25.1%)
		Age at baseline (days)	5 (1, 15)
		Sex	
		Male	1854 (57.9%)
		Female	1348 (42.1%)
		Indeterminate/intersex	2 (0.1%)
	Birth weight (grams)	2500 (1400, 3200)	
	Gestational age at birth (weeks)	37 (31, 39)	
	Estimated gestational age at birth, categories		
	Extremely preterm (<28 weeks)	227 (7.1%)	
	Very preterm (28 to <32 weeks)	607 (19.0%)	
	Moderate preterm 32 to <37 weeks	694 (21.7%)	
	Term (≥37 weeks)	1674 (52.3%)	
Birth history	Birth status		
		Hospitalized since birth	2215 (69.1%)
		Admitted from home/community	989 (30.9%)
	Time from admission to enrolment (hours)	22 (1, 126)	
Common sub-groups	Congenital anomalies	265 (8.3%)	
	Early-onset (age <48hrs)	1066 (39.3%)	
	Late onset community presenting, term	708 (26.1%)	
	Late onset healthcare associated, preterm	936 (34.5%)	
	Other	494 (15.4%)	
Supportive Care at baseline	IV fluid	2497 (77.9%)	
	Parenteral nutrition (TPN)	527 (16.4%)	
	Maximum respiratory support	None	1165 (36.4%)
		Oxygen supplementation	619 (19.3%)
		Non-invasive ventilation	719 (22.4%)
		Invasive ventilation	701 (21.9%)
		Nasogastric tube	1432 (44.7%)
Vital parameters	Oxygen Saturation (%)	96 (93, 98)	

	Respiratory Rate		52 (44, 60)
	Heart Rate		148 (138, 161)
	Temperature (°C)	<35.5	58 (1.8%)
		35.5-37.9	2737 (85.6%)
		≥38-<39	319 (10.0%)
		≥39	84 (2.6%)
Clinical signs at baseline (≥5%)	Respiratory signs*		2107 (65.8%)
	Difficulty feeding ☒		1465 (45.7%)
	Lethargy/reduced movement	Neither	2073 (64.7%)
		Lethargy only	803 (25.1%)
		Reduced/no movement	328 (10.2%)
	Abdominal distension		777 (24.3%)
	Evidence of shock		683 (21.3%)
	Apnoea		550 (17.2%)
	Jaundice requiring phototherapy		494 (15.4%)
	Grunting		435 (13.6%)
	Hypotonia/floppiness		435 (13.6%)
	Cyanosis		384 (12.0%)
	Irritability		322 (10.0%)
	Vomiting		287 (9.0%)
	Convulsions		236 (7.4%)

337

338 * Severe chest wall in-drawing, increased requirement for oxygen or respiratory support; ☒ Observed or reported, including feeding intolerance. All factors listed in this table were considered
 339 in the development of the baseline NeoSep severity score except WHO region, gestational age in categories (continuous gestational age considered instead), and common sepsis sub-groups
 340 (individual factors considered instead). Healthcare-associated=occurring >48h after hospital admission.

341

342

343 **Table 2: Predictors of mortality and risk score**

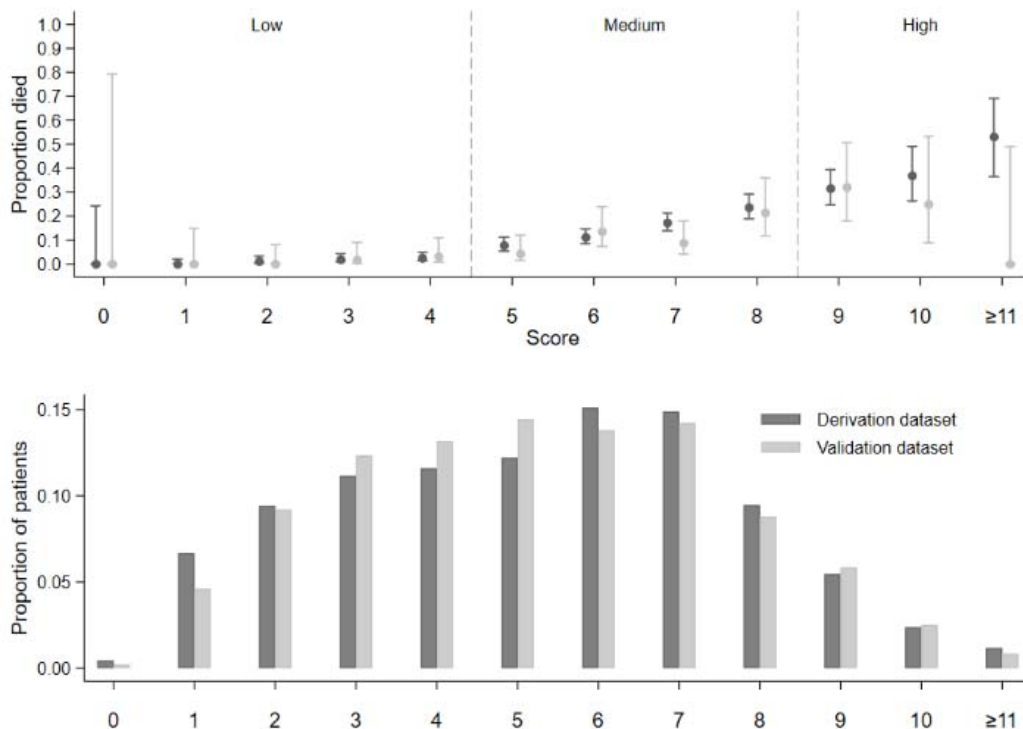
Factor	NeoSep Severity Score (at presentation)					NeoSep Recovery Score (daily on IV antibiotics)			
	HR (95% CI)	p-value	Coef	Severity score points [¶]	HR (95% CI)	p-value	Coef	Recovery score points [¶]	
Birth weight (kg) †	1.0	2.13 (1.55,2.92)	<0.001	0.67	<1 kg: 2				
	2.0	1.21 (1.12,1.31)		0.25	1-2 kg: 1				
	3.0	ref		0					
Time in hospital	per additional 24 hours	0.96 (0.94,0.98)	<0.001	-0.04	≤ 10 days: 1				
Gestational age	per additional week	0.94 (0.90,0.98)	0.004	-0.06	< 37 weeks: 1				
Congenital anomalies		2.81 (1.93,4.08)	<0.001	1.03	2				
Maximum respiratory support:	Oxygen supplementation	2.87 (1.74,4.74)	<0.001	1.06	2	9.21 (4.04,21.0)	<0.001	2.22	2
	CPAP, BiPAP, HFNC	4.72 (2.78,8.00)	<0.001	1.55	3	15.2 (6.65,34.8)	<0.001	2.72	3
	Invasive ventilation	10.2 (6.05,17.1)	<0.001	2.32	3	48.6 (21.5,110)	<0.001	3.88	3
Temperature (°C) †	35.5	1.17 (0.78,1.75)	<0.001	0.15	<35.5 °C: 1	1.32 (0.97,1.79)	0.005	0.26	<35.5 °C: 1
	37	ref		0		ref		0	
	38	1.41 (1.18,1.69)		0.35	≥38 – <39 °C: 1	1.43 (1.07,1.92)		0.34	≥38 – <39 °C: 1
	39	2.91 (1.76,4.81)		1.07	≥39°C: 2	3.07 (1.50,6.31)		1.11	≥39°C: 2
Abdominal distension		1.60 (1.24,2.07)	<0.001	0.47	1	1.50 (1.13,2.00)	0.005	0.41	1
Lethargy, no or reduced movement:	Lethargy only	1.28 (0.97,1.69)	0.083	0.25	1	1.85 (1.33,2.59)	<0.001	0.62	1
	No/reduced movement (± lethargy)	2.00 (1.36,2.95)	<0.001	0.70	2	3.10 (2.08,4.61)	<0.001	1.13	2
Difficulty feeding		1.50 (1.14,1.98)	0.004	0.41	1	2.15 (1.56,2.95)	<0.001	0.76	1
Evidence of shock		1.69 (1.28,2.24)	<0.001	0.53	1	2.51 (1.81,3.48)	<0.001	0.92	1
Cyanosis						1.87 (1.25,2.78)	0.002	0.62	1
Maximum number of score points possible					16				11

344 Notes: † Birth weight and temperature were analysed as continuous variables using fractional polynomials with powers -2 for birth weight, and powers 2 2 for temperature.
 345 For illustrative reasons, in this table we report HRs for specific values.[¶] For each model, score points for each factor were derived by dividing its coefficient by the smallest
 346 significant coefficient of any categorical factor, rounding to the nearest integer, and giving 1 score point for <3 raw points, 2 score points for 3-5 raw points, and 3 score
 347 points for ≥6 raw points. Shaded factors: not a significant predictor (cyanosis: baseline severity score) or deliberately not included in model (unmodifiable factors excluded
 348 from recovery score: birth weight, time in hospital, gestational age, congenital anomalies). CPAP = Continuous Positive Airway Pressure, BiPAP Bilevel Positive Airway
 349 Pressure, HFNC=High Flow Nasal Cannula.
 350

51 **Figure 1: NeoSep Severity Score at baseline**

52

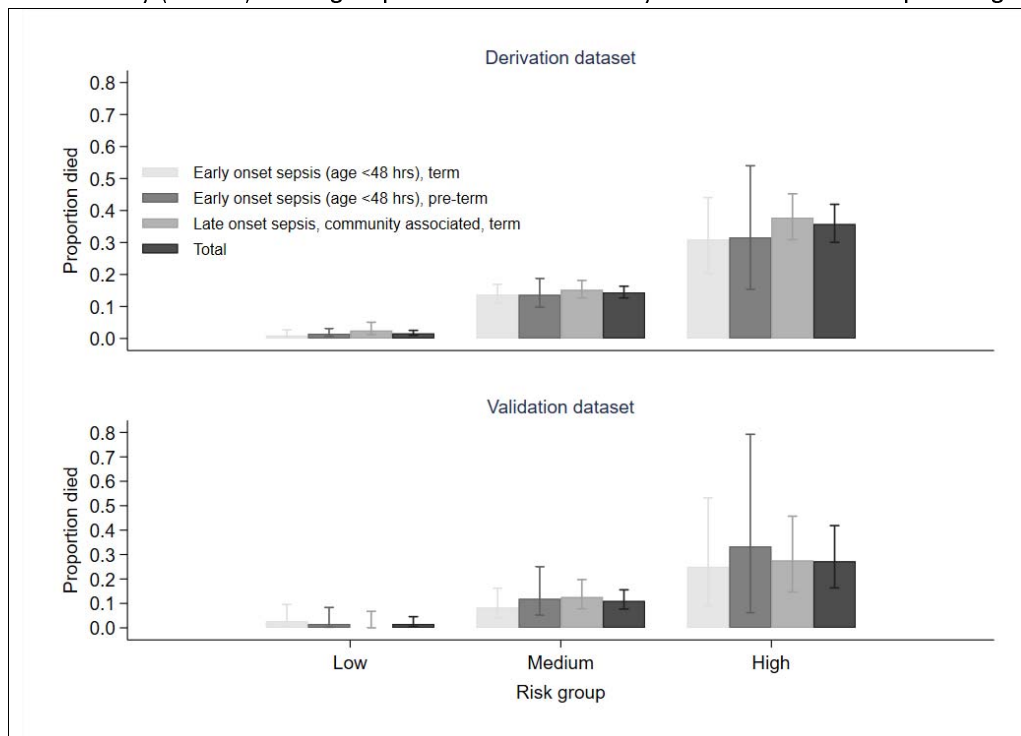
53 a. Proportion (95% CI) died per score point (top) and distribution of the Severity Score at baseline (bottom) in the
 54 derivation (dark grey) and in the validation dataset (light grey).



55

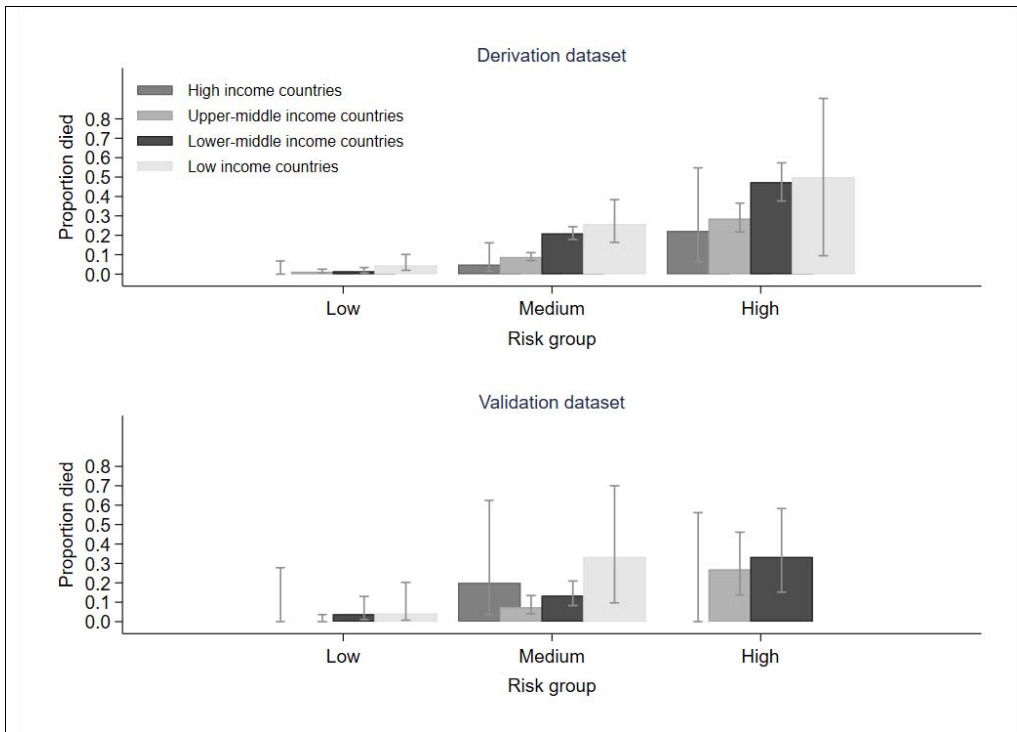
56

57 b. Mortality (95% CI) in risk groups based on the Severity Score and selected sepsis subgroups.



58

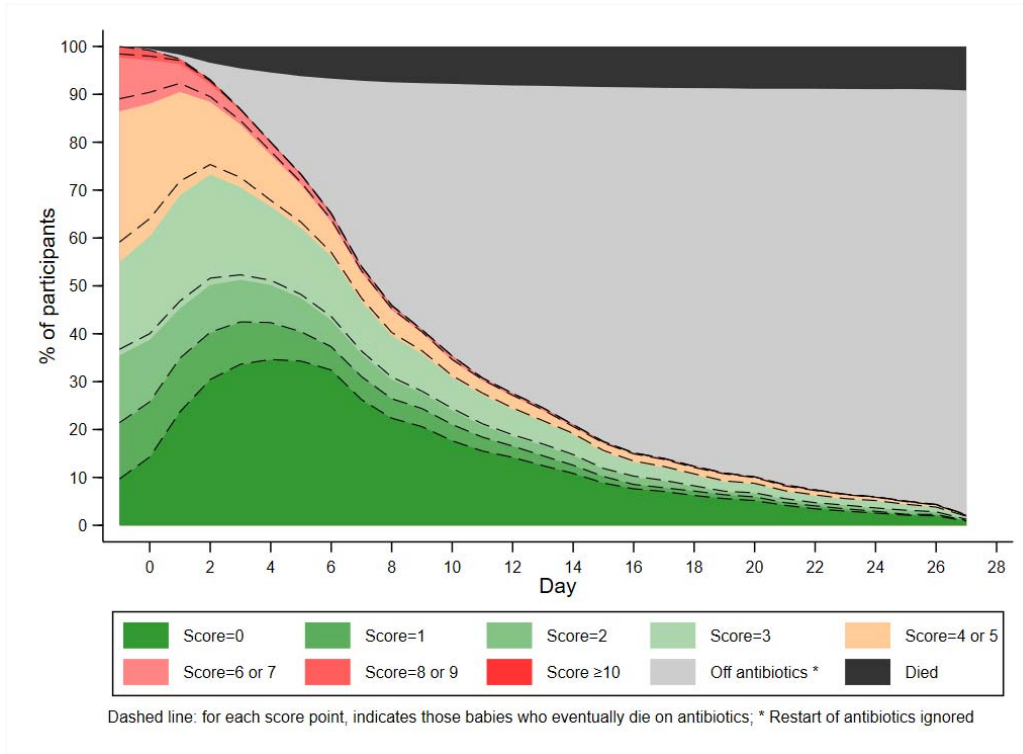
59 c. Mortality (95% CI) in risk groups based on the Severity Score and region.



50
51
52

53 **Figure 2: NeoSep Recovery Score over time (cross-sectional analysis).**

54

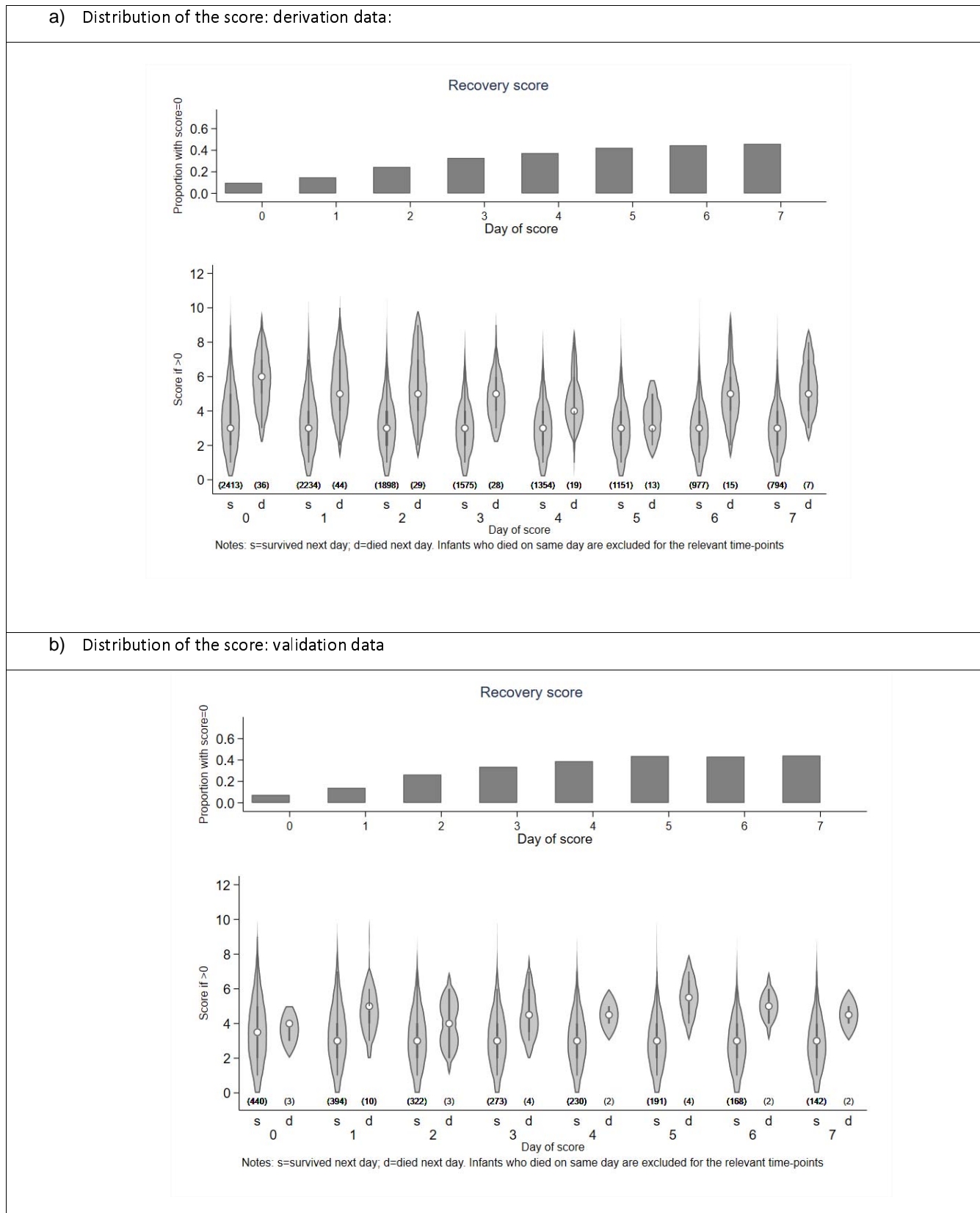


55

56

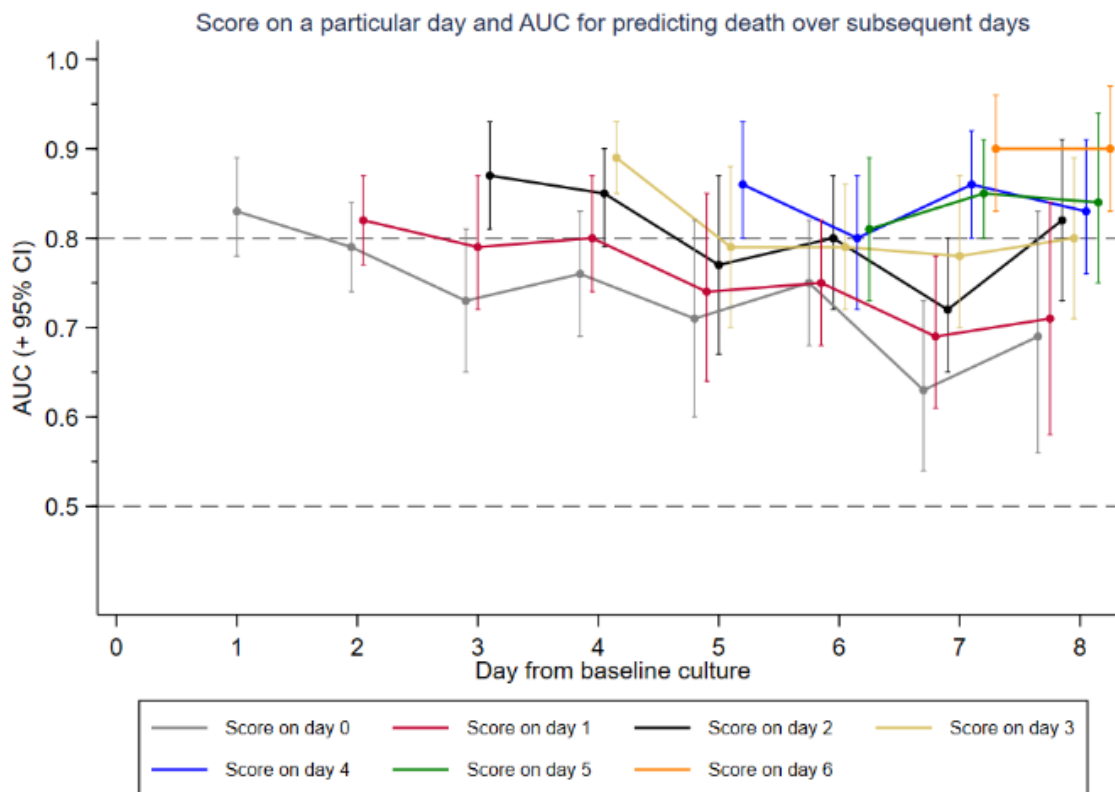
57 **Figure 3: NeoSep Recovery Score**

58



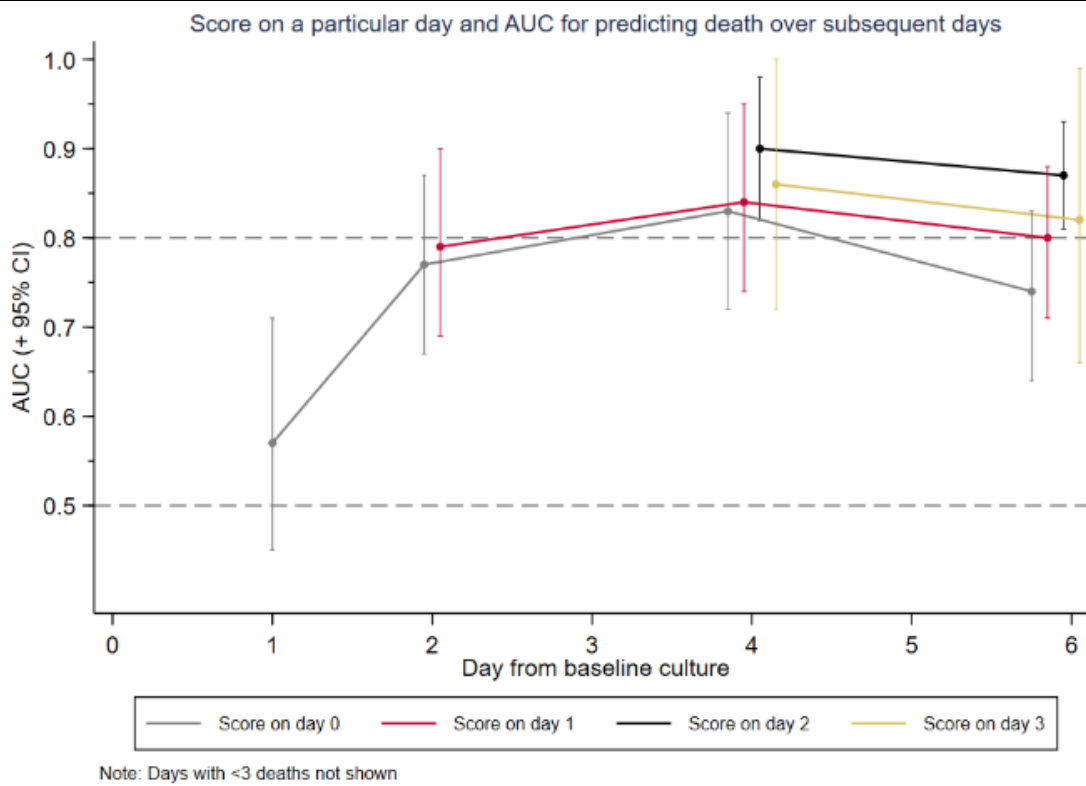
59

c) Time-updated AUROC: derivation data



*Lines represent the trend in predictive value over time of a score on a particular day for death of subsequent days

d) Time-updated AUROC: validation data



*Lines represent the trend in predictive value over time of a score on a particular day for death of subsequent days

70
71

References

1. James, S. L. *et al.* Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **392**, 1789–1858 (2018).
2. Roth, G. A. *et al.* Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **392**, 1736–1788 (2018).
3. Fuchs, A., Bielicki, J., Mathur, S., Sharland, M. & Van Den Anker, J. N. Reviewing the WHO guidelines for antibiotic use for sepsis in neonates and children. *Paediatr. Int. Child Health* **38**, S3–S15 (2018).
4. Okomo, U. *et al.* Aetiology of invasive bacterial infection and antimicrobial resistance in neonates in sub-Saharan Africa: a systematic review and meta-analysis in line with the STROBE-NI reporting guidelines. *Lancet Infect. Dis.* (2019). doi:10.1016/S1473-3099(19)30414-1
5. Chaurasia, S. *et al.* Neonatal sepsis in South Asia: huge burden and spiralling antimicrobial resistance. *BMJ* **364**, k5314 (2019).
6. Investigators of the Delhi Neonatal Infection Study (DeNIS) collaboration. Characterisation and antimicrobial resistance of sepsis pathogens in neonates born in tertiary care centres in Delhi, India: a cohort study. *Lancet Glob. Heal.* **4**, e752–e760 (2016).
7. Sands, K. *et al.* Characterization of antimicrobial-resistant Gram-negative bacteria that cause neonatal sepsis in seven low- and middle-income countries. *Nat. Microbiol.* **6**, 512–523 (2021).
8. Thomson, K. M. *et al.* Effects of antibiotic resistance, drug target attainment, bacterial pathogenicity and virulence, and antibiotic access and affordability on outcomes in neonatal sepsis: an international microbiology and drug evaluation prospective substudy (BARNARDS). *Lancet Infect. Dis.* (2021). doi:10.1016/S1473-3099(21)00050-5
9. Jackson, C. *et al.* Global Divergence From World Health Organization Treatment Guidelines for Neonatal and Pediatric Sepsis. *Pediatr. Infect. Dis. J.* **38**, (2019).
10. Fuchs, A., Bielicki, J., Mathur, S., Sharland, M. & Van Den Anker, J. N. Reviewing the WHO guidelines for antibiotic use for sepsis in neonates and children. *Paediatr. Int. Child Health* **38**, S3–S15 (2018).
11. Tshetu, A. *et al.* Oral amoxicillin compared with injectable procaine benzylpenicillin plus gentamicin for treatment of neonates and young infants with fast breathing when referral is not possible: a randomised, open-label, equivalence trial. *Lancet* **385**, 1758–1766 (2015).
12. Baqui, A. H. *et al.* Safety and efficacy of alternative antibiotic regimens compared with 7 day injectable procaine benzylpenicillin and gentamicin for outpatient treatment of neonates and young infants with clinical signs of severe infection when referral is not possible: a . *Lancet Glob. Heal.* **3**, e279–e287 (2015).
13. Zaidi, A. K. M. *et al.* Community-based treatment of serious bacterial infections in newborns and young infants: a randomized controlled trial assessing three antibiotic regimens. *Pediatr. Infect. Dis. J.* **31**, 667–72 (2012).
14. Oeser, C. *et al.* Clinical trials in neonatal sepsis. *J. Antimicrob. Chemother.* **68**, 2733–2745 (2013).
15. Tuzun, F. *et al.* Is European Medicines Agency (EMA) sepsis criteria accurate for neonatal sepsis diagnosis or do we need new criteria? *PLoS One* **14**, e0218002–e0218002 (2019).
16. Li, G. *et al.* Towards understanding global patterns of antimicrobial use and resistance in neonatal sepsis: insights from the NeoAMR network. *Arch. Dis. Child.* **105**, 26 LP – 31 (2020).
17. WHO. *Managing possible serious bacterial infection in young infants when referral is not feasible.* (2015).

18. Fitchett, E. J. A. *et al.* Strengthening the Reporting of Observational Studies in Epidemiology for Newborn Infection (STROBE-NI): an extension of the STROBE statement for neonatal infection research. *Lancet Infect. Dis.* **16**, e202–e213 (2016).
19. Harris, P. A. *et al.* Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *J. Biomed. Inform.* **42**, 377–381 (2009).
20. The Young Infants Clinical Signs Study Group. Clinical signs that predict severe illness in children under age 2 months: a multicentre study. *Lancet* **371**, 135–142 (2008).
21. Lutsar, I. *et al.* Meropenem vs standard of care for treatment of neonatal late onset sepsis (NeoMero1): A randomised controlled trial. *PLoS One* **15**, e0229380 (2020).
22. Walker, A. S. *et al.* Personalised randomised controlled trial designs—a new paradigm to define optimal treatments for carbapenem-resistant infections. *Lancet Infect. Dis.* **21**, e175–e181 (2021).
23. Liang, L. *et al.* Predictors of Mortality in Neonates and Infants Hospitalized With Sepsis or Serious Infections in Developing Countries: A Systematic Review. *Frontiers in Pediatrics* **6**, 277 (2018).
24. Aluvaala, J., Collins, G. S., Maina, M., Berkley, J. A. & English, M. A systematic review of neonatal treatment intensity scores and their potential application in low-resource setting hospitals for predicting mortality, morbidity and estimating resource use. *Syst. Rev.* **6**, 248 (2017).
25. Dorling, J. S., Field, D. J. & Manktelow, B. Neonatal disease severity scoring systems. *Arch. Dis. Child. - Fetal Neonatal Ed.* **90**, F11 LP-F16 (2005).
26. Garg, B., Sharma, D. & Farahbakhsh, N. Assessment of sickness severity of illness in neonates: review of various neonatal illness scoring systems. *J. Matern. Neonatal Med.* **31**, 1373–1380 (2018).
27. Medvedev, M. M. *et al.* Development and validation of a simplified score to predict neonatal mortality risk among neonates weighing 2000 g or less (NMR-2000): an analysis using data from the UK and The Gambia. *Lancet Child Adolesc. Heal.* **4**, 299–311 (2020).
28. Rosenberg, R. E. *et al.* Simplified Age-Weight Mortality Risk Classification for Very Low Birth Weight Infants in Low-Resource Settings. *J. Pediatr.* **153**, 519-524.e3 (2008).
29. Mansoor, K. P. *et al.* Modified Sick Neonatal Score (MSNS): A Novel Neonatal Disease Severity Scoring System for Resource-Limited Settings. *Crit. Care Res. Pract.* **2019**, 9059073 (2019).
30. Rosenberg, R. E. *et al.* Nosocomial sepsis risk score for preterm infants in low-resource settings. *J. Trop. Pediatr.* **56**, 82–89 (2010).
31. Pal, S., Jain, A., Garg, M. & Sekhar, J. C. Predicting Outcome in Neonates with Possible Clinical Sepsis by Estimating an Early Score for Neonatal Acute Physiology-II (SNAP-II). *J. Trop. Pediatr.* **66**, 377–384 (2020).
32. Wynn, J. L. & Polin, R. A. A neonatal sequential organ failure assessment score predicts mortality to late-onset sepsis in preterm very low birth weight infants. *Pediatr. Res.* **88**, 85–90 (2020).