

1 **Diagnosing early-onset neonatal sepsis in low-resource settings:**  
2 **development of a multivariable prediction model**

3  
4 Samuel R. Neal, MRes<sup>a,b</sup>, Felicity Fitzgerald, PhD<sup>b</sup>, Simbarashe  
5 Chimhuya, MMED<sup>c</sup>, Michelle Heys, MD(Res)<sup>a,d,\*</sup>, Mario Cortina-  
6 Borja, PhD<sup>a,†</sup>, Gwendoline Chimhini, MMED<sup>c,†</sup>

7  
8 **Affiliations:**

- 9 a. Population, Policy and Practice Research and Teaching Department, UCL Great  
10 Ormond Street Institute of Child Health, University College London, UK  
11 b. Infection, Immunity and Inflammation Research and Teaching Department, UCL  
12 Great Ormond Street Institute of Child Health, University College London, UK  
13 c. Child and Adolescent Health Unit, Faculty of Medicine and Health Sciences,  
14 University of Zimbabwe, Harare, Zimbabwe  
15 d. Specialist Children's and Young People's Services, East London NHS Foundation  
16 Trust, London, UK

17  
18 \* Corresponding author

19 † Contributed equally as last author

20  
21 **Address correspondence to:** Dr Michelle Heys, Population, Policy and Practice Research  
22 and Teaching Department, UCL Great Ormond Street Institute of Child Health, 30 Guilford  
23 Street, London, UK, WC1N 1EH; [m.heys@ucl.ac.uk](mailto:m.heys@ucl.ac.uk)

24  
25 **Keywords:** Decision Support Techniques; Infant, Newborn; Models, Statistical; Neonatal  
26 Sepsis; Zimbabwe

27  
28 **Abbreviations:** AIC, Akaike's information criterion; AUC, area under the curve; BIC,  
29 Bayesian information criterion; CI, confidence interval; EOS, early-onset neonatal sepsis;  
30 GBS, Group B streptococcus; HCP, healthcare professional; IQR, interquartile range; LRS,  
31 low-resource settings; OR, odds ratio; ROC, receiver operating characteristic; SD, standard  
32 deviation; SE, standard error; SMCH, Sally Mugabe Central Hospital.

33  
34 **Word count:** 3185; 2663

35 **ABSTRACT**

36

37 **Objective**

38 To develop a clinical prediction model to diagnose neonatal sepsis in low-resource settings.

39

40 **Design**

41 Secondary analysis of data collected by the Neotree digital health system from 01/02/2019 to  
42 31/03/2020. We used multivariable logistic regression with candidate predictors identified  
43 from expert opinion and literature review. Missing data were imputed using multivariate  
44 imputation and model performance was evaluated in the derivation cohort.

45

46 **Setting**

47 A tertiary neonatal unit at Sally Mugabe Central Hospital, Zimbabwe.

48

49 **Patients**

50 We included 2628 neonates aged <72 hours, gestation  $\geq 32^{+0}$  weeks and birth weight  $\geq 1500$   
51 grams.

52

53 **Interventions**

54 Participants received standard care as no specific interventions were dictated by the study  
55 protocol.

56

57 **Main outcome measures**

58 Clinical early-onset neonatal sepsis (within the first 72 hours of life), defined by the treating  
59 consultant neonatologist.

60

61 **Results**

62 Clinical early-onset sepsis was diagnosed in 297 neonates (11.3%). The optimal model  
63 included eight predictors: maternal fever, offensive liquor, prolonged rupture of membranes,  
64 neonatal temperature, respiratory rate, activity, chest retractions and grunting. Receiver  
65 operating characteristic analysis gave an area under the curve of 0.736 (95% confidence  
66 interval 0.701-0.772). For a sensitivity of 95% (92-97%), corresponding specificity was 11%  
67 (10-13%), positive predictive value 12% (11-13%), negative predictive value 95% (92-97%),  
68 positive likelihood ratio 1.1 (95% CI 1.0-1.1), and negative likelihood ratio 0.4 (95% CI 0.3-  
69 0.6).

70

71 **Conclusions**

72 Our clinical prediction model achieved high sensitivity with modest specificity, suggesting it  
73 may be suited to excluding early-onset sepsis. Future work will validate and refine this model  
74 before considering it for clinical use within the Neotree.

75

76 **What is already known on this topic**

77 Various clinical prediction models exist to diagnose neonatal sepsis. However, there is a  
78 paucity of literature on models in low-resource settings, particularly sub-Saharan Africa.

79

80 **What this study adds**

81 We developed a clinical prediction model to diagnose clinical early-onset neonatal sepsis  
82 with over 2,500 neonates in a lower middle-income, low-resource neonatal unit. Our model is  
83 easy to implement, does not require laboratory tests and achieved high sensitivity with  
84 modest specificity in the derivation cohort.

85

86 **How this study might affect research, practice or policy**

87 Our model could support less experienced healthcare professionals avoid unnecessary  
88 antibiotic therapy in the absence of immediate local senior support. Before implementation,  
89 this model must be externally validated and its impact on sepsis-related neonatal morbidity  
90 and mortality must be assessed in future studies.

91

92 **INTRODUCTION**

93

94 Neonatal sepsis caused 15% of the 2.5 million neonatal deaths worldwide in 2018 and has a  
95 mortality rate of 110-190 per 1000 livebirths.[1, 2] It can be difficult to diagnose as the  
96 clinical features overlap with non-infectious diseases.[3] Failing to treat sepsis with timely  
97 antimicrobials increases the risk of death or disability, but empirical antimicrobial therapy in  
98 non-infected neonates contributes to antimicrobial resistance and adverse outcomes.[4, 5]  
99

100 Isolating a pathogenic organism from a normally sterile site is the gold standard diagnostic  
101 method,[6] but has limitations. In low-resource settings (LRS), cultures and blood counts are  
102 often unavailable,[7] or turnaround times are too long to usefully inform management.[8, 9]  
103 Blood cultures have high sensitivity provided sufficient inoculate volume is obtained, but  
104 sampling can be difficult in unwell neonates.[10] Therefore, clinicians may diagnose sepsis  
105 and initiate empirical antimicrobial therapy despite negative cultures, based on clinical  
106 presentation, risk factors and/or raised inflammatory markers. This is often called ‘culture-  
107 negative’ sepsis and up to 16 times more neonates receive antibiotics for culture-negative  
108 sepsis than for sepsis with a positive culture.[11] Diagnostic challenges are increased in LRS  
109 where early neonatal care may be led by less experienced healthcare professionals (HCPs)  
110 without immediate local senior support.[8]

111

112 Clinical prediction models combine patient or disease characteristics to estimate the  
113 probability of a diagnosis or outcome.[12] Models to diagnose neonatal sepsis may improve  
114 diagnostic accuracy and rationalise antibiotic use. In LRS, they could provide clinical  
115 decision support for less experienced HCPs, especially if models do not require laboratory  
116 tests. Several existing models estimate the probability of neonatal sepsis,[13] for example, the

117 Kaiser Permanente Early-Onset Sepsis Calculator.[14] Unfortunately, few studies have  
118 investigated models in LRS, particularly in sub-Saharan Africa.[13]  
119  
120 Our primary objective was to develop a clinical prediction model to diagnose neonatal sepsis  
121 in a LRS neonatal unit, with the aim to support less experienced HCPs make this diagnosis.

122 **METHODS**

123

124 We describe methods according to the TRIPOD statement (Additional File 1).[15] Further

125 methods are in Additional File 2 and accompanying R code at

126 <https://doi.org/10.5281/zenodo.6969912>.

127

128 **Source of data**

129

130 We performed secondary analysis of data from the Neotree at the neonatal unit of Sally

131 Mugabe Central Hospital (SMCH), Zimbabwe. Data were collected over 14 months from

132 01/02/2019 to 31/03/2020.

133

134 The Neotree is an open-source digital health system for newborn care in LRS.[16, 17] It

135 combines evidence-based clinical decision support, education and digital data capture at

136 admission and discharge. It is embedded in routine practice at three neonatal units in sub-

137 Saharan Africa (Kamuzu Central Hospital, Malawi; SMCH, Zimbabwe; and Chinhoyi

138 Provincial Hospital, Zimbabwe).

139

140 On admission, HCPs complete an admission form using the Neotree application on an

141 Android tablet. The application guides assessment of the neonate and collects predefined

142 data. At discharge or after neonatal death, HCPs complete an outcome form, which includes

143 the final diagnoses or cause(s) of death after review by a consultant neonatologist (Additional

144 File 2, section 1).

145

146 **Participants**

147

148 SMCH has the largest of three tertiary neonatal units in Zimbabwe, with 100 cots. It admits

149 neonates born within the hospital and accepts national referrals for specialist surgical care.

150

151 We included neonates with chronological age <72 hours,  $\geq 32^{+0}$  weeks gestation at birth, and

152 birth weight  $\geq 1500$  grams. We excluded non-first-born multiples and those with a diagnosis

153 of major congenital anomaly, no outcome form completed, or anomalous admission durations

154 (e.g. date of discharge before date of admission).

155

## 156 **Outcome**

157

158 The primary outcome was clinical early-onset neonatal sepsis (EOS), defined as sepsis with

159 onset within the first 72 hours of life, as diagnosed by the treating consultant neonatologist

160 and recorded on the outcome form as one or more of: (i) primary discharge diagnosis, (ii)

161 additional problem during admission, (iii) primary cause of death, or (iv) contributory cause

162 of death. No specific actions were performed to blind outcome assessment.

163

## 164 **Predictors**

165

166 We identified candidate predictors through a modified Delphi method study,[18] and

167 literature review.[13] We mapped these predictors to available Neotree data, yielding 22

168 candidate predictors (Additional File 2, section 2). No specific actions were performed to

169 blind predictor assessment.

170

## 171 **Statistical analysis**

172

173 Analyses were performed in RStudio version 2022.02.0+443 (R version 4.1.3).[19, 20] No

174 specific sample size calculations were performed.

175

176 Data preparation

177

178 We linked admission and outcome forms using the Fellegi-Sunter framework of probabilistic

179 record linkage (Additional File 2, section 4).[21, 22] We imputed missing values using

180 multivariate imputation by chained equations.[23] Data were assumed to be missing at

181 random and we created 40 imputed datasets (Additional File 2, section 6).

182

183 Model development and specification

184

185 We used multivariable logistic regression to predict diagnosis of clinical EOS. For

186 convenience, model selection was performed in one dataset randomly selected from all

187 imputed datasets. First, we fitted a ‘full’ main effects model containing all candidate

188 predictors assuming linearity of continuous predictors and additivity at the predictor scale.

189 We excluded categorical variables with skewed distributions (<5% category prevalence in

190 either outcome group) if Fisher’s exact test was non-significant ( $p \geq 0.05$ ) for the  $m \times n$

191 contingency table. Otherwise, skewed categorical predictors were retained, and smaller

192 categories combined into an ‘other’ category. Next, we compared plausible variations to the

193 full model, selecting the model which minimised both Akaike’s information criterion (AIC)

194 and the Bayesian information criterion (BIC) as the ‘optimal’ model (Additional File 2,

195 section 8). We explored non-linear effects of continuous predictors with natural cubic spline

196 functions (2 to 10 degrees of freedom) and polynomial transformations (second-degree to



197 fifth-degree polynomials). We tested for the biologically plausible interaction between birth  
198 weight and gestational age. Finally, we fitted the optimal model across all imputed datasets  
199 and obtained pooled regression coefficients and their standard errors (SEs) using Rubin's  
200 rules.[24]

201

202 **Model performance**

203

204 We evaluated performance of the optimal model in the derivation cohort. Discrimination was  
205 quantified by plotting a receiver operating characteristic (ROC) curve in each imputed  
206 dataset. We pooled the area under the curve (AUC) and variance across imputed datasets  
207 using Rubin's rules.[24] Discrimination was visualised with a box plot and density plot of the  
208 distributions of predicted probabilities for each observed outcome group (in the single dataset  
209 used for model selection). We calculated Yates' discrimination slope as the absolute  
210 difference in mean predicted probabilities between the two observed outcome groups.[25]  
211 Sensitivity, specificity, predictive values, and likelihood ratios of the optimal model were  
212 estimated in the single dataset used for model selection. These metrics are presented for the  
213 'optimal' probability threshold according to Youden's *J* statistic,[26] and for thresholds  
214 corresponding to sensitivities of 80, 85, 90 and 95%. Confidence intervals for likelihood  
215 ratios and Yates' discrimination slope were obtained using bootstrap with 10,000 resamples  
216 (basic method or normal approximation).[27]

217

218 **Research ethics approval**

219

220 Research ethics approval was granted by the University College London Research Ethics  
221 Committee (16915/001, 5019/004), Medical Research Council Zimbabwe (MRCZ/A/2570),

222 and Sally Mugabe Central Hospital Ethics Committee (250418/48). The ‘session ID’ numbers  
223 shown in Additional File 2 were generated at the time of data import and, thus, were not  
224 known to anyone outside of the research group.

## 225 **RESULTS**

226

### 227 **Participants**

228

229 Of the 3577 neonates with matched admission and outcome records, 2628 (73.5%) were  
230 included (Figure 1). Mean gestational age was 38.0 (SD = 2.5) weeks, mean birth weight  
231 2889 (SD = 703) grams, and 221 (8.4%) neonates died (Table 1). In total, 297 had clinical  
232 EOS (11.3%, incidence 113 per 1000 admissions).

233

### 234 **Missing data**

235

236 In total, 14 variables had missing values. All variables had <1% missing values except  
237 temperature (31%) and birth weight (1.2%). Time since the start of data collection predicted  
238 missing temperature (odds ratio [OR] 0.96, 95% CI 0.96-0.96,  $p < 0.001$ ) as limited  
239 thermometers were available early in the study. Missing temperature was not associated with  
240 clinical EOS (OR 0.79, 95% CI 0.60-1.03,  $p = 0.084$ ).

241

### 242 **Model development**

243

244 From the set of 22 candidate predictors (Table 2), eight were excluded due to <5% category  
245 prevalence with a non-significant Fisher's exact test (cyanosis, seizures, fontanelle, colour,  
246 abdominal distention, omphalitis, abnormal skin appearance, and history of vomiting). Three  
247 of the five categories for activity had a prevalence of <5% in either outcome group but  
248 Fisher's exact test indicated a significant difference in the distribution between the two

249 groups ( $p < 0.001$ ). Activity was retained as a predictor and the three smaller categories were  
250 collapsed into one ‘other’ group.

251

252 Therefore, 14 candidate predictors were considered for model development. Of these, 12 had  
253 a significant univariable association with clinical EOS (Table 3). The strongest univariable  
254 predictor was maternal fever (OR 5.99, 95% CI 2.06-17.4). Neither birth weight (OR 1.14,  
255 95% CI 0.96-1.35) nor grunting at triage (OR 1.23, 95% CI 0.95-1.59) predicted clinical EOS  
256 in univariable models.

257

258 Among plausible multivariable models, a model containing eight of the 14 candidate  
259 predictors was selected as the optimal model (Additional File 2, section 8). Fitting non-linear  
260 effects for temperature or birth weight, or allowing for an interaction between birth weight  
261 and gestational age, did not improve fit.

262

### 263 **Model specification**

264

265 The optimal model included eight predictors: temperature at admission, respiratory rate,  
266 maternal fever during labour, offensive liquor, premature rupture of membranes, activity,  
267 chest retractions, and grunting (Table 4). It can be written as

268

$$\begin{aligned} \ln p(\text{EOS}) = & -39.4 + 0.99 * \text{temperature} + 0.06 * (\text{respiratory rate divided by } 5) + 1.44 \\ & * \text{maternal fever during labour} + 0.54 * \text{offensive liquor} + 0.36 \\ & * \text{prolonged rupture of membranes} + 0.59 * \text{lethargy} + 0.84 \\ & * \text{irritability, seizures or coma} + 0.41 * \text{chest retractions} + 0.18 * \text{grunting} \end{aligned}$$

269

270 where  $lp(EOS)$  is the logit transformation of the probability of clinical EOS. The probability  
271 of clinical EOS ( $Pr(EOS)$ ) is thus given by the inverse-logit function

272

$$Pr(EOS) = \frac{e^{lp(EOS)}}{1 + e^{lp(EOS)}}$$

273

## 274 **Model performance**

275

276 The pooled AUC was 0.736 (95% CI 0.701-0.772) (Figure 2). Median predicted probability  
277 was higher for observed cases with EOS than without EOS but there was significant overlap  
278 (Figure 3). Yates' discrimination slope was 0.11 (95% CI 0.064-0.13).

279

280 The 'optimal' classification threshold was 0.121 (i.e. 12.1% predicted probability of clinical  
281 EOS) yielding sensitivity 65% (95% CI 59-70%) and specificity 74% (95% CI 72-75%)  
282 (Table 5). For a sensitivity of 95%, the corresponding classification threshold was 0.034  
283 giving sensitivity 95% (95% CI 92-97%) and specificity 11% (95% CI 10-13%) (Table 5).

284 **DISCUSSION**

285

286 We developed a clinical prediction model to diagnose clinical EOS that can be applied in  
287 LRS. The optimal model included eight predictors: three perinatal risk factors (maternal fever  
288 during labour, offensive liquor, and prolonged rupture of membranes), and five clinical signs  
289 in the neonate (temperature, respiratory rate, activity on neurological examination, chest  
290 retractions, and grunting). Using a classification threshold for high sensitivity, this model had  
291 a relatively low specificity in the derivation cohort.

292

293 **Interpretation**

294

295 Incidence of clinical EOS in this cohort was 113 per 1000 admissions. This is greater than  
296 other estimates from low-income and middle-income countries, but there is marked  
297 heterogeneity between relatively few studies worldwide. For example, a 2019 meta-analysis  
298 estimated global EOS incidence of 31.1 per 1000 livebirths (95% CI 8.98-102.22;  $I^2$   
299 99.9%).[28]

300

301 Our model shares predictors with existing models for neonatal sepsis.[13] However, it does  
302 not need results from laboratory tests so is more applicable to LRS. Models exist for EOS  
303 that do not require laboratory tests (some of which have been validated in LRS), but data are  
304 limited to a few small studies. For example, Weber et al. developed a score with 14 clinical  
305 features to predict neonatal sepsis, meningitis, pneumonia or hypoxemia in neonates  
306 presenting to health facilities in four LRS countries.[29] Validation in the subgroup of 285  
307 neonates aged  $\leq 6$  days of life showed a sensitivity of 95% with a specificity of 26% if one or  
308 more clinical features were present.[29] Comparison of performance between models is

309 challenging as studies infrequently report readily comparable metrics (e.g. AUC) and often  
310 present performance at selected probability thresholds.

311

312 The Kaiser Permanente Early-Onset Sepsis Calculator has gained interest for managing  
313 neonates born at  $\geq 34$  weeks' gestation.[30, 31] It combines maternal perinatal risk factors  
314 with the neonate's clinical appearance to provide management recommendations based on the  
315 estimated probability of EOS. Meta-analyses suggest its use reduces rates of admission,  
316 antibiotic use and use of laboratory tests, without increased mortality (although some authors  
317 have voiced concerns about 'missed' or delayed diagnoses).[32-35] All included studies in  
318 these meta-analyses were performed in high-income countries.

319

320 The Kaiser Permanente calculator does not require results of laboratory tests but may be ill-  
321 suited to LRS. First, the baseline incidences of EOS used are lower than in most LRS: 0.1-4.0  
322 per 1000 livebirths for the calculator,[30] compared to 31.1 per 1000 livebirths (95% CI 8.98-  
323 102.22) in LRS.[28] Second, the calculator was developed in a population where Group B  
324 streptococcus (GBS) is the predominant organism in EOS and where antenatal GBS  
325 screening is performed routinely. The microbiology of EOS differs in LRS. *Staphylococcus*  
326 *aureus*, *Klebsiella* species and *Escherichia coli* are relatively more common isolates than  
327 GBS.[36, 37] Therefore, risk factors such as maternal GBS carriage and intrapartum  
328 antibiotic status are less relevant or unmeasured in LRS. Finally, descriptors used for  
329 categories of clinical presentation ("clinical illness", "equivocal" and "well appearing")  
330 include interventions such as mechanical ventilation, which are not useful measures of illness  
331 in neonatal units where these interventions are unavailable. Two studies have validated the  
332 Kaiser Permanente calculator in middle-income countries with variable results.[38, 39] No  
333 studies have validated the calculator in low-income countries or sub-Saharan Africa.

334

335 **Implications**

336

337 Our model is easy to apply in LRS. It includes clinical predictors and risk factors that are  
338 simple to identify by any cadre of HCP with minimal additional training. Acceptable  
339 thresholds of sensitivity and specificity will vary by clinical context. High sensitivity is  
340 important to avoid missing a true case of sepsis, but higher specificity would reduce risks  
341 associated with inappropriate antimicrobial therapy. During resource shortages, lower  
342 sensitivity might be favoured for higher specificity to allow treatment of neonates with the  
343 highest probability of EOS.

344

345 The high sensitivity with modest specificity of our model suggests it may be suited to  
346 excluding EOS. Approximately 300 neonates are admitted each month to SMCH.[40] At a  
347 sensitivity of 95% with our EOS incidence of 113 per 1000 admissions, we would expect one  
348 or two true cases of EOS to be missed per month. The corresponding negative predictive  
349 value of our model at this sensitivity was 95% (95% CI 92-97%). This might reassure HCPs  
350 antibiotics are not required.

351

352 **Limitations**

353

354 First, the Neotree collects data at admission and upon discharge or death. This limits granular  
355 analysis where the timing of clinical features or interventions is important. We restricted our  
356 analysis to EOS and to neonates admitted within 72 hours of birth so the association between  
357 clinical presentation at admission and final diagnoses was clearer. It is plausible neonates  
358 admitted for ‘safekeeping’ (while their mother received inpatient care) could have



359 unremarkable clinical appearance and vital signs on admission but develop symptoms of  
360 sepsis a few hours or days later.

361

362 Second, no specific actions were performed to blind assessment of the outcome. As we  
363 performed secondary analysis of data from a quality improvement project, the consultant  
364 neonatologist is unlikely to have been biased in their classification of EOS.

365

366 Third, although blood culture is the gold standard method for diagnosing EOS, erratic  
367 supplies of lab reagents meant we could not assess the correlation between positive blood  
368 cultures and the consultant neonatologists' diagnosis of EOS.

369

370 Fourth, relatively small sample size caused imprecise estimates of the effects of low  
371 prevalence predictors: history of maternal fever had the largest effect size in our optimal  
372 model with OR 4.21 and wide 95% CI 1.27-14.0. Similarly, we could not evaluate the  
373 predictive ability of some identified candidate predictors; for example, only 10 neonates  
374 (0.4%) had a bulging fontanelle. Large sample sizes would be needed to determine if  
375 inclusion of these features is beneficial.

376

377 Finally, we present model performance in the derivation data to maximise sample size for  
378 model development. Predictions made on the derivation cohort can be optimistic due to  
379 overfitting.[12]

380

381 **Conclusions**

382

383 Our prediction model to diagnose clinical EOS includes eight predictors: three perinatal risk  
384 factors (maternal fever during labour, offensive liquor, and prolonged rupture of membranes),  
385 and five clinical signs in the neonate (temperature, respiratory rate, activity on neurological  
386 examination, chest retractions, and grunting). With high sensitivity it achieved modest  
387 specificity in the derivation cohort suggesting it may be suited to excluding EOS, which  
388 could support HCPs' decisions to withhold antibiotics in non-septic neonates. Our future  
389 work will include (1) validating and refining this model; (2) evaluating the acceptability and  
390 feasibility of implementing this model via the Neotree; and (3) evaluating the impact of  
391 implementing this model on sepsis-related neonatal morbidity and mortality.

## 392 **ACKNOWLEDGMENTS**

393

394 We thank Dr David Musorowegomo, Dr Hannah Gannon and Ms Heather Chesters for  
395 assistance with the literature review. We thank Dr Liam Shaw and Mr Yali Sassoon for  
396 technical assistance exporting and manipulating Neotree data. We thank the wider Neotree  
397 team including Dr Caroline Crehan and Mr Tim Hull-Bailey for valuable discussions  
398 throughout the study. We also thank all the staff in the neonatal unit at Sally Mugabe Central  
399 Hospital, especially Dr Christopher Pasi (Chief Executive Officer), Dr Hopewell Mungani  
400 (Clinical Director), Ms Prisca Nyamapfeni, Matron Alice Mudzingwa and Matron Dade  
401 Pedzisai for local support. Finally, we are grateful to all the babies and families who  
402 participated.

403

## 404 **AUTHORS' CONTRIBUTIONS STATEMENT**

405

406 Dr Samuel Neal designed the study protocol, conducted the literature review, carried out the  
407 analyses, drafted the initial manuscript, and reviewed and revised the manuscript.

408

409 Drs Michelle Heys and Felicity Fitzgerald conceptualised the study, led the implementation  
410 of Neotree in Zimbabwe, supervised the analyses, and critically reviewed and revised the  
411 manuscript.

412

413 Drs Gwendoline Chimhini and Simbarashe Chimhuya led the implementation of Neotree in  
414 Zimbabwe, provided the data, contributed to study conception, and critically reviewed and  
415 revised the manuscript.

416

417 Professor Mario Cortina-Borja supervised the analyses, and critically reviewed and revised  
418 the manuscript.

419

420 All authors approved the final manuscript as submitted and agree to be accountable for all  
421 aspects of the work.

422

## 423 **CONFLICT OF INTEREST DISCLOSURES**

424

425 All authors declare they have no conflicts of interest.

426

## 427 **FUNDING STATEMENT**

428

429 This research was supported by the National Institute for Health Research (NIHR) Great  
430 Ormond Street Hospital Biomedical Research Centre. Funders of the wider Neotree project,  
431 past and present, include the Wellcome Trust Digital Innovation Award, RCPCH, Naughton-  
432 Cliffe Mathews, UCL Grand Challenges and Global Engagement Fund, and the Healthcare  
433 Infection Society. The funders had no role in the design of the study; in the collection,  
434 analyses, or interpretation of data; in the writing of the manuscript; or in the decision to  
435 publish the results.

436

## 437 **DATA SHARING STATEMENT**

438

439 An open-source, anonymised research database is planned as part of the wider Neotree  
440 project. Currently, sharing of deidentified individual participant data will be considered on a  
441 case-by-case basis.

## REFERENCES

1. United Nations Inter-agency Group for Child Mortality Estimation. Levels & Trends in Child Mortality: Report 2019. New York: United Nations Children's Fund; 2019.
2. Fleischmann-Struzek, C, Goldfarb, DM, Schlattmann, P, et al. The global burden of paediatric and neonatal sepsis: a systematic review. *The Lancet Respiratory Medicine*. 2018;6(3):223-30.
3. Shane, AL, Sanchez, PJ, Stoll, BJ. Neonatal sepsis. *Lancet*. 2017;390(10104):1770-80.
4. Esaiassen, E, Fjalstad, JW, Juvet, LK, et al. Antibiotic exposure in neonates and early adverse outcomes: a systematic review and meta-analysis. *Journal of Antimicrobial Chemotherapy*. 2017;72(7):1858-70.
5. Laxminarayan, R, Matsoso, P, Pant, S, et al. Access to effective antimicrobials: a worldwide challenge. *Lancet*. 2016;387(10014):168-75.
6. Iroh Tam, P-Y, Bendel, CM. Diagnostics for neonatal sepsis: current approaches and future directions. *Pediatr Res*. 2017;82(4):574-83.
7. Wiens, MO, Kumbakumba, E, Kissoon, N, et al. Pediatric sepsis in the developing world: challenges in defining sepsis and issues in post-discharge mortality. *Clinical epidemiology*. 2012;4(no issue):319-25.
8. Chimhini, G, Chimhuya, S, Madzudzo, L, et al. Auditing use of antibiotics in Zimbabwean neonates. *Infection Prevention in Practice*. 2020;2(2):100046.
9. Vamsi, SR, Bhat, RY, Lewis, LE, et al. Time to positivity of blood cultures in neonates. *Pediatric Infectious Disease Journal*. 2014;33(2):212-4.
10. Cantey, JB, Baird, SD. Ending the Culture of Culture-Negative Sepsis in the Neonatal ICU. *Pediatrics*. 2017;140(4):e20170044.
11. Klingenberg, C, Kornelisse, RF, Buonocore, G, et al. Culture-negative early-onset neonatal sepsis - at the crossroad between efficient sepsis care and antimicrobial stewardship. *Front Pediatr*. 2018;6(no issue):285.
12. Steyerberg, EW. *Clinical Prediction Models*. 2 ed. Switzerland: Springer International Publishing; 2019.
13. Neal, SR, Musorowegomo, D, Gannon, H, et al. Clinical prediction models to diagnose neonatal sepsis: a scoping review protocol. *BMJ Open*. 2020;10(8):e039712.
14. Kuzniewicz, MW, Walsh, EM, Li, S, et al. Development and Implementation of an Early-Onset Sepsis Calculator to Guide Antibiotic Management in Late Preterm and Term Neonates. *Joint Commission Journal on Quality & Patient Safety*. 2016;42(5):232-9.
15. Collins, GS, Reitsma, JB, Altman, DG, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement. *Annals of Internal Medicine*. 2015;162(1):55-63.

16. Heys, M, Kesler, E, Sassoon, Y, et al. Development and implementation experience of a learning healthcare system for facility based newborn care in low resource settings: The Neotree. *Learning Health Systems*. 2022;e10310.
17. Neotree. Neotree [GitHub repository]. 2022 [cited 2022 05 Jul]. Available from: <https://github.com/neotree/neotree>.
18. Evans, M, Corden, MH, Crehan, C, et al. Refining clinical algorithms for a neonatal digital platform for low-income countries: a modified Delphi technique. *BMJ Open*. 2021;11(5):e042124.
19. RStudio Team. RStudio: Integrated Development Environment for R. 2022.02.0+443 ed. Boston, MA: RStudio, Inc.; 2022.
20. R Core Team. R: A language and environment for statistical computing. 4.1.3 ed. Vienna, Austria: R Foundation for Statistical Computing; 2022.
21. Fellegi, IP, Sunter, AB. A theory for record linkage. *Journal of the American Statistical Association*. 1969;64(328):1183-210.
22. Enamorado, T, Fifield, B, Imai, K. fastLink: Fast probabilistic record linkage with missing data. 0.6.0 ed. CRAN: Enamorado, Ted; 2020.
23. van Buuren, S, Groothuis-Oudshoorn, K. mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*. 2011;45(3):1-67.
24. Rubin, DB. Inference and missing data. *Biometrika*. 1976;63(3):581-92.
25. Yates, JF. External correspondence: Decompositions of the mean probability score. *Organizational Behavior and Human Performance*. 1982;30(1):132-56.
26. Youden, WJ. Index for rating diagnostic tests. *Cancer*. 1950;3(1):32-5.
27. Davison, AC, Hinkley, DV. Bootstrap Methods and their Application. Cambridge: Cambridge University Press; 1997.
28. Fleischmann, C, Reichert, F, Cassini, A, et al. Global incidence and mortality of neonatal sepsis: a systematic review and meta-analysis. *Arch Dis Child*. 2021:archdischild-2020-320217.
29. Weber, MW, Carlin, JB, Gatchalian, S, et al. Predictors of neonatal sepsis in developing countries. *Pediatric Infectious Disease Journal*. 2003;22(8):711-7.
30. Kaiser Permanente. Neonatal Early-Onset Sepsis Calculator 2022 [cited 2022 05 Jul]. Available from: <https://neonatalsepsiscalculator.kaiserpermanente.org/>.
31. Kuzniewicz, MW, Puopolo, KM, Fischer, A, et al. A Quantitative, Risk-Based Approach to the Management of Neonatal Early-Onset Sepsis. *JAMA Pediatr*. 2017;171(4):365-71.

32. Achten, NB, Klingenberg, C, Benitz, WE, et al. Association of Use of the Neonatal Early-Onset Sepsis Calculator With Reduction in Antibiotic Therapy and Safety: A Systematic Review and Meta-analysis. *JAMA Pediatr.* 2019;173(11):1032-40.
33. Deshmukh, M, Mehta, S, Patole, S. Sepsis calculator for neonatal early onset sepsis – a systematic review and meta-analysis. *The Journal of Maternal-Fetal & Neonatal Medicine.* 2019:1-9.
34. Pettinger, KJ, Mayers, K, McKechnie, L, et al. Sensitivity of the Kaiser Permanente early-onset sepsis calculator: A systematic review and meta-analysis. *EClinicalMedicine.* 2020;19:100227.
35. Rajbhandari, S, La Gamma, EF. Early-Onset Sepsis Calculator—Risk of Delaying Treatment. *JAMA Pediatrics.* 2017;171(10):1015-.
36. Okomo, U, Akpalu, ENK, Le Doare, K, 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. *The Lancet Infectious Diseases.* 2019;19(11):1219-34.
37. Zaidi, AKM, Thaver, D, Ali, SA, et al. Pathogens associated with sepsis in newborns and young infants in developing countries. *The Pediatric infectious disease journal.* 2009;28(1 Suppl):S10-8.
38. He, Y, Chen, J, Liu, Z, et al. Efficacy and safety of applying a neonatal early-onset sepsis risk calculator in China. *J Paediatr Child Health.* 2020;56(2):237-43.
39. Al-Lawama, M, AlZaatreh, A, Elrajabi, R, et al. Prolonged Rupture of Membranes, Neonatal Outcomes and Management Guidelines. *J Clin Med Res.* 2019;11(5):360-6.
40. Gannon, H, Chimhuya, S, Chimhini, G, et al. Electronic application to improve management of infections in low-income neonatal units: pilot implementation of the NeoTree beta app in a public sector hospital in Zimbabwe. *BMJ Open Quality.* 2021;10(1):e001043.

**Table 1. Characteristics of the study participants**

Characteristics	Overall	No sepsis	Sepsis	<i>p</i> -value
	<i>n</i> 2628	2331	297	
<b>Admission</b>				
Gestational age, weeks	38.0 (2.5)	38.0 (2.5)	38.4 (2.3)	0.005
Birth weight, grams	2889 (703)	2881 (716)	2950 (595)	0.067
Sex				0.7
Male	1503 (57%)	1338 (57%)	165 (56%)	
Female	1122 (43%)	990 (42%)	132 (44%)	
Unsure	3 (0.1%)	3 (0.1%)	0 (0.0%)	
Type of birth				0.032
Singleton	2496 (95%)	2205 (95%)	291 (98%)	
First-born twin	127 (4.8%)	121 (5.2%)	6 (2.0%)	
First-born triplet	2 (<0.1%)	2 (<0.1%)	0 (0.0%)	
Mode of delivery				0.074
SVD	1889 (72%)	1663 (71%)	226 (76%)	
Elective C-section	136 (5.2%)	124 (5.3%)	12 (4.0%)	
Emergency C-section	561 (21%)	510 (22%)	51 (17%)	
Instrumental	42 (1.6%)	34 (1.5%)	8 (2.7%)	
Postnatal age				<0.001
< 2 hours of life	1001 (38%)	901 (39%)	100 (34%)	
2-24 hours of life	1257 (48%)	1136 (49%)	121 (41%)	
24-48 hours of life	235 (9.0%)	181 (7.8%)	54 (18%)	
48-72 hours of life	110 (4.2%)	91 (3.9%)	19 (6.5%)	
<b>Outcome</b>				
Admission duration	2.3 [1.3-4.9]	2.1 [1.2-4.1]	6.0 [3.5-8.8]	<0.001
Death	221 (8.4%)	184 (7.9%)	37 (12%)	0.008

Data are presented as mean (SD), *n* (%) or median [quartile 1 - quartile 3]. *P*-values are from Welch's two-sample *t*-test for gestational age and birth weight; the Wilcoxon-Mann-Whitney *U* test for admission duration; Pearson's chi-squared test for postnatal age at admission and death; and Fisher's exact test for sex, type of birth, and mode of delivery. C-section = caesarean section; SVD = spontaneous vaginal delivery.

**Table 2. Distributions of candidate predictors in the study cohort**

Candidate predictor	Overall	No sepsis	Sepsis	<i>p</i> -value
	<i>n</i> 2628	2331	297	
<b>Infant risk factors</b>				
Gestational age, weeks	38.0 [37.0-40.0]	38.0 [37.0-40.0]	38.0 [37.0-40.0]	0.032
Birth weight, grams	2950 [2400-3350]	2900 [2400-3350]	3000 [2600-3350]	0.035
<b>Maternal risk factors</b>				
Maternal fever	14 (0.5%)	8 (0.3%)	6 (2.0%)	0.003
Offensive liquor	163 (6.2%)	131 (5.6%)	32 (11%)	0.001
PROM	303 (12%)	257 (11%)	46 (15%)	0.027
<b>Infant clinical features</b>				
Grunting at triage	750 (29%)	654 (28%)	96 (32%)	0.13
Cyanosis at triage*	69 (2.6%)	60 (2.6%)	9 (3.0%)	0.6
Seizures at triage*	14 (0.5%)	10 (0.4%)	4 (1.3%)	0.064
Respiratory rate, breaths/min	56 [48-68]	56 [48-68]	60 [50-72]	<0.001
Heart rate, beats/min	138 [126-146]	138 [126-146]	139 [127-150]	0.011
Temperature, Celsius	36.5 [36.0-37.0]	36.5 [36.0-36.9]	36.9 [36.2-38.0]	<0.001
Fontanelle*				0.9
Flat	2,608 (99%)	2,312 (99%)	296 (100%)	
Sunken	10 (0.4%)	9 (0.4%)	1 (0.3%)	
Bulging	10 (0.4%)	10 (0.4%)	0 (0%)	
Activity†				<0.001
Alert	2,152 (82%)	1,933 (83%)	219 (74%)	
Lethargic	382 (15%)	327 (14%)	55 (19%)	
Irritable	62 (2.4%)	45 (1.9%)	17 (5.7%)	
Seizures	14 (0.5%)	9 (0.4%)	5 (1.7%)	
Coma	18 (0.7%)	17 (0.7%)	1 (0.3%)	
Nasal flaring	912 (35%)	791 (34%)	121 (41%)	0.023
Chest retractions	986 (38%)	848 (36%)	138 (46%)	<0.001
Grunting	421 (16%)	360 (15%)	61 (21%)	0.029
Work of breathing				<0.001
Normal	1,405 (54%)	1,263 (55%)	142 (48%)	
Mildly increased	413 (16%)	378 (16%)	35 (12%)	
Moderately increased	614 (24%)	529 (23%)	85 (29%)	
Severely increased	170 (6.5%)	139 (6.0%)	31 (11%)	
Colour*				0.11
Pink	2,507 (95%)	2,220 (95%)	287 (97%)	
Pale	10 (0.4%)	7 (0.3%)	3 (1.0%)	
Blue	62 (2.4%)	58 (2.5%)	4 (1.3%)	
Yellow	49 (1.9%)	46 (2.0%)	3 (1.0%)	
Abdominal distention*	28 (1.1%)	26 (1.1%)	2 (0.7%)	0.8
Omphalitis*	6 (0.2%)	4 (0.2%)	2 (0.7%)	0.14
Abnormal skin*	27 (1.0%)	23 (1.0%)	4 (1.3%)	0.5
Vomiting*				0.3
No	2,605 (99%)	2,309 (99%)	296 (100%)	
Yellow	7 (0.3%)	7 (0.3%)	0 (0%)	
Bilious	13 (0.5%)	13 (0.6%)	0 (0%)	
Blood-stained	3 (0.1%)	2 (<0.1%)	1 (0.3%)	

\*Eliminated from the final set of candidate predictors due to very skewed distributions. †The three smallest categories of activity were collapsed into one ‘other’ category for model development. Data are presented as median [quartile 1 - quartile 3] for continuous predictors or *n* (%) for categorical predictors. *P*-values are from the Wilcoxon-Mann-Whitney *U* test for continuous predictors and Fisher’s exact test for categorical predictors. Distributions are presented for the observed data only, before multiple imputation of missing values.



**Table 3. Univariable association between candidate predictors and outcome**

<b>Candidate predictor</b>	<b>Coeff</b>	<b>SE</b>	<b>OR</b>	<b>95% CI</b>	<b>p-value</b>
<b>Infant risk factors</b>					
Gestational age, weeks	0.067	0.026	1.07	1.02 – 1.12	0.009
Birth weight, kilograms	0.131	0.087	1.14	0.96 – 1.35	0.13
<b>Maternal risk factors</b>					
Maternal fever	1.79	0.544	5.99	2.06 – 17.4	0.001
Offensive liquor	0.707	0.208	2.03	1.35 – 3.05	0.001
PROM	0.391	0.173	1.48	1.05 – 2.08	0.024
<b>Infant clinical features</b>					
Grunting at triage	0.203	0.132	1.23	0.95 – 1.59	0.13
Respiratory rate, 5 breaths/min	0.093	0.022	1.10	1.05 – 1.14	<0.001
Heart rate, 5 beats/min	0.047	0.019	1.05	1.01 – 1.09	0.012
Temperature, degrees Celsius	0.886	0.087	2.42	2.04 – 2.88	<0.001
<b>Activity</b>					
Alert – Lethargic	0.395	0.162	1.48	1.08 – 2.04	0.015
Alert – Other	1.05	0.25	2.86	1.75 – 4.67	<0.001
Nasal flaring	0.290	0.126	1.34	1.04 – 1.71	0.021
Chest retractions	0.417	0.124	1.52	1.19 – 1.93	0.001
Grunting	0.346	0.155	1.41	1.04 – 1.92	0.025
<b>Work of breathing</b>					
Normal – Mildly increased	-0.207	0.197	0.813	0.552 – 1.20	0.29
Normal – Moderately increased	0.345	0.146	1.41	1.06 – 1.88	0.018
Normal – Severely increased	0.674	0.217	1.96	1.28 – 3.00	0.002

Analyses were performed on the complete data after multiple imputation of missing values. CI = confidence interval; OR = odds ratio; SE = standard error.

**Table 4. Predictors and their pooled regression coefficients and odds ratios for the optimal model**

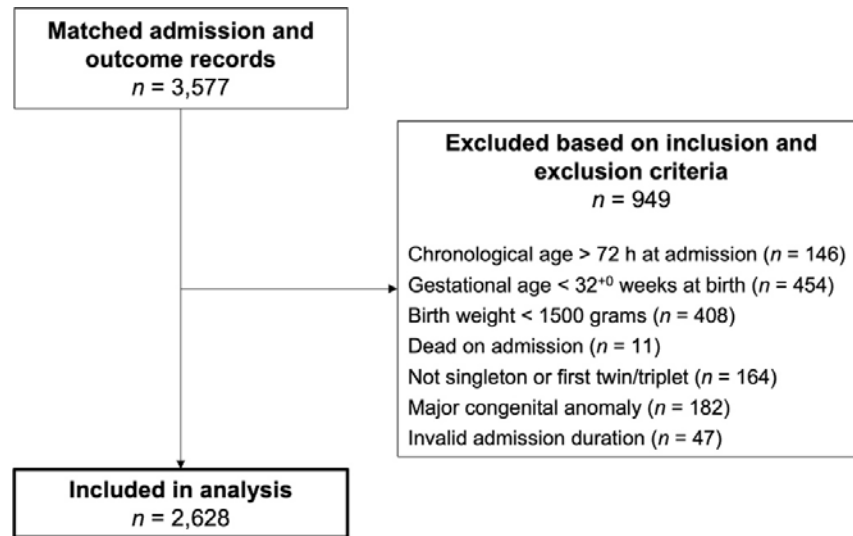
<b>Candidate predictor</b>	<b>Coeff</b>	<b>SE</b>	<b>OR</b>	<b>95% CI</b>	<b>p-value</b>
<i>Intercept</i>	-39.4	3.52			
Temperature, degrees Celsius	0.987	0.095	2.68	2.23 – 3.23	<0.001
Respiratory rate, 5 breaths/min	0.055	0.026	1.06	1.00 – 1.11	0.037
Maternal fever	1.44	0.612	4.21	1.27 – 14.0	0.019
Offensive liquor	0.543	0.228	1.72	1.10 – 2.69	0.017
PROM	0.360	0.192	1.43	0.98 – 2.09	0.06
Activity (Alert – Lethargic)	0.586	0.184	1.80	1.25 – 2.58	0.002
Activity (Alert – Other)	0.840	0.286	2.32	1.32 – 4.06	0.003
Chest retractions	0.406	0.172	1.50	1.07 – 2.10	0.019
Grunting	0.179	0.186	1.20	0.83 – 1.72	0.3

Analyses were performed on the complete data after multiple imputation of missing values. CI = confidence interval; OR = odds ratio; SE = standard error.

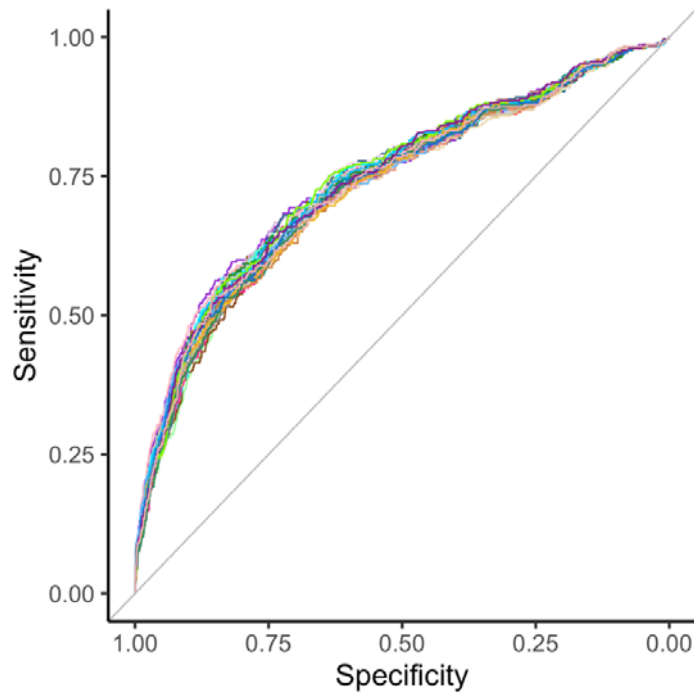
**Table 5. Model performance at several classification thresholds of predicted probability**

<b>Threshold</b>	<b>Sens</b>	<b>Spec</b>	<b>PPV</b>	<b>NPV</b>	<b>LR+</b>	<b>LR-</b>
0.121*	65 (59 – 70)	74 (72 – 75)	24 (21 – 27)	94 (93 – 95)	2.4 (1.6 – 2.9)	0.5 (0.4 – 0.6)
0.075	81 (76 – 85)	44 (42 – 46)	15 (14 – 17)	95 (93 – 96)	1.4 (1.0 – 1.6)	0.4 (0.4 – 0.5)
0.067	85 (80 – 88)	38 (36 – 40)	15 (13 – 17)	95 (94 – 96)	1.4 (1.2 – 1.6)	0.4 (0.2 – 0.5)
0.047	90 (86 – 93)	22 (20 – 24)	13 (12 – 14)	95 (92 – 96)	1.2 (0.9 – 1.2)	0.4 (0.3 – 0.6)
0.034	95 (92 – 97)	11 (10 – 13)	12 (11 – 13)	95 (92 – 97)	1.1 (1.0 – 1.1)	0.4 (0.3 – 0.6)

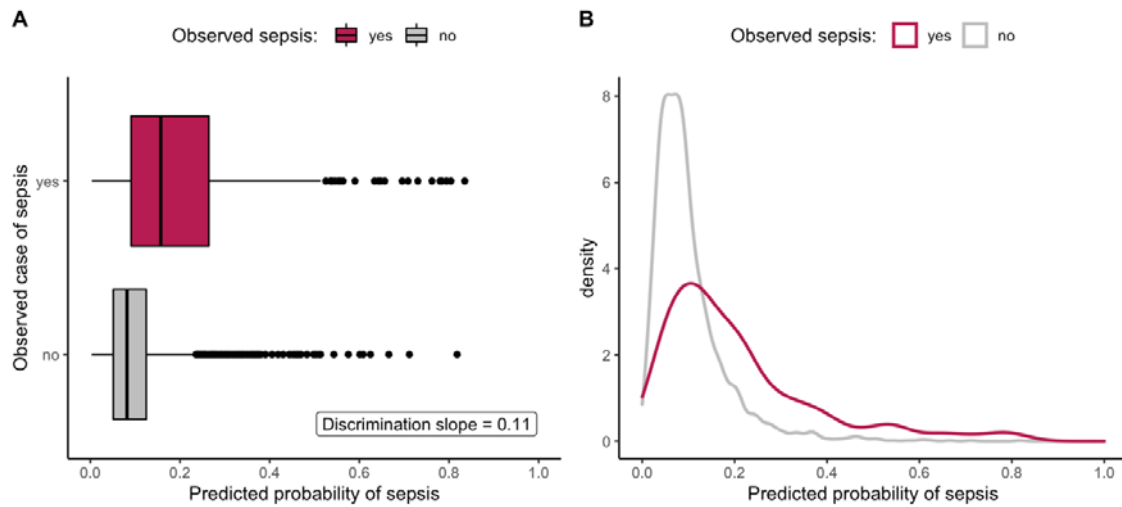
\*The ‘optimal’ threshold according to Youden’s *J* statistic. Data are presented for the single dataset used for model selection. Numbers in brackets represent the 95% confidence intervals. LR+ = positive likelihood ratio; LR- = negative likelihood ratio; PPV = positive predictive value; NPV = negative predictive value; sens = sensitivity; spec = specificity.



**Figure 1.** Flow diagram summarising participant inclusion and exclusion. Participants could fulfil multiple inclusion and/or exclusion criteria, therefore, the sum of participants excluded based on each criterion exceeds 949.



**Figure 2.** Receiver operating characteristic curve for the optimal model in each of the 40 imputed datasets. Pooled area under the curve (AUC) = 0.736 (95% confidence interval 0.701 – 0.772).



**Figure 3.** Boxplot (panel A) and density plot (panel B) of predicted probabilities of early-onset sepsis by observed outcome for the optimal model in the single dataset used for model selection.

**Matched admission and  
outcome records**

$n = 3,577$

**Excluded based on inclusion and  
exclusion criteria**

$n = 949$

Chronological age > 72 h at admission ( $n = 146$ )

Gestational age < 32<sup>+0</sup> weeks at birth ( $n = 454$ )

Birth weight < 1500 grams ( $n = 408$ )

Dead on admission ( $n = 11$ )

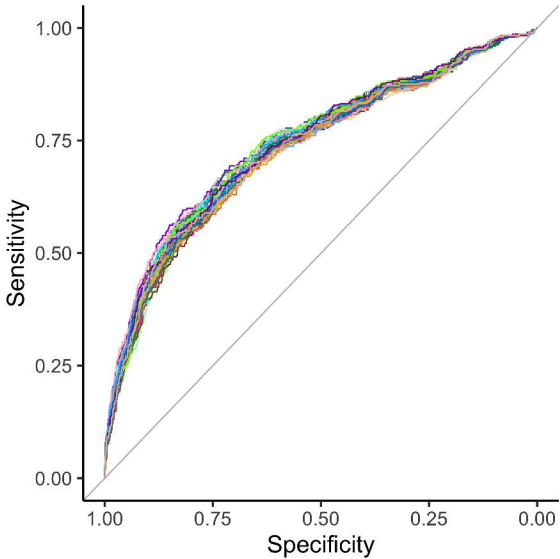
Not singleton or first twin/triplet ( $n = 164$ )

Major congenital anomaly ( $n = 182$ )

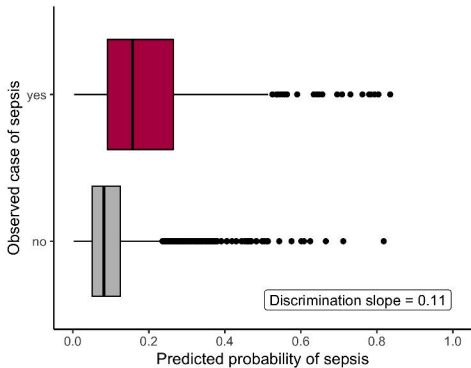
Invalid admission duration ( $n = 47$ )

**Included in analysis**

$n = 2,628$





**A**Observed sepsis: ■ yes ■ no**B**Observed sepsis: □ yes □ no