Characterisation of the Blood Cell Response in Neonatal Sepsis



Emily Hyde

Kellogg College

Nuffield Department of Women's and Reproductive Health
University of Oxford

A thesis presented for the degree of
Doctor of Philosophy in Women's and Reproductive Health

Trinity Term 2025

Abstract

Neonatal sepsis remains a major cause of mortality and morbidity, particularly in neonates admitted to Neonatal Intensive Care Units (NICUs). The current gold standard for diagnosis is a positive blood culture, which is limited by its low sensitivity and slow turnaround time. There remains a need for a rapid diagnostic tool capable of early detection of neonates at risk of developing neonatal sepsis.

This thesis investigates the diagnostic utility of the complete blood count (CBC) for neonatal sepsis, with a focus on late-onset sepsis. The main hypothesis was that specific CBC parameters, alone or in combination, could identify neonates with sepsis, and that multivariable diagnostic models would achieve sufficient accuracy to support clinical decision-making. A systematic review and meta-analysis assessing the current evidence found that, while some parameters demonstrated moderate potential, the evidence base was limited and yielded inconsistent results, highlighting the need for further investigation.

To explore this, CBC results from neonates admitted to the John Radcliffe Hospital NICU were analysed, classified based on clinical and laboratory signs of sepsis. Comparisons between septic and control neonates revealed that several parameters had strong diagnostic potential, including novel immune activation markers. Building upon these findings, I developed multivariable models that performed well across training, internal testing, and external validation cohorts. In longitudinal case studies, these models detected infections earlier than C-reactive protein and remained elevated throughout infection.

In conclusion, this work demonstrates that the neonatal CBC offers clinically valuable information for sepsis diagnosis. Individual parameters showed strong diagnostic potential, while multivariable models were highly accurate and generalisable. Although additional validation in larger, multicentre cohorts is necessary before clinical implementation, these findings suggest that the CBC, a widely available and inexpensive test, is a reliable diagnostic tool for neonatal sepsis.

Declaration

I, Emily Hyde, declare that this Thesis is my own work and has not been submitted, either wholly or substantially, for a degree in this university or elsewhere.

This thesis comprises approximately 28,180 words.



03/10/2025

Signature *Date*

Copyright ©2025 by Emily Hyde

All Rights Reserved

Statement of Contributions

This thesis presents original research undertaken by me as part of my DPhil. The overarching research theme – investigating the utility of the complete blood count in neonatal sepsis – was defined prior to my appointment to the project. Building upon this, I designed each study in collaboration with my supervisory team. I conducted all data analyses independently, interpreted the results, and wrote the thesis in full.

For Chapter 2, I developed the review protocol, registered the study with PROSPERO, conducted the literature searches, extracted data, performed the meta-analyses, and interpreted the findings. Two of my supervisors, Drs Mark Anthony and Jarob Saker, contributed to full-text screening, in line with best practice for systematic reviews. My supervisory team provided ongoing guidance and critically reviewed the manuscript for publication.

For Chapters 3 – 6, I defined all inclusion and exclusion criteria, collected and cleaned the clinical and laboratory datasets, performed all statistical analyses, developed and validated the diagnostic models, and interpreted all results. My supervisors provided ongoing guidance on study design, interpretation, and manuscript preparation.

Chapter 5 used an external validation cohort provided by the Erasmus University Medical Centre (EMC). Janno Schouten, a PhD student at the EMC, performed the data collection and cleaning, while I conducted all data analyses and interpreted the results independently. The EMC had no role in study design, analysis, or writing of the thesis.

Haematological data analysed in this thesis were generated using Sysmex haematology analysers as part of routine clinical care. Sysmex provided an XN-450 analyser and reagents to the John Radcliffe NICU for the duration of the study. Dr Jarob Saker, my industrial supervisor at Sysmex, provided technical and scientific guidance. Sysmex had no role in study design, data analysis, results interpretation, model development, or writing of this thesis.

Acknowledgements

I would like to begin by expressing my heartfelt thanks to my supervisors. To Professor Manu Vatish, without whom I would not have been able to study at Oxford, and who supported me through every milestone and obstacle during my DPhil journey. To Professor Stephen Kennedy, for his input and guidance during the course of this project. To Dr. Mark Anthony, from whom I learnt a great deal – your generosity with your time, ideas, and encouragement had a lasting impact on both this project and my growth as a researcher. And, last but not least, Dr. Jarob Saker – thank you for your unwavering support, for your tough love when I needed it, and for introducing me to your wonderful home city of Hamburg, which will always hold a special place in my heart. I could not have achieved this without each of you, and I will forever cherish everything you have taught me and the support you have given along the way.

To the past and present members of Team Manu, a wonderful, inspiring group of talented and friendly individuals: Gabriel Jones, Wei Zhang, Michelle Ma, William Cooke, Neva Kandzija, Prassana Logenthiran, Shuhan Jiang, Maryam Rahbar, Asselzhan Sarbalina, Angga Lokeswara, Katherine Dibbon, and Sofia Cerdeira. A special thank you to Morganne Wilbourne, Emily Mazey, Sarah Hutchinson, and Bryony Davies – the most wonderful group of women I could have had the pleasure of meeting. I am so thankful for your friendship and kindness, and I look forward to seeing all the fantastic things you will achieve. To my peers, Catarina Palma Dos Reis and Faheem Seedat – it was a pleasure undertaking my DPhil alongside you both. You have been an inspiration and a constant pillar of support through all the trials and tribulations, and I wish you both every success in the future.

I would also like to thank the staff at the NDWRH. To Lesa Levitt and Laura Goode, for their unwavering support during some of my most challenging times at Oxford – I thank you both immensely. To the research midwives, John Radcliffe NICU nurses, and John Radcliffe Haematology Laboratory staff – this project would not have been possible

without your hard work in data collection and your belief in my project. Thank you for believing in my project. I also extend a special thank you to Shirley Jenkins, Fenella Roseman, and Michael Jacobs, who went above and beyond to facilitate this work.

To the staff at Sysmex UK – John Bleby, Susan Beech, Jay Williams, Ryan Harlow, Ian Holding, Ben Hinchcliffe, and Isabel Milne – and to the Medical Science Division at Sysmex Europe – Jo Linssen, Marion Munster, Stefanie Klatte, Theresa Coetzer, and Talent Hwandih. Thank you for being so welcoming during my secondments and for sharing your wisdom. I am deeply grateful for your continued support.

Finally, to my family. To my Mum, Giselle, and Dad, Ian – I am who I am today because of you. Your unconditional love and support through every doubt and breakdown carried me through this journey. I truly would not be where I am without you both, and I love you always. To my beautiful sister, Poppy, and amazing brother, Freddie – you make me so proud every day, and it is an honour to grow and learn alongside you. To my wider family and friends who have my love and gratitude forever – Mert, Matt, Tracy, and everyone else who has patiently listened to me attempt to explain my project a million times! And to my wonderful boyfriend, Josh – I can never begin to thank you for your support over the last few years. I love you forever.

I dedicate this work to my beloved grandma, Sylvia Mary Putnam. Your example of love and kindness has guided me every step of the way. I carry you with me always.

Contents

Abstract i

List of Figures xii

List of Tables xvi

List of Abbreviations xix

1 Introduction and Background Review 1

1.1 Neonatal Sepsis . . . 1

1.1.1 Prevalence, Incidence, and Mortality . . . 1

1.1.2 Aetiology and Pathogenesis . . . 2

1.1.3 Diagnostic Techniques . . . 7

1.1.4 Antibiotic Treatment . . . 15

1.1.5 Summary of Neonatal Sepsis . . . 17

1.2 The Complete Blood Count . . . 18

1.2.1 The Composition of Blood . . . 18

1.2.2 Analysis Techniques . . . 23

1.2.3 Limitations of the Complete Blood Count . . . 27

1.2.4 Diagnostic Utility of the Complete Blood Count . . . 28

1.3 Rationale and Objectives . . . 32

2 The Diagnostic Accuracy of the Complete Blood Count for Neonatal Sepsis: A Systematic Review and Meta-Analysis 34

2.1 Introduction . . . 34

2.1.1 Chapter Aims . . . 36

2.2 Methods . . . 37

2.2.1 Search Strategy, Eligibility Criteria, and Study Selection 37
2.2.2 Data Extraction and Synthesis 37
2.2.3 Assessment of Risk of Bias 38
2.3 Results 39
2.3.1 Study Selection 39
2.3.2 Study Characteristics 39
2.3.3 Level of Applicability 43
2.3.4 Synthesis of Results 43
2.4 Discussion 56
2.4.1 Summary of Findings 56
2.4.2 Comparison with Existing Literature 57
2.4.3 Implications of Findings 58
2.4.4 Strengths and Limitations 59
2.5 Chapter Summary 61

3 Establishing a Baseline Profile for the Neonatal Complete Blood Count in a Cohort of Neonates Admitted to a Neonatal Intensive Care Unit 62

3.1 Introduction 62
3.1.1 The Challenges of Establishing Baseline Neonatal Profiles 62
3.1.2 Established Reference Ranges for the Neonatal CBC 64
3.1.3 Aims of the Chapter 64
3.2 Methods 66
3.2.1 Ethical Considerations 66
3.2.2 Data Collection 66
3.2.3 Data Analysis 69
3.3 Results 72
3.3.1 Patient Characteristics 72
3.3.2 Determination of Value Distributions 72

3.3.3 Effects of Corticosteroid Exposure . . . 72
3.3.4 Calculation of Reference Ranges . . . 76
3.3.5 Effects of Sex, Age, and Gestational Age at Birth . . . 76
3.4 Discussion . . . 84
3.4.1 Summary of Findings and Comparison with Existing Literature . . . 84
3.4.2 Clinical Implications of Findings . . . 85
3.4.3 Strengths and Limitations . . . 86
3.5 Chapter Summary . . . 88

4 Evaluation of the Diagnostic Accuracy of the Complete Blood Count for Late-Onset Neonatal Sepsis 89

4.1 Introduction . . . 89
4.1.1 Aims of the Chapter . . . 90
4.2 Methods . . . 91
4.2.1 Data Collection and Sample Stratification . . . 91
4.2.2 Establishing the Diagnostic Accuracy of CBC Parameters . . . 91
4.2.3 Validation of the Diagnostic Accuracy of CBC Parameters . . . 92
4.2.4 Parameter Selection for Algorithm Development . . . 92
4.3 Results . . . 93
4.3.1 Sample Characteristics . . . 93
4.3.2 Evaluation of CBC Parameters in the Training Dataset . . . 96
4.3.3 Evaluation of CBC Parameters in the Testing Dataset . . . 96
4.3.4 Parameter Selection for Algorithm Development . . . 102
4.4 Discussion . . . 103
4.4.1 Summary of Findings . . . 103
4.4.2 Comparison with Existing Literature . . . 104
4.4.3 Implications of Findings . . . 105
4.4.4 Strengths and Limitations . . . 106

4.5 Chapter Summary 108

5 Development and Validation of Diagnostic Algorithms for Late-Onset Neonatal Sepsis 109

5.1 Introduction 109

5.1.1 Pre-Existing Diagnostic Scores 109

5.1.2 Machine Learning-Based Tools 111

5.1.3 Aims of the Chapter 114

5.2 Methods 115

5.2.1 Data Collection 115

5.2.2 Diagnostic Algorithm Development 115

5.2.3 Diagnostic Algorithm Validation 118

5.3 Results 119

5.3.1 Sample Characteristics 119

5.3.2 Diagnostic Algorithm Development and Performance 119

5.3.3 Comparison of Diagnostic Performance 128

5.3.4 Longitudinal Case Studies 140

5.4 Discussion 152

5.4.1 Summary of Findings 152

5.4.2 Comparison with Existing Literature 153

5.4.3 Implications of Findings 155

5.4.4 Strengths and Limitations 156

5.5 Chapter Summary 158

6 The Application of Analytical Methods to Other Clinically Relevant Samples for the Diagnosis of Early-Onset Neonatal Sepsis 159

6.1 Introduction 159

6.1.1 CBC Results From the First Three Days of Life 159

6.1.2 CBC Results From Umbilical Cord Blood Samples 162

6.1.3 Aims of the Chapter . . . 165
6.2 Methods . . . 166
6.2.1 Data Collection . . . 166
6.2.2 Evaluation of the Baseline Profile . . . 168
6.2.3 Diagnostic Accuracy of Individual CBC Parameters . . . 169
6.2.4 Diagnostic Accuracy of Previously Developed Models . . . 170
6.3 Results . . . 171
6.3.1 Sample Characteristics . . . 171
6.3.2 Evaluation of the Baseline Profile . . . 174
6.3.3 Diagnostic Accuracy of Individual CBC Parameters . . . 185
6.3.4 Diagnostic Accuracy of Previously Developed Models . . . 191
6.4 Discussion . . . 200
6.4.1 Summary of Findings . . . 200
6.4.2 Comparison with Existing Literature . . . 201
6.4.3 Implications of Findings . . . 202
6.4.4 Strengths and Limitations . . . 203
6.5 Chapter Summary . . . 205

7 Discussion **206**

7.1 Summary of Chapters . . . 206
7.2 Clinical Implications of Findings . . . 208
7.3 Marketability of CBC-Based Diagnostic Tools . . . 209
7.4 Limitations and Suggestions for Future Work . . . 211
7.5 Concluding Remarks . . . 213

Bibliography **214**

Appendices **238**

A1 The Diagnosis Accuracy of the Complete Blood Count for Neonatal Sepsis: A Systematic Review and Meta-Analysis 238

A1.1 Search Strategies 238

A1.2 Level of Applicability 241

A1.3 Study Values 243

A2 Establishing a Baseline Profile for the Neonatal Complete Blood Count in a Cohort of Neonates Admitted to a Neonatal Intensive Care Unit 248

A2.1 Validation of the XN-450 Point-of-Care Analyser 248

A2.2 Clinical Observation Charts 253

A2.3 List of Surgical Interventions 254

A3 Evaluation of the Diagnostic Accuracy of the Complete Blood Count for Late-Onset Neonatal Sepsis 255

A3.1 Evaluation of the Training Dataset 256

A3.2 Evaluation of the Testing Dataset 259

A4 Development and Validation of Diagnostic Algorithms for Late-Onset Neonatal Sepsis 262

A4.1 Logistic Regression Model 263

A4.2 Computer-Generated Decision Tree 264

A4.3 Manually Constructed Decision Tree 265

A4.4 Random Forest Model 266

A4.5 Scoring System 268

A4.6 k-Nearest Neighbour Model 270

A4.7 Alternative Diagnostic Models 271

A5 Discussion 273

A5.1 Annual Reagent Consumption 274

List of Figures

1.1 Simplified diagram of haematopoiesis, showing the main cell lineages relevant to this study. . . . 19

1.2 Direct current sheath flow detection of red blood cells and platelets (A–C). 24

1.3 Fluorescence flow cytometry analysis of white blood cells, red blood cells, and platelets (A–C). . . . 26

2.1 PRISMA flowchart of study selection for the systematic review and meta-analysis. . . . 40

2.2 Level of applicability of the studies included in the systematic review and meta-analysis, adapted from QUADAS-2, the CASP diagnostic checklist, and the Newcastle-Ottawa scale. . . . 44

2.3 Meta-analysis of the diagnostic accuracy of white blood cell count (x 10^9/L) for neonatal sepsis (A–C). . . . 47

2.4 Meta-analysis of the diagnostic accuracy of neutrophil count (x 10^9/L) for neonatal sepsis (A–C). . . . 48

2.5 Meta-analysis of the diagnostic accuracy of neutrophil-to-lymphocyte ratio for neonatal sepsis (A–C). . . . 49

2.6 Meta-analysis of the diagnostic accuracy of immature-to-total neutrophil ratio for neonatal sepsis (A–C). . . . 50

2.7 Meta-analysis of the diagnostic accuracy of mean neutrophil volume (au) for neonatal sepsis (A–C). . . . 51

2.8 Meta-analysis of the diagnostic accuracy of platelet count (x 10^9/L) for neonatal sepsis (A–C). . . . 52

2.9 Meta-analysis of the diagnostic accuracy of mean platelet volume (fL) for neonatal sepsis (A–C). . . . 53

2.10 Meta-analysis of the diagnostic accuracy of immature-to-total neutrophil ratio for neonatal sepsis – subgroup analysis of studies reporting an optimal threshold of 0.20 (A–C). . . . 54

2.11 Meta-analysis of the diagnostic accuracy of platelet count (x 10^9/L) for neonatal sepsis – subgroup analysis of studies reporting an optimal threshold of 150 x 10^9/L (A–C). . . . 55

3.1 Criteria for the categorisation of complete blood count results for late-onset sepsis, based on clinical and laboratory signs of sepsis. 68

3.2 Flowchart illustrating the workflow for deriving reference ranges and analysing the effects of different factors on the neonatal complete blood count. 71

3.3 Analysis of white blood cell count in clinically stable neonates admitted to a Neonatal Intensive Care Unit (A–E). 74

3.4 Analysis of haemoglobin content in clinically stable neonates admitted to a Neonatal Intensive Care Unit (A–E). 75

4.1 Flowchart of complete blood count results included in the determination of the diagnostic utility of individual parameters for late-onset sepsis. . . 94

4.2 Evaluation of the diagnostic accuracy of white blood cell count for late-onset sepsis (A–D). 100

4.3 Evaluation of the diagnostic accuracy of neutrophil-to-lymphocyte ratio for late-onset sepsis (A–D). 101

5.1 Structure of a computer-generated decision tree diagnostic model for late-onset sepsis. 124

5.2 Structure of a manually calculated decision tree diagnostic model for late-onset sepsis. 125

5.3 Diagnostic performance of a logistic regression model for late-onset sepsis (A–C). 129

5.4 Diagnostic performance of a computer-generated decision tree model for late-onset sepsis (A–C). 130

5.5 Diagnostic performance of a manually constructed decision tree model for late-onset sepsis (A–C). 131

5.6 Diagnostic performance of a random forest model for late-onset sepsis (A–C). 132

5.7 Diagnostic performance of a scoring system for late-onset sepsis (A–C). . 133

5.8 Diagnostic performance of a k-nearest neighbour model for late-onset sepsis (A–C). 134

5.9 The diagnostic accuracy of six diagnostic models for late-onset sepsis. . . 135

5.10 Heatmap of DeLong Test p-values for the comparison of receiver operating characteristic curve analysis of six diagnostic models for late-onset sepsis. 136

5.11 The diagnostic performance of a logistic regression model for late-onset sepsis compared to traditional biomarkers. 141

5.12 The diagnostic performance of a computer-generated decision tree model for late-onset sepsis compared to traditional biomarkers. 142

5.13 The diagnostic performance of a manually constructed decision tree model for late-onset sepsis compared to traditional biomarkers. 143

5.14 The diagnostic performance of a random forest model for late-onset sepsis compared to traditional biomarkers. 144

5.15 The diagnostic performance of a scoring system for late-onset sepsis compared to traditional biomarkers. 145

5.16 The diagnostic performance of a k-nearest neighbour model for late-onset sepsis compared to traditional biomarkers. 146

5.17 Heatmap of DeLong Test p-values for the comparison of receiver operating characteristic curve analysis of six diagnostic models for late-onset sepsis with traditional haematological biomarkers. 147

5.18 Longitudinal case study of a neonate born at 30 weeks' gestation and admitted to the John Radcliffe Hospital Neonatal Intensive Care Unit (A–F). 149

5.19 Longitudinal case study of a neonate born at 27 weeks' gestation and admitted to the John Radcliffe Hospital Neonatal Intensive Care Unit (A–F). 150

5.20 Longitudinal case study of a neonate born at 26 weeks' gestation and admitted to the Erasmus University Medical Centre Neonatal Intensive Care Unit (A–F). 151

6.1 Criteria for the categorisation of complete blood count results for early-onset sepsis, based on clinical and laboratory signs of sepsis. 167

6.2 Analysis of the baseline white blood cell count in the first three days of life in clinically stable neonates admitted to a Neonatal Intensive Care Unit (A–E). 175

6.3 Analysis of the baseline umbilical cord white blood cell count (A–E). . . . 181

6.4 Evaluation of the diagnostic accuracy of white blood cell count for early-onset sepsis (A–D). 186

6.5 Evaluation of the diagnostic accuracy of the umbilical cord white blood cell count for admission to the Neonatal Intensive Care Unit and for suspected early-onset sepsis (A–D). 192

6.6 Diagnostic performance of a scoring system for late-onset sepsis in a cohort of neonates with early-onset sepsis (A–C). 197

6.7 Diagnostic performance of a k-nearest neighbour model for late-onset sepsis in a cohort of neonates with early-onset sepsis (A–C). 198

6.8 Diagnostic performance of a logistic regression model for early-onset sepsis (A–C). 199

A2.1 Validation of white blood cell count on a Sysmex XN-450 haematology analyser against an XN-2000 analyser in the haematology laboratory, for use as a point-of-care device in a Neonatal Intensive Care Unit (A–H). . 250

A2.2 Example of a daily observation chart from clinical paper records. 253

A4.1 Development of a random forest diagnostic model for late-onset sepsis (A–C). 266

List of Tables

1.1 Risk factors for early- and late-onset neonatal sepsis, adapted from the National Institute for Health and Care Excellence guidelines (2024). . . . 4

1.2 Clinical signs of early- and late-onset neonatal sepsis, adapted from the National Institute for Health and Care Excellence guidelines (2024). . . . 6

1.3 Diagnostic accuracy of neonatal sepsis biomarkers, as reported in meta-analyses and systematic reviews. . . . 10

1.4 Reference ranges for the complete blood count in healthy adults, adapted from van Pelt et al. (2022). . . . 29

2.1 Diagnostic accuracy of the complete blood count for neonatal sepsis, as evaluated by the NICE evidence review. . . . 35

2.2 Diagnostic accuracy of the complete blood count for neonatal sepsis, as evaluated by Celik et al. (2022). . . . 35

2.3 Characteristics of individual studies included in the systematic review and meta-analysis. . . . 41

2.4 Meta-analysis of the diagnostic accuracy of the complete blood count for neonatal sepsis. . . . 46

3.1 Published references values for parameters of the neonatal complete blood count. . . . 65

3.2 Sample characteristics of complete blood count results used for the evaluation of the baseline profile in clinically stable neonates admitted to the Neonatal Intensive Care Unit. . . . 73

3.3 P-values from multivariate analysis of variances assessing the effects of corticosteroids, sex, postnatal age, and gestational age at birth on parameters of the complete blood count in neonates admitted to a Neonatal Intensive Care Unit. . . . 78

3.4 Reference ranges for neonatal complete blood count parameters from the Neonatal Intensive Care Unit, with deviations from the average for parameters affected by corticosteroid exposure or sex. . . . 81

4.1 Sample characteristics of complete blood count results used for the evaluation of the diagnostic accuracy of complete blood count parameters for late-onset sepsis. . . . 95

4.2 Evaluation of the diagnostic accuracy of the neonatal complete blood count for late-onset sepsis using ROC curve analysis. 97

5.1 Diagnostic performance of published scoring systems for late-onset sepsis. 110

5.2 Diagnostic performance of published machine learning models for late-onset sepsis. 112

5.3 Sample characteristics of complete blood count results used for the development and validation of diagnostic models for late-onset sepsis. . . 120

5.4 Diagnostic accuracy of six diagnostic models for late-onset sepsis. 121

5.5 Structure of a logistic regression diagnostic model for late-onset sepsis. . 123

5.6 Structure of a scoring system for late-onset sepsis. 127

5.7 Review of the extreme outliers for each diagnostic model for late-onset sepsis. 137

6.1 Diagnostic accuracy of the complete blood count for early-onset sepsis from previously published studies. 161

6.2 Umbilical cord complete blood count reference ranges from previously published studies. 163

6.3 Sample characteristics for complete blood count results used for the evaluation of the baseline profile and diagnostic utility for early-onset sepsis in the first three days of life in neonates admitted to the Neonatal Intensive Care Unit. 172

6.4 Sample characteristics for umbilical cord complete blood count results used for the evaluation of the baseline profile and diagnostic utility for early-onset sepsis. 173

6.5 Evaluation of the baseline profile for complete blood count parameters in the first three days of life. 176

6.6 Evaluation of the baseline profile for umbilical cord complete blood count parameters. 182

6.7 Diagnostic potential of complete blood count parameters in the first few days of life for early-onset sepsis. 187

6.8 Diagnostic potential of umbilical cord complete blood count parameters for Neonatal Intensive Care Unit admission and suspected early-onset sepsis. 193

6.9 Diagnostic potential of diagnostic models developed for late-onset sepsis for early-onset sepsis 196

7.1 Estimated revenue for the sale of the complete blood count as a diagnostic test for neonatal sepsis 210

A1.1 Level of applicability of the studies included in the systematic review and meta-analysis, adapted from QUADAS-2, the CASP diagnostic checklist, and the Newcastle-Ottawa scale. 241

A1.2 Diagnostic accuracy values for parameters of the complete blood count for neonatal sepsis, from studies included in the systematic review and meta-analysis. 243

A2.1 Validation of a Sysmex XN-450 haematology analyser for use as a point-of-care device in a Neonatal Intensive Care Unit 251

A3.1 Evaluation of the diagnostic accuracy of the neonatal complete blood count for late-onset sepsis using receiver operating characteristic curve analysis of the training dataset. 256

A3.2 Validation of the diagnostic accuracy of the neonatal complete blood count for late-onset sepsis using receiver operating characteristic curve analysis of the testing dataset. 259

A4.1 Diagnostic accuracy of a logistic regression model for late-onset sepsis. . 263

A4.2 Diagnostic accuracy of a computer-generated decision tree model for late-onset sepsis. 264

A4.3 Diagnostic accuracy of a manually constructed decision tree model for late-onset sepsis. 265

A4.4 Diagnostic accuracy of a random forest model for late-onset sepsis. . . . 267

A4.5 Development of a diagnostic scoring system for late-onset sepsis. 268

A4.6 Diagnostic accuracy of a scoring system for late-onset sepsis. 269

A4.7 Diagnostic accuracy of a k-nearest neighbour model for late-onset sepsis. 270

A4.8 Comparison of five diagnostic models for late-onset sepsis with alternative models incorporating gestational age at birth and postnatal age at sampling across the training and testing datasets. 271

A5.1 Annual reagent consumption based on the average number of complete blood count results measured per day 274

List of Abbreviations

AIC	Akaike Information Criterion
ANOVA	Analysis of variance
AUC	Area under the curve
CASP	Critical Appraisal Skills Programme
CBC	Complete blood count
CI	Confidence interval
CLSI	Clinical and Laboratory Standards Institute
CoNS	Coagulase negative Staphylococcus
cp	Complexity parameter
CPAP	Continuous positive airway pressure
CRP	C-reactive protein
CS	Caesarean section
CSF	Cerebrospinal fluid
DC	Direct current
DNI	Delta neutrophil index
DOR	Diagnostic odds ratio
EMC	Erasmus University Medical Centre
EOS	Early-onset sepsis
EPR	Electronic Patient Records
FN	False negative
FP	False positive
FSC	Forward scattered light
GA	Gestational age at birth
GBS	Group B streptococcus
HFT	High flow therapy
HSS	Haematological Scoring System
ICIS	Intensive Care Infection Score

IL	Interleukin
IQR	Interquartile range
kNN	K-nearest neighbour
LOS	Late-onset sepsis
LR^+	Positive likelihood ratio
LR^-	Negative likelihood ratio
MNC	Mean neutrophil conductivity (au)
MNS	Mean neutrophil scatter (au)
MNV	Mean neutrophil volume (au)
MOD	Mode of delivery
MRSA	Methicillin-resistant *Staphylococcus aureus*
NEC	Necrotising enterocolitis
NICE	National Institute for Health and Care Excellence
NICU	Neonatal Intensive Care Unit
NPV	Negative predictive value
NRBC	Nucleated red blood cell
OR	Odds ratio
PCR	Polymerase chain reaction
pes	partial eta-squared
PPV	Positive predictive value
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QUADAS	Quality Assessment of Diagnostic Accuracy Studies
RBC	Red blood cell
ROC	Receiver operating characteristic
RR	Relative risk
SD	Standard deviation
SFL	Side fluorescence intensity
SLS	Sodium lauryl sulphate
SMS	Sepsis MetaScore

SSC	Side scattered light
sTREM-1	Soluble TREM-1
SVD	Spontaneous vaginal delivery
TN	True negative
TNF-α	Tumour necrosis factor alpha
TP	True positive
UK	United Kingdom
VIF	Variance Inflation Factor
WBC	White blood cell

Complete Blood Count Parameters

Red Blood Cell Parameters

RBC#	Red blood cell count (x 10^{12}/L)
HGB	Haemoglobin content (g/L)
HCT	Haematocrit (%)
MCV	Mean cell volume (fL)
MCH	Mean cell haemoglobin (pg)
MCHC	Mean cell haemoglobin concentration (g/L)
RDW-SD	Red cell distribution width (fL)
RDW-CV	Red cell distribution width (%)
NRBC#	Nucleated red blood cell count (x 10^{9}/L)
NRBC%	Nucleated red blood cell count (%)
RET#	Reticulocyte count (x 10^{12}/L)
RET%	Reticulocyte count (%)
IRF	Immature reticulocyte fraction (%)
RBC-HE	Red blood cell haemoglobin equivalent (pg)
RET-HE	Reticulocyte haemoglobin equivalent (pg)
DELTA-HE	Haemoglobin equivalent difference (pg)
HYPO-HE	Hypochromic red blood cell count (%)

HYPER-HE Hyperchromic red blood cell count (%)

MICRO-RBC Microcytic red blood cell count (%)

MACRO-RBC Macrocytic red blood cell count (%)

White Blood Cell Parameters

WBC# White blood cell count (x10^9/L)

NEUT# Neutrophil count (x 10^9/L)

NEUT% Neutrophil count (%)

LYMPH# Lymphocyte count (x 10^9/L)

LYMPH% Lymphocyte count (%)

MONO# Monocyte count (x 10^9/L)

MONO% Monocyte count (%)

EOSI# Eosinophil count (x 10^9/L)

EOSI% Eosinophil count (%)

BASO# Basophil count (x 10^9/L)

BASO% Basophil count (%)

IG# Immature granulocyte count (x 10^9/L)

IG% Immature granulocyte count (%)

NLR Neutrophil-to-lymphocyte ratio

ITR Immature-to-total neutrophil ratio

NEUT-RI Neutrophil reactive index (fluorescence intensity, FI)

NEUT-GI Neutrophil granularity index (scatter intensity, SI)

NEUT-TW Neutrophil distribution width (au)

RE-LYMPH Reactive lymphocyte count (% of total lymphocytes, %L)

AS-LYMPH Antibody-synthesising lymphocyte count (% of total lymphocytes, %L)

HFLC High fluorescing lymphocyte count (% of total lymphocytes, %L)

RE-MONO Reactive monocyte count (% of total monocytes, %M)

Platelet Parameters

PLT#	Platelet count (x 10^9/L)
PDW	Platelet distribution width (%)
MPV	Mean platelet volume (fL)
P-LCR	Platelet large cell ratio
PCT	Plateletcrit (%)
IPF	Immature platelet fraction (%)
PLR	Platelet-to-lymphocyte ratio

Chapter 1

Introduction and Background Review

1.1 Neonatal Sepsis

The term "neonatal sepsis" traditionally refers to a systemic bloodstream infection that occurs in infants less than 28 days of age [1]. Clinicians classify cases by the timing of onset: early-onset sepsis (EOS) occurs within the first 72 hours of life, and late-onset sepsis (LOS) occurs thereafter [2]. This distinction is clinically significant, as it reflects differing sources of infection, pathogen profiles, and approaches to prevention and treatment. Given the physiological immaturity and developmental delays in preterm infants, some clinical definitions extend the definition of the neonatal period, and thus the definition of LOS, up to 90 days of life [3, 4]. This thesis adopts this broader definition.

1.1.1 Prevalence, Incidence, and Mortality

Neonatal sepsis remains a significant cause of morbidity and mortality in newborns worldwide [1, 2]. Reported statistics vary considerably, primarily due to differences in diagnostic criteria and reporting practices [2]. In some regions, diagnosis relies on clinical signs, such as fever or lethargy, while others incorporate laboratory confirmation, including blood culture results [2]. The absence of a unified global diagnostic standard contributes to wide variability in reported statistics, complicating efforts to compare data across populations. Understanding when infections occur helps tailor prevention strategies, such as intrapartum antibiotic prophylaxis for EOS and improved aseptic practices for LOS.

A 2021 meta-analysis of 26 studies estimated a global incidence rate of 39.3 per 1,000 live births between 2009 and 2018 [5]. Subgroup analyses revealed a higher incidence rate for EOS (3.11 per 1,000 live births, $n = 4$) compared to LOS (0.66 per 1,000 live births, $n = 1$). However, the reported I^2 statistic of 0.99 across studies reporting on EOS indicated substantial heterogeneity, likely due to variations in diagnostic criteria.

The reported data for the incidence of culture-confirmed sepsis in the United Kingdom (UK) are markedly lower, with reported incidence rates of 8 per 1,000 live births or 71 per 1,000 neonatal admissions between 2006 and 2008 [6]. Excluding coagulase negative Staphylococcus (CoNS) infections, which are frequently regarded as contaminants, the incidence rate for EOS (defined as infection within the first 48 hours of life) was 0.9 per 1,000 live births [6]. In contrast, the reported incidence rate for LOS was 3 per 1,000 live births in the same period [6]. Overall, EOS constituted 23% of cases, while 85% of LOS cases were observed in preterm neonates [6].

Unsurprisingly, neonatal sepsis is more prevalent in low- to middle-income countries. The reported incidence in South Asia is 15.8 per 1,000 live births, and 99% of neonatal deaths attributed to sepsis occur in low-income countries, including an estimated 300,000 annual deaths in sub-Saharan Africa [7]. In contrast, there were 0.14 deaths per 1,000 live births in the United States during the same period [7]. These regional differences highlight how disparities in diagnostic criteria, healthcare infrastructure, and data quality complicate global epidemiological comparisons and may lead to under- or overestimation of the burden of neonatal sepsis.

1.1.2 Aetiology and Pathogenesis

The mechanisms of infection and causative pathogens differ considerably between EOS and LOS. As such, the UK guidelines, published by the National Institute for Health and Care Excellence (NICE), address their diagnosis and treatment separately. However, whether they should be regarded as distinct conditions remains a matter of debate due to their biological and clinical similarities [8]. Because maternal factors strongly influence EOS and such data are limited in the study cohort, this thesis concentrates mainly on LOS. Nonetheless, both classifications of neonatal sepsis are discussed here for completeness.

Mechanisms of Infection

The mechanism of infection varies depending on the timing of onset. EOS is typically acquired through vertical transmission from the mother [9]. Pathogens may be transmitted to the neonate during vaginal delivery or may ascend the cervix to infect the fetus *in utero*

[10]. In cases of intrauterine infection, transmission across the placenta can also occur [10]. Conversely, LOS is generally acquired via horizontal transmission from environmental sources, such as healthcare workers, caregivers, or contaminated medical equipment, or from opportunistic commensal microorganisms, including those found in the gut or on the skin [10]. Therefore, EOS prevention focuses on maternal screening and intervention with intrapartum antibiotics, while LOS prevention concentrates on hospital hygiene.

Risk Factors

Maternal infections and prolonged rupture of membranes constitute significant risk factors for EOS, as they increase fetal exposure to ascending pathogens (Table 1.1) [11]. To mitigate the occurrence of EOS, intrapartum antibiotics administration may be considered for women with prolonged rupture of membranes; in preterm labour; with colonisation of group B streptococcus (GBS) or bacteriuria during the current or previous pregnancy; or whose previous newborn was affected by GBS infection [11]. Additionally, earlier delivery is recommended when membrane rupture occurs alongside GBS infection to reduce the risk of an ascending infection [11].

The Kaiser-Permanente EOS risk calculator is extensively employed to estimate the probability of EOS in neonates born after 34 weeks of gestation [12]. It assesses factors such as gestational age at birth (GA), membrane rupture, and the neonate's clinical presentation at birth to generate a risk estimation. The utilisation of the calculator has been demonstrated to decrease empiric antibiotic administration and reduce the frequency of blood cultures [12]. Although some research indicates delays or missed treatments in 19 – 31% of EOS cases [13], the calculator continues to be included in the NICE guidelines owing to its effectiveness in identifying neonates with evolving EOS and in minimising unnecessary antibiotic usage among unaffected infants [11].

The most significant risk factor for LOS is admission to a Neonatal Intensive Care Unit (NICU). Neonates admitted to the NICU are exposed to invasive procedures and medical devices as part of their routine care, including the insertion of chest drains, urinary catheterisation, and intubation for ventilation. Preterm neonates, who are

Table 1.1: Risk factors for early- and late-onset neonatal sepsis, adapted from the National Institute for Health and Care Excellence guidelines (2024) [11].

<table>
<tr><th colspan="2">Early-Onset Sepsis</th><th>Late-Onset Sepsis</th></tr>
<tr><td>Maternal Factors</td><td>Maternal group B streptococcus colonisation, bacteriuria, or infection

Intrapartum fever higher than 38°C if there is suspected/confirmed bacterial infection

Clinical diagnosis of chorioamnionitis</td><td rowspan="3">Prematurity

Mechanical ventilation

History of surgery

Invasive procedures e.g. central catheter</td></tr>
<tr><td>Perinatal Factors</td><td>Rupture of membranes for more than 18 hours before a preterm birth

Rupture of membranes at term for more than 24 hours before the onset of labour

Preterm birth following spontaneous labour</td></tr>
<tr><td>Other Factors</td><td>Suspected/confirmed infection in another baby (in the case of multiple pregnancy)

Invasive group B streptococcus infection in a previous baby</td></tr>
</table>

regularly admitted to the NICU following birth and tend to remain hospitalised for extended durations, possess immature immune systems that diminish their capacity to mount an effective immune response (Table 1.1) [11]. Approximately one in seven neonates is admitted to NICUs in the UK each year, often due to prematurity, suspected infection, congenital anomalies, or traumatic birth [14]. This constitutes over 90,000 neonates at risk of developing LOS each year.

Causative Pathogens

The most common causative pathogen for EOS is GBS, which is a commensal bacterium found in the female genitourinary and gastrointestinal tracts [7]. Around 50% of EOS cases are due to GBS, with other common isolates including *Escherichia coli*, *Staphylococcus aureus*, *Haemophilus influenzae*, and *Listeria monocytogenes* [6, 9, 15].

By contrast, the causative pathogens in most cases of LOS are CoNS (e.g. *S. epidermidis*), *E. coli*, and *S. aureus*, although *Enterococcus*, *Enterobacter*, and *Klebsiella* species are also frequently isolated [6, 15]. While CoNS is the most frequently isolated organism in LOS, distinguishing between actual infection and contamination remains challenging because it is a normal skin commensal [7]. Some guidelines suggest confirming diagnosis only after two positive CoNS cultures or one positive culture combined with abnormal C-reactive protein (CRP) levels [16]. However, no universal consensus exists.

Neonatal sepsis may also be caused by non-bacterial pathogens, such as the herpes simplex virus, *Candida*, and other fungal species, accounting for less than 0.1% of cases [9, 15]. While treatments for viral and fungal infections differ, their presentation is often very similar to a bacterial infection.

Disease Progression and Outcomes

Early clinical signs of neonatal sepsis include temperature instability, poor feeding, or respiratory distress (Table 1.2) [11]. A recent study assessed the prevalence of clinical signs in neonatal sepsis, reporting cardiopulmonary dysfunction (tachypnoea, bradycardia, or apnoea) in 68% of cases and fever in 6.5% [17]. However, many of these symptoms

Table 1.2: Clinical signs of early- and late-onset neonatal sepsis, adapted from the National Institute for Health and Care Excellence guidelines (2024) [11].

Early-Onset Sepsis		Late-Onset Sepsis	
Red flags	Apnoea Seizures Cardiopulmonary resuscitation Mechanical ventilation Signs of shock	Behaviour	Change in behaviour Appears ill Does not wake or stay awake Weak high-pitched or continuous cry
Other	Altered behaviour or responsiveness Altered muscle tone (e.g. floppiness) Feeding difficulties Feeding intolerance (e.g. vomiting or abdominal distension) Abnormal heart rate (bradycardia or tachycardia) Sign of respiratory distress (e.g. grunting or tachypnoea) Hypoxia (e.g. central cyanosis) Persistent pulmonary hypertension Jaundice within 24 hours of birth Neonatal encephalopathy Temperature abnormality Unexplained excessive bleeding, thrombocytopenia, or abnormal coagulation Altered glucose homeostasis Metabolic acidosis	Respiratory	Raised respiratory rate (> 60 bpm) Grunting or apnoea Oxygen saturation less than 90% or increased requirement above baseline
		Circulation	Persistent tachycardia (> 160 bpm) Persistent bradycardia (< 100 bpm)
		Skin	Mottle or ashen appearance Cyanosis of skin, lips, or tongue Non-blanching skin rash
		Other	Temperature $> 38°C$ or $< 36°C$ Alterations in feeding patterns Abdominal distension Seizures Bulging fontanelle

are non-specific, appearing in cases of infection, trauma, or metabolic disease, or manifest late or not at all during disease progression, often leading to the sudden onset of systemic inflammatory response syndrome [18, 19]. The non-specific nature of early signs may be due to the immaturity of the neonatal immune system, which produces a muted or delayed inflammatory response [20].

Neonates with neonatal sepsis face a higher relative risk (RR) of developing respiratory distress (RR: 1.34), intraventricular haemorrhage (1.59), meningitis (2.08), and necrotising enterocolitis (NEC) (1.62) [21]. Without prompt and targeted antibiotic intervention, neonatal sepsis may result in multiple organ failure and, ultimately, death [21]. In the long term, neonates with LOS demonstrate a higher RR (1.14) of experiencing neurodevelopmental impairments later in life compared to those without, irrespective of culture positivity [22]. A meta-analysis examining the long-term impacts of neonatal sepsis identified higher RR for cerebral palsy (1.98), cognitive delay (1.55), and auditory impairment (3.19) [21]. These adverse outcomes have become increasingly apparent as survival rates have improved due to advances in neonatal care.

1.1.3 Diagnostic Techniques

Neonatal sepsis remains a clinical emergency that requires prompt diagnosis and targeted intervention. The clinical signs of sepsis described above are predominantly non-specific, frequently occurring in neonates without obvious infection. Additionally, neonates may not display signs of deterioration until a bacterial infection has become well-established. Reliance solely on clinical signs can delay critical treatment for neonates with subtle or late-onset signs. It may also contribute to the overuse of antibiotics in neonates presenting with non-specific signs.

Instead, laboratory techniques are used to diagnose and monitor at-risk neonates. Current guidelines recommend obtaining a blood culture before initiating antibiotics, along with a baseline CRP measurement to assess treatment response [11]. In cases of strong clinical suspicion of sepsis or meningitis, a lumbar puncture may also be necessary to obtain a cerebrospinal fluid (CSF) sample for analysis [11].

Microbiology Cultures

Blood cultures remain the definitive diagnostic standard for neonatal sepsis [11]. With a reported sensitivity (the ability to correctly identify true positives) ranging from 97% to 100% after 48 hours, blood cultures serve as a highly effective rule-in tool for neonatal sepsis [7]. Nonetheless, several factors limit their clinical utility.

Microbiology results are greatly influenced by sample volume, representing a notable limitation in neonatal care [2]. The minimum recommended sample volume for a blood culture is 1 mL, although positive results may be obtained with as little as 0.5 mL [19]. Obtaining adequate sample volumes, particularly in extremely preterm or small-for-gestational-age neonates or those needing repeated sampling, may induce iatrogenic anaemia and increase the need for blood transfusions, each of which entails additional risks [2]. Neonates are also at an increased risk of intermittent bacteraemia, observed in up to 70% of septic neonates, thereby further diminishing blood culture sensitivity [2, 19]. Owing to the combined impact of limited sample volumes and intermittent bacteraemia, false-negative results may occur in as many as 60% of cases [16, 19].

The clinical utility of the blood culture is further constrained by its turnaround time, which can extend up to 48 hours [15]. This delay in pathogen identification often necessitates empirical treatment with broad-spectrum antibiotics, thereby contributing to the rising antibiotic resistance among common neonatal pathogens. Delays in administering targeted therapy may elevate the risk of septic shock and additional harm to the neonate, particularly in cases where the causative organism exhibits resistance to the commonly employed broad-spectrum first-line antibiotics [2]. Conversely, non-infected neonates with non-specific clinical signs frequently receive antibiotics unnecessarily while waiting for microbiology culture results [2].

Laboratory Tests

In addition to microbiological cultures, blood samples may be analysed using additional laboratory techniques, such as flow cytometry, aimed at identifying alternative infection-related biomarkers. The diagnostic precision of the most extensively researched biomarkers,

including the pooled sensitivity, specificity, positive likelihood ratio (LR^+), negative likelihood ratio (LR^-), and diagnostic odds ratio (DOR), is summarised in Table 1.3.

An optimal diagnostic biomarker for neonatal sepsis would be rapid, accurate, cost-effective, and minimally invasive, providing prompt responses to infection [18]. Proposed thresholds for acceptable diagnostic performance include a sensitivity approaching 100%, greater than 85% specificity (the ability to identify true negatives correctly), a positive predictive value (PPV; the likelihood that a positive test result indicates the actual presence of disease) approaching 100%, and greater than 85% negative predictive value (NPV; the likelihood that a negative test result indicates the absence of disease) [18, 19]. While these thresholds represent ideal diagnostic parameters, no currently available test consistently fulfils all these criteria. Nonetheless, these benchmarks serve as a framework for evaluating the comparative strengths of emerging and established markers.

CRP remains the most widely used biomarker for diagnosing and monitoring neonatal sepsis, and is the only biomarker explicitly recommended in the NICE guidelines [11]. It is an acute-phase reactant protein whose production is stimulated by proinflammatory cytokines, such as interleukin (IL)-6 and tumour necrosis factor alpha (TNF-α) [19]. CRP levels begin to rise 10 to 12 hours after the onset of infection and peak within 48 hours [2, 19], which limits its usefulness for early diagnosis. CRP also lacks specificity, as its levels increase in non-infectious conditions, such as meconium aspiration, surgery, and pneumothorax [2, 18, 19]. Overall, it may be most suitable for ruling out neonatal sepsis or for monitoring recovery [19].

Procalcitonin is another acute-phase protein synthesised by the liver and monocytes in response to bacterial infection [2]. In healthy individuals, serum concentrations rarely exceed 0.01 µg/L [19]. Its levels increase within two to four hours following the onset of infection and reach their peak at six to eight hours, thus serving as a valuable biomarker for early diagnosis, particularly of LOS [2]. Its production is suppressed during viral infections, which facilitates differentiation between bacterial and viral pathogens [2, 19].

Table 1.3: Diagnostic accuracy of neonatal sepsis biomarkers, as reported in meta-analyses and systematic reviews.

Study	No of Studies	Onset	AUC	Optimal Threshold	Sensitivity (%)	Specificity (%)	LR^+	LR^-	DOR	I^2
C-reactive protein (mg/L)										
Xu et al., 2025 [23]	4	NS	0.945	0.82 – 5.00	85.0	84.0	6.00	0.23	33.34	0.80 <
Liu et al., 2019 [24]	10	NS	0.900	5.82 – 10.00	70.0	89.0	5.63	0.36	17.99	0.78 <
Ruan et al., 2018 [25]	22	NS	0.850	0.50 – 2.00	71.0	88.0	6.10	0.33	19.00	0.98
Xu et al., 2016 [26]	31	NS	0.845	2.50 – 21.00	69.0	77.0	3.83	0.38	12.65	0.68 <
Yuan et al., 2013 [27]	6	NS	0.910	–	67.0	92.0	–	–	77.16	–
Eschborn and Weitkamp, 2019[a] [28]	15	NS	–	–	70.0	94.0	–	–	–	–
	12	EOS	–	–	68.5	85.5	–	–	–	–
	12	LOS	–	–	85.0	85.7	–	–	–	–
Hedegaard, Wisborg, and Hvas, 2015[a] [29]	10	NS	–	4.00 – 15.00	30.0 – 80.0	83.0 – 100.0	–	–	–	–
Procalcitonin (ng/mL)										
Ruan et al., 2018 [25]	22	NS	0.910	0.50 – 2.00	85.0	84.0	5.40	0.18	31.00	0.96
Yuan et al., 2013 [27]	2	NS	–	–	89.0	87.0	–	–	60.17	–
Yu et al., 2010 [30]	2	EOS	–	0.50	76.0	75.0	–	–	10.00	< 0.01
	8	LOS	–	–	62.0	92.0	5.10	0.53	49.30	< 0.01

Continued on next page

Continued from last page

Study	No of Studies	Onset	AUC	Optimal Threshold	Sensitivity (%)	Specificity (%)	LR^+	LR^-	DOR	I^2
Vouloumanou et al., 2011 [31]	16	NS	0.870	0.50 – 5.75	81.0	79.0	3.90	0.24	16.00	0.96
	6	EOS	0.780	–	76.0	76.0	3.20	0.32	10.00	0.89
	5	LOS	0.950	–	90.0	88.0	7.70	0.11	67.00	0.93
Eschborn and Weitkamp, 2019[a] [28]	15	NS	–	–	77.0	83.9	–	–	–	–
	12	EOS	–	–	80.0	87.3	–	–	–	–
	12	LOS	–	–	91.7	80.0	–	–	–	–
Hedegaard, Wisborg, and Hvas, 2015[a] [29]	2	EOS	–	–	72.0 – 79.0	72.0 – 90.0	–	–	–	–
Interleukin-6 (pg/mL)										
Xing et al., 2023 [32]	6	EOS	–	2.4 – 40	81.0	77.0	3.17	0.27	13.90	0.86 <
Shahkar et al., 2011 [33]	13	NS	0.890	10 – 160	79.0	84.0	4.55	0.26	20.74	–
Eichberger and Resch, 2022[a] [34]	31	EOS	–	7.6 – 300	76.0	79.0	–	–	–	–
Boskabadi and Zakerihamidi, 2018[a] [35]	11	NS	–	10 – 100	71.0 – 100.0	47.0 – 95.0	–	–	–	–
Hedegaard, Wisborg, and Hvas, 2015[a] [29]	8	NS	–	–	61.0 – 89.0	65.0 – 96.0	–	–	–	–

Continued on next page

Continued from last page

Study	No of Studies	Onset	AUC	Optimal Threshold	Sensitivity (%)	Specificity (%)	LR^+	LR^-	DOR	I^2
Interleukin-8 (pg/mL)										
Xing et al., 2023 [32]	2	EOS	–	50 – 300	91.0	94.0	12.87	0.10	143.74	< 0.46
Xing et al., 2023 [32]	1	LOS	–	60	100.0	5.0	1.06	0.13	8.29	–
Zhou et al., 2015 [36]	8	NS	0.891	50 – 300	78.0	84.0	4.58	0.25	21.64	$0.83 <$
Boskabadi and Zakerihamidi, 2018[a] [35]	4	NS	–	54 – 900	50.0 – 95.0	10.0 – 88.0	–	–	–	–
Serum amyloid A (mg/L)										
Yuan et al., 2013 [27]	9	NS	0.960	8 – 68	84.0	89.0			91.84	$0.86 <$
Hedegaard, Wisborg, and Hvas, 2015[a] [29]	2	NS	–	–	> 93.0	> 93.0	–	–	–	–
Tumor necrosis factor alpha (pg/mL)										
Hedegaard, Wisborg, and Hvas, 2015[a] [29]	3	NS	–	–	54.0 – 82.0	60.0 – 92.0	–	–	–	–
Presepsin (pg/mL)										
Ruan et al., 2018 [25]	6	NS	0.990	539 – 722	94.0	98.0	50.80	0.06	864.00	0.80
Bellos et al., 2017 [37]	5	NS	0.975	650 – 850	91.0	91.0	10.67	0.09	170.28	–
sTREM-1 (pg/mL)										
Bellos et al., 2018 [38]	8	NS	–	78 – 1700	95.0	87.0	7.20	0.05	132.49	

Continued on next page

Continued from last page

Study	No of Studies	Onset	AUC	Optimal Threshold	Sensitivity (%)	Specificity (%)	LR^+	LR^-	DOR	I^2
Neutrophil CD64										
Shi Tang, and Chen, 2016 [39]	21	NS	0.866	–	77.0	74.0	3.58	0.29	15.18	0.50 <
	7	EOS	0.842	–	75.0	74.0	3.57	0.31	12.40	–
	6	LOS	0.926	–	79.0	71.0	3.42	0.24	18.86	–
Hedegaard, Wisborg, and Hvas, 2015[a] [29]	3	LOS	–	–	75.0 – 95.0	77.0 – 88.0	–	–	–	–
Transcriptomics										
Sweeney et al., 2018[a] [40]	3	NS	0.930	–	90.0	84.7	–	0.12	–	–

[a] *Systematic review. AUC values >0.8 and sensitivities/specificities >85% are highlighted in green, suggesting strong diagnostic performance. Abbreviations: AUC – area under the curve; DOR – diagnostic odds ratio; EOS – early-onset sepsis; LOS – late-onset sepsis; LR^+ – positive likelihood ratio; LR^- – negative likelihood ratio; NS – neonatal sepsis.*

However, physiological elevations occur shortly after birth, peaking at 24 hours and returning to baseline by 72 hours [2], thereby limiting the diagnostic utility of procalcitonin for EOS. Similar to CRP, elevated procalcitonin levels may also be observed in non-infectious conditions such as respiratory distress and perinatal asphyxia [19]. Furthermore, procalcitonin is significantly more costly, averaging £8 per test compared to roughly £1.50 for CRP (unpublished data from the John Radcliffe hospital, Oxford, UK).

Precursors to CRP and procalcitonin, particularly IL-6 and IL-8, are also frequently assessed. These cytokines are secreted by monocytes and macrophages in response to infection, with serum levels rising within two hours and reaching their peak around six hours [2, 19]. Notably, elevated levels may precede clinical symptoms by several days [19]. However, both IL-6 and IL-8 have short half-lives and generally return to baseline within 24 hours of antibiotic initiation [18]. Similar to other biomarkers, their levels can be influenced by non-infectious stressors, including hypoxia, prematurity, and mechanical ventilation [19]. Consequently, they are most effectively used in conjunction with longer-acting markers rather than as standalone tests.

Among the currently available and routinely used biomarkers, CRP and procalcitonin are the most comprehensively studied, although their diagnostic performance varies (Table 1.3). Meta-analyses report areas under the curve (AUCs) ranging from 0.845 to 0.945 for CRP [23–27] and up to 0.950 for procalcitonin [25, 31], with values exceeding 0.8 indicating strong diagnostic utility. IL-6 and IL-8 demonstrate more consistent performance; however, their short half-lives and vulnerability to confounding factors limit their utility as standalone biomarkers.

Emerging biomarkers, such as presepsin, a soluble CD14 subtype indicative of monocyte activation in response to infection, may outperform traditional markers, with sensitivities and specificities nearing 95% and AUCs approaching 0.99 [25, 38]. Unlike CRP and procalcitonin, presepsin is unaffected by non-infectious inflammation, thereby offering greater specificity [19]. Similarly, soluble TREM-1 (sTREM-1), released by activated neutrophils and monocytes, has demonstrated high DORs and likelihood ratios in recent

research [38]. Despite their prospective utility, these biomarkers face limitations related to cost, assay standardisation, and the lack of validated neonatal reference ranges. Compared to more readily accessible tests such as CRP or procalcitonin, assays for presepsin or sTREM-1 remain substantially more expensive and necessitate specialised equipment.

Future diagnostic approaches for neonatal sepsis increasingly explore transcriptomic technologies, such as microarrays and next-generation sequencing [41]. Unlike traditional biomarkers, these methods analyse RNA expression patterns to capture the host immune response at a systems level, offering the potential for earlier and more precise diagnosis [41]. A recent review highlighted several studies applying transcriptomic profiling in neonatal sepsis, identifying reproducible gene signatures that differentiate septic from non-septic neonates [40, 42–46]. One such study validated the Sepsis MetaScore (SMS), which incorporates 11 genes linked to neutrophil immaturity and downregulation of the adaptive immune response [40]. The SMS achieved an AUC greater than 0.9 across three independent cohorts, suggesting excellent diagnostic potential.

However, implementation in clinical practice remains challenging [41]. Transcriptomic assays are technically complex, costly, and require specialised infrastructure, while standardisation and prospective validation in large neonatal populations are still lacking. As such, although transcriptomics represents a promising avenue, it is not yet suitable for routine clinical use.

The complete blood count (CBC) has also been examined as a diagnostic tool for neonatal sepsis [2]. However, it is not currently recommended by NICE guidelines [11]. The evidence supporting its clinical utility will be discussed in detail in Chapter 2.

1.1.4 Antibiotic Treatment

Prompt initiation of antibiotic treatment remains essential for managing neonatal sepsis, although it must be carefully balanced against the risks of overtreatment. According to NICE guidelines, benzylpenicillin and gentamicin are the first-line antibiotics for EOS, while flucloxacillin and gentamicin are recommended for LOS [11]. When Gram-negative

organisms are isolated, targeted treatment with cefotaxime may be necessary. For cases of suspected NEC, anaerobe-targeting antibiotics, such as metronidazole, are also recommended.

The NICE guidelines advise re-evaluating neonates after 36 to 48 hours, and discontinuing antibiotics if the blood culture result is negative, and both the clinical condition and CRP levels are reassuring [11]. In cases where blood cultures yield positive results or strong clinical suspicion persists, antibiotic therapy should be extended for at least seven days [11]. Shorter courses may be considered appropriate if the neonate demonstrates rapid clinical improvement or if a commensal organism, such as CoNS, is identified [11].

A combination of distrust in negative microbiology results and the absence of a gold standard test and standardised definition for neonatal sepsis has contributed to the overuse of antibiotics in neonatal care [47]. However, both intrapartum and postnatal antibiotic exposure have been shown to significantly disrupt the neonatal gut microbiome [48], increasing the risk of adverse outcomes such as NEC and fungal infections [2, 49].

Prolonged exposure to antibiotics has also been linked to chronic disease later in life. A meta-analysis examining the association between antibiotic use and childhood asthma determined that longer treatment durations were correlated with higher DORs: 1.29 for one to two days versus 1.79 for five or more days [50]. Earlier antibiotic exposure was also associated with higher risks, with a DOR of 1.82 when antibiotics were administered during the first week of life, compared to 1.43 at six months of age [50]. Similarly, Canova et al. (2020) reported an adjusted DOR of 1.38 for inflammatory bowel disease among children who received antibiotics within their first six months [51].

This overuse not only contributes to adverse health outcomes but also accelerates the emergence of antibiotic-resistant pathogens [19, 52]. Surveillance data from the NeonIN network illustrate the scale of this problem in UK NICUs: 16% of LOS pathogens were resistant to flucloxacillin and gentamicin, while 21% were resistant to amoxicillin and cefotaxime [6]. CoNS demonstrated the highest levels of resistance, with 76% (100/136

cases) of isolates resistant to flucloxacillin and gentamicin, and 100% (14/14) resistant to amoxicillin and cefotaxime. Additionally, eight instances of methicillin-resistant *Staphylococcus aureus* (MRSA), accounting for 2% of isolates, were identified. These resistance patterns pose significant challenges for empirical treatment in NICUs and emphasise the critical importance of antimicrobial stewardship in these settings.

1.1.5 Summary of Neonatal Sepsis

Neonatal sepsis remains a significant global health concern, presenting considerable risks during both the acute phase and long-term periods. While NICE guidelines recommend blood cultures as the definitive gold standard for diagnosis, their clinical usefulness is constrained by slow turnaround times and diminished sensitivity in cases of intermittent or low-level bacteraemia. Alternative biomarkers, such as CRP and procalcitonin, are frequently used to aid clinical decision-making; however, their limitations include delayed responses to infection and low specificity.

This diagnostic uncertainty has resulted in the widespread use of empirical antibiotics, which presents direct risks to neonates, including disruption of the gut microbiome and an elevated risk of chronic diseases, as well as contributing to the increasing issue of antimicrobial resistance. Furthermore, delays in the initiation of targeted therapy may worsen clinical outcomes, increase mortality rates, and cause long-term neurodevelopmental impairment. Collectively, these challenges highlight a critical unmet need for rapid, accurate, and minimally invasive diagnostic tools to facilitate the early detection and appropriate management of neonates at risk of sepsis.

Given these diagnostic challenges, there is growing interest in alternative tools that can provide more rapid, accessible, and dependable indicators of neonatal sepsis. One such tool is the CBC – a routine, affordable test that provides comprehensive information regarding the cellular components of blood. Although NICE guidelines do not currently recommend the CBC for the diagnosis of neonatal sepsis, it remains widely utilised in clinical practice and may have unrecognised diagnostic value.

1.2 The Complete Blood Count

The CBC is among the most frequently ordered investigations in neonatal care, owing to its accessibility, cost-effectiveness, and capacity to quantify multiple haematological parameters from a minimal blood volume. Although it is not definitive for diagnosing sepsis, individual components may indicate early haematological changes associated with infection. To assess the potential utility of the CBC in diagnosing neonatal sepsis, it is necessary first to understand the composition and function of blood cells.

1.2.1 The Composition of Blood

Neonates typically have between 85 and 105 mL of blood per kg of body weight, amounting to approximately 280 mL in a term infant [53]. Blood comprises plasma (55%), which contains soluble gases and dissolved salts, as well as cellular components (45%), including red blood cells (RBCs), white blood cells (WBCs), and platelets [54].

Haematopoeisis

All haematological cell types originate from pluripotent haematopoietic stem cells within the bone marrow through the process of haematopoiesis [54]. During fetal development, haematopoiesis predominantly occurs in the liver and vertebral column, subsequently transitioning to the bone marrow from five months' gestation [54, 55]. The initial stage of haematopoiesis (Figure 1.1) involves the division of a single haematopoietic stem cell into two daughter cells – one remains pluripotent, thereby preserving the stem cell pool, while the other differentiates into a lineage-restricted precursor cell [54]. Myeloid precursor cells give rise to RBCs, platelets, monocytes, and granulocytes, whereas lymphoid precursor cells differentiate into lymphocytes [54].

Erythrocytes

Myeloid precursor cells committed to the erythroid lineage develop into mature RBCs through the process of erythropoiesis (Figure 1.1). The earliest identifiable erythroid precursors capable of haemoglobin synthesis are nucleated red blood cells (NRBCs), which are large cells characterised by a nucleus and basophilic cytoplasm [54, 57]. NRBCs are

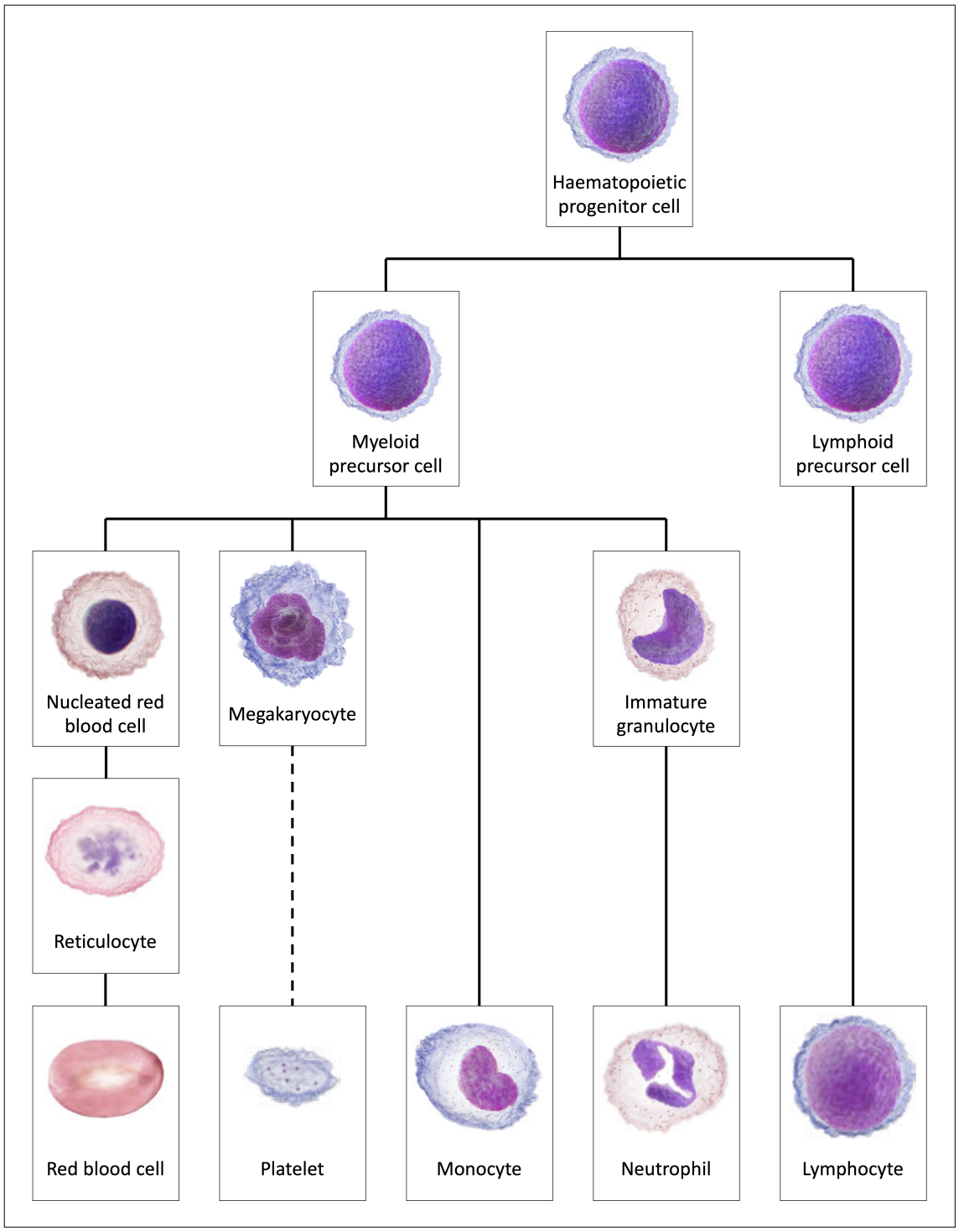


Figure 1.1: Simplified diagram of haematopoiesis, showing the main cell lineages relevant to this study. Cell images reproduced with permission from Sysmex Europe [56].

typically absent from the circulation of healthy adults, and their presence may indicate severe physiological stress or dysregulation [58]. In neonates, NRBCs are commonly observed following birth during the transition from fetal to adult haemoglobin. However, their levels generally return to baseline within the first week of life [58].

A final mitotic division within the bone marrow produces two daughter cells – a nucleated pyrenocyte, which is phagocytosed by macrophages, and an anucleate reticulocyte, which subsequently enters the bloodstream [54, 57]. Over the next 24 to 48 hours, the reticulocyte expels its remaining internal organelles and undergoes cytoskeleton reorganisation to develop the characteristic biconcave shape of a mature RBC [54, 57]. RBCs primarily transport oxygen from the lungs to peripheral tissues and constitute approximately 40% of the circulating blood volume [54]. The lifespan of mature RBCs in circulation is typically around 120 days [55].

Leucocytes

WBCs constitute about 1% of the circulating blood volume and are mainly responsible for immune defence. They are broadly classified into granulocytes, which differentiate from myeloid precursor cells and comprise the innate immune system, and lymphocytes, which originate from lymphoid precursor cells and facilitate the adaptive immune response [20]. The innate immune response is characterised by rapid and non-specific action, while the adaptive response is slower and more targeted, ultimately providing long-term immunity. Although these responses are distinct, they are functionally interconnected such that the activation of the adaptive response depends on signals from the innate system [20].

Although immune cells begin developing in utero, neonates are born with immature immune systems which undergo rapid maturation within the first three months [59, 60]. While the adaptive immune system remains functionally limited until approximately 18 months of age, thereby causing an apparent dependence on the innate immune system, recent research indicates that neonatal adaptive immune cells may demonstrate innate-like functions throughout this period [59, 60].

Monocytes

Monocytes are the first responders of the innate immune response [55]. They originate from myeloid precursor cells via monocytopoiesis and constitute approximately 10% of circulating WBCs (Figure 1.1) [61, 62]. After a period of 8 to 72 hours in circulation, monocytes migrate into tissues, where they differentiate into macrophages that can survive for up to three months [61]. Tissue-specific macrophage populations are present in the lungs, liver, spleen and central nervous system, where they carry out site-specific functions [61].

Upon pathogen recognition, monocytes and macrophages release cytokines and chemokines to attract other immune cells to the site of infection [55]. They remove pathogens via phagocytosis and can differentiate into dendritic cells, which present antigens to lymphocytes to trigger the adaptive immune response [54].

Neutrophils

Neutrophils represent the most prevalent WBC subtype, constituting up to 70% of circulating leukocytes [62]. They originate from myeloid precursor cells undergoing granulocytopoiesis, initially characterised by the production of primary (azurophilic) granules (Figure 1.1) [61]. Subsequent maturation involves nuclear segmentation and the synthesis of secondary (specific) and tertiary (gelatinase) granules [54]. This differentiation process takes seven to ten days, with mature neutrophils circulating for up to 24 hours before migrating into tissues, where they persist for up to three days [61].

In response to infection, neutrophils migrate to sites of inflammation via chemotaxis, guided by cytokine and chemokine gradients released by monocytes [54]. At these sites, they engulf pathogens by phagocytosis and degranulate to release neutrophil extracellular traps, which immobilise and eliminate pathogens [54].

Eosinophils and Basophils

Eosinophils and basophils are less common granulocytes that differentiate from immature granulocytes after forming cell-specific secondary granules [61]. They mainly reside within

tissues, resulting in low circulating numbers [62]. It is believed that they function in defending against parasites or in allergic responses. Although their exact role is not yet fully understood, it is thought that they do not function in the immune response to bacterial infection [55].

Lymphocytes

Lymphocytes mediate the adaptive immune response. B cells and T cells, the two main subtypes, originate from lymphoid precursor cells in the bone marrow (Figure 1.1) [54]. Following differentiation, naïve B cells migrate to lymph nodes and secondary lymphoid tissues, while naïve T cells undergo further maturation in the thymus [54]. Natural killer cells are a less prevalent subtype that contributes to early antiviral defence [54].

Upon encountering antigen-presenting cells, such as dendritic cells, naïve T cells proliferate and differentiate into effector subtypes [54]. Helper T cells secrete cytokines that recruit monocytes, macrophages, and neutrophils to the site of infection, and also induce B cell activation in the lymph tissues [54]. Activated B cells differentiate into plasma cells, which proliferate rapidly before migrating to infection sites and secreting immunoglobulins to neutralise pathogens [54]. Once an infection has been cleared, a subset of B cells develops into memory cells, enabling a more rapid response upon re-exposure [54].

Platelets

Myeloid precursor cells committed to the thrombotic lineage undergo thrombocytopoiesis, resulting in the formation of large megakaryocytes (Figure 1.1) [63]. These polyploid cells develop from repeated cycles of DNA replication without cell division [54]. Megakaryocytes are primarily located within the bone marrow, with smaller populations found in the spleen, lungs, and circulation [63]. They extend proplatelet processes into blood vessels that fragment to produce between 1,000 and 2,000 platelets per megakaryocyte [63].

Under normal physiological conditions, platelets circulate for up to ten days without adhering to the vessel walls [55]. In the event of endothelial damage, platelets adhere to exposed extracellular matrix components and undergo shape changes to facilitate

aggregation, thereby forming a haemostatic platelet plug [54]. Activated platelets also secrete chemokines, growth factors, and immune mediators to mediate inflammation, tissue repair, and immune responses [54].

1.2.2 Analysis Techniques

Modern haematology analysers, such as the Sysmex XN-series, use three main analytical techniques to measure the CBC: direct current (DC) detection, sodium lauryl sulphate (SLS) colourimetry, and fluorescence flow cytometry. These techniques not only support routine haematological assessment by providing absolute and relative counts of the main haematological cell types but also provide a range of secondary parameters that may serve as early indicators of infection, inflammation, or bone marrow response.

This section focuses on outputs from Sysmex analysers, as these were used throughout this thesis. Although many additional parameters may be generated, the following subsections focus on the parameters that were investigated in this thesis.

Direct Current Detection

DC sheath flow detection is used to quantify RBCs and platelets [64]. Diluted blood is directed through an aperture, passing one cell at a time while an electric current is applied. The change in electric resistance correlates to the volume of each cell (Figure 1.2A). Based on cell size, particles are classified as either RBCs (25 – 250 fL) or platelets (2 – 30 fL), which allows for the calculation of the red blood cell count (RBC#), platelet count (PLT#), and cell-to-plasma ratios: haematocrit (HCT) and plateletcrit (PCT).

Analysis of the distribution curves provides further insights into the RBC and platelet populations. Mean cell volume (MCV) is derived from the RBC histogram (Figure 1.2B), alongside red cell distribution width based on standard deviation (RDW-SD) and coefficient of variation (RDW-CV). Proportions of macrocytic (MACRO-RBC) and microcytic (MICRO-RBC) RBCs are also estimated. Similarly, the platelet histogram (Figure 1.2C) yields mean platelet volume (MPV), platelet distribution width (PDW), and platelet large cell ratio (P-LCR), which indicates the ratio of platelets with a volume exceeding 12 fL.

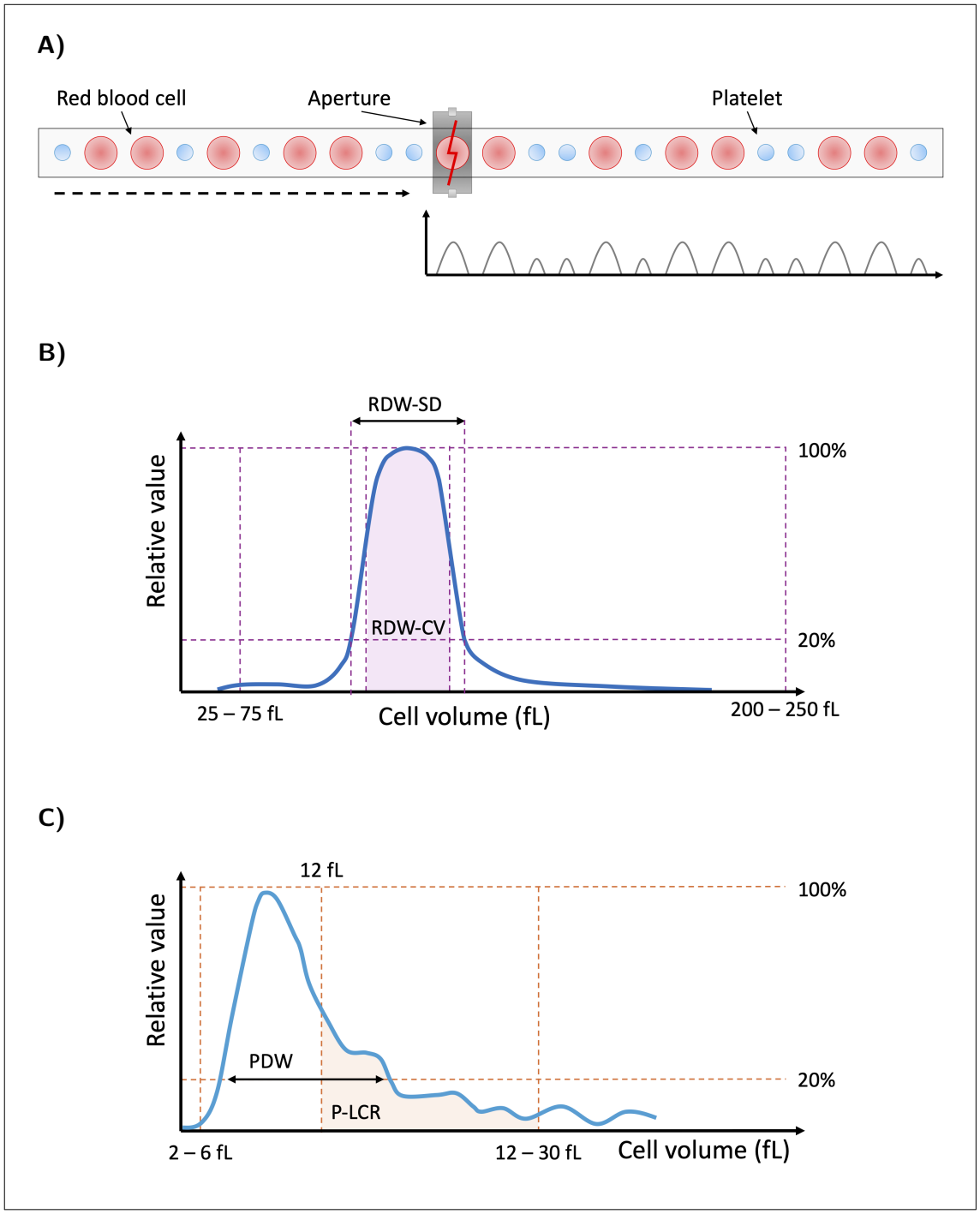


Figure 1.2: Direct current sheath flow detection of red blood cells and platelets. **A)** Simplified diagram of direct current detection, showing red blood cells and platelets passing through an aperture, with representative current pulses generated by each cell type. **B)** Distribution curve of red blood cells. **C)** Distribution curve of platelets.

Abbreviations: PDW – platelet distribution width; P-LCR – platelet large cell ratio; RDW-CV; RDW-SD – red cell distribution width.

SLS Haemoglobin Method

Haemoglobin content (HGB) is assessed using SLS, which lyses RBCs and forms a stable coloured complex with haemoglobin [65]. This complex is measured through photometry at a wavelength of 555 nm. Additional indices are subsequently derived from HGB, RBC#, and HCT values: mean cell haemoglobin (MCH) indicates the average haemoglobin content per RBC, calculated by dividing HGB by RBC#, and mean cell haemoglobin concentration (MCHC) reflects the average concentration of haemoglobin in RBCs, determined by dividing HGB by HCT.

Fluorescence Flow Cytometry

Fluorescence flow cytometry is the primary technique used for categorising leukocytes and assessing cell function (Figure 1.3A) [66]. Fluorescent dyes penetrate cells and bind to nucleic acids and intracellular granules. As stained cells pass through a laser, three signals are detected: forward scattered light (FSC), which correlates with cell size; side scattered light (SSC), which indicates internal complexity and granularity; and side fluorescence intensity (SFL), which is proportional to nucleic acid content.

WBC subpopulations are classified based on SSC and SFL parameters (Figure 1.3B). Analysers report total white blood cell count (WBC#), along with both absolute and relative counts of neutrophils (NEUT# and NEUT%), lymphocytes (LYMPH# and LYMPH%), monocytes (MONO# and MONO%), basophils (BASO# and BASO%), eosinophils (EOSI# and EOSI%), and immature granulocytes (IG# and IG%).

Functional changes in WBCs are assessed using several parameters (Figure 1.3B). Neutrophil reactive index (NEUT-RI) measures nucleic acid content, whereas neutrophil granularity index (NEUT-GI) measures the abundance of granules. Neutrophil distribution width (NEUT-TW) captures heterogeneity within neutrophil populations by measuring the size of the neutrophil cloud. Lymphocyte activation is indicated by the reactive lymphocyte count (RE-LYMPH), which encompasses the antibody-synthesising lymphocyte count (AS-LYMPH), and the high fluorescing lymphocyte count (HFLC), representing atypical lymphocytes. Similarly, reactive monocyte count (RE-MONO) is reported based on

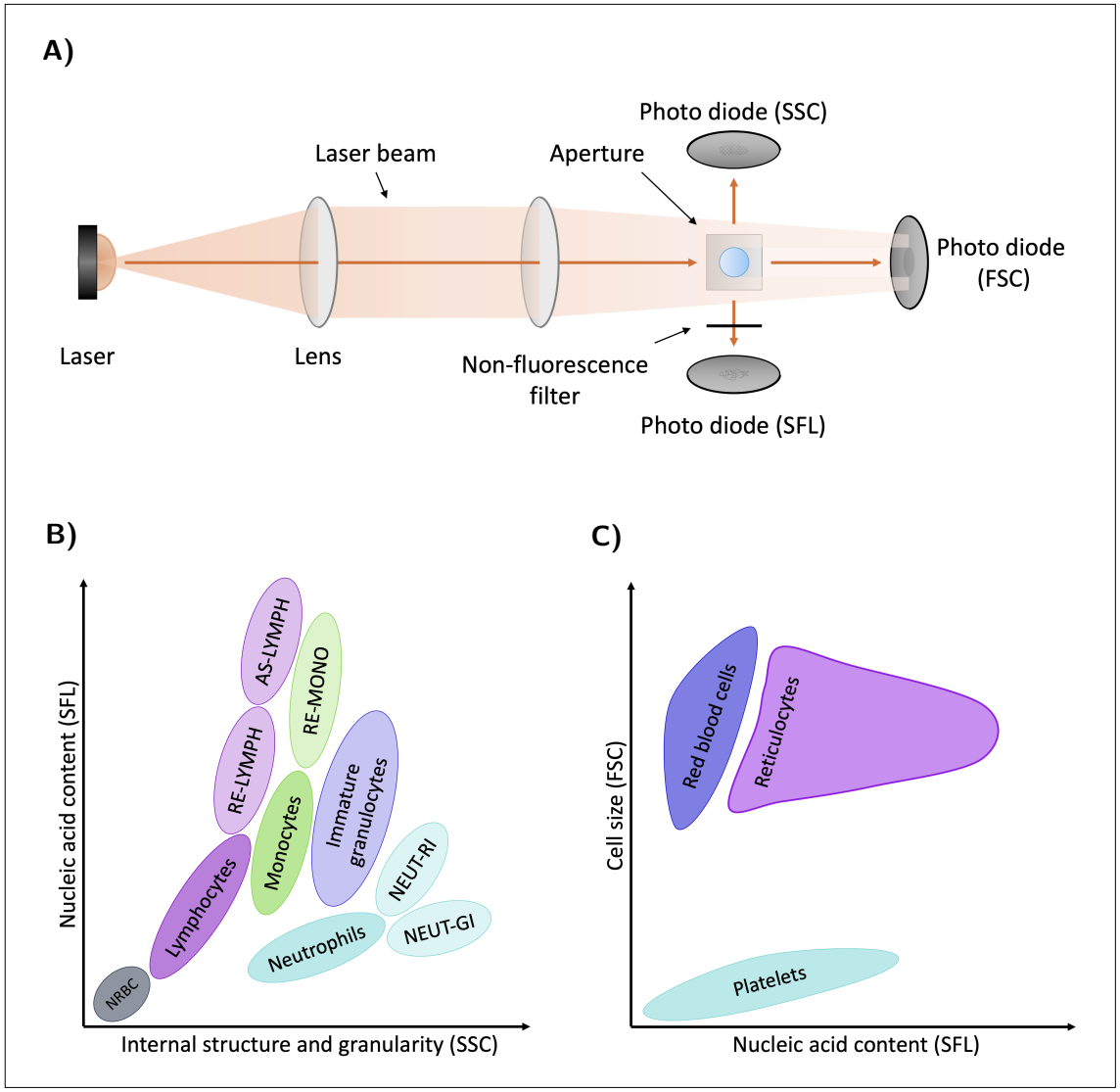


Figure 1.3: Fluorescence flow cytometry analysis of white blood cells, red blood cells, and platelets.

A) Simplified diagram of fluorescence flow cytometry showing the detection of FSC, SSC, and SFL. **B)** SSC/SFL scattergram for analysis of white blood cells. **C)** SFL/FSC scattergram for analysis of red blood cells, reticulocytes, and platelets.

Abbreviations: AS-LYMPH – antibody-synthesising lymphocyte count; FSC – forward scattered light; NEUT-GI – neutrophil granularity index; NEUT-RI – neutrophil reactive index; NRBC – nucleated red blood cell; RE-LYMPH – reactive lymphocyte count; SFL – side fluorescence intensity; SSC – side scattered light.

elevated monocyte fluorescence levels. NRBC counts (NRBC# and NRBC%) are also calcualted from SSC/SFL analysis.

The FSC/SFL scatterplot enables further analysis of erythroid and platelet populations (Figure 1.3C), including red blood cell haemoglobin equivalent (RBC-HE), hypochromic red blood cell count (HYPO-HE; the proportion of cells with a volume less than 17 pg), and hyperchromic red blood cell count (HYPER-HE; the proportion of cells with a volume greater than 49 pg). Platelet-related metrics derived from this scatterplot include an alternative PLT# and the immature platelet fraction (IPF).

Reticulocytes are evaluated based on fluorescence intensity, which is indicative of their RNA content. Analysers provide both absolute and relative reticulocyte counts (RET# and RET%), and calculate immature reticulocyte fraction (IRF), representing the proportion of highly fluorescent, and therefore less mature, reticulocytes (Figure 1.3C). Reticulocyte haemoglobin equivalent (RET-HE) and haemoglobin equivalent difference (DELTA-HE) between RBCs and reticulocytes offer insights into recent erythropoietic activity and iron availability.

1.2.3 Limitations of the Complete Blood Count

Despite its widespread clinical utility, the CBC has several limitations that should be considered when interpreting results. Pre-analytical errors, occurring before sample analysis, account for up to 70% of total laboratory errors [67]. Factors such as patient posture and strenuous activity can temporarily affect blood parameters, including spurious anaemia or leukopenia, which can lead to incorrect diagnoses [67]. Sample clotting is another source of pre-analytical error, accounting for 40 – 60% of rejected samples [67]. Errors may also result from improper mixing or incorrect dilution with EDTA [67]. These issues are exacerbated in neonatal patients due to their small circulating blood volume, making it difficult to obtain adequate samples.

Ensuring proper transport is vital to preserve sample integrity, as CBC analysis should ideally be completed within six hours of collection to prevent morphological changes

in RBCs and WBCs [67]. At room temperature, statistically significant changes in the haematological profile may occur after 6, 24, and 48 hours, including a decrease in PLT# and an increase in HCT [68]. Refrigeration for over 24 hours leads to an increase in MPV, while NEUT# decreases after 48 hours. Therefore, blood samples must be delivered and analysed promptly to ensure accurate results.

Analytical errors that occur during measurement account for an additional 7 – 13% of total laboratory errors [67]. These include falsely low PLT# due to platelet clumping or a falsely high HGB caused by interference from elevated bilirubin or lipid concentrations, which is common in neonatal samples [67, 69]. Of particular concern for neonatal samples, NRBCs may be misclassified as lymphocytes due to their internal structure, leading to a falsely increased WBC# and LYMPH# [67, 69].

While modern haematology analysers are often able to identify and flag such anomalies, not all flagged results undergo manual review. Larger Sysmex analysers detect and correct for NRBC interference; however, this feature is not available on smaller point-of-care models, which are more often used in neonatal settings. To ensure reliability, analysers are subject to strict quality control protocols and external validation programmes [70].

1.2.4 Diagnostic Utility of the Complete Blood Count

Despite these limitations, the CBC remains an essential and indispensable diagnostic tool. Specific abnormalities in cell counts and morphologies, detectable by the CBC, may serve as early indicators of infection, inflammation, or bone marrow dysfunction [55]. Reference ranges for many CBC parameters are well established in adult populations (Table 1.4) [62], and values that fall outside of these established ranges are indicative of various medical conditions or diseases.

Anaemia, characterised by a reduction in HGB values, may result from iron deficiency, inflammatory processes, thalassaemia, or infectious diseases such as malaria, and affects approximately one-third of the global population [55]. Clinical manifestations vary according to the severity and speed of onset, but common symptoms include shortness

Table 1.4: Reference ranges for the complete blood count in healthy adults, adapted from van Pelt et al. (2022) [62].

Parameter	Units	Sex	Lower	Average	Upper
Red blood cell parameters					
RBC#	x 10^{12}/L	Male	4.40	5.10	5.70
		Female	4.00	4.50	5.20
HGB	g/L	Male	134	152	170
		Female	118	136	152
HCT	%	Male	40.6	45.4	50.2
		Female	36.5	41.1	45.8
MCV	fL	All	82.5	90.3	97.4
MCH	pg	All	26.8	30.0	32.6
MCHC	g/L	Male	317	336	352
		Female	311	330	346
RDW-CV	fL	All	37.9	42.5	48.3
RDW-SD	%	All	11.8	12.8	14.3
NRBC#	x 10^{9}/L	All	0.00	0.00	0.01
NRBC%	%	All	0.0	0.0	0.2
RET#	x 10^{12}/L	All	0.328	0.578	0.977
RET%	%	All	0.7	1.2	2.0
IRF	%	All	2.7	6.9	13.8
RBC-HE	pg	All	27.2	30.2	32.5
RET-HE	pg	All	29.3	32.8	35.4
DELTA-HE	pg	All	1.2	2.6	3.6
HYPO-HE	%	All	0.0	0.1	0.4
HYPER-HE	%	All	0.4	0.6	0.8
MICRO-RBC	%	All	0.3	1.1	3.3
MACRO-RBC	%	All	3.1	3.6	4.5
White blood cell parameters					
WBC#	x 10^{9}/L	All	3.70	5.80	9.20
NEUT#	x 10^{9}/L	All	1.60	3.10	5.80
NEUT%	%	All	37.6	54.1	69.3
LYMPH#	x 10^{9}/L	All	1.10	1.90	3.30

Continued on next page

Continued from last page

Parameter	Units	Sex	Lower	Average	Upper
LYMPH%	%	All	20.0	33.5	48.8
MONO#	x 10^9/L	All	0.30	0.50	0.80
MONO%	%	All	5.3	8.2	12.4
EOSI#	x 10^9/L	All	0.05	0.16	0.53
EOSI%	%	All	0.9	2.8	8.3
BASO#	x 10^9/L	All	0.02	0.04	0.10
BASO%	%	All	0.3	0.8	1.6
IG#	x 10^9/L	All	0.01	0.03	0.07
IG%	%	All	0.2	0.6	1.0
NEUT-RI	FI	All	42.0	46.1	50.6
NEUT-GI	SI	All	142.5	149.4	157.0
RE-LYMPH	%L	All	1.3	3.3	7.8
AS-LYMPH	%L	All	0.0	0.0	0.0
HFLC	%L	All	0.0	0.0	0.3
RE-MONO	%M	All	0.0	2.0	4.4
Platelet parameters					
PLT#	x 10^9/L	All	164	254	369
PDW	%	All	10.0	12.8	17.4
MPV	fL	All	9.3	10.7	12.7
P-LCR	%	All	19.3	31.2	47.1
PCT	%	All	0.2	0.3	0.4
IPF	%	All	1.2	3.1	8.9

For parameter abbreviations, see List of Complete Blood Count Parameter Abbreviations (p. xxi).

of breath, lethargy, tachycardia, headaches, chest pain, and cognitive impairment [55]. RET# increases in response to heightened erythropoietic activity, correlating with anaemia severity. A low RET# despite anaemia may indicate bone marrow dysfunction [55].

Leukocytosis, particularly an increase in NEUT#, is frequently associated with acute infection or inflammation but may also indicate tissue necrosis, metabolic stress, or malignancy [55, 71]. Peripheral blood smears may be necessary to evaluate morphological abnormalities, such as toxic granulation, which can aid in distinguishing between reactive and malignant causes [71]. Treatment with corticosteroids and epinephrine also induces transient neutrophilia [71]. Conversely, neutropenia may result from systemic inflammatory responses, viral infections, or chemotherapy [71]. Typically, neither monocytosis nor monocytopenia occurs in isolation, but instead in conjunction with broader changes in the WBC differential [71].

The inclusion of traditional CBC parameters (RBC#, HGB, HCT, MCV, MCH, MCHC, RDW-SD or RDW-CV, WBC# and its differentials, and PLT#) in sepsis evaluations is well established [72]. However, emerging evidence supports the utility of advanced CBC parameters for early detection of sepsis and risk stratification. For example, extended activation markers, such as NEUT-RI and RE-LYMPH, show strong diagnostic accuracy in identifying patients at risk of developing sepsis [72]. Similarly, ratios such as neutrophil-to-lymphocyte ratio (NLR), platelet-to-lymphocyte ratio (PLR), and immature-to-total neutrophil ratio (ITR) provide more detailed insights into WBC dynamics during infection, which absolute or relative counts alone cannot offer [72].

More complex diagnostic models incorporating multiple CBC parameters may offer improved clinical utility over individual markers. For example, the Intensive Care Infection Score (ICIS) identifies septic adult patients by scoring four CBC parameters based on pre-defined diagnostic thresholds [73–75]. ICIS demonstrated an AUC of 0.852, a sensitivity of 82.9%, and a specificity of 75.1%, outperforming CRP, procalcitonin, and WBC# [74]. CBC-based diagnostic models for neonatal sepsis will be discussed in Chapter 5.

1.3 Rationale and Objectives

Neonatal sepsis continues to be a globally prevalent and clinically significant condition, associated with both immediate and long-term complications in affected neonates. The current gold standard, a positive blood culture, is constrained by prolonged turnaround times and decreased sensitivity, particularly in instances of low-level or intermittent bacteraemia. While CRP is valuable for monitoring recovery, it is not suitable for early detection due to its delayed response to infection.

The CBC is a rapid, inexpensive, and readily available test routinely performed in NICU settings. Beyond its traditional metrics, the modern CBC provides a wide range of often underreported parameters that may serve as indicators of early immune activation, including dynamic changes in neutrophil, monocyte, and lymphocyte populations. Several of these emerging parameters have demonstrated strong diagnostic potential in adult sepsis, suggesting their potential utility within neonatal populations.

The main aim of this thesis was to evaluate the diagnostic value of the CBC, including both conventional and less frequently reported parameters, in the identification of neonatal sepsis, with an emphasis on LOS. The central hypothesis is that specific parameters of the neonatal CBC, individually and in combination, discriminate between neonates with and without sepsis, and that multivariable diagnostic models incorporating these parameters achieve diagnostic accuracy sufficient to inform clinical decision-making.

This thesis is organised into seven chapters, beginning with the present Chapter 1, which provides the clinical background and rationale for the study. Chapter 2 presents a meta-analysis of the existing research evaluating the diagnostic accuracy of CBC parameters for neonatal sepsis. Chapter 3 defines the baseline CBC profile in a cohort of neonates admitted to the NICU, establishing a foundation for Chapter 4, which evaluates the diagnostic performance of individual CBC parameters in cases of confirmed LOS. Chapter 5 describes the development of multivariable diagnostic models using the parameters with the highest diagnostic performance.

Chapter 6 applies these analytical methods to CBC data collected within the first three days of life and from umbilical cord blood samples, evaluating their utility in the context of EOS. Finally, Chapter 7 synthesises the findings of this research, contextualises them within current clinical practices and policies, evaluates the commercial potential of the developed models, and proposes directions for future research.

Through the utilisation of routinely collected haematological data, this thesis aims to facilitate earlier and more accurate detection of neonatal sepsis, with the ultimate goal of reducing unnecessary antibiotic exposure and improving clinical outcomes.

Chapter 2

The Diagnostic Accuracy of the Complete Blood Count for Neonatal Sepsis: A Systematic Review and Meta-Analysis

2.1 Introduction

The results presented in this chapter have been published by *Archives of Disease in Childhood*, a peer-reviewed academic journal [76].

As discussed in Chapter 1, a positive blood culture is the gold standard for diagnosing neonatal sepsis, and CRP is suggested as a tool for monitoring recovery [11]. In the development of the NICE guidelines, the evidence for four parameters of the CBC was evaluated (summarised in Table 2.1) [77]. ITR provided the highest pooled sensitivity (70%) and the lowest specificity (55%) while WBC# had the highest specificity (87%) and lowest sensitivity (46%). Overall, the quality of evidence for WBC#, NEUT#, ITR, and PLT# was determined to be "very low", and so the CBC was excluded from the guidelines.

However, this evaluation did not include all parameters available from the routine CBC. As previously outlined in Chapter 1, modern CBC analysers provide both traditional metrics and details on the functionality of different haematological cell types. Although the NICE evidence review evaluated the evidence for the diagnostic utility of the ITR, it did not consider the efficacy of novel immune cell activation parameters, which have previously been shown to have utility in diagnosing adult sepsis [72].

In contrast, some guidelines do advise including the CBC in neonatal sepsis diagnostics. The European Medicines Agency (2010) concluded that a WBC# less than 4 x 10^9/L or greater than 20 x 10^9/L, ITR greater than 0.2, and PLT# less than 100 x 10^9/L are

Table 2.1: Diagnostic accuracy of the complete blood count for neonatal sepsis, as evaluated by the NICE evidence review [77].

Parameter	Threshold	Sensitivity (%; 95% CI)	Specificity (%; 95% CI)
WBC# (x 10^9/L)	< 5.00 / > 20.00	46 (32 – 60)	87 (66 – 96)
NEUT# (x 10^9/L)	< 1.80 / > 5.40	60 (48 – 70)	62 (51 – 72)
ITR	> 0.12 / > 0.20 / > 0.65	70 (39 – 89)	55 (26 – 81)
PLT# (x 10^9/L)	< 100 / < 150	53 (34 – 71)	63 (19 – 92)

Table 2.2: Diagnostic accuracy of the complete blood count for neonatal sepsis, as evaluated by Celik et al. (2022) [2].

Parameter	Threshold	Sensitivity (%)	Specificity (%)
WBC# (x 10^9/L)	< 5.00 / > 20.00	10 – 23	80 – 99
NEUT# (x 10^9/L)	< 1.00 / > 5.00	8 – 68	95 – 99
ITR	> 0.20	33 – 54	62 – 100
DNI	> 4.60	85	80
MNV[a] (au)	> 157	79	82
MNC[a] (au)	< 159	66	64
MNS[a] (au)	< 128	60	65
PLT# (x 10^9/L)	Unspecified	8 – 48	89 – 98

[a] *MNV, MNC, and MNS are measures of neutrophil activation specific to Beckman Coulter analysers. AUC values >0.8 and sensitivities/specificities >85% are highlighted in green, suggesting strong diagnostic performance. Abbreviations: CI – confidence interval; DNI – delta neutrophil index; ITR – immature-to-total neutrophil ratio; MNC – mean neutrophil conductivity; MNS – mean neutrophil scatter; MNV – mean neutrophil volume; NEUT# – neutrophil count; PLT# – platelet count; WBC# – white blood cell count.*

indicators of neonatal sepsis, but should be considered in the context of the broader condition [78]. Similarly, the CBC is mentioned in the World Health Organisation Pocket Book of Hospital Care for Children as a possible investigative tool for neonatal sepsis [79]. However, the Pocket Book does not detail its interpretation.

The diagnostic utility of the CBC for neonatal sepsis was recently reviewed by Boscarino et al. (2023) [19]. The authors draw attention to the prevalence of confounders within the literature, specifically the lack of a definitive case definition, which has led to overlaps between cases and controls across studies. They also comment on the possibility of combining CBC parameters with clinical measures to produce more sophisticated models, although they note that their interpretation is limited.

Celik et al. (2022) present a more comprehensive review of the current literature (summarised in Table 2.2) [2]. This evaluation included neutrophil activation parameters specific to Beckman Coulter analysers: mean neutrophil volume (MNV), mean neutrophil conductivity (MNC), and mean neutrophil scatter (MNS). WBC#, NEUT#, ITR, and PLT# provided high specificity but low sensitivity for neonatal sepsis, while MNV and delta neutrophil index (DNI; the percentage of circulating immature to mature neutrophils) exhibited more well-rounded performance. However, the included case definitions were not specified, so it is unclear whether there was overlap between the cases and controls of different studies.

2.1.1 Chapter Aims

This chapter aimed to evaluate the evidence for the diagnostic accuracy of the CBC for neonatal sepsis. This was achieved by conducting a systematic review and meta-analysis of the available literature. The literature search targeted traditional markers, including WBC# and NEUT#, as well as functional parameters such as ITR and neutrophil activation parameters, ensuring that the full extent of the CBC was evaluated. Studies that did not define sepsis using the gold standard of a blood culture were excluded, as were studies in which the control group exhibited clinical signs of sepsis without a positive culture. This avoided overlap between the cases and controls of the included studies.

2.2 Methods

This review follows the Cochrane Library guidelines for conducting a diagnostic accuracy systematic review [80] and was registered prospectively with PROSPERO on 5$^{\text{th}}$ December 2023 (ID: CRD42023476510) [81].

2.2.1 Search Strategy, Eligibility Criteria, and Study Selection

I conducted a comprehensive electronic search for studies reporting the diagnostic accuracy of the CBC for neonatal sepsis. Eligibility criteria included prospective, retrospective, and cross-sectional observational studies conducted on humans and published in English. Abstracts, case reports, and conference proceedings were excluded because they lacked sufficient detail on diagnostic accuracy metrics.

Searches were performed in MEDLINE, Embase, Cochrane, and CINAHL from database inception to 28$^{\text{th}}$ October 2024. Search terms included references to neonatal sepsis, blood cell types, and diagnostic accuracy metrics such as sensitivity and specificity. Full search strategies for each database are provided in Appendix A1 (p. 238). To identify additional studies, the reference lists of included studies and published reviews were manually screened. All search results were imported into Covidence [82] for screening and evaluation.

Two reviewers (EH, MA) independently screened the titles and abstracts of all identified studies for relevance, with discrepancies resolved by a third reviewer (JS). Full texts of selected studies were reviewed to confirm eligibility, and data on study design, participant characteristics, and diagnostic accuracy of individual parameters were extracted.

2.2.2 Data Extraction and Synthesis

For each study, diagnostic accuracy measures including sensitivity, specificity, PPV, and NPV were recorded in a spreadsheet (Microsoft Excel). Where PPV and NPV were not directly reported, they were calculated using sensitivity, specificity, and the prevalence of sepsis in each study. True positives (TPs), true negatives (TNs), false positives (FPs),

and false negatives (FNs) were also derived. Case definitions for neonatal sepsis (i.e. positive culture with or without clinical signs of sepsis) and control definitions were also documented.

Patient characteristics and biomarker values were summarised as the mean $\pm$ standard deviation (SD) or median with the interquartile range (IQR), according to underlying distributions. Optimal thresholds for each parameter were identified by calculating the Youden index value for individual studies and selecting the threshold corresponding to the highest Youden index value. Diagnostic accuracy was assessed using a bivariate model (R package *mada*, function *reitsma*), incorporating sensitivity, specificity, and AUC. This model was employed to account for study-level differences in diagnostic thresholds. Heterogeneity was quantified using the I^2 statistic, where values closer to zero denote lower heterogeneity.

A minimum of three studies was required to perform a meta-analysis for each parameter. Subgroup analyses were not conducted to explore variability across study designs and sampling times due to the limited number of studies. Publication bias was assessed visually through funnel plots.

2.2.3 Assessment of Risk of Bias

Applicability was assessed using a modified tool derived from QUADAS-2 [83], the CASP diagnostic checklist [84], and the Newcastle-Ottawa scale [85]. Questions were selected to ensure studies were relevant to the research objective. Two reviewers (EH, MA) independently applied the tool, with disagreements resolved by a third reviewer (JS). Studies scoring less than 6/8 were considered to have lower applicability; however, all studies were included in the meta-analysis, regardless of their score. All statistical analyses were performed using RStudio (R version 4.3.0) [86].

2.3 Results

2.3.1 Study Selection

A total of 428 studies were identified through database searches (Figure 2.1), of which 259 (61%) were excluded after title and abstract screening. Ultimately, 24 studies were included in this systematic review: 14 from database screening [87–100] and 10 additional studies identified through reference list screening of previously published reviews (n = 5) [101–105] and included studies (n = 5) [106–110]. The reasons for exclusion are presented in Figure 2.1.

2.3.2 Study Characteristics

Among the 24 studies included in this review, 16 were prospective [87–90, 92, 93, 95–98, 100, 104, 106–108, 110] and eight were retrospective [91, 94, 99, 101–103, 105, 109] (Table 2.3). All studies defined neonatal sepsis based on a positive blood culture. Nineteen studies also included clinical signs in their case definition [88–93, 95–98, 100–105, 108–110], and four included additional laboratory indicators [89–91, 101]. The median prevalence of sepsis across the included studies was 50.0% (IQR: 40.9% – 51.7%).

Fifteen studies defined the control group as an absence of clinical signs of sepsis [87–90, 92, 93, 96–98, 101–103, 105, 108, 109], six included the absence of laboratory markers [87, 88, 92, 96, 98, 101], and nine did not detail the criteria used to define the control group beyond "healthy neonates" [91, 94, 95, 99, 100, 104, 106, 107, 110].

In total, 1,377 neonates with sepsis were included across the 24 studies (Table 2.3). The median GA was 34.3 weeks (IQR: 31.1 – 38.0 weeks), and the median postnatal age at sampling was 8 days (IQR: 7 – 15 days). The control group consisted of 1,865 neonates, with a median GA of 35.6 weeks (IQR: 31.3 – 38.0 weeks) and a median postnatal age at sampling of 7 days (IQR: 3 – 10 days). No significant differences were observed in GA ($p = 0.71$) or age at sampling ($p = 0.36$) between the sepsis and control groups.

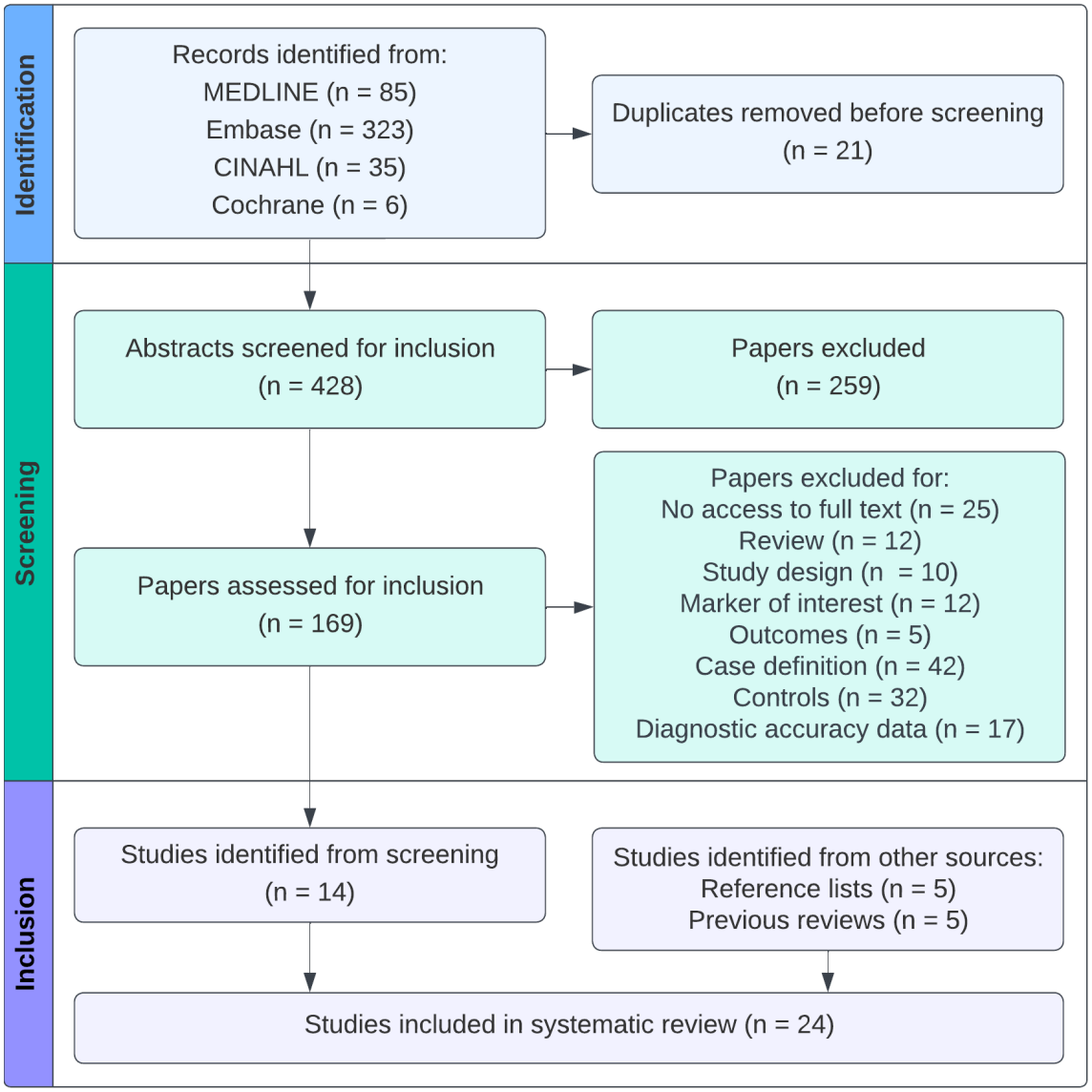


Figure 2.1: PRISMA [111] flowchart of study selection for this systematic review and meta-analysis.

Table 2.3: Characteristics of individual studies included in the systematic review and meta-analysis.

			Cases					Controls					
No	First Author and Year	Design	Definition	N	Males	Age	GA	Definition	N	Males	Age	GA	Parameters
#1	Abiramalatha 2016 [87]	Pro	BC	38	25	4	35.5	No CS/LS	360	195	1	36.5	WBC#, NEUT#, ITR, MNV, PLT#
#2	Aboud 2010 [88]	Pro	BC/CS	25	13	9	36.4	No CS/LS	22	8	10	37.3	WBC#
#3	Arcagok 2019 [101]	Retro	BC/CS/LS	67	33	–	39.1	No CS/LS	92	43	–	40.1	ITR, PLR
#4	Buyukeren 2021 [89]	Pro	BC/CS/LS	77	38	15	34.3	No CS	131	60	9	34.8	DNI
#5	Çelik 2016 [90]	Pro	BC/CS/LS	40	26	5	32.5	No CS	111	70	1	34.4	WBC#, NEUT#, ITR, MNV, MNC, MNS
#6	Celik 2019 [91]	Retro	BC/CS/LS	110	64	8	28.0	No SUS	87	40	7	30.0	DNI
#7	Elawady 2013 [92]	Pro	BC/CS	25	17	–	–	No CS/LS	25	15	–	–	WBC#, NEUT#, ITR, PLT#
#8	Guney Varal 2020 [102]	Retro	BC/CS	76	40	15	29.0	No CS	40	22	15	28.0	NLR
#9	Guney Varal 2023 [106]	Pro	BC	64	36	–	28.0	No SUS	44	24	–	29.0	PLT#, MPV
#10	Hanaganahalli 2018 [103]	Retro	BC/CS	64	41	7	39.6	No CS	71	46	4	39.8	MPV
#11	Kumar 2015 [93]	Pro	BC/CS	101	58	–	–	No CS	96	52	–	–	WBC#
#12	Leibovitch 2024 [94]	Retro	BC	63	34	–	31.1	No SUS	63	28	–	31.7	MPV
#13	Madani 2019 [104]	Pro	BC/CS	20	10	8	–	No SUS	20	12	3	–	MPV
#14	Mahmoud 2022 [95]	Pro	BC/CS	80	–	2	38.7	No SUS	80	–	2	38.8	NLR, PLR
#15	Manucha 2002 [107]	Pro	BC	21	–	–	–	No SUS	40	–	–	–	WBC#, NEUT#, ITR, PLT#
#16	Nesargi 2020 [96]	Pro	BC/CS	84	44	7	37.0	No CS/LS	160	88	3	38.0	MNV, MNC, MNS
#17	Panda 2021 [105]	Retro	BC/CS	41	27	–	34.2	No CS	52	25	–	33.8	NLR
#18	Panda 2022 [108]	Pro	BC/CS	43	23	–	33.29	No CS	54	24	–	34.2	PLT#, MPV

Continued on next page

Continued from last page

			Cases					Controls					
No	First Author and Year	Design	Definition	N	Males	Age	GA	Definition	N	Males	Age	GA	Parameters
#19	Schrama 2008 [97]	Pro	BC/CS	24	8	–	31.0	No CS	55	27	–	30.1	ITR
#20	Tosson 2024 [98]	Pro	BC/CS	45	25	15	38.1	No CS/LS	45	25	13	38.3	ITR, MPV
#21	Vardar 2023 [99]	Retro	BC	77	38	–	28.0	No SUS	77	39	–	28.0	NLR
#22	Yang 2016 [109]	Retro	BC/CS	60	31	18	34.3	No CS	60	35	18	37.5	WBC#
#23	Zaki 2009 [110]	Pro	BC/CS	58	32	9	39.3	No SUS	30	16	9	36.9	WBC#, NEUT#, ITR, PLT#
#24	Zhang 2021 [100]	Pro	BC/CS	74	45	–	38.0	No SUS	50	25	–	38.0	NLR, PLR

Abbreviations: Age – median postnatal age at sampling (days); BC – positive blood culture; CS – clinical signs of sepsis; DNI – delta neutrophil index; GA – median gestational age at birth (weeks); ITR – immature-to-total neutrophil ratio; LS – laboratory signs of sepsis; MNC – mean neutrophil conductivity; MNS – mean neutrophil scatter; MNV – mean neutrophil volume; MPV – mean platelet volume; NEUT# – neutrophil count; NLR – neutrophil-to-lymphocyte ratio; PLR – platelet-to-lymphocyte ratio; PLT# – platelet count; Pro – prospective study; Retro – retrospective study; SUS – suspicion of sepsis; WBC# – white blood cell count.

2.3.3 Level of Applicability

Sixteen of the 24 included studies scored 6/8 or higher on the quality assessment tool [87, 89–91, 94–98, 101–103, 105, 106, 108, 110] (Table A1.1, p. 241). Nine studies did not report when samples were taken [91–93, 95, 99, 101, 103, 104, 110], and nine did not specify how the CBC was measured [88, 93, 94, 97, 99–101, 103, 109]. Only five studies included a sample size or power calculation [87, 95, 101, 104, 106]. While the overall applicability of the included studies was high (Figure 2.2), the lack of detailed reporting in some studies made assessing certain aspects of applicability more challenging.

2.3.4 Synthesis of Results

The included studies assessed the diagnostic accuracy of multiple parameters for neonatal sepsis, including WBC# (n = 8) [87, 88, 90, 92, 93, 107, 109, 110], NEUT# (n = 5) [87, 90, 92, 107, 110], ITR (n = 8) [87, 90, 92, 97, 98, 101, 107, 110], NLR (n = 5) [95, 99, 100, 102, 105], DNI (n = 2) [89, 91], neutrophil activation markers (MNV (n = 3) [87, 90, 96], MNC (n = 2) [90, 96], and MNS (n = 2) [90, 96]), PLT# (n = 6) [87, 92, 106–108, 110], MPV (n = 6) [94, 98, 103, 104, 106, 108], and PLR (n = 3) [95, 100, 101] (Table A1.2, p. 243).

Analyses were conducted for WBC#, NEUT#, ITR, NLR, MNV, PLT#, and MPV (Table 2.4; Figures 2.3 – 2.11). Due to insufficient studies reporting diagnostic accuracy data, meta-analyses were not performed for DNI, MNC, and MNS. Additionally, a reliable meta-analysis could not be conducted for PLR due to substantial heterogeneity in reported values across studies (15.00, 62.40, and 99.57), which prevented meaningful statistical pooling. One study [95] was excluded from the meta-analysis of NLR due to extreme variability and non-overlapping values compared to other studies (0.80 versus 2.10 – 4.16 in different studies), which suggested methodological differences or population heterogeneity. Therefore, three studies were not included in any meta-analyses [90, 91, 95].

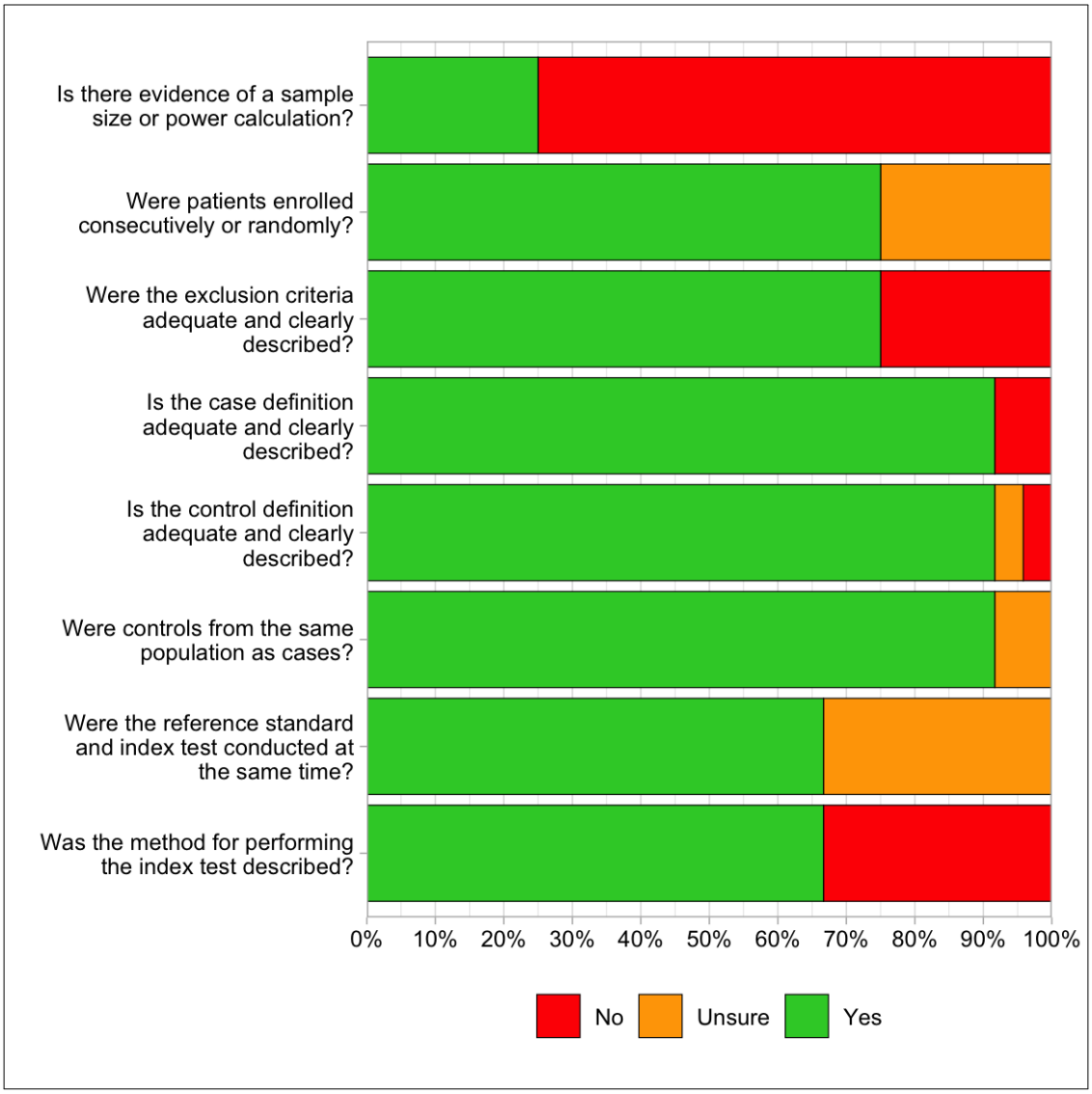


Figure 2.2: Level of applicability of the studies included in the systematic review and meta-analysis, adapted from QUADAS-2 [83], the CASP diagnostic checklist [84], and the Newcastle-Ottawa scale [85].

MNV showed the highest diagnostic accuracy for neonatal sepsis, with an optimal threshold of 158 au [96], 76.6% sensitivity and 90.9% specificity (Figure 2.7). However, there was a great deal of heterogeneity among the included studies (I^2: 0.85).

In contrast, studies reporting on ITR exhibited much less heterogeneity ($I^2 < 0.01$) (Figure 2.6). The diagnostic accuracy for ITR was also high, with 66.3% sensitivity, 85.4% specificity, and an optimal threshold of 0.2 [92]. A subgroup analysis of studies only evaluating ITR greater than 0.2 ($n = 5$) [87, 92, 97, 101, 110] revealed a higher sensitivity compared to the full cohort (72.5% versus 66.3%) while specificity remained similar (85.3% versus 85.4%) and the AUC was higher (0.816 versus 0.773) (Figure 2.10).

PLT# demonstrated moderate diagnostic accuracy, with a sensitivity of 66.5%, a specificity of 81.4%, and an optimal threshold of 278 x 10^9/L [106] (Figure 2.8). Subgroup analysis of studies reporting an optimal threshold of 150 x 10^9/L reduced the I^2 statistic (< 0.1) and resulted in a slightly lower overall diagnostic accuracy (sensitivity: 51.2%; specificity: 82.0%) (Figure 2.11).

NLR had the lowest diagnostic accuracy, with a sensitivity of 69.1%, a specificity of 72.6%, and an AUC of 0.685 with an optimal threshold of 3.17 [100] (Figure 2.5). Analyses for WBC# (Figure 2.3), NEUT# (Figure 2.8), and MPV (Figure 2.9) are presented in Table 2.4. Subgroup analyses based on study design (e.g., prospective, retrospective) or onset of sepsis (e.g., EOS, LOS) could not be conducted due to the low number of studies available for each parameter.

Table 2.4: Meta-analysis of the diagnostic accuracy of the complete blood count for neonatal sepsis.

Parameter	No of Studies	Cases[a]	Controls[a]	Optimal Threshold	Sensitivity[a] (%)	Specificity[a] (%)	AUC	I^2	DOR
WBC# (x 10^9/L)	8	15.55 (11.19 – 16.54)	10.89 (9.02 – 12.76)	$>$ 10.50	55.1 (38.1 – 71.0)	83.3 (73.4 – 90.0)	0.786	0.29	4.08
NEUT# (x 10^9/L)	5	6.90 (6.90 – 7.76)	6.51 (4.90 – 6.51)	$>$ 4.13	67.2 (46.8 – 82.6)	75.9 (63.2 – 85.2)	0.783	$<$ 0.01	1.54
NLR	4	3.20 (2.10 – 4.16)	1.40 (0.70 – 2.34)	$>$ 3.17	69.1 (63.3 – 74.4)	72.6 (54.8 – 85.3)	0.685	0.03	1.18
ITR	8	0.30 (0.18 – 0.33)	0.10 (0.09 – 0.11)	$>$ 0.20	66.3 (58.6 – 73.3)	85.4 (75.0 – 92.0)	0.773	$<$ 0.01	2.97
MNV (au)	3	180 (161 – 180)	147 (147 – 150)	$>$ 158	76.6 (27.5 – 96.6)	90.9 (70.2 – 97.7)	0.925	0.85	3.04
PLT# (x 10^9/L)	6	187 (109 – 220)	285 (251 – 289)	$<$ 278	66.5 (45.7 – 82.3)	81.4 (66.5 – 90.6)	0.815	0.52	2.21
MPV (fL)	6	9.7 (9.6 – 10.0)	8.9 (8.9 – 9.2)	$>$ 11.6	67.7 (58.7 – 75.6)	73.7 (55.8 – 86.2)	0.735	0.39	1.34
ITR $>$ 0.20	5	0.33 (0.30 – 0.33)	0.10 (0.09 – 0.10)	$>$ 0.20	72.5 (60.7 – 81.7)	85.3 (64.4 – 94.9)	0.816	$<$ 0.01	2.20
PLT# $<$ 150	4	187 (187 – 220)	251 (251 – 285)	$<$ 150	51.2 (38.1 – 64.2)	82.0 (62.2 – 92.7)	0.653	$<$ 0.01	4.34

[a] *Median (IQR). AUC values >0.8 and sensitivities/specificities >85% are highlighted in green, suggesting strong diagnostic performance. Abbreviations: AUC – area under the curve; DOR – diagnostic odds ratio; ITR – immature-to-total neutrophil ratio; IQR – interquartile range; MNV – mean neutrophil volume; MPV – mean platelet volume; NEUT# – neutrophil count; NLR – neutrophil-to-lymphocyte ratio; PLT# – platelet count; WBC# – white blood cell count.*

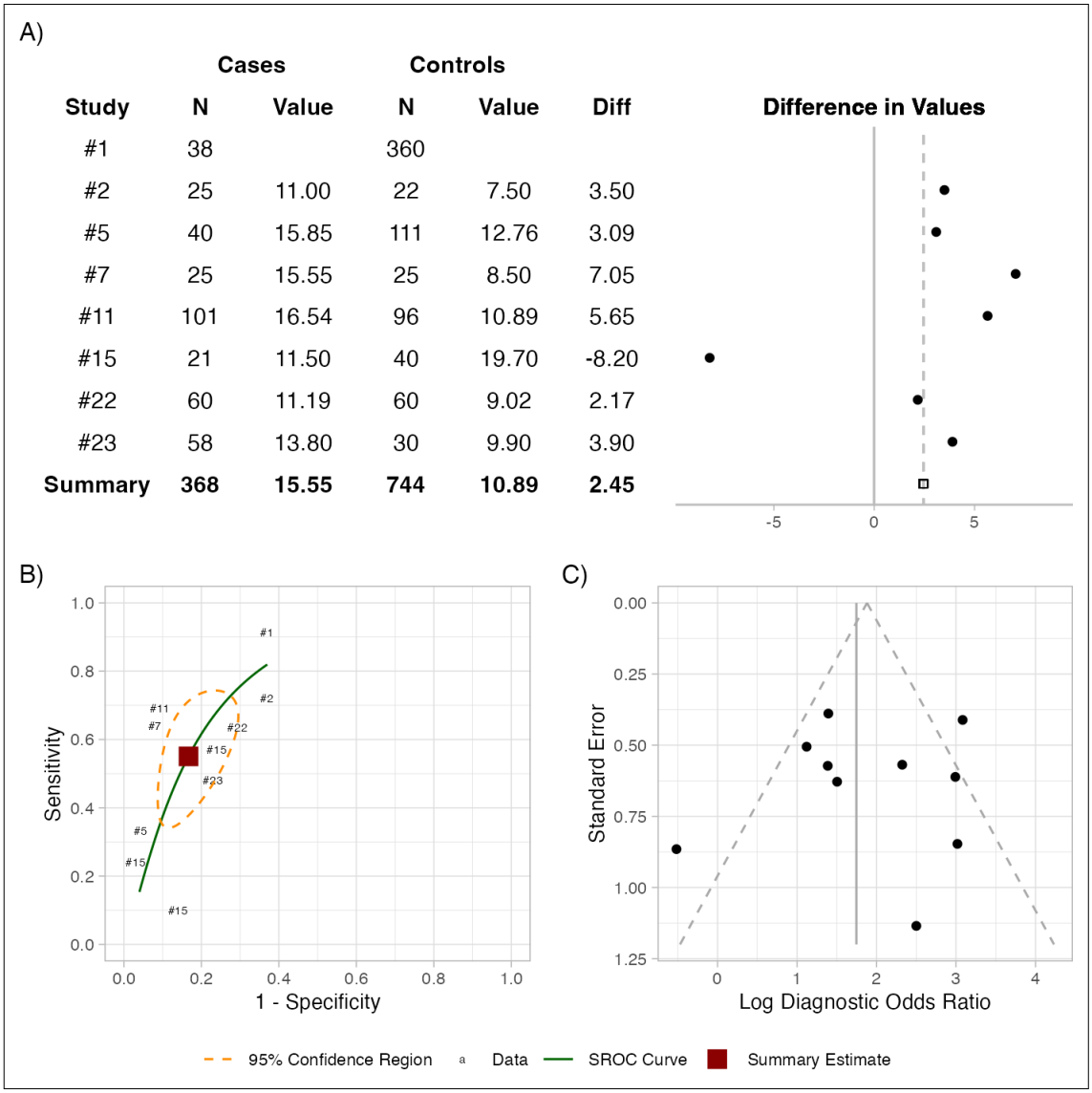


Figure 2.3: Meta-analysis of the diagnostic accuracy of white blood cell count (x 10^9/L) for neonatal sepsis.

A) Forest plot showing differences in values between cases and controls across individual studies. **B)** Summary receiver operating characteristic curve showing overall diagnostic performance across included studies – individual studies are labelled by their study number. **C)** Funnel plot providing a visual assessment of publication bias.

Abbreviations: Diff – difference between values from cases and controls; SROC – summary receiver operating characteristic; Value – mean or median value.

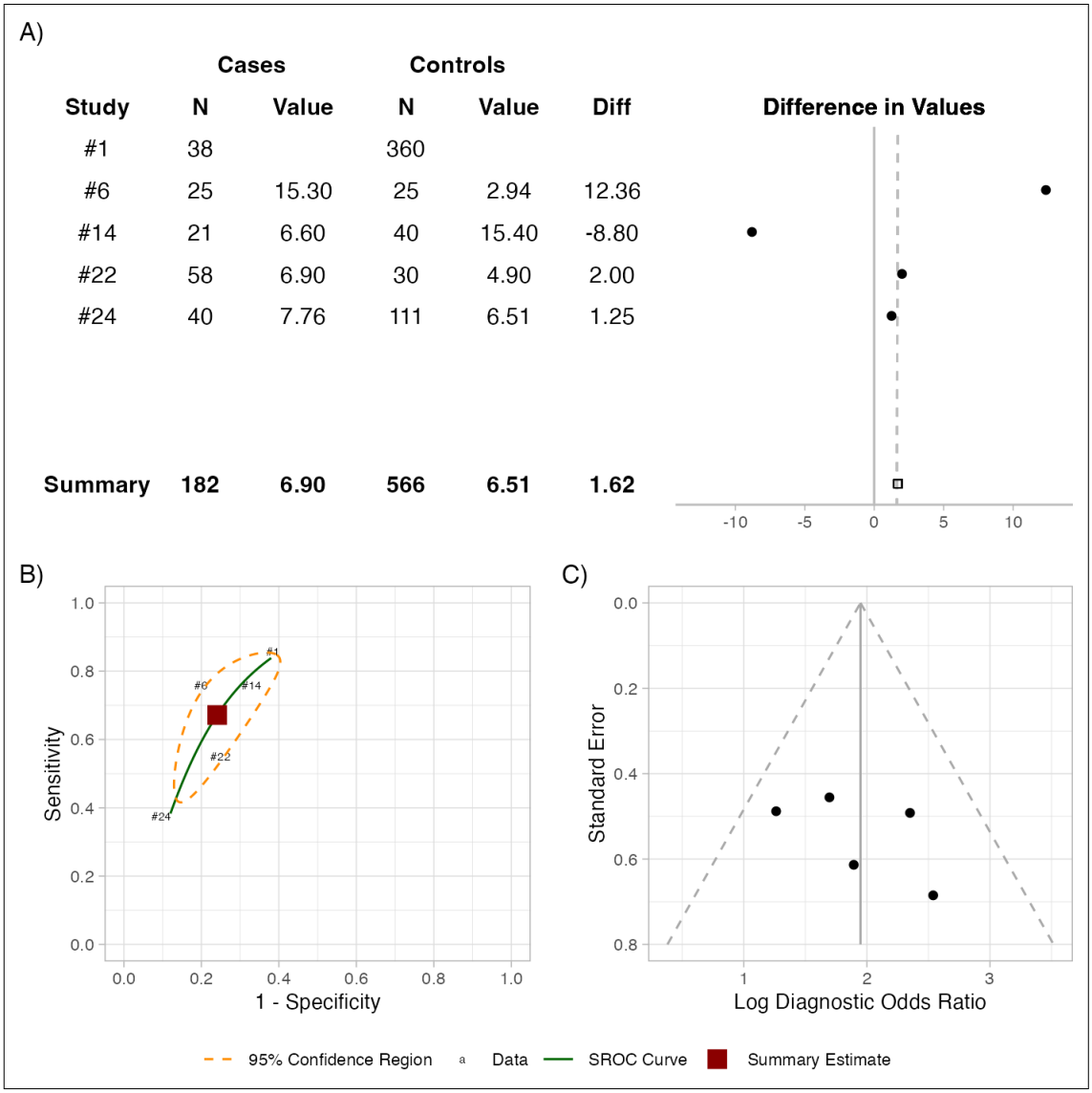


Figure 2.4: Meta-analysis of the diagnostic accuracy of neutrophil count (x 10^9/L) for neonatal sepsis.

A) Forest plot showing differences in values between cases and controls across individual studies, **B)** Summary receiver operating characteristic curve showing overall diagnostic performance across included studies – individual studies are labelled by their study number. **C)** Funnel plot providing a visual assessment of publication bias.

Abbreviations: Diff – difference between values from cases and controls; SROC – summary receiver operating characteristic; Value – mean or median value.

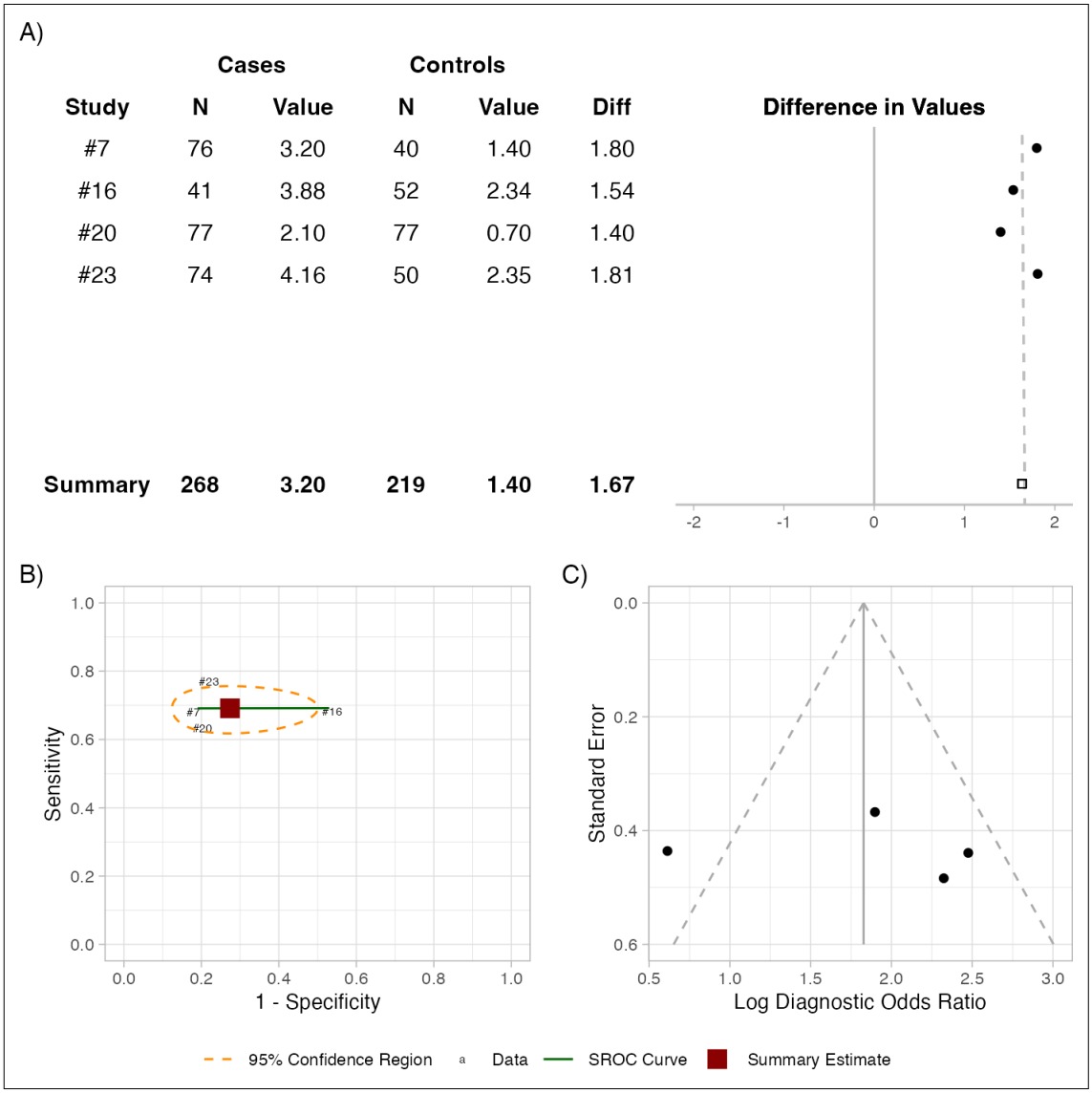


Figure 2.5: Meta-analysis of the diagnostic accuracy of neutrophil-to-lymphocyte ratio for neonatal sepsis.

A) Forest plot showing differences in values between cases and controls across individual studies. **B)** Summary receiver operating characteristic curve showing overall diagnostic performance across included studies – individual studies are labelled by their study number. **C)** Funnel plot providing a visual assessment of publication bias.

Abbreviations: Diff – difference between values from cases and controls; SROC – summary receiver operating characteristic; Value – mean or median value.

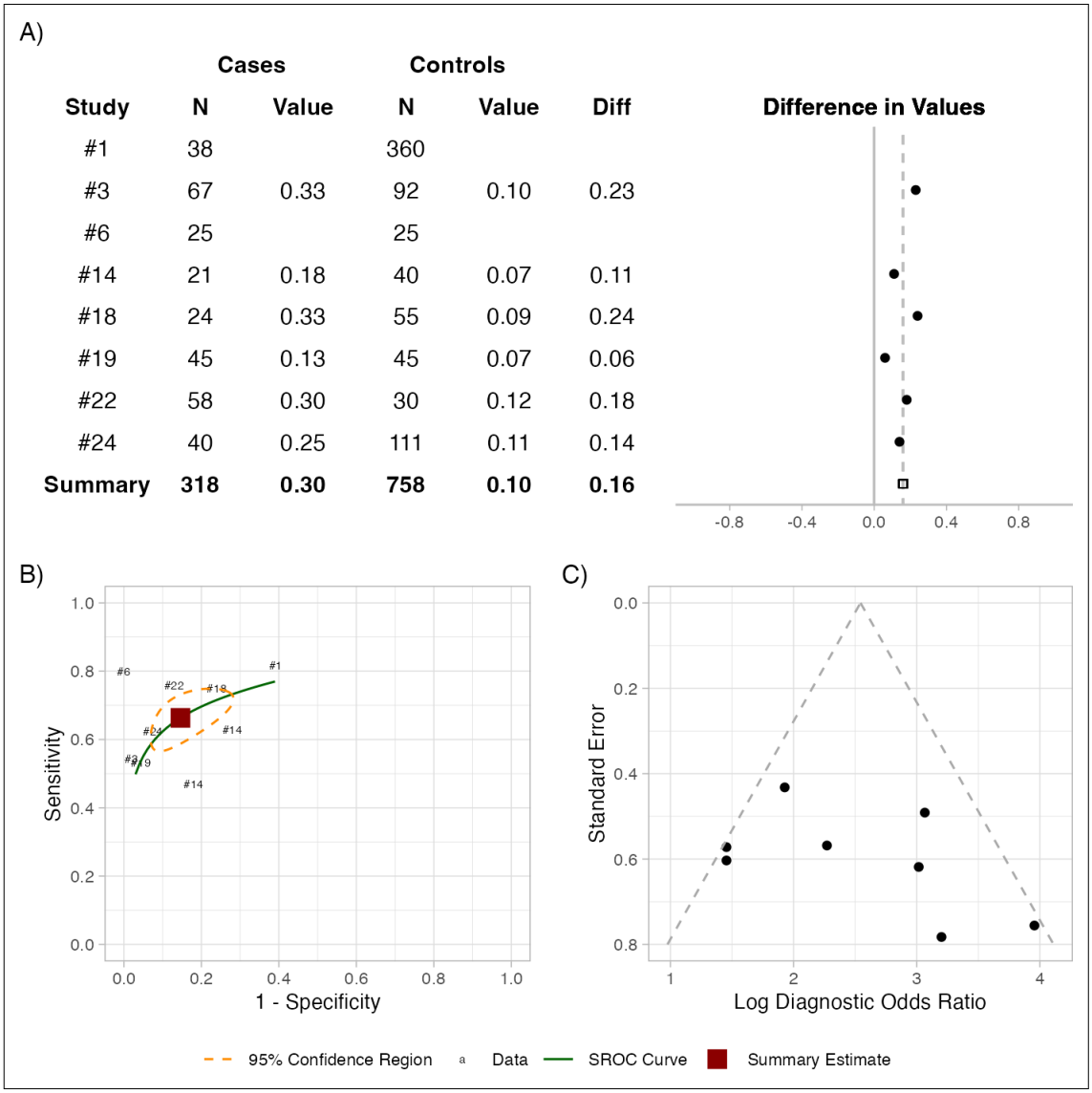


Figure 2.6: Meta-analysis of the diagnostic accuracy of immature-to-total neutrophil ratio for neonatal sepsis.

A) Forest plot showing differences in values between cases and controls across individual studies. **B)** Summary receiver operating characteristic curve showing overall diagnostic performance across included studies – individual studies are labelled by their study number. **C)** Funnel plot providing a visual assessment of publication bias.

Abbreviations: Diff – difference between values from cases and controls; SROC – summary receiver operating characteristic; Value – mean or median value.

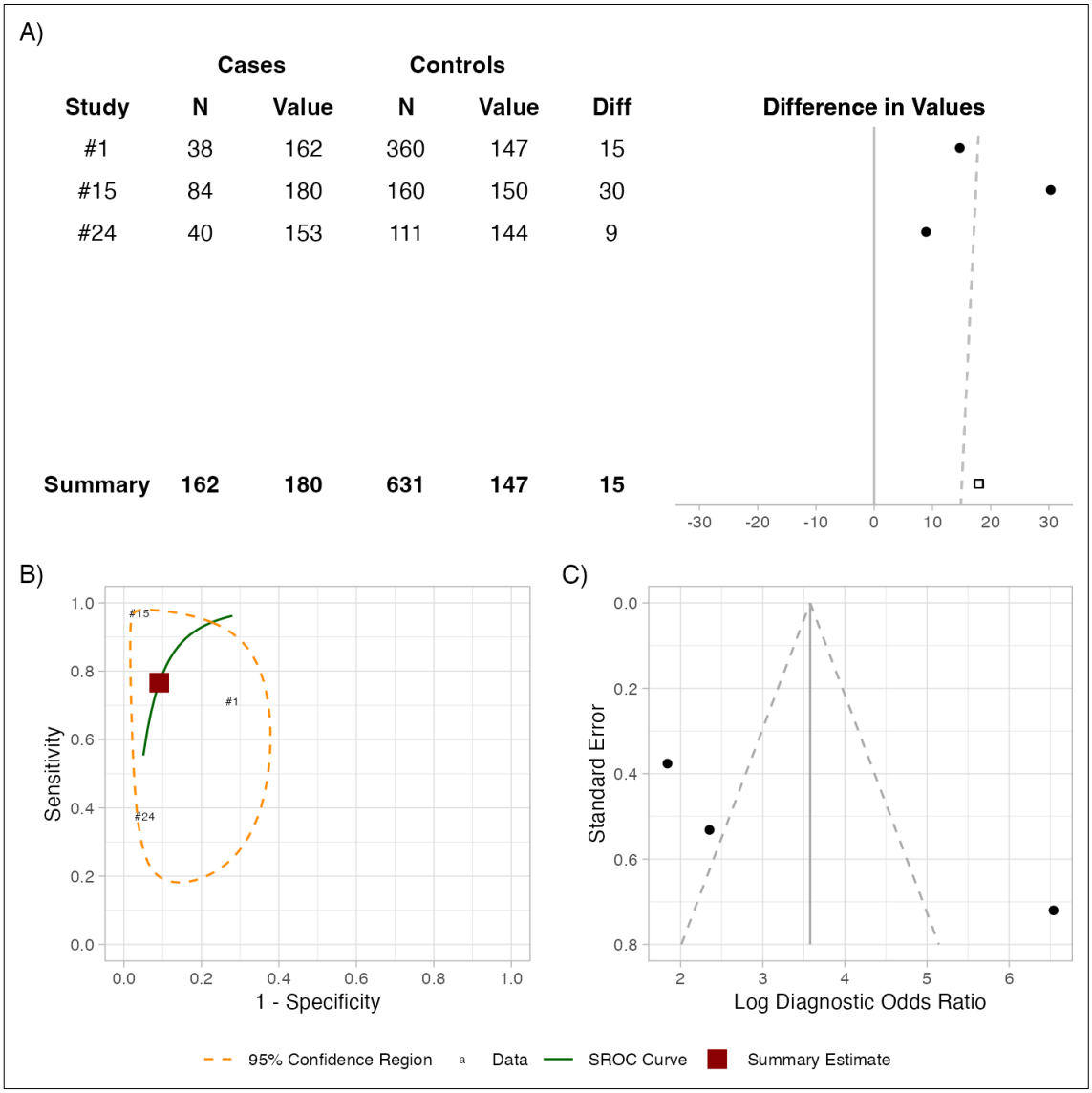


Figure 2.7: Meta-analysis of the diagnostic accuracy of mean neutrophil volume (au) for neonatal sepsis.

A) Forest plot showing differences in values between cases and controls across individual studies. **B)** Summary receiver operating characteristic curve showing overall diagnostic performance across included studies – individual studies are labelled by their study number. **C)** Funnel plot providing a visual assessment of publication bias.

Abbreviations: Diff – difference between values from cases and controls; SROC – summary receiver operating characteristic; Value – mean or median value.

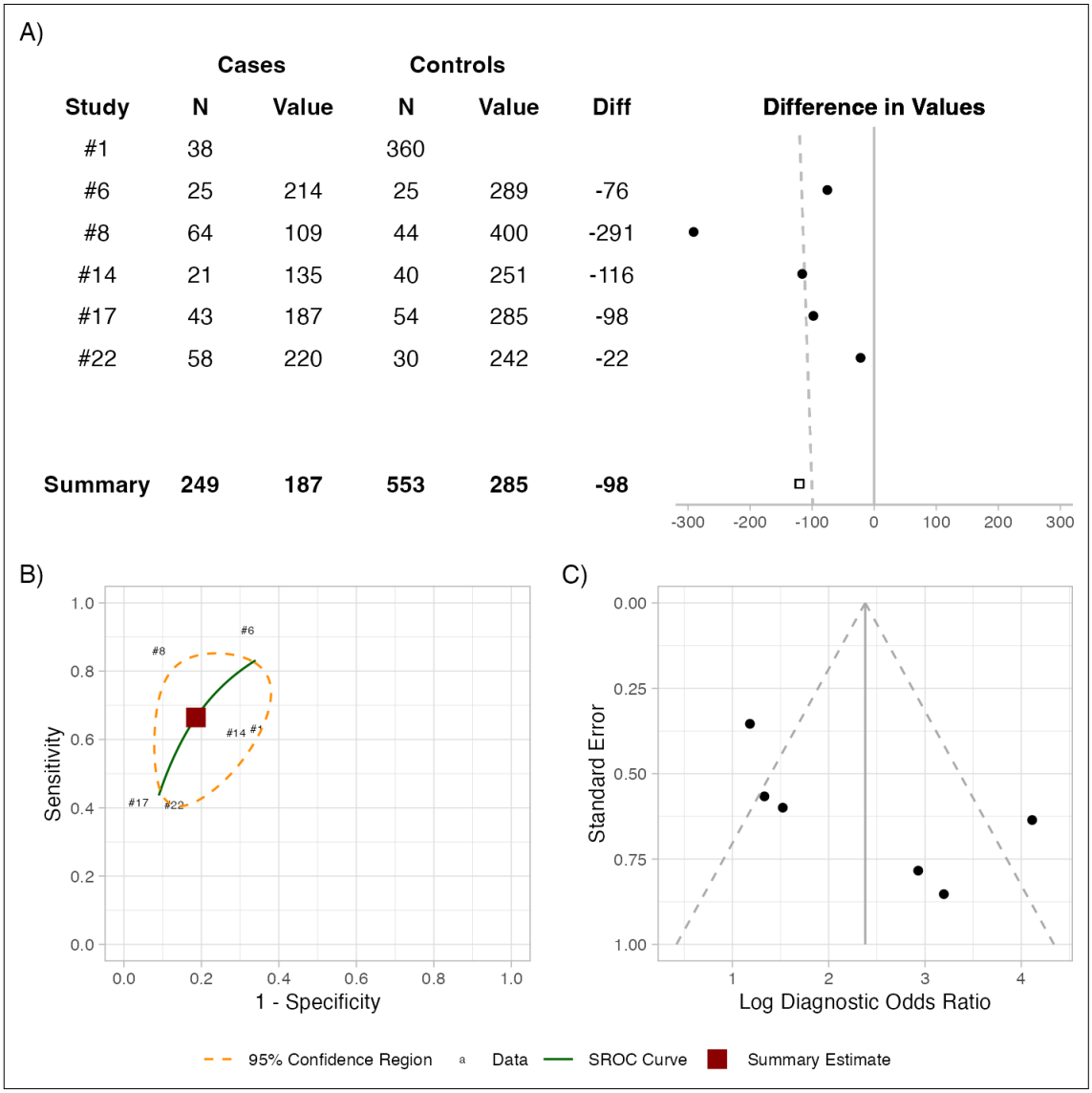


Figure 2.8: Meta-analysis of the diagnostic accuracy of platelet count (x 10^9/L) for neonatal sepsis.

A) Forest plot showing differences in values between cases and controls across individual studies. **B)** Summary receiver operating characteristic curve showing overall diagnostic performance across included studies – individual studies are labelled by their study number. **C)** Funnel plot providing a visual assessment of publication bias.

Abbreviations: Diff – difference between values from cases and controls; SROC – summary receiver operating characteristic; Value – mean or median value.

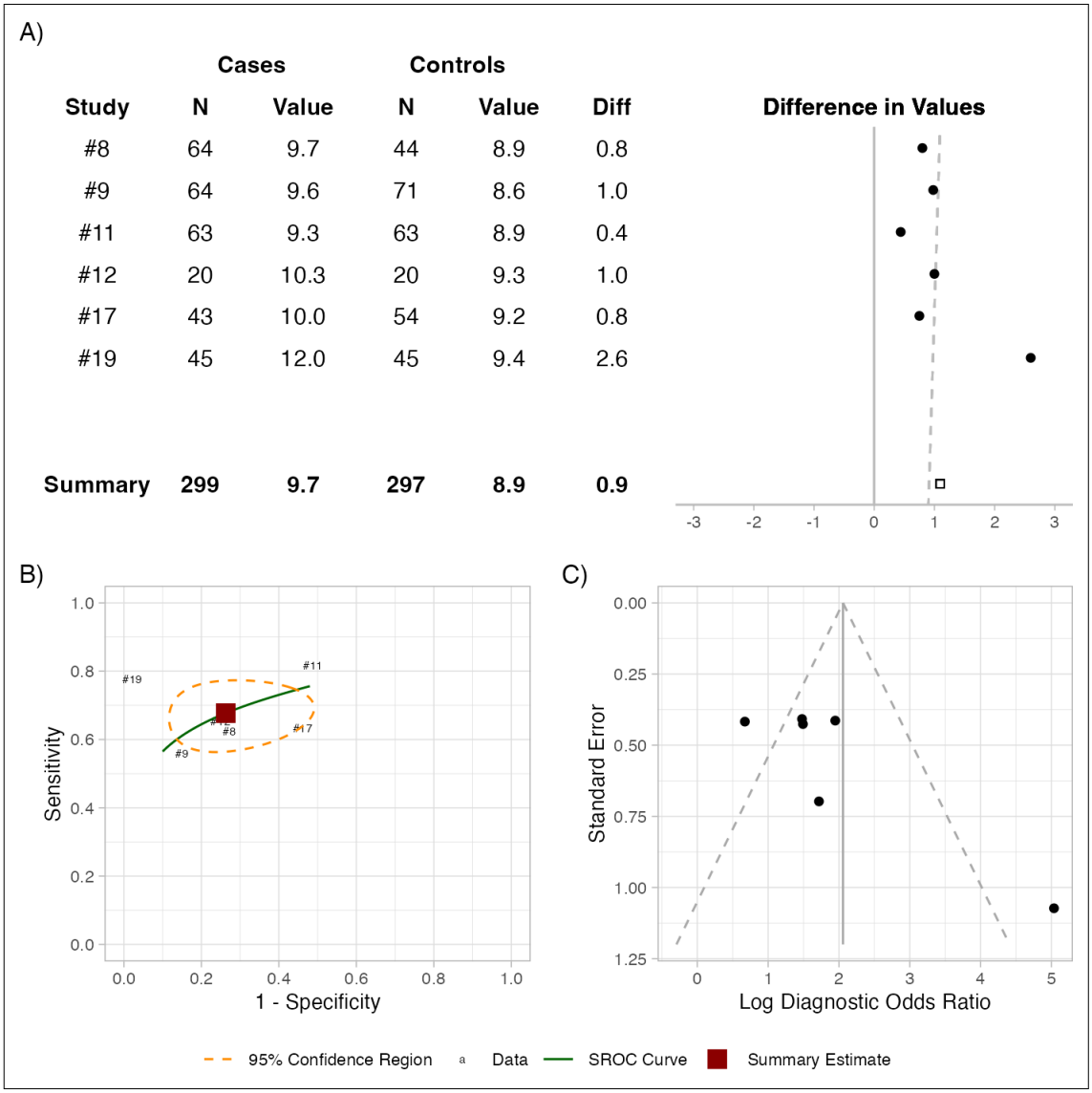


Figure 2.9: Meta-analysis of the diagnostic accuracy of mean platelet volume (fL) for neonatal sepsis.

A) Forest plot showing differences in values between cases and controls across individual studies. **B)** Summary receiver operating characteristic curve showing overall diagnostic performance across included studies – individual studies are labelled by their study number. **C)** Funnel plot providing a visual assessment of publication bias.

Abbreviations: Diff – difference between values from cases and controls; SROC – summary receiver operating characteristic; Value – mean or median value.

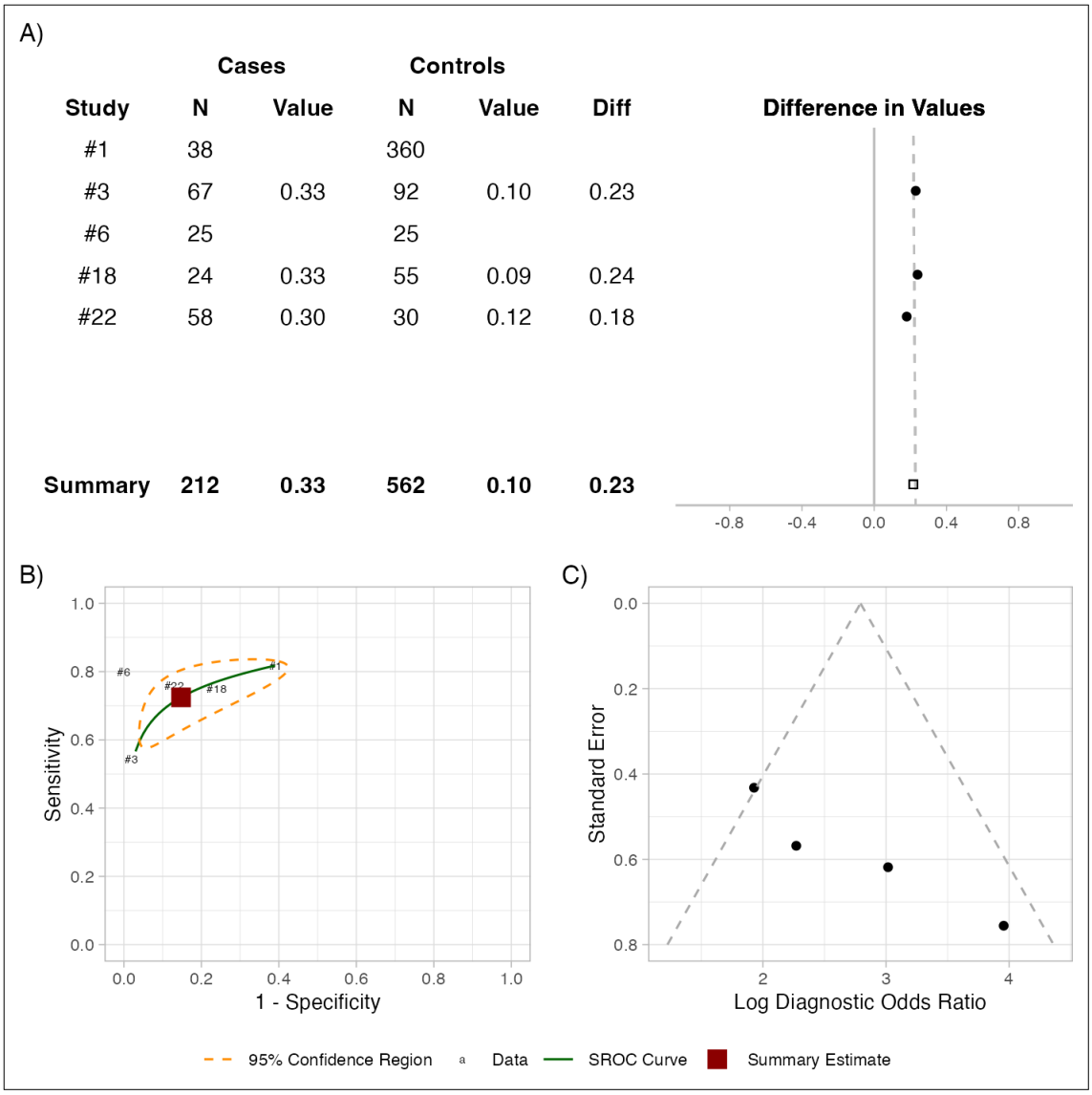


Figure 2.10: Meta-analysis of the diagnostic accuracy of immature-to-total neutrophil ratio for neonatal sepsis – subgroup analysis of studies reporting an optimal threshold of 0.20.

A) Forest plot showing differences in values between cases and controls across individual studies. **B)** Summary receiver operating characteristic curve showing overall diagnostic performance across included studies – individual studies are labelled by their study number. **C)** Funnel plot providing a visual assessment of publication bias.

Abbreviations: Diff – difference between values from cases and controls; SROC – summary receiver operating characteristic; Value – mean or median value.

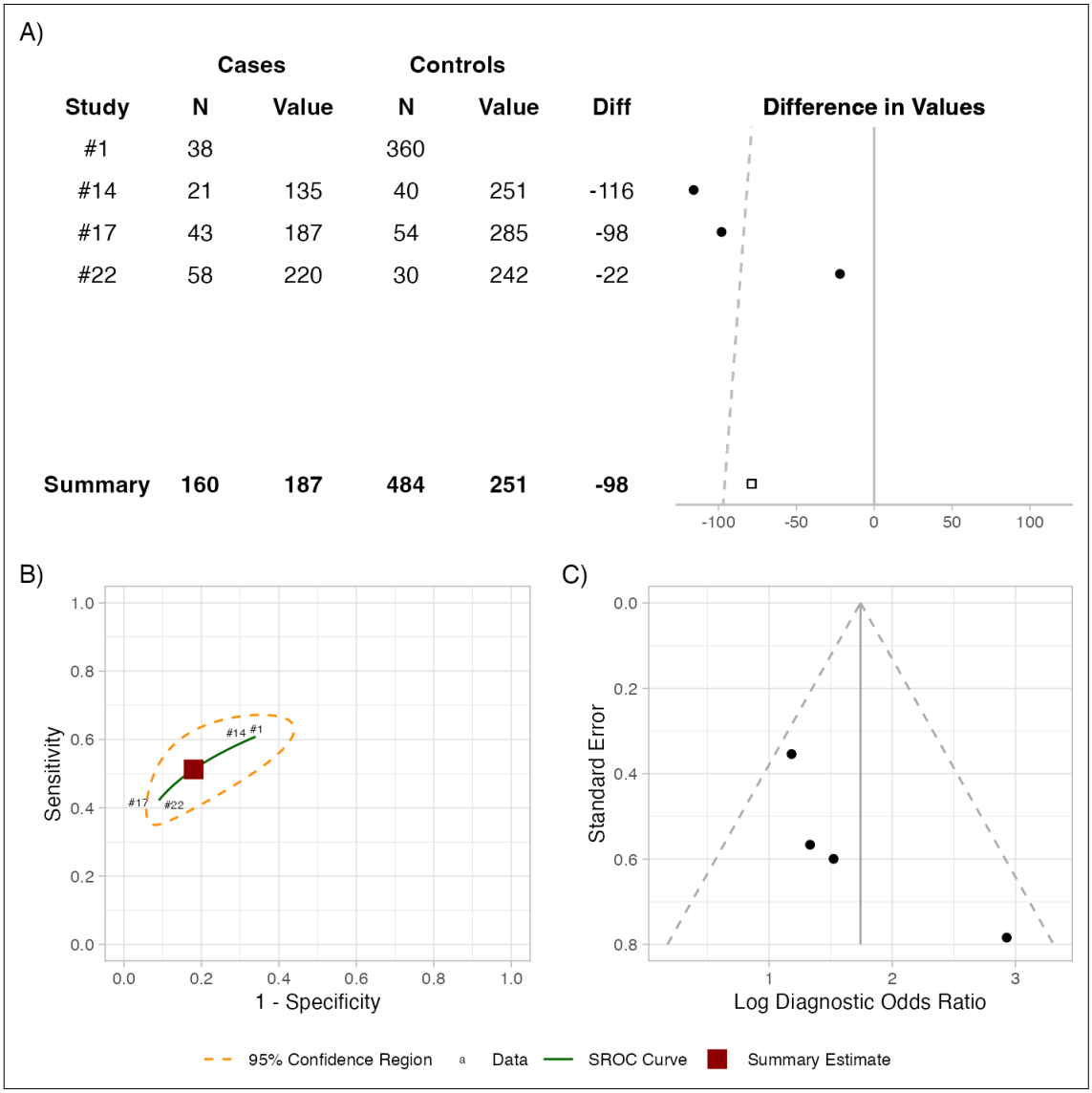


Figure 2.11: Meta-analysis of the diagnostic accuracy of platelet count for neonatal sepsis – subgroup analysis of studies reporting an optimal threshold of 150 x 10^9/L.

A) Forest plot showing differences in values between cases and controls across individual studies. **B)** Summary receiver operating characteristic curve showing overall diagnostic performance across included studies – individual studies are labelled by their study number. **C)** Funnel plot providing a visual assessment of publication bias.

Abbreviations: Diff – difference between values from cases and controls; SROC – summary receiver operating characteristic; Value – mean or median value.

2.4 Discussion

2.4.1 Summary of Findings

This systematic review and meta-analysis investigated the diagnostic utility of the CBC for neonatal sepsis and identified moderate diagnostic accuracy in several parameters. Functional parameters, such as ITR and MNV, demonstrated greater diagnostic accuracy compared to traditional measures, such as WBC# and NEUT#.

ITR emerged as the most reliable parameter, with a threshold of 0.2 achieving high diagnostic accuracy and low inter-study heterogeneity ($I^2 < 0.01$). ITR reflects neutrophil consumption and regeneration, processes prevalent during bacterial infection [112]. No established reference interval exists for neonates, although the average ITR for control groups in this review (0.10; IQR: 0.09 – 0.11) was consistently lower than all reported optimal thresholds. MNV also showed high diagnostic accuracy, but significant heterogeneity among studies ($I^2 = 0.85$) prevented accurate analysis. Differences in study design, diagnostic thresholds, and analytical techniques likely caused this heterogeneity. As a measure of neutrophil activation and a known marker in adult sepsis [73], MNV represents a promising avenue for future research.

WBC# and NEUT# demonstrated low diagnostic accuracy, consistent with the NICE evidence review (Table 2.1) [77]. Septic neonates may present with either elevated or reduced values, complicating the determination of appropriate thresholds [16]. Only two studies in this review investigated both high and low values for these parameters [88, 110]. The omission of lower thresholds in other studies may have resulted in underestimated accuracy metrics. Furthermore, WBC# and NEUT# are influenced by factors unrelated to infection, such as gut colonisation by non-infectious bacteria, which can drive NEUT# above 15 x 10^9/L in the first few days of life [113]. Non-infectious sterile inflammation [72] and corticosteroids [114], which are commonly administered to preterm neonates to treat respiratory distress [115], can also confound these measures.

PLT# and MPV showed moderate diagnostic accuracy for neonatal sepsis in this review. Thrombocytopenia is frequently observed in neonates older than three days with sepsis or NEC [116]. Conversely, early-onset thrombocytopenia commonly reflects maternal conditions such as gestational hypertension, gestational diabetes, or intrauterine growth restriction [116]. MPV, a marker of platelet activation, demonstrated moderate diagnostic potential, though significant inter-study heterogeneity hindered reliable analysis (I^2: 0.39).

Other CBC parameters, such as LYMPH# and MONO#, were not evaluated due to insufficient detail from included studies. This represents an area for future research, as lymphopenia and monocytosis are known features of adult sepsis [72]. Similarly, NLR could not be evaluated due to high inter-study variability, though emerging evidence suggests it may provide greater utility than raw counts alone [72].

2.4.2 Comparison with Existing Literature

The NICE evidence review of the diagnostic accuracy of four CBC parameters for neonatal sepsis concluded that there was "very low" evidence to support including the CBC in the guidelines (Table 2.1) [11, 77]. However, the present meta-analysis found that ITR demonstrated high diagnostic accuracy, with an AUC exceeding 0.7, and highlighted additional CBC parameters that were not evaluated in the NICE evidence review but showed strong diagnostic potential in this meta-analysis, including MNV and MPV.

The ideal diagnostic tool for neonatal sepsis would provide a sensitivity approaching 100% and a specificity of at least 85% [18, 19]. Unfortunately, no parameter exhibited this level of diagnostic precision within the scope of the current meta-analysis. Nonetheless, a few studies reported diagnostic accuracy metrics that align with these criteria. For instance, Nesargi et al. (2020) reported a sensitivity of 97% and a specificity of 96% for MNV [96]. Similarly, Arcagok and Karabulut (2019) reported a sensitivity of 91.3% and a specificity of 97.6% for PLR [101]. These findings further support the conclusion that novel immune activation parameters may provide superior utility compared to conventional counts.

Several reviews have previously evaluated CBC parameters for the diagnosis of neonatal sepsis. For example, an ITR greater than 0.2 has been shown to achieve 79% sensitivity and 91% specificity [117], results slightly outperforming those presented in this systematic review. Similarly, NLR has been investigated in two reviews and found to demonstrate high diagnostic accuracy [118, 119]. However, these reviews did not specify whether the gold standard of blood culture was used for case definitions, which could potentially limit the robustness of their findings. MPV has also been previously assessed, but those reviews included studies using clinical features rather than blood culture for diagnosing sepsis, limiting the utility of their findings [120].

Another avenue for improving the utility of the CBC is its integration into more complex diagnostic models, which has already been achieved in adults with the development of ICIS [73]. In the database search for the current review, four studies were identified that explored similar approaches in neonates [121–124]. Three employed scoring systems that evaluated abnormalities in parameters such as WBC#, PLT#, or ITR [121–123]. These models often lacked direct comparability due to differing methodologies, although they all showed promising results.

Additionally, two studies evaluated clinical signs in conjunction with CBC parameters for diagnosing sepsis, one of which combined CBC results and clinical features to train multiple complex algorithms [124]. The random forest model achieved a sensitivity of 38% and a specificity of 88%, suggesting its utility as a rule-in tool. Again, this study included both culture-proven and clinically diagnosed sepsis cases to train its models, which may confound the findings and limit their applicability. Despite these limitations, these studies demonstrate the potential for including the CBC in more sophisticated models, potentially enhancing sepsis diagnostics.

2.4.3 Implications of Findings

The CBC is a rapid, inexpensive test that provides clinically valuable information, such as HGB for anaemia monitoring [125]. It requires minimal blood volumes, typically as little as 20 µL, making it ideal for use in neonates where repeated sampling can induce

anaemia. Additionally, most analysers automatically calculate sophisticated activation parameters, such as ITR and MNV, without the need for additional sampling.

On their own, no single parameter provided sufficient diagnostic accuracy to detect all cases of neonatal sepsis. However, this meta-analysis provides a case for including the CBC in conjunction with current diagnostic methods. The evidence presented in this review emphasises the need for a re-evaluation of the current guidelines. Future research should aim to standardise the use of these parameters, explore their role in combined diagnostic algorithms, and validate their performance in prospective clinical settings.

2.4.4 Strengths and Limitations

Decision-making in the diagnosis and evaluation of neonatal sepsis is a critical challenge for neonatologists. A strong evidence base is essential, given the vulnerability of this patient cohort due to their immature immune systems. One of the key challenges is a lack of a standard definition for neonatal sepsis, with many studies relying solely on clinical signs. However, signs such as tachycardia, tachypnoea, and jaundice are non-specific and commonly present in healthy neonates [18, 19]. To improve the accuracy of the findings, this review excluded all studies that did not include a positive culture as part of the sepsis diagnosis, which accounted for 42 of the excluded studies.

The breadth of definitions for control groups in the literature also posed a significant challenge. Thirty-two studies were excluded because their control groups were not adequately defined, often including neonates with clinical signs of sepsis but negative blood cultures. This highlights the overlap between case and control definitions in neonatal sepsis research. By excluding these studies, this systematic review provides the most accurate picture of the CBCs diagnostic utility for neonatal sepsis to date. However, even within the included studies, it is likely that control groups contained culture-negative septic neonates without clinical signs, and the septic groups may have included non-septic neonates with blood culture contaminants, such as CoNS. The lack of a standardised definition for neonatal sepsis continues to weaken the literature on this subject.

A notable limitation of this review is the insufficient number of relevant studies available for comprehensive analysis. This constraint hindered the assessment of specific parameters, such as NLR and DNI. Additionally, it precluded the evaluation of differences in diagnostic accuracy between EOS and LOS cases, despite their differing aetiologies and pathogenesis. Moreover, the limited number of studies for each evaluated parameter weakened the overall robustness of the findings. There is a clear need for further research on all the CBC parameters evaluated in this systematic review to confirm these findings.

2.5 Chapter Summary

This systematic review and meta-analysis evaluated the diagnostic utility of several CBC parameters for neonatal sepsis, with a specific focus on the functional parameters of the CBC. These parameters, which have not been previously evaluated for inclusion in diagnostic guidelines, provided moderate diagnostic accuracy. While these findings are not sufficient to suggest that the CBC replace current diagnostic techniques, they highlight the need for a re-evaluation of the entirety of the CBC for inclusion in these guidelines.

This review also highlighted issues within the literature regarding the definition of neonatal sepsis and the selection of control cohorts for neonatal research. Building on these findings, the next chapter will focus on establishing a strict definition for a control cohort of neonates admitted to the NICU and evaluating a baseline profile for the CBC in this cohort, which is a critical step before advancing to the development and validation of a diagnostic algorithm for neonatal sepsis.

Chapter 3

Establishing a Baseline Profile for the Neonatal Complete Blood Count in a Cohort of Neonates Admitted to a Neonatal Intensive Care Unit

---

3.1 Introduction

The results presented in this chapter have been published by the *International Journal of Laboratory Hematology*, a peer-reviewed academic journal [126].

3.1.1 The Challenges of Establishing Baseline Neonatal Profiles

Accurate reference ranges for neonatal CBC parameters are essential for the detection and monitoring of haematological abnormalities in newborns. However, defining these ranges poses significant challenges. Unlike adult reference values, neonatal samples are rarely obtained from healthy newborns. Instead, reference ranges are typically estimated retrospectively using data from hospitalised but clinically stable neonates. This approach is hindered by the absence of a universally accepted definition of neonatal clinical stability, as discussed in the previous chapter [19]. Consequently, variations in study inclusion criteria contribute to inconsistencies in reported reference ranges.

Neonatal haematological parameters fluctuate significantly in the early neonatal period, further complicating the establishment of reference ranges. For example, NEUT# values surge within the first 72 hours of life, reaching up to 14.4 x 10^9/L in response to bacterial colonisation of the gut [127, 128]. NRBC# and RET# also exhibit substantial variability at birth as adult haemoglobin replaces fetal haemoglobin, with reported NRBC# values ranging from 0 – 3 x 10^9/L before stabilising below 1 x 10^9/L in older neonates [129, 130].

In contrast, LYMPH# values are lower at birth but progressively increase over the first two months, though the underlying mechanisms for this trend remain unclear [131].

GA also influences the neonatal haematological profile. Preterm neonates exhibit considerably higher NRBC# at birth (up to 8 x 10^9/L) than term neonates (up to 3 x 10^9/L) and maintain higher levels for the first week [129]. In contrast, values in term neonates stabilise within four days. These observations highlight the intricate relationship between GA and postnatal age in influencing neonatal CBC values.

Sex-based differences introduce additional variability, though they remain underexplored in neonatal populations. Female neonates exhibit higher NEUT# than males, with differences of approximately 2 x 10^9/L [132]. In adults, sex-based differences are well-documented across multiple CBC parameters, including RBC#, HGB, and HCT [62]. Understanding these differences in neonates is essential for refining reference ranges and ensuring accurate clinical interpretation.

In addition to physiological variability, external factors such as corticosteroid exposure further complicate the interpretation of neonatal CBC values. Corticosteroids are frequently administered to preterm neonates to treat respiratory distress or facilitate extubation [115, 133]. While the haematological effects of corticosteroids in neonates remain unclear, studies in adults demonstrate dose-dependent increases in WBC#, NEUT#, and MONO# within 24 hours of administration, thereby mimicking haematological changes associated with bacterial infections [114, 134] and raising concerns about potential misinterpretation in neonatal care.

Characterising the neonatal CBC profile and establishing accurate reference ranges requires a thorough assessment of both independent and combined effects of factors such as corticosteroid exposure, sex, postnatal age, and GA. Addressing these complexities will enable the development of baseline comparators to assist in identifying neonates with abnormal haematological profiles. This, in turn, will facilitate more informed interpretation of the neonatal CBC and enhance clinical decision-making in neonatal care.

3.1.2 Established Reference Ranges for the Neonatal CBC

Despite these challenges, several clinical textbooks provide reference ranges for some neonatal CBC parameters. For instance, reference values from Rennie and Robertson's Textbook of Neonatology (2012) [135] and the Oxford Handbook of Neonatology (2017) [136] are summarised in Table 3.1. A comparison of these references highlights considerable variability in the reported values for several key parameters, such as WBC#, NEUT#, and RET#.

These discrepancies have significant clinical implications. Variability in reference ranges can lead to differing interpretations of the same haematological results, potentially affecting diagnoses and treatment decisions. For example, a neonate's CBC result that falls within the reference range of one source might be flagged as abnormal by another. Such inconsistencies underscore the need for a comprehensive analysis of reference ranges in this population.

3.1.3 Aims of the Chapter

This chapter aims to characterise the baseline profile and establish relevant reference ranges for all neonatal CBC parameters using data from neonates admitted to the NICU at the John Radcliffe Hospital (Oxford, UK). I applied strict inclusion criteria to select CBC results from clinically stable neonates, providing a reliable foundation for analysis.

In addition to defining reference ranges, this chapter examines the independent and interactive effects of corticosteroid exposure, sex, postnatal age, and GA on CBC parameters. I used multivariable analyses of variance (ANOVAs) to assess these factors, providing insight into how they shape the neonatal haematological profile.

Table 3.1: Published references values for parameters of the neonatal complete blood count. **A)** Adapted from Roberts and Murray [135][a]. **B)** Adapted from Oski et al. (in Roberts and Murray [135]). **C)** Adapted from Fox et al. [136].

	A) Term Neonates			**B)** Preterm Neonates at Birth				**C)** All Neonates		
Parameter	Birth	Two Weeks	Two Months	24 – 25 Weeks	26 – 27 Weeks	28 – 29 Weeks	30 – 31 Weeks	Cut-Off	Before Cut-Off	After Cut-Off
HGB (g/L)	149 – 237	134 – 198	94 – 130	194 $\pm$ 15	190 $\pm$ 25	193 $\pm$ 18	191 $\pm$ 21	Preterm vs Term	110 – 180	140 – 210
HCT (%)	0.47 – 0.75	0.41 – 0.65	0.28 – 0.42	0.63 $\pm$ 0.04	0.62 $\pm$ 0.08	0.60 $\pm$ 0.07	0.60 $\pm$ 0.08	Preterm vs Term	0.30 – 0.50	0.40 – 0.65
RET# (x 10^{12}/L)	0.110 – 0.450	0.010 – 0.085	0.035 – 0.200	0.279 $\pm$ 0.023	0.454 $\pm$ 0.015	0.347 $\pm$ 0.012	0.278 $\pm$ 0.010	4-6 Weeks	$<$ 0.005	0.005 – 0.250
NRBC# (x 10^{9}/L)	$<$ 5.00	$<$ 0.10	$<$ 0.10	–	–	–	–	–	–	–
WBC# (x 10^{9}/L)	10 – 26	6 – 21	5 – 15	–	–	–	–	48 Hours	20.0 – 40.0	5.0 – 15.0
NEUT# (x 10^{9}/L)	2.7 – 14.4	1.5 – 5.4	0.7 – 4.8	–	–	–	–	48 Hours	6.0 – 26.0	1.8 – 10.0
LYMPH# (x 10^{9}/L)	2.0 – 7.3	2.8 – 9.1	3.3 – 10.3	–	–	–	–	One Week	2.0 – 6.0	2.0 – 15.0
MONO# (x 10^{9}/L)	0.0 – 1.9	0.1 – 1.7	0.4 – 1.2	–	–	–	–	–	–	–
PLT# (x 10^{9}/L)	150 – 450	150 – 450	150 – 450	150 – 450	150 – 450	150 – 450	150 – 450	All	150 – 400	150 – 400

[a] *Data obtained from multiple sources, selected for their relevance to interpreting haematological results. Abbreviations: HCT – haematocrit, HGB – haemoglobin content, LYMPH# – lymphocyte count, MONO# – monocyte count, NEUT# – neutrophil count, NRBC# – nucleated red blood cell count, PLT# – platelet count, RET# – reticulocyte count, WBC# – white blood cell count.*

3.2 Methods

3.2.1 Ethical Considerations

The Oxford Research Ethics Committee provided ethical approval for this research (REC 08/H0606/139). An opt-out consent process was used, whereby parents were informed about the study through patient information leaflets. Consent was presumed unless explicitly declined, and parents were assured that opting out at any stage would not impact their child's medical care. All data were anonymised before analysis to ensure confidentiality.

3.2.2 Data Collection

This retrospective study was conducted within the NICU at the John Radcliffe Hospital (Oxford, UK) between 13$^{\text{th}}$ January 2022 and 13$^{\text{th}}$ January 2023. Neonatal CBC results are routinely measured on a Sysmex XN-2000 (Kobe, Japan) analyser in the haematology laboratory as part of the clinical workflow. Before the study began, an XN-450 point-of-care analyser was installed in the NICU to allow for more routine CBC measurements, which I validated against the XN-2000 analyser (Appendix A2, p. 248).

During the study period, raw data files for all CBC results measured on both analysers were collected. Using Sysmex software, the output of all relevant CBC results was extracted – this included all reported and unreported parameters. If multiple CBC results were measured on the same neonate in a single day, daily means were calculated to minimise the influence of repeated sampling. Matched clinical data were extracted for the duration of each neonate's stay in the NICU.

Neonatal BadgerNet Records

Admission and discharge dates, gestational and postnatal ages, and reasons for admission were collected from BadgerNet, alongside daily summary data on:

- Administration of antibiotics, inotropes, and corticosteroids.
- Diagnoses and management issues (e.g. central line placement).
- Surgical procedures.

Cerner PowerChart EPR

Relevant clinical data were extracted from the Electronic Patient Records (EPR), including:

- Maternal details: age at delivery, gravidity, and parity.
- Delivery details: birth date, GA, maternal and neonatal ICD-10 codes, delivery onset and method, antibiotics and corticosteroids administered, and umbilical cord blood gas results.
- Neonatal details: sex, ethnicity, longitudinal biochemistry results (e.g. CRP and lactate), microbiology results (blood, respiratory, and CSF), and viral polymerase chain reaction (PCR) results. Where multiple biochemistry results were available on the same day, the mean values were calculated.

Clinical Observation Charts

Daily clinical observation charts provided hourly data (Figure A2.2, p. 253), including:

- Ventilation methods (e.g. high flow therapy (HFT) or continuous positive airway pressure (CPAP)).
- Ventilator oxygen, HFT/CPAP flow, and mean airway pressure values.
- Overall ventilator trends (improving, steady, or worsening).
- Skin temperature, heart rate, and respiratory rate measurements.

Categorisation of CBC Results

To minimise potential confounding, CBC results obtained within two days of a surgical intervention (listed in Appendix A2.3, p. 254) or blood transfusions were excluded. Records with incomplete documentation, such as missing clinical records or paper charts, were also excluded from the analysis. Excluding records with incomplete documentation ensured the accuracy and reliability of the dataset.

Remaining CBC results were categorised into clinical groups based on predefined laboratory and clinical criteria (Figure 3.1). Results obtained within two days before or after a positive blood, respiratory, or CSF culture were classified as *probable sepsis* (case group). CBC results measured during periods with no clinical suspicion of infection (e.g. low CRP values, no cultures requested, no antibiotics administered, and no clinical signs

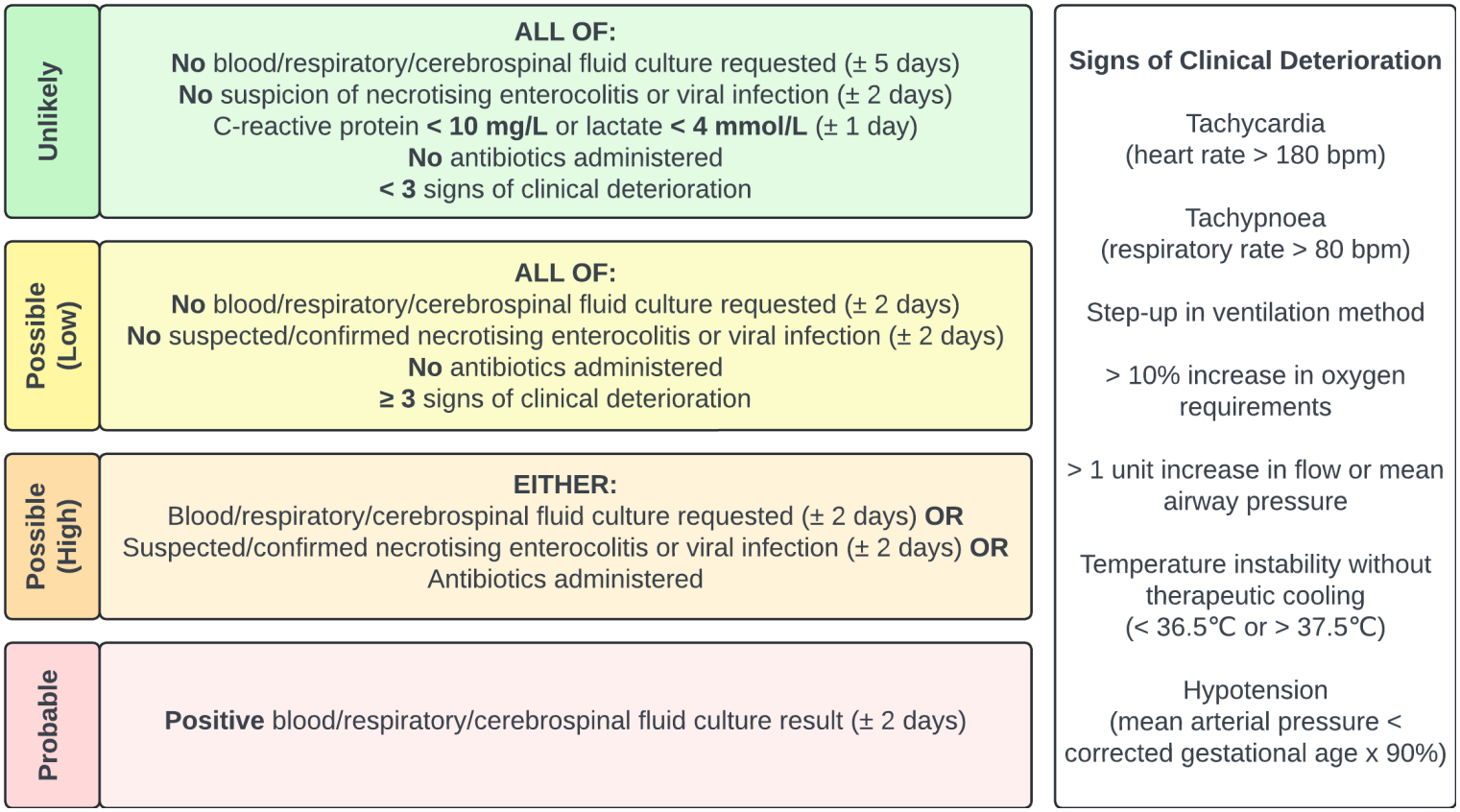


Figure 3.1: Criteria for the categorisation of complete blood count results for late-onset sepsis, based on clinical and laboratory signs of sepsis.

of deterioration) were classified as *unlikely sepsis* (control group). Results not meeting either criterion were assigned to one of two possible sepsis groups (*possible low* and *possible high*) depending on the level of clinical suspicion.

3.2.3 Data Analysis

All statistical analyses were conducted using RStudio (R version 4.3.0) [86]. This analysis focused exclusively on CBC results categorised as *unlikely sepsis*, representing clinically stable neonates without suspected infection. CBC results taken within the first three days of life were excluded due to the unknown confounding effects of maternal or perinatal factors.

Determination of Value Distributions

Distributions of individual CBC parameters were assessed using the Shapiro-Wilk test. Parameters with non-parametric distributions ($p < 0.05$) were transformed using an ordered quantile transformation, which successfully normalised most parameters. Once transformed, values were re-tested to verify normality assumptions for subsequent analyses.

Effects of Corticosteroid Exposure

To evaluate the impact of corticosteroid exposure, a multivariable ANOVA model was applied using the following formula. Significant corticosteroid effects ($p < 0.05$) were reported as mean or median differences, depending on the distribution of each parameter. CBC results affected by corticosteroid exposure were excluded from subsequent analyses to reduce bias.

$$value \sim corticosteroid\ exposure \times (sex + postnatal\ age + GA)$$

Calculation of Reference Ranges

Reference ranges were derived following the Clinical and Laboratory Standards Institute (CLSI) guidelines (EP28 A3C; workflow summarised in Figure 3.2) [137]. For normally distributed parameters, reference ranges were defined as the mean $\pm$ 1.96 times the SD, with 95% confidence intervals (CIs) for the bounds calculated as 1.96 times the SD divided by the square root of the sample size. For non-normally distributed parameters,

reference ranges were defined by the 2.5^{th} and 97.5^{th} percentiles, and CIs were computed using non-parametric bootstrap methods.

Interactive Effects of Sex, Postnatal Age, and Gestational Age at Birth

The interactive effects of sex, postnatal age, and GA were assessed using a multivariable ANOVA with the following model. When a significant effect of sex was identified ($p < 0.05$), sex-specific deviations were calculated relative to the overall mean or median, based on the parameter's distribution.

$$value \sim sex \times postnatal\ age \times GA$$

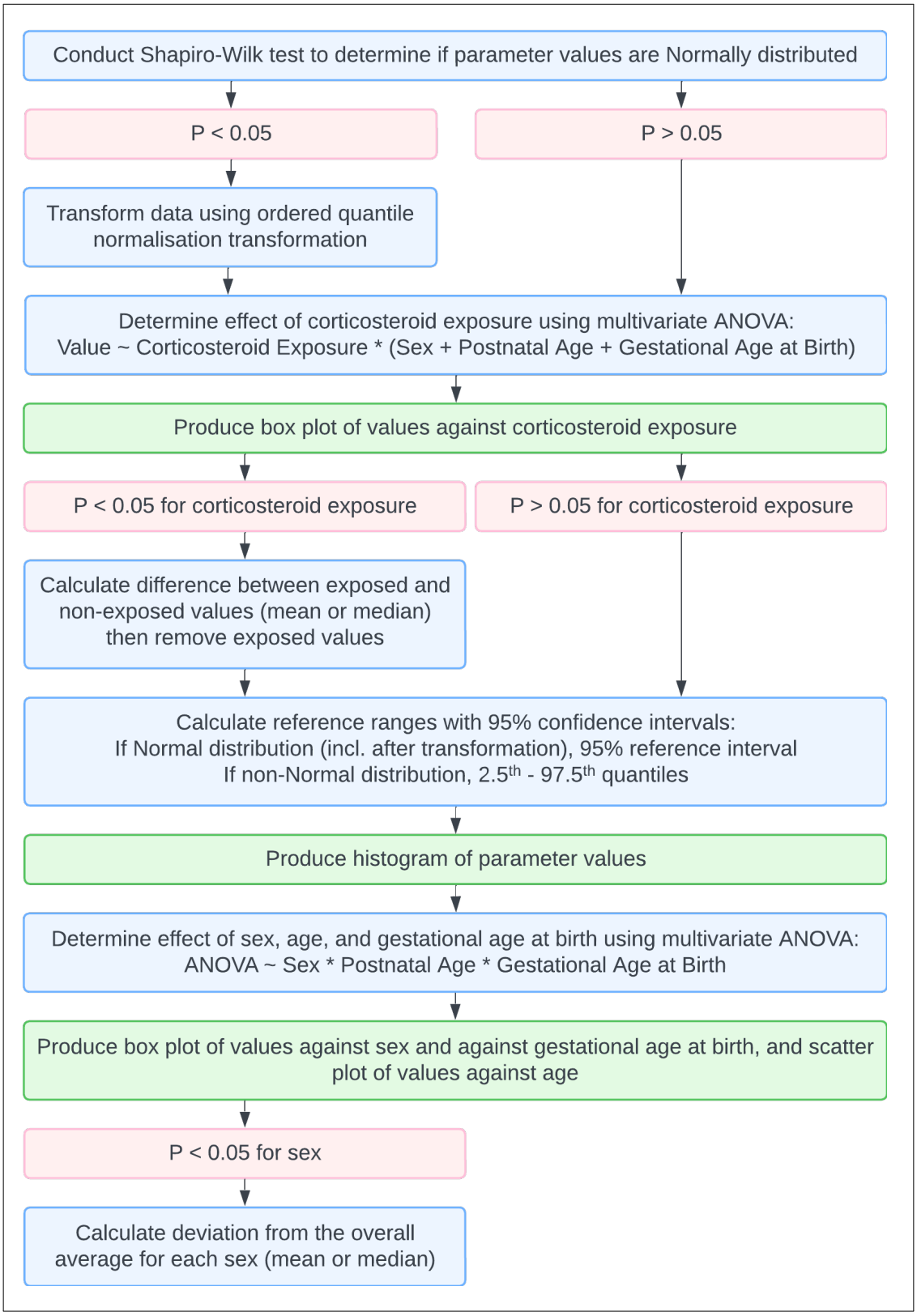


Figure 3.2: Flowchart illustrating the workflow for deriving reference ranges and analysing the effects of different factors on the neonatal complete blood count.

3.3 Results

This section summarises the results from the analysis of 49 CBC parameters. All supplementary figures are available online at `https://4fsqon-emily-hyde.shinyapps.io/ThesisFigures/` under the tab *Neonatal CBC Baseline Profile* [138]. Representative figures for the analysis of WBC# (Figure 3.3) and HGB (Figure 3.4) are shown below, as these parameters are particularly relevant in clinical settings for assessing infection and anaemia.

3.3.1 Patient Characteristics

During the study period, 3,490 CBC results were collected from 587 neonates admitted to the NICU. Of these, 332 results from 159 neonates met the inclusion criteria for clinically stable, non-septic patients. Statistically significant differences were observed between the postnatal ages of male and female neonates (see Table 3.2 for full summary statistics). In addition, postnatal age varied significantly with GA, corticosteroid exposure, admission indications, and discharge outcomes. These findings highlight the presence of clinically distinct subgroups within the cohort.

3.3.2 Determination of Value Distributions

Of the 49 CBC parameters evaluated, 47 required transforming using an ordered quantile transformation before analysis. Only IRF and P-LCR were normally distributed without transformation. The transformed values were reassessed using the Shapiro-Wilk test and, in most cases, achieved normality. However, seven parameters, including NRBC#, BASO#, and AS-LYMPH, could not be normalised due to a high proportion of zero values. These were subsequently analysed using non-parametric methods.

3.3.3 Effects of Corticosteroid Exposure

Corticosteroid exposure significantly affected 29 CBC parameters (Table 3.3). Notably, there was a significant decrease in nine RBC-related parameters, including NRBC# and RET#. In contrast, both RET-HE and DELTA-HE were significantly increased. WBC-

Table 3.2: Sample characteristics of complete blood count results used for the evaluation of the baseline profile in clinically stable neonates admitted to the Neonatal Intensive Care Unit.

		All	Females	Males	Sex P-Value	Age (Days)	Age P-Value
CBC Results		332	137	195	0.002[a]	–	–
Neonates		159	66	93	0.03[a]	–	–
Postnatal Age (Days)		25 (10 – 51)	22 (9 – 43)	29 (11 – 54)	0.03[b]	–	–
GA (weeks)		29.7 (26.0 – 34.7)	30.0 (26.7 – 36.6)	29.1 (26.4 – 34.1)	0.30[b]	–	–
GA Categories	Term ($\geq$ 37 weeks)	29	15	14	0.55[a]	9 (7 – 13)	$<$ 0.001[c]
	Moderate (32 – 36 weeks)	32	14	18		11 (7 – 15)	
	Very (28 – 32 weeks)	32	11	21		24 (11 – 33)	
	Extreme ($<$ 28 weeks)	66	26	40		38 (23 – 65)	
Corticosteroid CBC Results		56	20	36	0.44[a]	44 (29 – 69)	$<$ 0.001[c]
Admission	Prematurity	100	40	60	0.93[a]	30 (12 – 52)	0.08[c]
	Acute Condition	29	13	16		10 (7 – 14)	
	Chronic Condition	18	8	10		16 (11 – 35)	
	Planned Surgery	10	3	7		63 (14 – 101)	
Discharge	Home	155	62	93	0.64[a]	25 (10 – 49)	0.002[c]
	Transferred	7	3	4		12 (10 – 50)	
	Died	3	2	1		66 (50 – 77)	

Values are presented as counts or median (interquartile range). [a] *Chi-squared test.* [b] *Wilcoxon rank sum test.* [c] *Kruskal-Wallis rank sum test. Significant p-values ($<$ 0.05) are highlighted in red. Abbreviations: CBC – complete blood count; GA – gestational age at birth.*

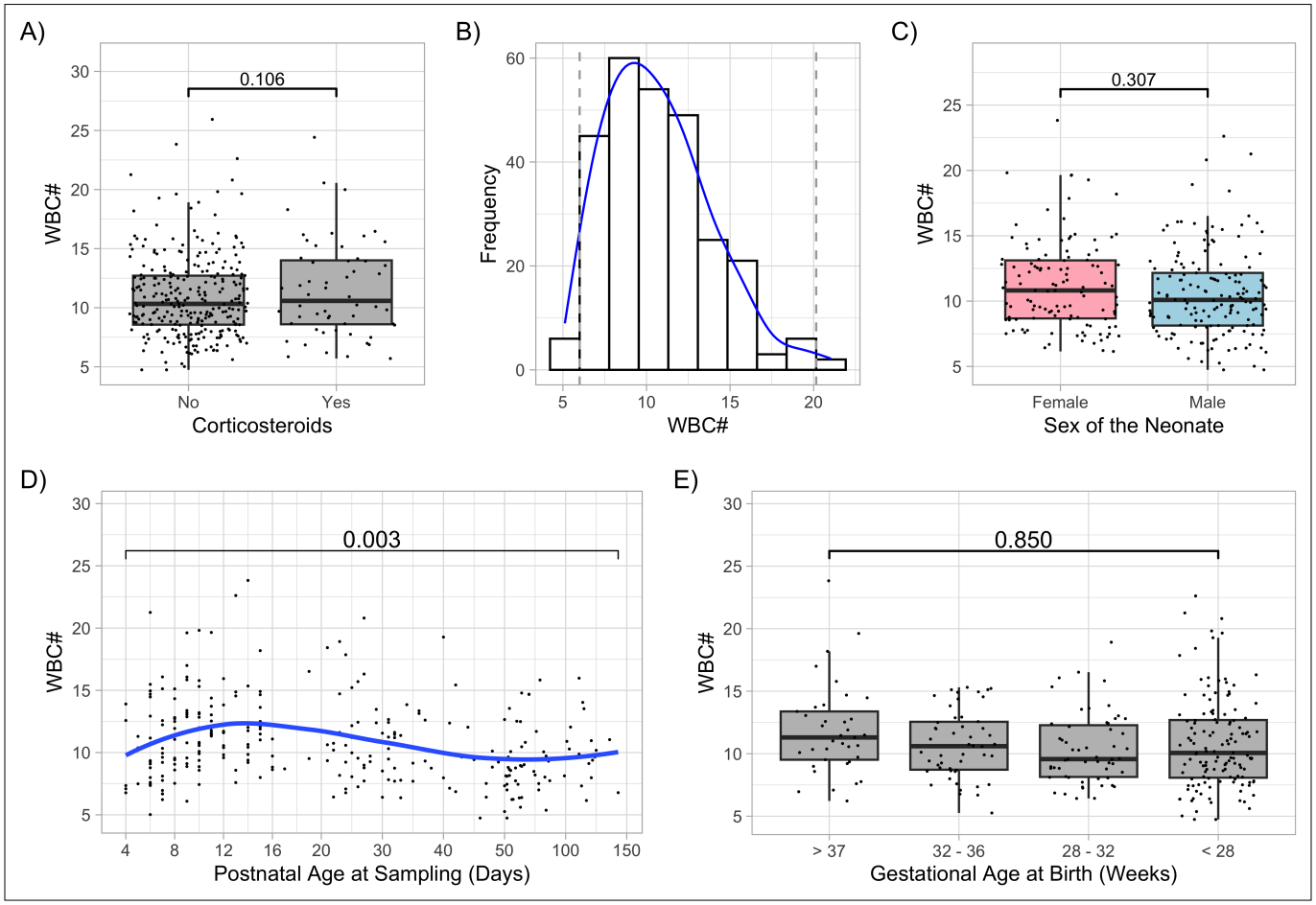


Figure 3.3: Analysis of white blood cell count (WBC#; x 10^9/L) in clinically stable neonates admitted to a Neonatal Intensive Care Unit. **A)** Box plot by corticosteroid exposure **B)** Histogram with reference ranges (grey dotted lines) and a distribution curve (blue line). **C)** Box plot by sex. **D)** Scatter plot of postnatal age with a fitted curve (blue line). **E)** Box plot by gestational age at birth.

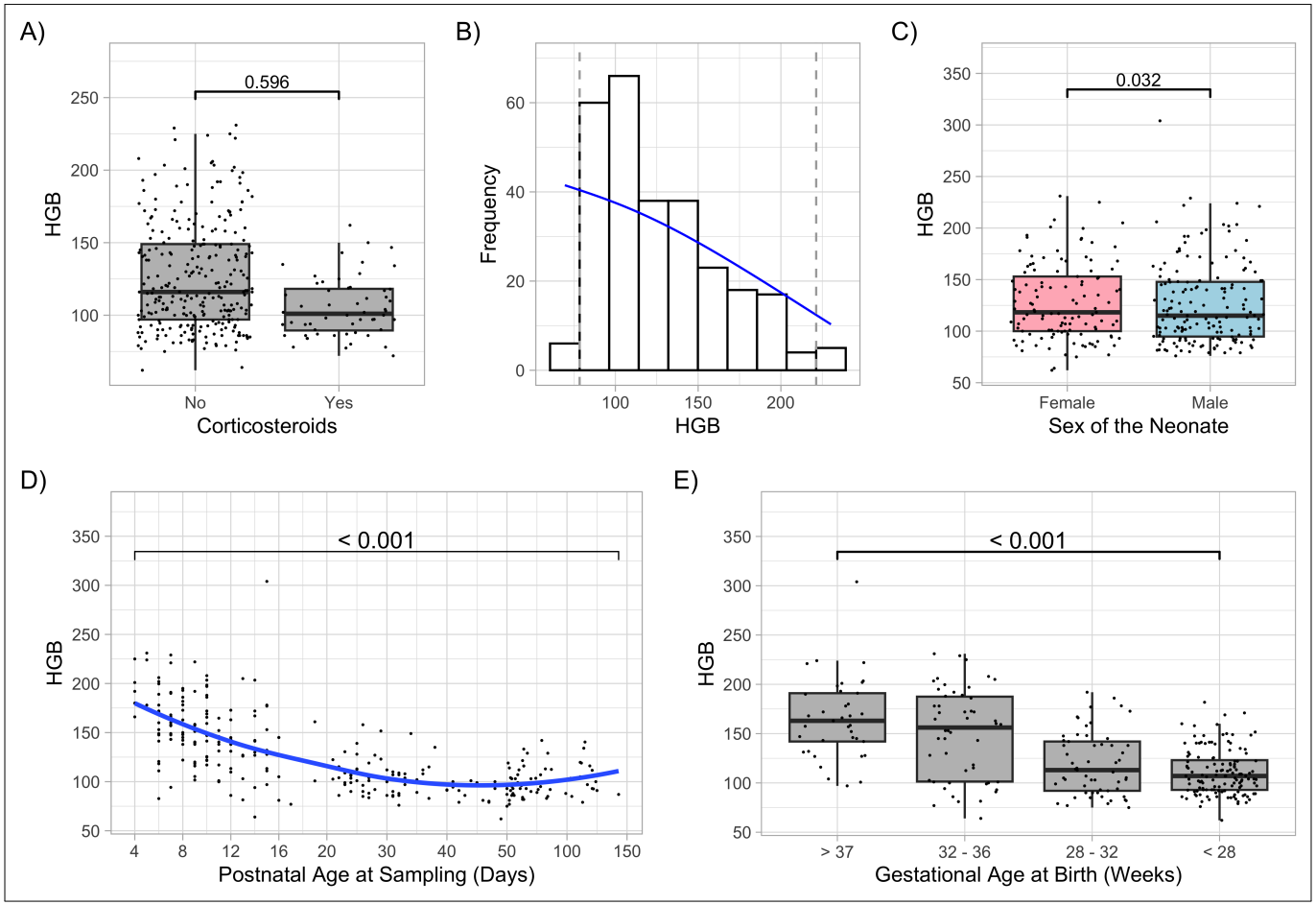


Figure 3.4: Analysis of haemoglobin content (HGB; g/L) in clinically stable neonates admitted to a Neonatal Intensive Care Unit.

A) Box plot by corticosteroid exposure **B)** Histogram with reference ranges (grey dotted lines) and a distribution curve (blue line). **C)** Box plot by sex. **D)** Scatter plot of postnatal age with a fitted curve (blue line). **E)** Box plot by gestational age at birth.

related parameters were also significantly affected: WBC# (Figure 3.3A), NEUT#, and MONO# increased, while LYMPH# decreased. Four immune cell activation parameters (NLR, NEUT-GI, RE-LYMPH, and AS-LYMPH) were significantly elevated with exposure to corticosteroids. PLT# was also significantly increased.

Effect sizes were estimated using partial eta-squared (pes) from the multivariable ANOVA model. Five CBC parameters demonstrated large effect sizes (pes > 0.14), indicating that corticosteroids explained a substantial proportion of variance in these values. Changes in parameter values associated with corticosteroid exposure are presented in Table 3.4. To minimise bias, values from the 29 CBC parameters significantly affected by corticosteroid exposure were excluded from further analysis.

3.3.4 Calculation of Reference Ranges

Reference ranges with 95% CIs were calculated for all 49 CBC parameters (Table 3.4). For parameters that were normally distributed without transformation, reference ranges were calculated as the mean $\pm$ 1.96 times the SD. Values for parameters requiring transformation were calculated from the transformed data and then back-transformed to the original scale. For parameters that could not be normalised, reference ranges were defined using the 2.5^{th} and 97.5^{th} percentiles, with CIs estimated by non-parametric bootstrap methods.

3.3.5 Effects of Sex, Age, and Gestational Age at Birth

In the multivariable ANOVA, seven CBC parameters were significantly influenced by sex (Table 3.3). RBC# and RE-MONO were higher in males than in females (although the difference was slight), whereas PLT# was higher in females. However, sex did not account for a substantial proportion of variance in any parameter (pes > 0.14).

Postnatal age significantly influenced 33 parameters, including NRBC# and WBC# (Figure 3.3D), as well as five neutrophil activation parameters (NLR, ITR, NEUT-RI, NEUT-GI, and NEUT-TW; Table 3.3). Large effect sizes (pes > 0.14) were observed for seven parameters, including RBC-HE, indicating strong developmental trends over

time. Changes with GA were also significant in 30 parameters (Table 3.3). RBC# was lower and NRBC# was higher in more preterm neonates. GA exerted large effects (pes $>$ 0.14) on six parameters, including NRBC# and NEUT-GI.

Significant interactions were observed between sex and postnatal age in nine parameters, sex and GA in six parameters, and postnatal age and GA in 24 parameters (Table 3.3). The three-way interaction between sex, postnatal age, and GA was significant in seven parameters. Most interactions were modest (pes $<$ 0.14), though interactions between postnatal age and GA had significant effects on RBC#, HGB, and HCT.

Table 3.3: P-values from multivariate analyses of variance assessing the effects of corticosteroids, sex, postnatal age, and gestational age at birth on parameters of the complete blood count in neonates admitted to a Neonatal Intensive Care Unit.

	Individual Effects								Interactive Effects							
	Corticosteroids		Sex		Age		GA		Sex and Age		Sex and GA		Age and GA		Sex, Age, and GA	
Parameter	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect
Red blood cell parameters																
RBC#	0.62	0.00	0.01	0.02	0.48	0.00	< 0.001	0.14	0.89	0.00	0.02	0.02	< 0.001	0.18	< 0.001	0.06
HGB	0.60	0.00	0.03	0.02	< 0.001	0.10	< 0.001	0.12	0.46	0.00	0.003	0.03	< 0.001	0.19	< 0.001	0.06
HCT	0.70	0.00	0.06	0.01	< 0.001	0.06	< 0.001	0.11	0.46	0.00	0.005	0.03	< 0.001	0.20	< 0.001	0.05
MCV	0.004	0.03	0.51	0.00	< 0.001	0.35	0.58	0.00	0.04	0.02	0.18	0.01	0.006	0.03	0.94	0.00
MCH	0.02	0.02	0.87	0.00	< 0.001	0.47	0.77	0.00	0.14	0.01	0.17	0.01	0.03	0.02	0.47	0.00
MCHC	0.19	0.01	0.22	0.01	< 0.001	0.24	0.13	0.01	0.89	0.00	0.63	0.00	0.32	0.00	0.20	0.01
RDW-CV	0.60	0.00	0.14	0.01	< 0.001	0.13	0.005	0.03	0.34	0.00	0.13	0.01	0.02	0.02	0.55	0.00
RDW-SD	0.30	0.00	0.16	0.01	0.26	0.01	< 0.001	0.05	0.94	0.00	0.19	0.01	0.10	0.01	0.56	0.00
NRBC#	< 0.001	0.05	0.23	0.01	< 0.001	0.08	< 0.001	0.23	0.04	0.02	0.41	0.00	0.28	0.00	0.34	0.00
NRBC%	< 0.001	0.05	0.20	0.01	< 0.001	0.07	< 0.001	0.23	0.03	0.02	0.30	0.00	0.19	0.01	0.34	0.00
RET#	< 0.001	0.04	0.39	0.00	0.36	0.00	< 0.001	0.05	0.19	0.01	0.36	0.00	0.65	0.00	0.92	0.00
RET%	< 0.001	0.04	0.81	0.00	0.25	0.01	< 0.001	0.10	0.24	0.01	0.16	0.01	0.004	0.03	0.32	0.00
IRF	< 0.001	0.08	0.28	0.00	< 0.001	0.06	< 0.001	0.15	0.32	0.00	0.80	0.00	0.003	0.03	0.05	0.01
RBC-HE	0.81	0.00	0.16	0.01	< 0.001	0.35	< 0.001	0.06	0.24	0.01	0.25	0.01	0.07	0.01	0.44	0.00
RET-HE	< 0.001	0.15	0.12	0.01	0.20	0.01	< 0.001	0.17	0.92	0.00	0.31	0.00	0.94	0.00	0.03	0.02
DELTA-HE	< 0.001	0.17	0.57	0.00	< 0.001	0.18	< 0.001	0.06	0.31	0.00	0.81	0.00	0.08	0.01	0.08	0.01

Continued on next page

Continued from last page

	Individual Effects								Interactive Effects							
	Corticosteroids		Sex		Age		GA		Sex and Age		Sex and GA		Age and GA		Sex, Age, and GA	
Parameter	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect
HYPO-HE	< 0.001	0.11	< 0.001	0.04	0.16	0.01	< 0.001	0.20	0.47	0.00	0.55	0.00	0.51	0.00	0.61	0.00
HYPER-HE	0.14	0.01	0.99	0.00	< 0.001	0.38	0.52	0.00	0.07	0.01	0.02	0.02	0.008	0.03	0.09	0.01
MICRO-RBC	0.22	0.01	0.03	0.02	< 0.001	0.12	< 0.001	0.09	0.79	0.00	0.19	0.01	0.03	0.02	0.80	0.00
MACRO-RBC	0.03	0.02	0.76	0.00	< 0.001	0.21	0.37	0.00	0.01	0.02	0.04	0.02	0.002	0.04	0.75	0.00
White blood cell parameters																
WBC#	0.11	0.01	0.31	0.00	0.003	0.03	0.85	0.00	0.55	0.00	0.13	0.01	0.69	0.00	0.83	0.00
NEUT#	< 0.001	0.09	0.35	0.00	< 0.001	0.06	0.68	0.00	0.42	0.00	0.13	0.01	0.009	0.03	0.25	0.01
NEUT%	< 0.001	0.19	0.43	0.00	< 0.001	0.05	0.65	0.00	0.31	0.00	0.25	0.01	< 0.001	0.05	0.08	0.01
LYMPH#	< 0.001	0.08	0.84	0.00	0.12	0.01	0.005	0.03	1.00	0.00	0.45	0.00	0.003	0.03	0.12	0.01
LYMPH%	< 0.001	0.13	0.93	0.00	< 0.001	0.14	< 0.001	0.07	0.90	0.00	0.70	0.00	< 0.001	0.07	0.06	0.01
MONO#	0.04	0.01	0.21	0.01	< 0.001	0.14	< 0.001	0.06	0.96	0.00	0.43	0.00	0.35	0.00	0.81	0.00
MONO%	0.09	0.01	0.35	0.00	< 0.001	0.10	< 0.001	0.07	0.87	0.00	0.99	0.00	0.45	0.00	0.98	0.00
EOSI#	< 0.001	0.14	0.91	0.00	0.82	0.00	0.18	0.01	0.01	0.02	0.18	0.01	0.91	0.00	0.47	0.00
EOSI%	< 0.001	0.17	0.81	0.00	0.33	0.00	0.15	0.01	0.02	0.02	0.08	0.01	0.99	0.00	0.45	0.00
BASO#	0.03	0.02	0.68	0.00	< 0.001	0.12	0.17	0.01	0.88	0.00	0.55	0.00	0.007	0.03	0.21	0.01
BASO%	0.002	0.03	0.12	0.01	< 0.001	0.11	0.65	0.00	0.85	0.00	0.32	0.00	< 0.001	0.05	0.20	0.01
IG#	0.10	0.01	0.89	0.00	< 0.001	0.06	0.45	0.00	0.71	0.00	0.82	0.00	0.007	0.03	0.22	0.01
IG%	0.25	0.00	0.85	0.00	< 0.001	0.05	0.60	0.00	0.37	0.00	0.16	0.01	0.006	0.03	0.75	0.00

Continued on next page

Continued from last page

	Individual Effects								Interactive Effects							
	Corticosteroids		Sex		Age		GA		Sex and Age		Sex and GA		Age and GA		Sex, Age, and GA	
Parameter	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect
NLR	< 0.001	0.15	0.61	0.00	< 0.001	0.09	0.009	0.03	0.50	0.00	0.46	0.00	< 0.001	0.05	0.03	0.02
ITR	0.02	0.02	0.73	0.00	0.005	0.03	0.53	0.00	0.14	0.01	0.04	0.02	0.12	0.01	0.13	0.01
NEUT-RI	0.41	0.00	0.86	0.00	< 0.001	0.09	0.45	0.00	0.50	0.00	0.46	0.00	0.98	0.00	0.50	0.00
NEUT-GI	0.001	0.03	0.009	0.03	0.03	0.02	< 0.001	0.18	0.21	0.01	0.62	0.00	0.30	0.00	0.57	0.00
NEUT-TW	0.46	0.00	0.93	0.00	0.005	0.03	0.002	0.04	0.23	0.01	0.07	0.01	0.32	0.00	0.05	0.02
RE-LYMPH	< 0.001	0.07	0.08	0.01	0.01	0.02	< 0.001	0.06	0.91	0.00	0.65	0.00	0.14	0.01	0.70	0.00
AS-LYMPH	0.04	0.01	0.72	0.00	0.54	0.00	0.002	0.04	0.53	0.00	0.60	0.00	0.90	0.00	0.04	0.02
HFLC	0.29	0.00	0.61	0.00	0.75	0.00	0.001	0.04	0.34	0.00	0.68	0.00	0.64	0.00	0.06	0.01
RE-MONO	0.19	0.01	0.05	0.02	0.15	0.01	0.002	0.04	0.61	0.00	0.75	0.00	0.62	0.00	0.10	0.01
Platelet parameters																
PLT#	< 0.001	0.05	0.02	0.02	0.46	0.00	0.90	0.00	0.37	0.00	0.67	0.00	< 0.001	0.07	0.85	0.00
PDW	0.19	0.01	0.44	0.00	0.01	0.03	0.07	0.02	0.07	0.01	0.15	0.01	0.05	0.02	0.82	0.00
MPV	0.18	0.01	0.20	0.01	0.001	0.05	< 0.001	0.05	0.04	0.02	0.21	0.01	0.12	0.01	0.83	0.00
P-LCR	0.20	0.01	0.15	0.01	0.003	0.04	< 0.001	0.05	0.02	0.02	0.11	0.01	0.10	0.01	0.70	0.00
PCT	0.002	0.03	0.34	0.00	0.98	0.00	0.01	0.03	0.70	0.00	0.84	0.00	0.007	0.03	0.72	0.00
IPF	0.03	0.02	0.50	0.00	0.29	0.00	0.003	0.03	0.008	0.03	0.10	0.01	< 0.001	0.12	0.53	0.00
PLR	< 0.001	0.11	0.05	0.01	0.51	0.00	0.06	0.01	0.45	0.00	0.84	0.00	0.008	0.03	0.37	0.00

Effect refers to partial eta-squared (pes) values from multivariable analyses of variance. Significant p-values ($p < 0.05$) and effect sizes ($pes > 0.14$) are highlighted in red. Abbreviations: GA – gestational age at birth. For parameter abbreviations, see List of Complete Blood Count Parameter Abbreviations (p. xxi).

Table 3.4: Reference ranges for neonatal complete blood count parameters from the Neonatal Intensive Care Unit, with deviations from the average for parameters affected by corticosteroid exposure or sex.

Parameter	Lower Limit (95% CI)	Average (95% CI)	Upper Limit (95% CI)	Corticosteroids	Males	Females
Red blood cell parameters						
RBC#	2.42 (2.33 – 2.51)	3.49 (3.41 – 3.60)	5.70 (5.47 – 6.00)		+0.01	–0.04
HGB	78 (77 – 81)	116 (112 – 122)	221 (204 – 225)		–1	+2
HCT	21.7 (21.2 – 22.4)	32.4 (31.4 – 33.7)	57.0 (54.5 – 59.6)			
MCV	82.2 (81.7 – 83.0)	93.9 (92.5 – 95.0)	113.6 (113.0 – 116.5)	–6.0		
MCH	28.8 (28.0 – 29.1)	34.1 (33.3 – 34.6)	43.8 (42.9 – 44.4)	–3.2		
MCHC	336 (331 – 338)	361 (360 – 363)	388 (385 – 392)			
RDW-CV	40.8 (40.0 – 42.3)	56.2 (55.1 – 57.2)	81.6 (76.5 – 84.4)			
RDW-SD	12.5 (12.3 – 13.5)	16.2 (15.9 – 16.7)	22.3 (21.6 – 23.3)			
NRBC#[a]	0.00 (0.00 – 0.00)	0.06 (0.04 – 0.08)	2.35 (1.64 – 2.57)	–0.03		
NRBC%[a]	0.0 (0.0 – 0.0)	0.5 (0.4 – 0.8)	20.2 (12.6 – 22.8)	–0.2		
RET#	0.023 (0.019 – 0.034)	0.113 (0.101 – 0.120)	0.320 (0.293 – 0.346)	–0.023		
RET%	0.7 (0.5 – 0.9)	3.2 (2.9 – 3.4)	10.4 (9.0 – 10.7)	–0.3		
IRF	10.2 (7.8 – 12.5)	32.5 (31.1 – 33.8)	54.7 (52.4 – 57.1)	–5.5		
RBC-HE	26.0 (25.2 – 26.9)	31.0 (30.5 – 31.5)	36.7 (36.4 – 37.0)			
RET-HE	22.1 (21.3 – 22.5)	29.7 (29.3 – 30.0)	35.1 (34.8 – 35.4)	+1.4		
DELTA-HE	-8.1 (-9.2 – -7.5)	-1.9 (-2.3 – -1.6)	3.0 (2.4 – 3.5)	+3.2		
HYPO-HE	0.4 (0.3 – 0.5)	1.8 (1.6 – 2.0)	6.7 (6.2 – 8.9)	–0.5	+0.3	–0.3
HYPER-HE	0.1 (0.1 – 0.2)	1.1 (0.9 – 1.2)	12.8 (10.2 – 15.1)			

Continued on next page

Continued from last page

Parameter	Lower Limit (95% CI)	Average (95% CI)	Upper Limit (95% CI)	Corticosteroids	Males	Females
MICRO-RBC	1.1 (1.0 – 1.2)	3.3 (3.1 – 3.7)	10.1 (9.6 – 10.7)		+0.6	–0.5
MACRO-RBC	1.8 (1.4 – 2.1)	5.6 (5.0 – 6.1)	31.6 (28.4 – 35.4)	–1.9		
White blood cell parameters						
WBC#	6.00 (5.57 – 6.23)	10.29 (9.81 – 10.76)	20.15 (18.88 – 22.12)			
NEUT#	1.15 (1.00 – 1.28)	3.33 (3.07 – 3.65)	10.30 (9.04 – 11.87)	+1.66		
NEUT%	16.4 (13.3 – 17.3)	33.1 (31.7 – 34.6)	56.8 (55.4 – 58.7)	+17.2		
LYMPH#	2.37 (2.19 – 2.55)	4.19 (4.05 – 4.34)	8.52 (7.99 – 8.99)	–0.75		
LYMPH%	22.6 (21.3 – 23.6)	43.4 (42.0 – 44.6)	63.5 (62.2 – 69.1)	–12.2		
MONO#	0.60 (0.54 – 0.73)	1.61 (1.52 – 1.71)	3.54 (3.28 – 3.88)	+0.23		
MONO%	6.4 (5.8 – 7.8)	15.3 (14.9 – 15.9)	26.2 (24.1 – 27.7)			
EOSI#	0.05 (0.03 – 0.07)	0.43 (0.40 – 0.47)	1.77 (1.59 – 2.06)	–0.22		
EOSI%	0.7 (0.5 – 0.9)	4.4 (4.1 – 4.7)	16.1 (15.0 – 17.4)	–2.2		
BASO#[a]	0.02 (0.01 – 0.02)	0.09 (0.08 – 0.10)	0.22 (0.16 – 0.25)	–0.04		
BASO%[a]	0.2 (0.2 – 0.2)	0.7 (0.6 – 0.7)	1.5 (1.3 – 1.9)	–0.3		
IG#	0.02 (0.01 – 0.02)	0.09 (0.08 – 0.10)	1.74 (1.12 – 2.02)			
IG%	0.2 (0.2 – 0.3)	1.0 (0.9 – 1.1)	9.3 (7.9 – 9.8)			
NLR	0.26 (0.22 – 0.28)	0.98 (0.88 – 1.02)	2.47 (2.32 – 2.65)	+0.76		
ITR	0.01 (0.01 – 0.01)	0.03 (0.03 – 0.03)	0.17 (0.16 – 0.30)	–0.01		
NEUT-RI	36.3 (36.1 – 37.0)	42.7 (42.4 – 43.4)	53.2 (52.2 – 55.0)			
NEUT-GI	137.2 (136.1 – 138.6)	147.5 (146.9 – 148.2)	156.5 (156.0 – 156.9)	+0.6	+0.6	–0.7

Continued on next page

Continued from last page

Parameter	Lower Limit (95% CI)	Average (95% CI)	Upper Limit (95% CI)	Corticosteroids	Males	Females
NEUT-TW	24.0 (23.0 – 25.0)	31.0 (30.8 – 31.9)	67.0 (60.1 – 71.5)			
RE-LYMPH	0.4 (0.3 – 0.8)	7.9 (6.5 – 9.2)	28.5 (26.1 – 32.7)	+7.7		
AS-LYMPH[a]	0.0 (0.0 – 0.0)	0.4 (0.2 – 0.5)	1.2 (0.9 – 1.6)	+0.4		
HFLC[a]	0.0 (0.0 – 0.0)	0.0 (0.0 – 0.0)	0.0 (0.0 – 0.0)			
RE-MONO[a]	0.0 (0.0 – 0.0)	0.2 (0.2 – 0.3)	5.0 (2.7 – 7.6)		+0.1	–0.1
Platelet parameters						
PLT#	74 (67 – 98)	311 (294 – 328)	730 (662 – 862)	+118	–11	+17
PDW	9.7 (9.4 – 10.0)	13.6 (13.2 – 14.1)	23.5 (22.7 – 24.3)			
MPV	9.7 (9.5 – 10.0)	11.6 (11.5 – 11.8)	13.7 (13.5 – 13.9)			
P-LCR	21.4 (19.6 – 23.1)	37.1 (36.0 – 38.1)	52.8 (51.0 – 54.6)			
PCT	0.2 (0.1 – 0.2)	0.4 (0.4 – 0.4)	0.8 (0.7 – 0.9)	+0.1		
IPF	2.2 (2.0 – 2.6)	8.3 (7.6 – 9.1)	28.5 (26.6 – 29.8)	–1.1		
PLR	0.018 (0.015 – 0.019)	0.067 (0.063 – 0.070)	0.214 (0.181 – 0.263)	+0.047		

All values were calculated parametrically unless transformation failed to normalise values ([a]). Average refers to the mean or median, depending on the underlying distribution. Lower Limit and Upper Limit refer to the lower and upper 95% reference limits or 2.5$^{\text{th}}$ and 97.5$^{\text{th}}$ percentiles, depending on the underlying distribution. Abbreviations: CI – confidence interval. For parameter abbreviations, see List of Complete Blood Count Parameter Abbreviations (p. xxi).

3.4 Discussion

3.4.1 Summary of Findings & Comparison with Existing Literature

This chapter established reference ranges for 49 neonatal CBC parameters and evaluated the individual and interactive effects of corticosteroid exposure, sex, postnatal age, and GA. All parameters yielded narrow 95% CIs, which underscores the robustness of the calculated reference ranges, enhancing their potential utility in clinical decision-making.

Effects of Corticosteroid Exposure

Corticosteroid exposure significantly influenced 29 parameters, including the WBC differentials – findings that align with adult studies, as reported by Barden et al. (2018) [134]. For example, both the current research and Barden et al. observed an increase in NEUT# and a decrease in LYMPH# following corticosteroid exposure. However, while the authors reported an increase in WBC# with corticosteroid exposure in adults, the current study observed no change in WBC# among corticosteroid-exposed neonates. This discrepancy may be attributable to confounding factors, such as the older postnatal age of corticosteroid-exposed neonates in this cohort. Nonetheless, the consistent patterns in NEUT# and LYMPH# across neonates and adults support the reliability of the current neonatal findings, particularly the novel insights into immune activation markers.

Interactive Effects of Sex, Postnatal Age, and Gestational Age at Birth

Sex-based differences were observed in seven parameters, consistent with the findings of Zhang et al. [139], who reported differences in some RBC-related parameters but no significant variations in the WBC differentials. In the current study, a significant sex-based difference in PLT# was observed, with males averaging 11 x 10^9/L below and females 17 x 10^9/L above the cohort mean. Although statistically significant, these differences are unlikely to be clinically relevant, given their small magnitude relative to the wide reference ranges and small effect size (pes < 0.14). This mirrors Zhang et al.'s [139] conclusion that sex has minimal impact on PLT#.

Postnatal age-related changes were significant in 33 CBC parameters, corroborating existing neonatal studies [140]. For example, the observed decrease in HGB and HCT from day 4 – 28 of life is consistent with prior findings [140]. Similarly, the observed decline in NRBC# with increasing postnatal age and GA aligns with previously published research [129].

Comparison with Established Reference Ranges

The reference ranges calculated in this study closely align with those in many previously published studies. For example, reported reference values for HGB at two months of age typically range from 94 – 130 g/L [135]. In comparison, the current study produced a broader reference interval of 78 to 221 g/L, reflecting the wide biological variability in HGB across GA and postnatal age. Similarly, the calculated WBC# reference range of 6.00 to 20.15 $\times$ 10^9/L was comparable to published intervals of 5 to 15 $\times$ 10^9/L for neonates older than 48 hours [135, 136]. One notable divergence lies in the reference interval for PLT#. Published values consistently suggest a narrower interval (150 – 450 $\times$ 10^9/L) [135, 136], whereas the wider interval observed here (74 – 730 $\times$ 10^9/L) may better reflect the broader physiological variability in a NICU population.

3.4.2 Clinical Implications of Findings

This study presents robust reference ranges for parameters of the neonatal CBC, which could be utilised in clinical settings to identify neonates at risk of adverse clinical outcomes. For example, results that fall outside the presented reference ranges for the immune cell activation markers (e.g., NEUT-RI and NEUT-GI) may be indicative of immune activation commonly associated with bacterial infection.

This study also highlights the importance of considering additional factors when interpreting neonatal CBC results. Postnatal age has long been recognised as a key confounder, and this study illustrates its impact on CBC parameters. Furthermore, these findings demonstrate that sex, GA, and corticosteroid exposure also influence haematological variability, reinforcing the need for individualised interpretation.

3.4.3 Strengths and Limitations

This research successfully established comprehensive reference ranges for the neonatal CBC while evaluating the effects of key factors such as corticosteroid exposure, sex, postnatal age, and GA on the neonatal haematological profile. The study provides novel insights by addressing significant gaps in the literature, particularly regarding the impact of corticosteroids on the neonatal CBC, which has not been previously explored. Corticosteroids are widely used to treat respiratory distress in neonates, and my findings represent an essential contribution to understanding their effects.

Furthermore, this study expands on the limited literature regarding the effects of sex, postnatal age, and GA. While prior research has identified their influence on a few CBC parameters, this study provides a more comprehensive analysis encompassing 49 parameters, including activation markers such as NEUT-RI and NEUT-GI, which have not been widely studied in neonates. This extensive analysis ensures a more nuanced understanding of the neonatal haematological profile.

The study cohort was carefully selected to include hospitalised neonates without clinical suspicion of pathology, making the findings directly relevant to hospitalised, but clinically stable, neonates. As blood samples are rarely obtained from entirely healthy neonates, these results serve as a critical baseline for assessing neonates with potential clinical concerns in hospital settings.

Despite its strengths, this study has certain limitations. While strict inclusion criteria were applied, the possibility remains that some non-clinically stable neonates were inadvertently included. To mitigate this risk, outliers in key clinical parameters, such as WBC#, HGB, and PLT#, were carefully examined to assess whether the inclusion criteria were sufficient. This evaluation found that the initial exclusion period of two days either side of a culture request was insufficient to exclude results affected by pre- or post-infection leucocytosis, and so the exclusion period was extended from two to five days.

As a retrospective study, there is an inherent risk of bias due to incomplete or unrecorded data. For instance, averaging results by day of life masks nuanced changes occurring within shorter timeframes. Additionally, the timing of CBC measurements relative to corticosteroid exposure was not recorded, which may have led to misclassification of results between corticosteroid-exposed and unexposed groups. Future studies could mitigate this by recording the precise timing of corticosteroid administration and measuring subsequent haematological changes.

Another potential limitation is the absence of data on additional confounding factors. Variables such as mode of delivery (MOD), maternal health at birth, or weight in older neonates might influence haematological parameters and could be explored in future analyses. Ethnicity may also influence certain haematological indices; however, meaningful evaluation of ethnicity was not feasible in the present study due to limited sample sizes within individual ethnic groups and the complexity of classifying ethnicity in a neonatal population, where ancestry is often mixed and recorded using broad categorical labels. While the current study met the minimum requirement of 120 samples for developing reference ranges, as determined by the CLSI guidelines [137], subgroup analyses were not possible due to low sample numbers in most subgroups. Robust assessment of the effects of these confounding factors would require substantially larger cohorts and carefully defined classification frameworks to avoid oversimplification or misinterpretation.

Finally, as a single-centre study, the generalisability of these findings is limited. While the study provides valuable data for its specific population, external validation across multiple centres and diverse neonatal populations is necessary to confirm the broader applicability of the results. Future research should focus on prospective, multicentre studies to validate and expand upon these findings. Such studies would enable real-time data collection, ensuring the inclusion of clinically stable neonates and allowing for precise tracking of changes in haematological parameters following corticosteroid treatment. By addressing these limitations, subsequent research could provide more reliable and universally applicable reference intervals for neonatal CBC results.

3.5 Chapter Summary

In this chapter, I produced robust reference ranges for 49 parameters of the neonatal CBC and comprehensively analysed the individual and interactive effects of sex, postnatal age, and GA. To the best of my knowledge, many of these parameters have not been previously evaluated in neonates. This study also provides novel insights into the effect of corticosteroid exposure on the CBC, for which no data in neonates currently exist.

While the interactive effects of sex, postnatal age, and GA on CBC parameters remain underexplored, the findings of this study align well with the limited available literature. These consistencies lend credence to the reliability of the novel results presented here and their potential to inform future clinical and research applications.

Establishing a robust baseline neonatal CBC profile is critical for identifying abnormal results and ensuring accurate clinical interpretation. Building on this, the next chapter will explore how individual CBC parameters vary in cases of neonatal sepsis and assess their diagnostic accuracy.

Chapter 4

Evaluation of the Diagnostic Accuracy of the Complete Blood Count for Late-Onset Neonatal Sepsis

4.1 Introduction

The gold standard for diagnosing LOS, as defined by the NICE guidelines, is a positive blood culture [11]. The evidence review informing these guidelines evaluated the available literature on various sepsis biomarkers and concluded that the diagnostic utility of the CBC was limited [77]. As a result, the CBC was excluded from clinical recommendations. Notably, the review evaluated only four traditional parameters (WBC#, NEUT#, ITR, and PLT#) and did not include immune activation markers. Instead, the guidelines recommend CRP as a monitoring tool, citing a sensitivity of 62% and specificity of 73% at a threshold of 10 mg/L.

Numerous studies have investigated the diagnostic accuracy of CBC parameters for neonatal sepsis, as discussed and referenced in Chapter 2. While most studies have focused on conventional markers, more recent work has evaluated novel haematological indicators. For instance, a meta-analysis of ITR reported an AUC of 0.816, with 72% sensitivity and 85% specificity at an optimal threshold of 0.2 [76]. However, the overall diagnostic value of CBC parameters for LOS remains uncertain due to inconsistent findings, heterogeneity in study designs, and the lack of standardised reference definitions, underscoring the need for further evaluation using well-defined cohorts and consistent diagnostic criteria.

Chapter 3 established reference intervals for CBC parameters in neonates without clinical or laboratory signs of sepsis, providing a baseline against which abnormal results may be interpreted. Building on this, the current chapter evaluates the ability of these same

CBC parameters to distinguish between septic and non-septic neonates in a clinically stratified, real-world cohort.

4.1.1 Aims of the Chapter

This chapter aimed to evaluate the diagnostic accuracy of CBC parameters for LOS. This was achieved by comparing CBC results from neonates without clinical or laboratory signs of infection to those from neonates with culture-confirmed sepsis. All routinely measured CBC parameters, including both traditional indices and novel immune activation markers, were assessed using receiver operating characteristic (ROC) curve analysis and standard diagnostic accuracy metrics. The findings from this analysis informed the selection of parameters for inclusion in the diagnostic algorithms developed in Chapter 5.

4.2 Methods

4.2.1 Data Collection and Sample Stratification

The dataset used in this chapter is the same as that described in Chapter 3. Ethical approval was obtained from the Oxford Research Ethics Committee (REC 08/H0606/139), as outlined in Chapter 3. An opt-out consent process was used, and data were anonymised before analysis to ensure confidentiality.

In summary, neonatal CBC data obtained between 13$^{\text{th}}$ January 2022 and 13$^{\text{th}}$ January 2023 using a Sysmex XN-2000 analyser (Kobe, Japan) in the haematology laboratory at the John Radcliffe Hospital (Oxford, UK) and an XN-450 point-of-care analyser in the NICU were analysed retrospectively. Where statistically significant differences were observed between paired samples, results from the XN-450 were transformed to align with XN-2000 outputs (Appendix A2, p. 248). CBC results were then matched to clinical data extracted from BadgerNet, EPR, and clinical observation charts. Results were subsequently categorised based on the presence of clinical or laboratory signs of sepsis, as described in Chapter 3 (Figure 3.1, p. 68).

For the current chapter, CBC results from the first three days of life were excluded to focus specifically on LOS. The full dataset was stratified into training and testing subsets based on sex, GA, and duration of NICU stay, ensuring statistical comparability between the resulting groups. Stratification was performed before applying exclusion criteria, as clinical data for stratification were more readily available. The effectiveness of the stratification was assessed using Student's t-tests, with p-values greater than 0.05 indicating no significant differences between subsets.

4.2.2 Establishing the Diagnostic Accuracy of CBC Parameters

All statistical analyses were performed using RStudio (R version 4.3.0) [86]. For each CBC parameter, the training dataset was filtered to include only CBC results categorised as *unlikely* or *probable sepsis*, to ensure comparison between clearly defined clinical categories. The normality of parameter distributions was assessed using the Shapiro-Wilk

test. Parameters that were not normally distributed ($p < 0.05$) were transformed using ordered quantile normalisation to meet the assumptions required for subsequent analyses.

Parameters that were normally distributed (after transformation, if applicable) were compared between sepsis categories using Student's t-test, and group means were reported. Parameters that remained non-normally distributed were analysed using the Wilcoxon rank-sum test, with group medians reported.

Diagnostic performance was assessed using ROC curve analysis. The following diagnostic accuracy metrics were calculated:

- AUC, sensitivity, and specificity.
- Optimal threshold, determined using Youden's Index.
- TP, FP, TN, and FN.
- PPV, NPV, LR^+, and LR^-.

4.2.3 Validation of the Diagnostic Accuracy of CBC Parameters

To validate findings from the training dataset, the testing dataset was analysed using the previously identified optimal thresholds. CBC results categorised as *unlikely* or *probable sepsis* were reclassified based on these thresholds. Diagnostic accuracy metrics were recalculated using the TP, TN, FP, and FN values from the testing dataset. To assess generalisability, ROC curves from the training and testing datasets were compared using DeLong's test - a non-parametric approach to comparing AUCs.

4.2.4 Parameter Selection for Algorithm Development

Parameters included in algorithm development (Chapter 5) were selected using the following criteria to ensure clinical relevance and methodological robustness:

- Statistical significance: parameters demonstrating strong diagnostic accuracy.
- Biological rationale: parameters with a recognised role in sepsis pathophysiology, as supported by the literature.
- Redundancy avoidance: where both absolute and relative counts were available, absolute counts were preferred.

4.3 Results

This section summarises the results from the analysis of 49 CBC parameters. All supplementary figures are available online at `https://4fsqon-emily-hyde.shinyapps.io/ThesisFigures/` under the tab *Neonatal CBC for LOS Diagnosis* [138]. Shown below are representative figures for the analysis of WBC# (Figure 4.2), which is commonly assessed in neonatal care, and NLR (Figure 4.3), which was the best performing parameter in the current analyses.

4.3.1 Sample Characteristics

During the study period, 3,490 CBC results were collected from 587 neonates (Figure 4.1). Stratified sampling was performed before exclusion, using sex ($p = 0.60$), GA ($p = 0.78$), and duration of NICU stay ($p = 0.63$) as stratification factors. These variables were readily available for all neonates, allowing manual clinical data collection (e.g. from BadgerNet) to be prioritised for the training subset. After stratification, ineligible samples were excluded (e.g. missing clinical data or within the first three days of life), and CBC results recorded on the same day were averaged, ensuring that each day contributed only one representative sample per neonate. This process yielded a final dataset of 2,330 CBC results from 304 neonates (Figure 4.1).

The training dataset consisted of 1,164 CBC results from 168 neonates, and the testing dataset comprised 1,166 CBC results from 136 neonates (Figure 4.1, Table 4.1). There were no statistically significant differences between the training and testing datasets in terms of sex, GA, or postnatal age ($p > 0.05$). However, the duration of NICU stay was significantly longer in the testing dataset (median: 24 days; IQR: 12 – 55) than in the training dataset (median: 20 days; IQR: 10 – 42; $p = 0.04$). There were no significant differences in the proportions of CBC results classified as *unlikely* or *probable sepsis* between the datasets (Table 4.1). However, the age at sampling for *probable sepsis* CBC results was significantly higher in the testing dataset (median: 24 days; IQR: 14 – 43) than in the training dataset (median: 19 days; IQR: 10 – 33; $p < 0.001$).

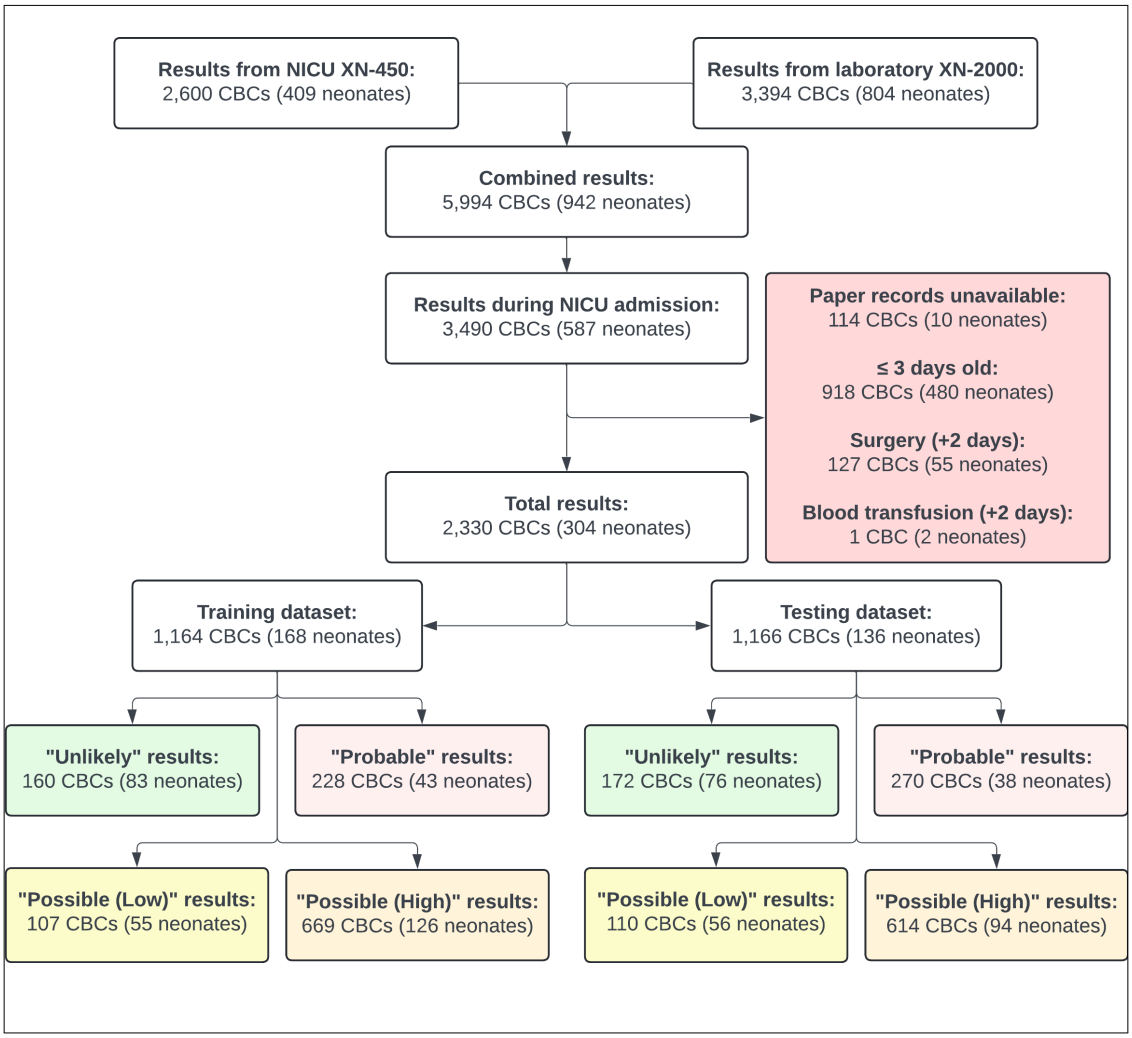


Figure 4.1: Flowchart of complete blood count (CBC) results included in the determination of the diagnostic utility of individual parameters for late-onset sepsis.

Table 4.1: Sample characteristics of complete blood count results used for the evaluation of the diagnostic accuracy of complete blood count parameters for late-onset sepsis.

		Training Dataset			Testing Dataset			Train vs Test P-Value		
	Full Dataset	All Values	*Unlikely Sepsis*	*Probable Sepsis*	All Values	*Unlikely Sepsis*	*Probable Sepsis*	All Values	*Unlikely Sepsis*	*Probable Sepsis*
CBC Results	2,330	1,164	160	228	1,166	172	270	0.97[a]	0.51[a]	0.06[a]
Neonates	304	168	83	43	136	76	38	0.07[a]	0.58[a]	0.58[a]
GA (weeks)	32.0 (27.3 – 37.1)	32.1 (27.4 – 37.6)	29.6 (26.2 – 35.1)	26.7 (24.6 – 30.4)	31.1 (26.8 – 36.2)	29.9 (26.7 – 34.5)	26.0 (24.6 – 28.8)	0.15[b]	0.93[b]	0.49[b]
Males (%)	174 (57%)	94 (56%)	46 (55%)	22 (51%)	80 (59%)	47 (62%)	25 (66%)	0.70[a]	0.51[a]	0.27[a]
Duration of Stay (days)	22 (10 – 47)	20 (10 – 42)	30 (14 – 50)	47 (30 – 72)	24 (12 – 55)	32 (15 – 71)	60 (38 – 84)	0.04[b]	0.76[b]	0.29[b]
Postnatal Age (days)	20 (10 – 38)	19 (9 – 37)	26 (11 – 48)	19 (10 – 33)	21 (10 – 40)	24 (10 – 52)	24 (14 – 43)	0.10[b]	0.82[b]	0.001[b]

[a] *Chi-squared test.* [b] *Wilcoxon rank sum test. Significant p-values (< 0.05) are highlighted in red. Abbreviations: CBC – complete blood count; GA – gestational age at birth.*

4.3.2 Evaluation of CBC Parameters in the Training Dataset

A total of 49 CBC parameters were evaluated for their diagnostic accuracy in identifying LOS (Table 4.2; Table A3.1, p. 256). Normality was tested separately within the *unlikely* and *probable sepsis* groups. RET-HE was the only parameter that was normally distributed in both groups ($p > 0.05$). The remaining 48 parameters showed non-normality in at least one group, so were transformed using the ordered quantile transformation. Following the transformation, 38 parameters were found to approximate normal distributions. Ten parameters, including NRBC# and AS-LYMPH, could not be successfully normalised, primarily due to a high proportion of zero values.

Statistically significant differences were observed between the *unlikely* or *probable sepsis* groups in 43 parameters, including WBC differential counts, immune cell activation markers, and platelet indices (Figure 4.2; Table 4.2). Of these, 22 parameters showed moderate to large effect sizes (absolute Cohen's $d > 0.50$ or Wilcoxon $r > 0.30$). Five parameters demonstrated an AUC greater than 0.8, indicating strong diagnostic accuracy: NEUT%, LYMPH%, NLR, NEUT-RI, and NEUT-TW (Table 4.2). Among these, NLR (with an optimal threshold of 1.46) provided the highest sensitivity (76.7%; Figure 4.3). In contrast, LYMPH% less than 30.5% yielded the highest specificity (85%).

4.3.3 Evaluation of CBC Parameters in the Testing Dataset

Statistical significance was retained in 36 parameters in the testing dataset, including the WBC differential counts and most immune cell activation parameters (Figure 4.2; Table 4.2; Table A3.2, p. 259). Seven parameters that were statistically significant in the training dataset lost significance in the testing dataset, including RBC# and ITR. Conversely, five parameters that were not significant in the training dataset became significant in the testing dataset, including MONO#. Four parameters maintained an AUC greater than 0.8 in the testing dataset: NEUT#, NEUT%, LYMPH%, and NLR (Figure 4.3).

Table 4.2: Evaluation of the diagnostic accuracy of the neonatal complete blood count for late-onset sepsis using ROC curve analysis.

	Training Dataset								Testing Dataset							
Parameter	*Unlikely* Average	*Probable* Average	P-Value	Effect	Threshold	AUC	Sens (%)	Spec (%)	*Unlikely* Average	*Probable* Average	P-Value	Effect	AUC	Sens (%)	Spec (%)	DeLong's P-Value
Red blood cell parameters																
RBC#	3.44	3.62	0.02	-0.25	3.23	0.588	77.6	45.6	3.60	3.59	0.70	0.04	0.502	73.0	33.7	0.05
HGB	110[a]	117[a]	0.02	0.12	104	0.569	80.7	44.4	118[a]	115[a]	0.10	0.08	0.547	73.7	34.9	0.62
HCT	30.6[a]	32.9[a]	0.02	0.12	28.2	0.572	87.7	37.5	32.9[a]	32.1[a]	0.09	0.08	0.547	78.9	30.8	0.57
MCV	91.9	90.7	0.07	0.19	96.4	0.552	75.4	35.0	92.0	88.0	< 0.001	0.60	0.666	88.1	39.5	0.005
MCH	33.3	32.9	0.09	0.18	36.8	0.545	92.1	21.9	33.3[a]	31.4[a]	< 0.001	0.25	0.649	96.3	27.3	0.01
MCHC	360	359	0.34	0.10	368	0.530	76.8	36.3	360	358	0.05	0.19	0.562	84.1	21.5	0.43
RDW-SD	55.0	58.8	< 0.001	-0.46	52.7	0.624	80.3	40.0	55.9	55.9	0.97	0.00	0.523	70.7	28.5	0.01
RDW-CV	16.5	18.5	< 0.001	-0.74	16.9	0.710	75.4	58.8	16.7	17.8	< 0.001	-0.51	0.635	67.4	51.2	0.06
NRBC#	0.06[a]	0.30[a]	< 0.001	0.33	0.17	0.694	61.0	69.4	0.04[a]	0.27[a]	< 0.001	0.31	0.686	57.2	70.9	0.82
NRBC%	0.6[a]	1.9[a]	< 0.001	0.26	0.6	0.650	73.3	51.3	0.4[a]	1.5[a]	< 0.001	0.24	0.642	66.4	58.1	0.83
RET#	0.103	0.084	0.005	0.30	0.130	0.586	80.2	39.6	0.114	0.077	< 0.001	0.58	0.653	75.7	38.4	0.10
RET%	3.0	2.3	< 0.001	0.37	4.3	0.602	87.7	34.0	3.1	2.1	< 0.001	0.54	0.642	88.1	32.0	0.33
IRF	33.1	30.5	0.09	0.18	16.7	0.542	20.7	91.2	32.5	27.5	< 0.001	0.34	0.597	25.7	89.0	0.17
RBC-HE	30.6	29.7	< 0.001	0.40	31.0	0.626	74.9	50.9	30.2	29.1	< 0.001	0.70	0.695	88.1	43.6	0.09
RET-HE	29.5	26.9	< 0.001	0.72	29.6	0.703	77.5	56.6	29.9	26.7	< 0.001	0.89	0.737	78.4	55.8	0.34
DELTA-HE	-1.7	-2.9	< 0.001	0.39	-2.4	0.615	56.4	64.8	-1.3	-2.5	< 0.001	0.37	0.604	51.9	65.1	0.78
HYPO-HE	1.6	2.3	< 0.001	-0.42	1.4	0.619	76.7	45.3	1.8	2.3	< 0.001	-0.33	0.592	73.1	38.4	0.51

Continued on next page

Continued from last page

	Training Dataset								Testing Dataset							
Parameter	*Unlikely* Average	*Probable* Average	P-Value	Effect	Threshold	AUC	Sens (%)	Spec (%)	*Unlikely* Average	*Probable* Average	P-Value	Effect	AUC	Sens (%)	Spec (%)	DeLong's P-Value
HYPER-HE	0.9	0.9	0.92	-0.01	0.8	0.521	64.3	46.5	1.0[a]	0.6[a]	< 0.001	0.27	0.661	36.6	39.0	< 0.001
MICRO-RBC	3.6	4.7	< 0.001	-0.40	4.3	0.630	62.3	66.9	4.1	5.4	< 0.001	-0.66	0.682	74.1	55.2	0.20
MACRO-RBC	4.8	5.5	0.009	-0.28	4.3	0.582	73.7	46.9	5.5	4.1	< 0.001	0.49	0.628	44.8	36.6	0.26
White blood cell parameters																
WBC#	10.44	14.43	< 0.001	-0.79	15.26	0.721	50.4	88.8	9.89[a]	15.54[a]	< 0.001	0.44	0.758	50.9	87.2	0.27
NEUT#	3.74[a]	8.35[a]	< 0.001	0.50	5.46	0.796	70.6	78.8	3.53[a]	9.49[a]	< 0.001	0.56	0.830	77.0	75.6	0.26
NEUT%	36.0	54.2	< 0.001	-1.17	45.5	0.806	72.4	77.5	35.8	58.7	< 0.001	-1.42	0.845	79.1	71.5	0.17
LYMPH#	4.17	3.63	< 0.001	0.59	3.84	0.659	62.1	64.4	3.94	3.45	< 0.001	0.65	0.679	70.0	54.7	0.60
LYMPH%	43.2	23.9	< 0.001	1.34	30.5	0.836	71.8	85.0	41.0	22.1	< 0.001	1.49	0.861	75.3	77.3	0.35
MONO#	1.59	2.30	< 0.001	-0.67	2.70	0.687	43.9	90.0	1.64	2.09	< 0.001	-0.52	0.654	34.7	91.9	0.38
MONO%	15.3	15.5	0.40	-0.09	21.0	0.523	18.9	91.3	15.9	14.8	0.01	0.24	0.571	13.1	84.9	0.23
EOSI#	0.40[a]	0.29[a]	0.01	0.13	0.20	0.576	42.1	85.0	0.38	0.29	0.02	0.22	0.561	41.4	77.3	0.72
EOSI%	4.1	2.5	< 0.001	0.51	2.1	0.658	50.0	84.4	3.8[a]	1.9[a]	< 0.001	0.22	0.630	52.2	77.3	0.45
BASO#	0.08[a]	0.11[a]	< 0.001	0.24	0.06	0.640	85.5	40.0	0.08[a]	0.11[a]	< 0.001	0.19	0.611	82.5	36.6	0.47
BASO%	0.6	0.7	0.002	-0.32	0.7	0.591	34.2	80.6	0.7	0.7	0.38	-0.09	0.522	26.9	74.4	0.09
IG#	0.09	0.27	< 0.001	-0.92	0.19	0.745	64.0	76.3	0.08[a]	0.26[a]	< 0.001	0.36	0.713	57.8	72.7	0.37
IG%	1.0	2.0	< 0.001	-0.80	1.4	0.723	69.3	68.8	1.0	1.9	< 0.001	-0.56	0.662	58.2	66.3	0.10
NLR	0.98	2.40	< 0.001	-1.28	1.46	0.825	76.7	76.3	1.05	2.62	< 0.001	-1.47	0.857	85.0	70.3	0.25

Continued on next page

Continued from last page

	Training Dataset								Testing Dataset							
Parameter	*Unlikely* Average	*Probable* Average	P-Value	Effect	Threshold	AUC	Sens (%)	Spec (%)	*Unlikely* Average	*Probable* Average	P-Value	Effect	AUC	Sens (%)	Spec (%)	DeLong's P-Value
ITR	0.03	0.04	< 0.001	-0.37	0.06	0.611	36.4	83.8	0.03	0.03	0.24	-0.11	0.539	31.3	77.3	0.07
NEUT-RI	43.2	49.7	< 0.001	-1.25	46.1	0.812	75.0	74.4	42.9	48.9	< 0.001	-1.13	0.788	67.5	74.4	0.42
NEUT-GI	147.7	145.9	< 0.001	0.34	144.2	0.598	40.4	76.9	147.4	144.1	< 0.001	0.58	0.658	51.5	76.2	0.13
NEUT-TW	31.0	42.4	< 0.001	-1.17	37.3	0.811	68.4	81.9	31.5	43.1	< 0.001	-1.00	0.762	64.6	72.1	0.13
RE-LYMPH	9.9	13.4	< 0.001	-0.61	14.2	0.675	49.6	77.5	9.5	14.1	< 0.001	-0.75	0.710	50.4	79.1	0.35
AS-LYMPH	0.4[a]	0.9[a]	< 0.001	0.26	1.0	0.649	32.5	93.8	0.5[a]	0.9[a]	< 0.001	0.27	0.659	34.0	93.0	0.77
HFLC	0.0[a]	0.0[a]	< 0.001	0.33	0.0	0.696	41.0	91.9	0.0[a]	0.0[a]	< 0.001	0.31	0.684	43.1	88.4	0.74
RE-MONO	0.3[a]	0.8[a]	< 0.001	0.28	0.9	0.667	47.4	82.5	0.2[a]	0.8[a]	< 0.001	0.34	0.704	45.1	81.4	0.32
Platelet parameters																
PLT#	370	211	< 0.001	0.87	257	0.731	65.4	75.6	306	217	< 0.001	0.56	0.647	60.4	58.7	0.02
PDW	13.3	15.6	< 0.001	-0.55	13.7	0.656	71.0	57.0	14.0	15.0	0.02	-0.26	0.598	66.5	51.7	0.20
MPV	11.5	12.2	< 0.001	-0.47	11.5	0.640	69.6	54.9	11.8	12.1	0.02	-0.27	0.594	72.6	45.5	0.32
P-LCR	35.5	41.6	< 0.001	-0.51	35.2	0.646	73.9	50.7	38.1	40.3	0.006	-0.31	0.604	75.4	40.0	0.36
PCT	0.4	0.4	< 0.001	0.52	0.3	0.633	44.9	80.3	0.4	0.3	< 0.001	0.44	0.610	49.7	66.2	0.62
IPF	7.5	14.4	< 0.001	-0.91	7.5	0.740	84.6	52.8	9.3	13.7	< 0.001	-0.62	0.678	83.6	39.0	0.09
PLR	0.075	0.060	< 0.001	0.40	0.059	0.610	52.0	70.0	0.073	0.067	0.20	0.12	0.529	43.8	62.8	0.04

Effect refers to Cohen's d or Wilcoxon r effect size. All values were calculated parametrically unless transformation failed to normalise values ([a]). Average refers to the mean or median, depending on the underlying distribution. Significant p-values (< 0.05) and effect sizes (absolute Cohen's $d > 0.50$ and Wilcoxon $r > 0.3$) are highlighted in red. AUC values > 0.8 and sensitivities/specificities $> 85\%$ are highlighted in green, suggesting strong diagnostic performance. Abbreviations: AUC – area under the curve, ROC – receiver operating characteristic, Sens – sensitivity, Spec – specificity. For parameter abbreviations, see List of Complete Blood Count Parameter Abbreviations (p. xxi).

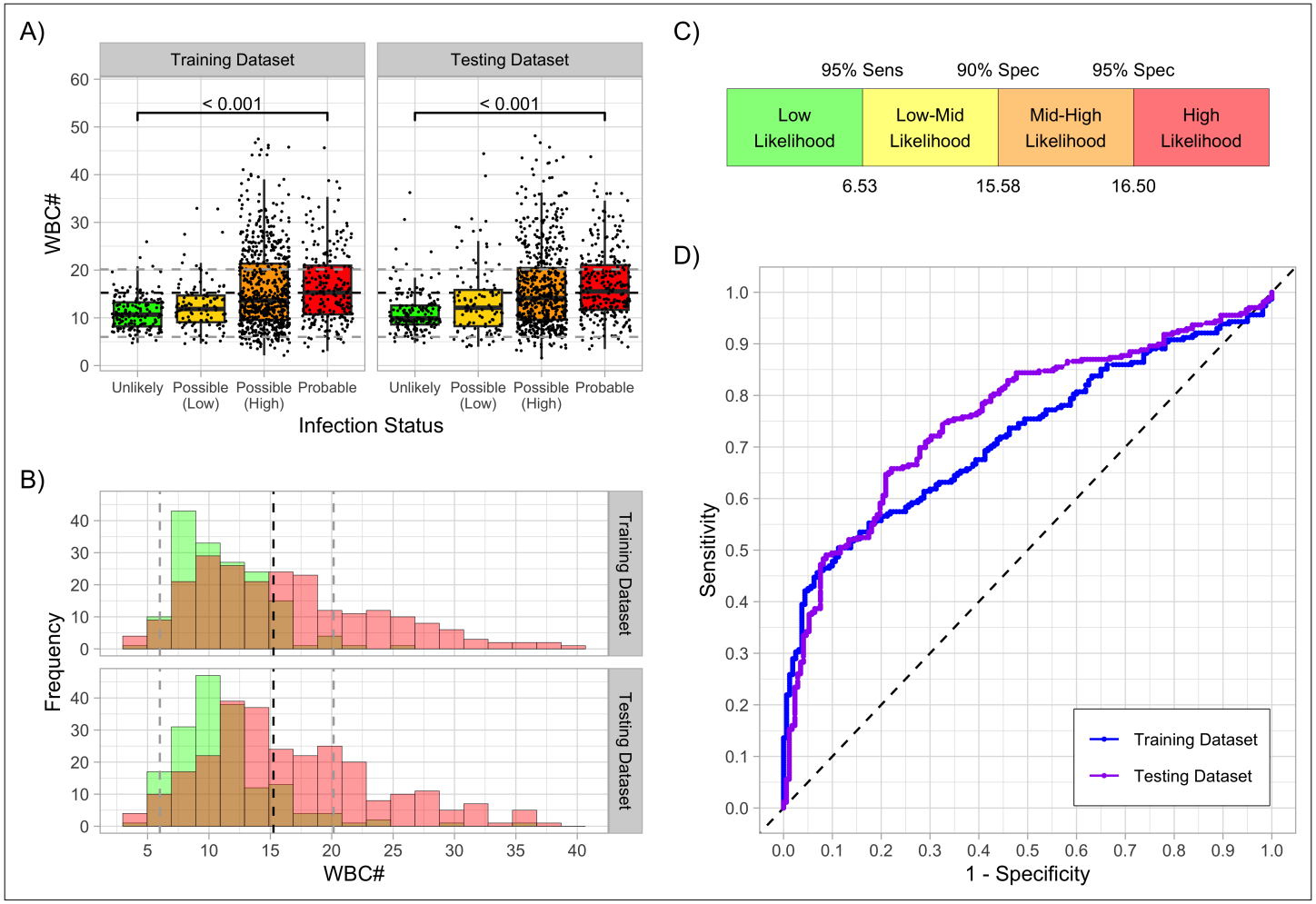


Figure 4.2: Evaluation of the diagnostic accuracy of white blood cell count (WBC#; x 10^9/L) for late-onset sepsis.

A) Box plot of values with clinical suspicion of sepsis. **B)** Histogram of *unlikely* (green) and *probable sepsis* (red) values with the optimal threshold (dashed black) and reference ranges (dashed grey). **C)** Clinically relevant thresholds calculated from ROC curve analysis. **D)** ROC curves.

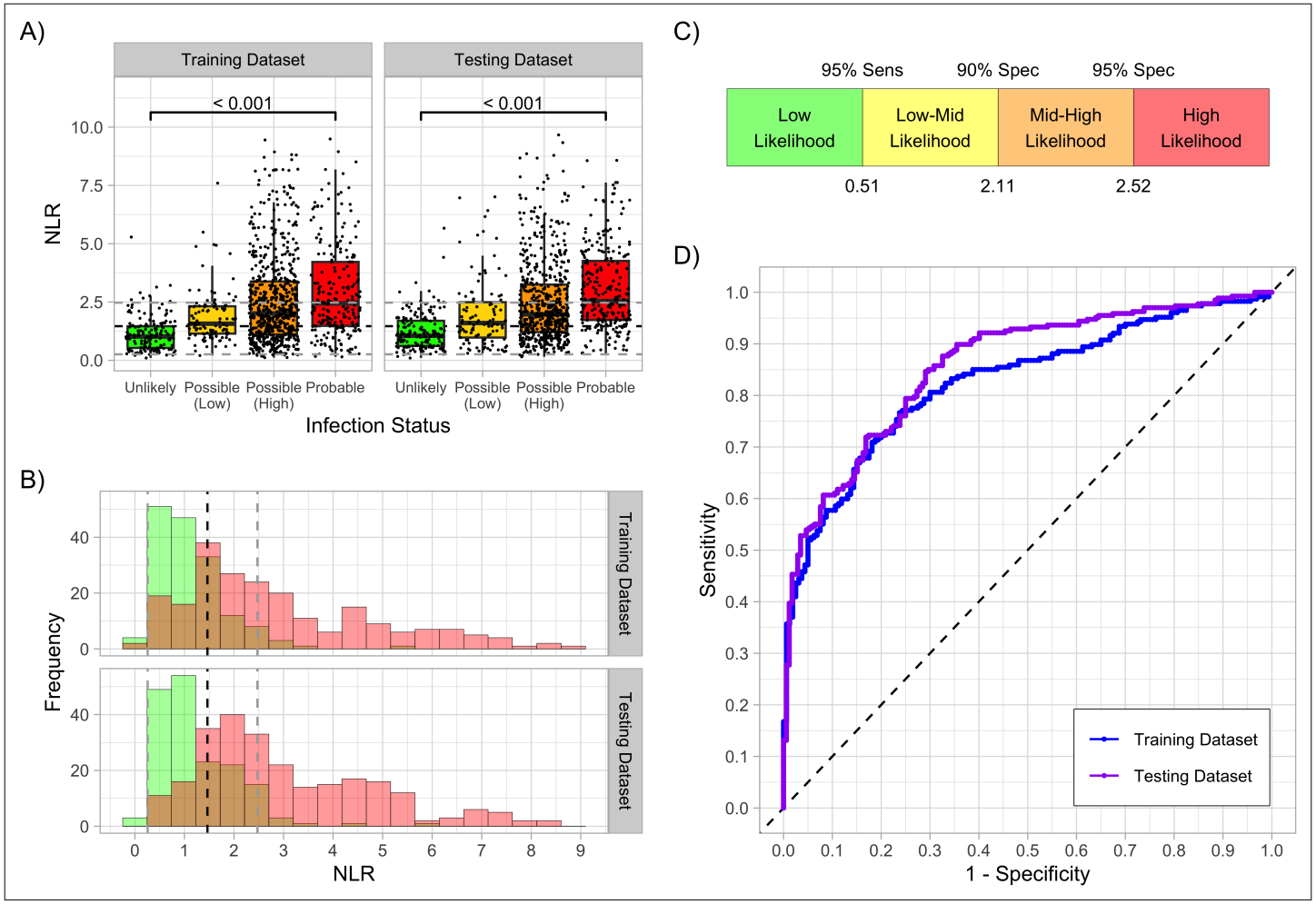


Figure 4.3: Evaluation of the diagnostic accuracy of neutrophil-to-lymphocyte ratio for late-onset sepsis.

A) Box plot of values with clinical suspicion of sepsis. **B)** Histogram of *unlikely* (green) and *probable sepsis* (red) values with the optimal threshold (dashed black) and reference ranges (dashed grey). **C)** Clinically relevant thresholds calculated from ROC curve analysis. **D)** ROC curves.

To validate the predictive power of the optimal thresholds from the training dataset, sensitivity and specificity were recalculated (Figure 4.2; Table 4.2). Eight parameters demonstrated a sensitivity greater than 80%, including RET# and NLR (Figure 4.3), and seven parameters demonstrated a specificity greater than 80%, including WBC# (Figure 4.2) and MONO#. Only seven parameters showed significant differences in AUCs between the training and testing datasets (DeLong's $p < 0.05$), indicating that most thresholds were generalisable.

4.3.4 Parameter Selection for Algorithm Development

Parameters were selected for inclusion in algorithm development based on their assumed biological relevance in sepsis pathophysiology, statistical significance in the training dataset, and avoidance of redundancy. After applying these criteria, MONO# and IRF were removed due to a lack of significance. NRBC%, RET%, NEUT%, LYMPH%, and IG% were excluded in favour of their absolute counterparts to minimise redundancy. DELTA-HE was removed in favour of RET-HE, and HFLC in favour of AS-LYMPH. This process resulted in 18 parameters selected for algorithm development (Chapter 5):

- NRBC#
- RET#
- RET-HE
- NEUT#
- LYMPH#
- MONO#
- IG#
- NLR
- ITR
- NEUT-RI
- NEUT-GI
- NEUT-TW
- RE-LYMPH
- AS-LYMPH
- RE-MONO
- PLT#
- IPF
- PLR

4.4 Discussion

4.4.1 Summary of Findings

In this chapter, I evaluated the diagnostic utility of 49 CBC parameters for identifying LOS. Of these, 36 parameters demonstrated statistically significant differences between results classified as *unlikely* and *probable sepsis* in both the training and testing datasets, indicating consistent discriminatory potential across temporally distinct cohorts. NEUT%, LYMPH%, NLR, NEUT-RI, and NEUT-TW demonstrated strong diagnostic performance in both datasets, with AUCs exceeding 0.8. Overall, NLR emerged as the most diagnostically accurate individual parameter, with an AUC of 0.825 in the training dataset. At its optimal threshold, NLR correctly identified 76.7% of sepsis cases (sensitivity) and excluded 76.3% of non-sepsis cases (specificity).

Comparison of AUCs between the training and testing datasets using DeLong's test confirmed the generalisability of these findings, with only seven parameters exhibiting a statistically significant reduction in AUC in the testing dataset (DeLong's $p < 0.05$). These included RBC#, PLT#, and PLR, which may reflect differences in patient characteristics or other clinical variability. In contrast, the most diagnostically accurate parameters, such as NLR and NEUT-RI, retained their performance across both datasets, supporting their potential clinical value in sepsis detection.

Underlying biological mechanisms support these findings. The systemic inflammatory response to sepsis involves the activation of the innate immune system and dysregulation of the adaptive immune response, resulting in neutrophilia and lymphopenia [72]. This shift explains the high diagnostic accuracy of NLR in the current study. Similarly, increases in NEUT-RI, a marker of neutrophil activation, and NEUT-TW, which describes the heterogeneity of the neutrophil population, align with the established evidence of a push towards the innate immune response during a septic episode.

Conversely, the lower diagnostic accuracy of PLT# in the testing dataset is surprising, as thrombocytopenia is a usual indicator of sepsis [141]. One possible explanation for this

observation could be the differences in patient characteristics between the training and testing datasets. CBC results categorised as *probable sepsis* were from older neonates in the testing dataset than the training dataset (median age at sampling: 19 and 24 days, respectively). In Chapter 3, PLT# was significantly impacted by corticosteroid exposure, sex, and the interaction between postnatal age and GA. The current findings further highlight the complex and nuanced interpretation of CBC values in neonates.

4.4.2 Comparison with Existing Literature

These findings align partially with previous research. For example, the diagnostic metrics calculated here for WBC# and NEUT# are consistent with those reported in the earlier systematic review and meta-analysis (Chapter 2) [76]. Previously, NEUT# demonstrated a pooled sensitivity and specificity of 67% and 76%, respectively, with an AUC of 0.783 and an optimal threshold of 4.13 x 10^9/L [76, 92]. In comparison, the current analysis identified a higher optimal threshold of 5.46 x 10^9/L, which yielded over 70% sensitivity and approximately 75% specificity in both datasets, with AUCs approaching 0.8, suggesting a higher diagnostic performance in the current study.

Neutrophil activation markers also performed well in this context, consistent with previous findings. NEUT-RI, a Sysmex-derived activation marker, achieved AUCs of 0.812 (training dataset) and 0.788 (testing dataset), comparable to the high diagnostic accuracy previously reported for MNV, a Beckman Coulter equivalent. The meta-analysis found that MNV had an AUC of 0.925, with a pooled sensitivity and specificity of 77% and 91%, respectively. However, substantial heterogeneity was noted ($I^2 = 0.85$), which limited the generalisability of the pooled estimates [76]. The current research confirms that these immune markers may be strong tools for the detection of LOS.

By contrast, NLR demonstrated substantially higher diagnostic utility in the present cohort compared to the earlier meta-analysis (Chapter 2), which reported a pooled AUC of 0.685, with 69% sensitivity and 73% specificity at an optimal threshold of 3.17 [76, 100]. In the current analysis, an optimal threshold of 1.46 yielded balanced diagnostic performance, with both sensitivity and specificity near 76%, and a significantly improved AUC of 0.825.

This lower threshold may reflect the detection of earlier or subtler inflammatory responses in this cohort, possibly due to differences in timing or population characteristics.

Comparisons with individual studies further support this theory. One such study reported an AUC of 0.764 for NLR and a sensitivity of 63% [99], both of which are notably lower than those observed in this study. Conversely, MPV, which was previously reported to have excellent diagnostic performance (AUC: 0.89; sensitivity: 78%; specificity: 98%) [98], showed considerably lower accuracy in this study (AUC: 0.64; sensitivity: 69.6%; specificity: 54.9%). Such discrepancies likely reflect differences in sample size (e.g., 45 cases and 45 controls in the prior study), population demographics, or methodological aspects such as timing of blood sampling or diagnostic definitions. The present analysis offers improved statistical power due to its larger and more representative cohort.

Overall, the current findings reinforce the diagnostic relevance of several CBC parameters, particularly NLR and NEUT-RI, while highlighting the variability in reported performance across the literature. This underscores the importance of validating proposed biomarkers in representative, temporally distinct cohorts before considering their clinical adoption.

4.4.3 Implications of Findings

The current study illustrates the potential of the CBC as an adjunctive tool in diagnosing neonatal sepsis. While no single parameter demonstrated sufficient accuracy for use as a standalone diagnostic test, as concluded in the meta-analysis (Chapter 2) [76], several showed strong and consistent performance. Immune activation markers such as NLR and NEUT-RI in particular offer promising diagnostic utility, suggesting that routinely collected haematological profiles could play a valuable role in early sepsis evaluation.

Despite their omission from the NICE guidelines [11, 77], several CBC parameters demonstrated diagnostic accuracy comparable to established biomarkers. For instance, a recent meta-analysis reported that a CRP threshold of 10 mg/L identified neonatal sepsis cases with a sensitivity of 69%, a specificity of 81%, and an AUC of 0.81 [142]. In the current analysis, NLR, NEUT-RI, and four other immune activation markers achieved

AUCs of at least 0.8, reinforcing the conclusion from Chapter 2 that current guidelines may overlook valuable diagnostic information available within routine CBC panels [76]. Unfortunately, no CBC parameters met the criteria for the ideal diagnostic tool for neonatal sepsis [18, 19].

Overall, these findings suggest that the diagnostic value of the CBC lies in its ability to capture dynamic immune responses through activation-based metrics, rather than traditional parameters such as absolute counts. Parameters like NLR and NEUT-RI encapsulate this shift, offering opportunities for earlier and more nuanced detection of sepsis. Further research is required to assess the longitudinal utility of these parameters and to evaluate whether continuous monitoring of these parameters using point-of-care devices could provide early detection of LOS.

4.4.4 Strengths and Limitations

Many of the strengths and limitations of this research reflect those described in Chapter 3. The CBC results analysed here were derived from a large, real-world dataset, enhancing generalisability and clinical relevance. Stratification into training and testing subsets enabled temporal validation, thereby strengthening the validity of the conclusions.

Strict classification criteria based on clinical suspicion of sepsis helped to separate *unlikely* and *probable sepsis* cases, addressing a common limitation in previous studies that used more subjective, inconsistent, or overlapping definitions. Only CBC results taken within 48 hours of a positive culture were classified as *probable sepsis*, while results from hospitalised neonates without signs of infection were classified as *unlikely sepsis*. This approach created delineated groups, reducing overlap between cases and controls.

As in the previous chapter, the primary limitation is the retrospective study design, which introduces the risk of bias due to incomplete or inconsistently recorded clinical information. For example, it was not feasible to review every positive culture individually to rule out contamination (e.g., in cases with a single positive CoNS result). Additionally, the completeness of clinical documentation, such as antibiotic administration or the presence

of infection-related clinical signs, could not be verified, which may have introduced misclassification in a minority of cases.

Although it was not feasible in the present study due to the limited sample size, future research should investigate whether alternative diagnostic thresholds are required for different subpopulations. As demonstrated in Chapter 3, most CBC parameters are influenced by age, GA, sex, or a combination of these factors. Tailoring thresholds to account for these variables may enhance the diagnostic utility of the CBC for LOS and improve clinical implementation.

4.5 Chapter Summary

This chapter established the diagnostic accuracy of 49 CBC parameters for identifying LOS and confirmed these findings using a temporally distinct validation cohort. Several parameters, including NLR and NEUT-RI, demonstrated strong and consistent diagnostic performance, offering new evidence in support of their potential use in sepsis screening. These findings represent the first evaluation of many of these parameters in this context and extend the conclusions of the earlier systematic review (Chapter 2) that the CBC has an underappreciated clinical utility in the diagnosis of LOS. Building on this foundation, the next chapter will integrate the most diagnostically valuable parameters into multivariable models aimed at improving the detection of neonatal sepsis.

Chapter 5

Development and Validation of Diagnostic Algorithms for Late-Onset Neonatal Sepsis

5.1 Introduction

The results presented in this chapter have been submitted to a peer-reviewed journal [Reference to be added upon publication].

5.1.1 Pre-Existing Diagnostic Scores

While the current diagnostic approach for LOS often relies on individual laboratory or clinical markers, growing evidence suggests that combining multiple parameters can improve diagnostic accuracy. One prominent example is the Kaiser Permanente EOS calculator, introduced in Chapter 1, which integrates maternal risk factors to estimate a newborn's probability of EOS [143]. Although it is the most widely used composite model in clinical settings [11], numerous other tools have been proposed in the literature that incorporate both clinical and laboratory data to support the diagnosis of LOS.

The Töllner score, developed in 1982, was one of the earliest attempts to formalise such a diagnostic tool for neonatal sepsis [144]. It combines clinical, biochemical, and haematological findings into a single composite score (Table 5.1). Reported performance metrics vary, with sensitivity ranging from 41.7% to 91.7% and specificity from 69.0% to 87.5%, depending on the threshold applied [145, 146]. Despite requiring inputs from multiple sources, the Töllner score established an essential precedent for multivariable diagnostic tools.

In 1988, Rodwell et al. introduced the Haematological Scoring System (HSS), a tool based solely on CBC and blood smear features, including WBC#, IG#, neutrophil ratios, and PLT# (Table 5.1) [122]. Using a threshold of ≥ 3, the model demonstrated high

Table 5.1: Diagnostic performance of published scoring systems for late-onset sepsis.

Diagnostic Model	Parameters	Sepsis Definition	Study	AUC	Threshold	Sens (%)	Spec (%)	PPV (%)	NPV (%)
Tollner Score [144]	11 variables, including WBC#, left shift, and PLT#	Positive culture	[145]	0.969	11.8	91.7	87.5	91.7	87.5
			[146]	–	5.0	41.7	69.0	31.3	77.8
HSS [122]	WBC#, NEUT#, IG#, ITR, IMR, PLT#, and neutrophil changes	Positive culture	[122]	–	3.0	96.0	78.0	31.0	99.0
			[123]	–	3.0	79.0	39.0	48.0	72.0
			[146]	–	5.0	29.4	100.0	100.0	77.8
NOSEP Score [147]	5 variables, including NEUT# and PLT#	Positive culture	[147]	0.820	11.0	60.0	84.0	72.0	75.0
			[148]	0.660	11.0	73.0	57.0	67.0	63.0
Predictive Clinical Score [149]	12 clinical variables	Positive culture	[149]	–	1.0	87.0	29.0	38.0	85.0
			[150]	–	1.0	90.0	22.5	30.3	85.7
			[151]	–	1.0	56.6	52.1	78.1	28.4
Bedside Prediction Scoring Model [152]	6 variables, including band count and PLT#	Suspected	[152]	0.800	4.0	92.0	56.0	56.0	90.0
		Positive culture	[153]	0.780	4.0	97.0	39.0	43.0	96.0
Predictive Score [154]	6 variables, including WBC#	Positive culture	[154]	–	2.0	88.5	90.4	75.4	95.9
				–	3.0	82.7	93.6	81.1	94.2
Sepscore [123]	WBC#, NEUT#, IG#, ITR, PLT#, NRBC#, and neutrophil changes	Suspected with SIRS	[123]	–	3.0	68.0	61.0	56.0	73.0

AUC values > 0.8 and sensitivities/specificities $> 85\%$ are highlighted in green, suggesting strong diagnostic performance. Abbreviations: AUC – area under the curve; HSS – Haematological Scoring System; IG# – immature granulocyte count; IMR – immature-to-mature neutrophil ratio; ITR – immature-to-total neutrophil ratio; NEUT# – neutrophil count; NPV – negative predictive value; NRBC# – nucleated red blood cell count; PLT# – platelet count; PPV – positive predictive value; Sens – sensitivity; Spec – specificity; SIRS – systemic inflammatory response syndrome; WBC# – white blood cell count.

sensitivity (96%) and moderate specificity (78%). However, performance varied by population and GA, with the best performance in preterm neonates on day one of life (sensitivity: 100%; specificity: 82%).

This model was further refined in 2023 to produce the Sepscore – a modified version of the HSS that eliminates redundant parameters and incorporates NRBC# to improve specificity (Table 5.1) [123]. Although both models performed modestly, the Sepscore demonstrated higher specificity (61% versus 54%) and overall better performance in the evaluated cohort. However, its sensitivity was slightly lower (68% versus 71%). Notably, only 16.4% (24/146) of the case group were culture-positive, raising questions about the actual diagnostic validity of these scores.

In addition to the Töllner score, HSS, and Sepscore, several other diagnostic tools have been proposed for neonatal sepsis [147, 149, 152, 154]. However, despite promising results, most lack external validation and have not been adopted into clinical practice. A summary of these tools is presented in Table 5.1.

5.1.2 Machine Learning-Based Tools

Alongside traditional scoring systems, recent efforts have focused on machine learning approaches that leverage CBC and clinical data to enhance patient care. One study, identified in the database search in Chapter 2, developed five models based on 60 CBC parameters and nine clinical features (Table 5.2) [124]. While three models achieved AUCs greater than 0.7, none demonstrated sensitivities greater than 65%. Furthermore, although these models outperformed those based on clinical features alone, the lack of transparency regarding which haematological variables were retained in the final models limits both their reproducibility and interpretability.

A recent systematic review of machine learning-based diagnostic models for neonatal sepsis reported a pooled AUC of 0.94, with a pooled sensitivity and specificity of 87% and 89%, respectively (Table 5.2) [155]. The most frequently used machine learning technique employed to develop these models was the random forest, and WBC#, PLT#,

Table 5.2: Diagnostic performance of published machine learning models for late-onset sepsis.

Study	Parameters	Model Structure	AUC	Sens (%)	Spec (%)	PPV (%)	NPV (%)	Note
[155]	Meta-analysis of 19 studies reporting on 76 diagnostic models		0.940	87.0	89.0	–	–	–
[124]	Initial: 69 variables, including 60 CBC parameters	Logistic Regression	0.740	62.0	73.0	48.0	83.0	Includes culture-proven and suspected sepsis
		k-Nearest Neighbour	0.620	45.0	73.0	40.0	77.0	
		Random Forest	0.730	38.0	88.0	57.0	78.0	
		Extreme Gradient Boost	0.710	40.0	85.0	53.0	78.0	
		Support Vector Mechanism	0.670	13.0	95.0	55.0	73.0	
[156]	Initial: 93 variables, including 18 CBC parameters Final: 14 – 69 variables (model dependent)	Naïve Bayes	0.780	73.0	47.0	59.0	62.0	Sensitivity fixed at 75.0% (physicians sensitivity)
		Random Forest	0.650	73.0	47.0	59.0	62.0	
		Support Vector Mechanism	0.680	73.0	26.0	51.0	47.0	
		k-Nearest Neighbour	0.620	73.0	29.0	52.0	50.0	
		Classification and Regression Trees	0.770	73.0	30.0	52.0	51.0	
		Logistic Regression	0.610	73.0	30.0	52.0	51.0	
		Lazy Bayesian Rules	0.580	73.0	30.0	52.0	51.0	
		Averaged One Dependence Estimators	0.530	73.0	28.0	52.0	49.0	
		Tree Augmented Naïve Bayes	0.530	73.0	23.0	50.0	45.0	

Continued on next page

Continued from last page

Study	Parameters	Model Structure	AUC	Sens (%)	Spec (%)	PPV (%)	NPV (%)	Note
[157]	Initial: 36 variables, including 4 CBC parameters	AdaBoost	0.830	80.0	71.0	23.0	97.0	Sensitivity fixed at 80.0% to facilitate comparisons
		Gradient Boost	0.800	80.0	69.0	23.0	98.0	
		Gaussian Process	0.750	80.0	53.0	16.0	97.0	
		k-Nearest Neighbour	0.730	80.0	20.0	13.0	29.0	
		Logistic Regression	0.830	80.0	71.0	23.0	97.0	
		Naïve Bayes	0.810	80.0	69.0	22.0	98.0	
		Random Forest	0.820	80.0	71.0	23.0	98.0	
		Support Vector Mechanism	0.820	80.0	72.0	23.0	98.0	
[158]	Initial: 14 variables, including 6 CBC parameters	Logistic Regression	0.766	60.4	84.2	32.2	94.5	Testing dataset results presented
		Artificial Neural Network	0.795	68.8	79.3	29.2	95.3	
		Multiple Discriminant Analysis	0.751	64.6	78.8	27.4	94.7	
[159]	Final: 25 variables, including 6 CBC parameters	Artificial Neural Network	0.944	93.3	80.0	82.4	92.3	–

AUC values $>$ 0.8 and sensitivities/specificities $>$ 85% are highlighted in green, suggesting strong diagnostic performance. Abbreviations: AUC – area under the curve; CBC – complete blood count; NPV – negative predictive value; PPV – positive predictive value; Sens – sensitivity; Spec – specificity.

and ITR were the most commonly retained CBC features. Despite encouraging results, many models displayed overfitting, poor external validation, or a lack of transparency regarding outcome calculation, which are substantial barriers to clinical implementation.

5.1.3 Aims of the Chapter

This chapter aims to develop and validate multiple diagnostic models for LOS using a range of statistical and machine learning techniques. Building on findings from Chapter 4, these models integrate CBC parameters with strong diagnostic performance into multivariable frameworks. Six models were constructed and tested for both temporal and external validity to evaluate diagnostic performance, robustness, and potential clinical utility.

5.2 Methods

5.2.1 Data Collection

The dataset used in this chapter is the same as described in Chapter 3 and Chapter 4. Ethical approval was obtained from the Oxford Research Ethics Committee (REC 08/H0606/139), as described in Chapter 3. An opt-out consent process was used, and data were anonymised before analysis to ensure confidentiality.

In summary, neonatal CBC results measured on a Sysmex XN-2000 analyser (Kobe, Japan) in the haematology laboratory at the John Radcliffe Hospital (Oxford, UK) and an XN-450 point-of-care analyser in the NICU between 13$^{\text{th}}$ January 2022 and 13$^{\text{th}}$ January 2023 were retrospectively collected. CBC results were matched to clinical data extracted from BadgerNet, EPR, and clinical observation charts. Results were subsequently categorised based on the presence of clinical or laboratory signs of sepsis, as described in Chapter 3 (Figure 3.1, p. 68).

In the previous chapter, CBC results from the first three days of life were excluded to focus specifically on LOS. Before applying the exclusion criteria, the dataset was stratified into training and testing subsets based on sex, GA, and duration of NICU stay, ensuring statistical comparability between the resulting groups. The effectiveness of the stratification was assessed using Student's t-tests, with p-values greater than 0.05 indicating no significant differences between subsets.

5.2.2 Diagnostic Algorithm Development

All statistical analyses were performed using RStudio (R version 4.3.0) [86]. Six models were developed using the parameters selected in Chapter 4 and values from CBC results categorised as *unlikely* and *probable sepsis* in the training dataset. For each model, the predictive variables were the selected CBC parameters, and the outcome variable was the presence (*probable sepsis*) or absence (*unlikely sepsis*) of infection. Diagnostic accuracy was tested using ROC curve analysis, with key diagnostic accuracy metrics calculated.

Five modelling techniques were selected to reflect different priorities in diagnostic algorithm development. For example, simpler models, such as a logistic regression or decision tree, prioritise transparency and interpretability, while more complex models, such as the random forest, offer improved predictive power. Similarly, including both automated and manually constructed decision trees allowed for assessment of the trade-offs between statistical optimisation and clinical interpretability.

Each model was constructed using the selected parameters, followed by an alternative version incorporating postnatal age and GA. Each pair of models was compared using DeLong's Test to determine whether the inclusion of clinical information improved the overall accuracy of the model.

Logistic Regression

A multivariable logistic regression model was developed using the *glm* function and the *stepAIC* function from the *MASS* package. Initially, all parameters were included, and a stepwise selection process iteratively removed or reintroduced variables. The model was optimised by minimising the Akaike Information Criterion (AIC). Variables that contributed to the lowest AIC were retained in the final model. The outcome variable was binary (*probable* versus *unlikely sepsis*), and the model assumed a binomial distribution. Multicollinearity was assessed using Variance Inflation Factors (VIFs) for all predictors included in the final logistic regression model, with VIF values less than five indicating acceptable levels of independence between predictors.

Computer-Generated Decision Tree

A classification model was developed using the *rpart* function from the *rpart* package, with the method specified as "class" for classification. The model recursively partitions the data into subsets based on the values of the predictor variables, aiming to maximise homogeneity of the outcome within each subgroup. To prevent overfitting, the tree was pruned by selecting the complexity parameter (cp) that minimised the cross-validation error, identified from the model's cp table. Pruning reduces model complexity by removing branches with minimal contributions.

Manually Constructed Decision Tree

A second decision tree model was developed manually, with predictor variables and their respective cut-offs at each node selected based on exploratory data analysis. This approach enabled greater alignment with reference ranges. At each decision point, multiple parameters were sampled to identify the parameter producing the best split between groups. Larger end nodes were introduced to prevent the model from overfitting to the training dataset.

Random Forest

A random forest model was developed using the *randomForest* function from the *randomForest* package. The model was configured to generate 100 decision trees ($ntree = 100$), with three predictor variables randomly sampled at each node ($mtry = 3$), and a minimum terminal node size of 25 ($nodesize = 25$). By aggregating predictions from multiple decision trees, the random forest offers improved classification accuracy and robustness compared to single-tree models.

Scoring System

A scoring system was developed based on the values for various parameters. The thresholds for scoring were determined using ROC analysis. For each parameter, three levels of contribution to the overall score were specified:

- Values exceeding the optimal threshold received a score equal to the AUC.
- Values exceeding the threshold for 85% specificity received a score of 2 $\times$ AUC.
- Values exceeding the threshold for 95% specificity received a score of 4 $\times$ AUC.

The final diagnostic score for each observation was the sum of the individual scores across all parameters. Initially, all 18 parameters were included, and an iterative selection process was used to remove those that did not improve the AUC, optimising the model's overall performance.

k-Nearest Neighbour

A k-nearest neighbour (kNN) model was developed to classify new observations based on their similarity to points in the training dataset. The model calculates the distance between the latest observation and each known data point. The new observation is classified according to the majority class of the k nearest neighbours. The value of k was set to the square root of the number of observations in the training dataset, rounded to the nearest integer, and the Euclidean distance was used to measure how close observations were to each other. This approach relies on comparing new data directly to the training dataset without explicit training. While kNN models often benefit from feature standardisation, preliminary analysis indicated that standardising CBC parameters reduced model performance. This likely reflects the inherent clinical importance of absolute parameter values. Consequently, all features were used in their original scales.

5.2.3 Diagnostic Algorithm Validation

The algorithms developed with the training dataset were then validated internally with the testing dataset. For external validation, a longitudinal dataset from the Erasmus University Medical Centre (EMC; Rotterdam, The Netherlands) was used. This cohort included neonates born before 32 weeks' gestation and admitted to the NICU between January 2018 and October 2021 for at least three days. CBC results and matched clinical data from postnatal day four onwards were extracted from EPR. Physiological parameters (e.g., heart rate) were averaged hourly to calculate the clinical deterioration score. CBC results were categorised according to the same criteria as the original dataset.

CBC results categorised as *unlikely* or *probable sepsis* were reclassified by applying the model and comparing predicted values to the optimal diagnostic thresholds. Standard diagnostic accuracy metrics (e.g. sensitivity, specificity, PPV, and NPV) were recalculated using the TP, TN, FP, and FN values from each dataset. To assess the generalisability of each algorithm's diagnostic performance, ROC curves from the training and validation datasets were compared using DeLong's test.

5.3 Results

5.3.1 Sample Characteristics

This chapter utilises a subset of the dataset described in Chapter 4, specifically focusing on CBC results with complete data for all 18 selected parameters. Eighteen CBC results were excluded due to missing data: six from the training dataset and twelve from the testing dataset.

The patient characteristics in this subset closely matched those of the entire cohort. No significant differences were observed between the training and testing datasets in terms of sex, GA, or postnatal age at sampling (Table 5.3). However, neonates in the testing dataset had longer NICU stays ($p = 0.04$), and *probable sepsis* CBC results were from younger neonates ($p = 0.002$), as seen in the previous chapter.

The external validation dataset comprised 2,638 CBC results from 474 neonates (Table 5.3). The median GA was lower than the training dataset ($p < 0.001$), while postnatal age at sampling was higher ($p < 0.001$). These differences reflect the inclusion criteria of the external cohort, which was limited to neonates born before 32 weeks' gestation.

5.3.2 Diagnostic Algorithm Development and Performance

Six diagnostic models were developed using the 18 selected CBC parameters: logistic regression model, computer-generated decision tree, manually constructed decision tree, random forest model, scoring system, and kNN model. Table 5.4 summarises key diagnostic metrics, including AUC, sensitivity, specificity, PPV, NPV, DOR, and Youden Index. Detailed model metrics, including 95% CIs and dataset-specific optimal thresholds, are presented in Appendix A4.

Logistic Regression

The logistic regression model included six predictors selected via stepwise AIC optimisation (Table 5.5). NEUT# (OR: 1.40) and NEUT-RI (1.23) were the strongest positive predictors, while PLT# (0.99) and RET-HE (0.91) were inversely associated with sepsis

Table 5.3: Sample characteristics of complete blood count results used for the development and validation of diagnostic models for late-onset sepsis, collected from neonates admitted to the Neonatal Intensive Care Units at the John Radcliffe Hospital (Oxford, United Kingdom) and Erasmus University Medical Centre (Rotterdam, The Netherlands).

	Training Dataset (JR)			Testing Dataset (JR)			Validation Dataset (EMC)		
	All Values	*Unlikely Sepsis*	*Probable Sepsis*	All Values	*Unlikely Sepsis*	*Probable Sepsis*	All Values	*Unlikely Sepsis*	*Probable Sepsis*
CBC Results	1,158	159	226	1,154	172	266	2,638	1,738	399
Neonates	168	83	43	136	76	38	474	412	142
GA (weeks)	32.1 (27.4 – 37.6)	29.6 (26.2 – 35.1)	26.7 (24.6 – 30.4)	31.1 (26.8 – 36.2)	29.9 (26.7 – 34.5)	26.0 (24.6 – 28.8)	28.0[a] (26.3 – 29.7)	28.1 (26.3 – 29.7)	26.6 (25.3 – 28.7)
Males (%)	94 (56.0%)	46 (55.4%)	22 (51.2%)	80 (58.8%)	47 (61.8%)	25 (65.8%)	277 (58.4%)	243 (59.0%)	86 (60.6%)
Duration of Stay (days)	20 (10 – 42)	30 (14 – 50)	47 (30 – 72)	24[a] (12 – 55)	32 (15 – 71)	60 (38 – 84)	–	–	–
Postnatal Age (days)	19 (9 – 37)	26 (11 – 48)	19 (10 – 33)	21 (10 – 40)	24[a] (10 – 52)	24 (14 – 43)	29[a] (12 – 68)	31 (13 – 71)	31 (13 – 65)

Values are presented as counts or median (interquartile range). [a]*Statistically significant difference ($p < 0.05$) from the training dataset, highlighted in red. Abbreviations: CBC – complete blood count; EMC – Erasmus University Medical Centre; GA – gestational age at birth; JR – John Radcliffe Hospital.*

Table 5.4: Diagnostic accuracy of six diagnostic models for late-onset sepsis.

Model	Dataset	AUC	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	LR^+	LR^-	DOR	Youden Index	DeLong P-Value
Logistic Regression	Training	0.923	82.7	89.3	91.7	78.5	7.74	0.19	40.05	0.72	–
	Testing	0.914	83.8	85.5	89.9	77.4	5.77	0.19	30.49	0.69	0.63
	Validation	0.818	86.5	63.6	50.5	91.6	2.38	0.21	11.18	0.50	< 0.001
Computer-Generated Decision Tree	Training	0.918	90.3	83.7	88.7	85.8	5.52	0.12	47.43	0.74	–
	Testing	0.807	85.0	57.0	75.3	71.0	1.97	0.26	7.48	0.42	< 0.001
	Validation	0.786	89.1	53.0	44.9	91.9	1.90	0.21	9.19	0.42	< 0.001
Manually Constructed Decision Tree	Training	0.864	75.2	83.0	86.3	70.2	4.43	0.30	14.84	0.58	–
	Testing	0.864	77.4	82.0	86.9	70.1	4.30	0.28	15.62	0.59	1.00
	Validation	0.788	77.1	66.5	49.7	87.1	2.30	0.34	6.70	0.44	< 0.001
Random Forest	Training	0.988	95.1	95.0	96.4	93.2	18.91	0.05	368.92	0.90	–
	Testing	0.911	86.8	71.5	82.5	77.8	3.05	0.18	16.57	0.58	< 0.001
	Validation	0.852	91.0	59.6	49.1	93.9	2.25	0.15	14.85	0.51	< 0.001
Scoring System	Training	0.906	77.0	91.2	92.6	73.6	8.74	0.25	34.66	0.68	–
	Testing	0.879	77.8	83.7	88.1	70.9	4.78	0.26	18.04	0.62	0.23
	Validation	0.813	80.9	68.6	52.5	89.3	2.57	0.28	9.26	0.50	< 0.001
K-Nearest Neighbour	Training	0.836	72.1	78.6	82.7	66.5	3.37	0.35	9.51	0.51	–
	Testing	0.716	65.4	64.0	73.7	54.5	1.81	0.54	3.36	0.29	< 0.001
	Validation	0.751	71.6	66.0	47.5	84.4	2.10	0.43	4.89	0.38	< 0.001

Significant p-values (< 0.05) are highlighted in red. AUC values > 0.8 and sensitivities/specificities $> 85\%$ are highlighted in green, suggesting strong diagnostic performance. Abbreviations: AUC – area under the curve; DOR – diagnostic odds ratio; LR^+ – positive likelihood ratio; LR^- – negative likelihood ratio; NPV – negative predictive value; PPV – positive predictive value.

risk. PLR produced an enormous odds ratio (OR: 6,706), likely reflecting skewed scaling or distributional artefacts. Despite moderate collinearity for PLT# (VIF: 2.98) and NEUT-RI (3.08), all predictors were retained.

The model performed well internally, with AUCs of 0.923 (training) and 0.914 (testing). External validation showed a moderate decline in performance (AUC: 0.818) (Table 5.4; Figure 5.3). The DOR remained elevated across all datasets, suggesting consistent discriminative ability (40.05, 30.49, and 11.18, respectively). Although sensitivity and PPV declined in the external cohort, the NPV remained high (91.6%), supporting its potential use as a rule-out tool. Re-evaluation using a dataset-specific threshold yielded minimal improvements in the validation dataset (Table A4.1, p. 263).

Computer-Generated Decision Tree

The computer-generated decision tree model consisted of nine internal splits and ten terminal nodes, based on eight parameters (Figure 5.1). The most influential predictors were NLR (55.92), NEUT-TW (41.43), and NEUT# (41.38), while NEUT-RI (19.28) and IPF (7.46) had lower influence. IG# and LYMPH# were excluded despite high initial importance scores (22.93 and 22.43, respectively).

Initial model performance was high (AUC: 0.918 in training), but fell sharply (AUC: 0.807 in testing and 0.786 in validation; DeLong's $p < 0.001$), primarily due to reduced specificity (83.7%, 57.0%, and 53.0%, respectively) (Table 5.4; Figure 5.4). Adjusting thresholds improved specificity in the testing dataset from 57% to 75%, and DOR from 7.48 to 11.00 (Table A4.2, p. 264). Nonetheless, the model remained prone to overfitting.

Manually Constructed Decision Tree

This model was constructed using six parameters selected through exploratory analysis and clinical reasoning. The model prioritised generalisability by retaining large terminal nodes (Figure 5.2). Performance was consistent in the internal datasets (AUC: 0.864 for both training and testing) but dropped in validation (AUC: 0.788) due to reduced specificity (83.0%, 82.0%, and 66.5%, respectively; Table 5.4).

Table 5.5: Structure of a logistic regression diagnostic model for late-onset sepsis.

Parameter	Estimate	Standard Error	P-Value	OR (95% CI)	VIF
Intercept	-8.28	2.282	< 0.001	–	–
RET-HE	-0.09	0.044	0.04	0.91 (0.83 – 1.00)	1.44
NEUT#	0.29	0.058	< 0.001	1.34 (1.23 – 1.45)	1.40
MONO#	0.35	0.201	0.08	1.42 (1.03 – 1.82)	1.09
NEUT-RI	0.21	0.037	< 0.001	1.23 (1.16 – 1.31)	3.08
PLT#	-0.01	0.001	< 0.001	0.99 (0.99 – 1.00)	2.98
PLR	8.81	3.416	0.01	6,706 (6,699 – 6,713)	1.19

Significant p-values (< 0.05) are highlighted in red. Abbreviations: CI – confidence interval; MONO# – monocyte count; NEUT# – neutrophil count; NEUT-RI – neutrophil reactive index; OR – odds ratio; PLR – platelet-to-lymphocyte ratio; PLT# – platelet count; RET-HE – reticulocyte haemoglobin equivalent; VIF – Variance Inflation Factor.

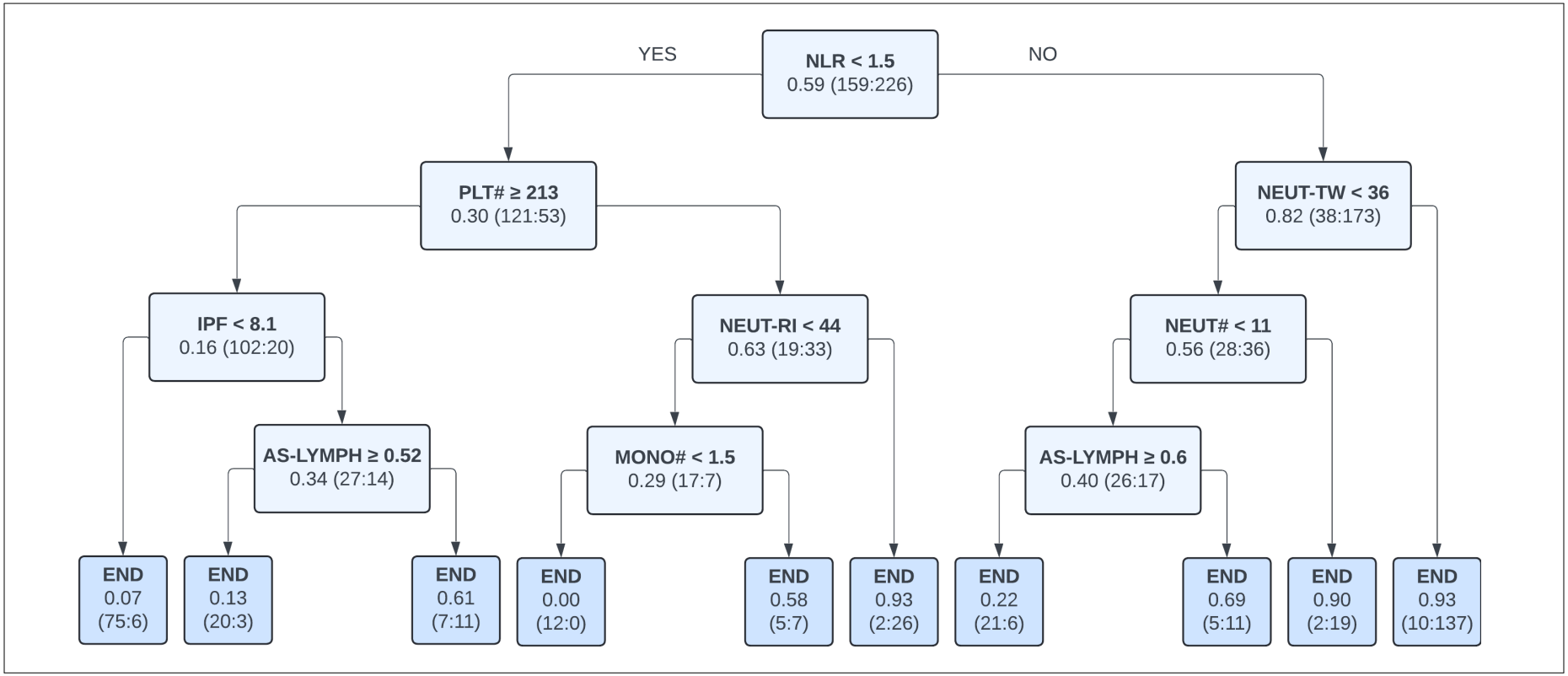


Figure 5.1: Structure of a computer-generated decision tree diagnostic model for late-onset sepsis. Each node displays the proportion of included results classified as *probable sepsis*, along with the ratio of *unlikely* to *probable sepsis* results.

Abbreviations: AS-LYMPH – antibody-synthesising lymphocyte count; IPF – immature platelet fraction; MONO# – monocyte count; NEUT# – neutrophil count; NEUT-RI – neutrophil reactive index; NEUT-TW – neutrophil distribution width; NLR – neutrophil-to-lymphocyte ratio; PLT# – platelet count.

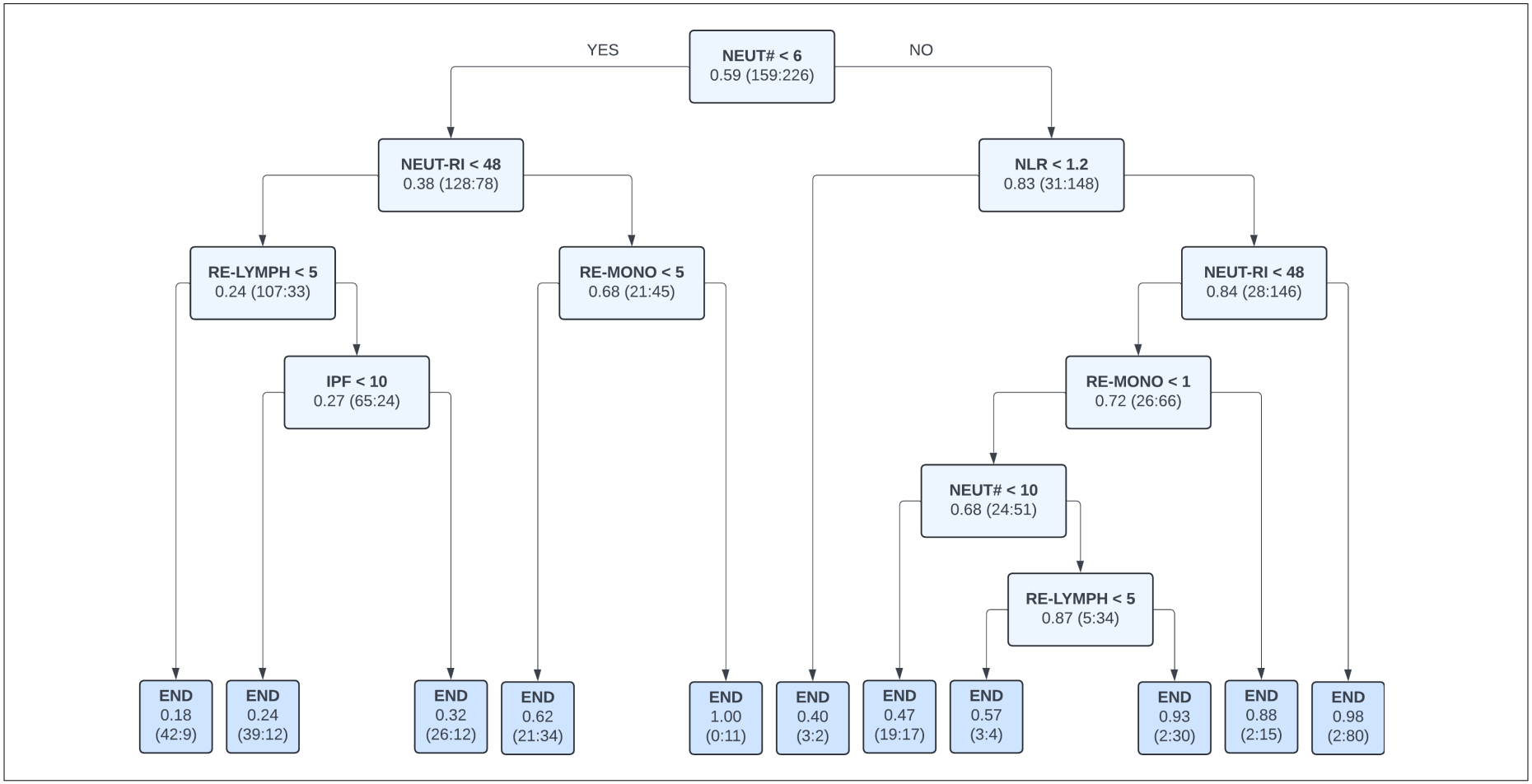


Figure 5.2: Structure of a manually calculated decision tree diagnostic model for late-onset sepsis. Each node displays the proportion of included results classified as *probable sepsis*, along with the ratio of *unlikely* to *probable sepsis* results.

Abbreviations: IPF – immature platelet fraction; NEUT# – neutrophil count; NEUT-RI – neutrophil reactive index; NLR – neutrophil-to-lymphocyte ratio; RE-LYMPH – reactive lymphocyte count; RE-MONO – reactive monocyte count.

Threshold optimisation improved specificity in the validation dataset from 66.5% to 81.8% but at the cost of reduced sensitivity, resulting in only modest overall improvement (DOR: 6.70 to 7.87; Table A4.3, p. 265). Despite its design to enhance generalisability compared to the computer-generated tree, external reliability remained limited.

Random Forest

The initial random forest model showed evidence of overfitting, likely due to the smaller sample size, resulting in poor generalisability. To address this, the model was reconfigured to 100 decision trees ($ntree = 100$), three predictors sampled per split ($mtry = 3$), and a minimum terminal node size of 25 ($nodesize = 25$). NLR (node purity: 9.19), NEUT# (8.63), and NEUT-RI (8.46) were the most informative predictors, while NEUT-GI (0.86) contributed least (Figure A4.1, p. 266). Performance stabilised after 30-40 trees (ΔMSE < 0.001 over 10 trees).

Performance remained strong across datasets, with AUCs of 0.988 (training), 0.911 (testing), and 0.852 (validation; Table 5.4; Figure 5.6). Sensitivity was consistently high, but specificity dropped in validation from 95.0% to 59.6%, reducing PPV from 96.4% to 49.1%. Threshold adjustment in the external validation dataset improved PPV (61.1%) and DOR (12.77; Table A4.4, p. 267). Despite declining specificity, the model demonstrated strong discrimination and generalisability.

Scoring System

All iterations of the scoring system performed well (AUCs > 0.8; Table A4.5, p. 268). The final model included eight parameters (Table A4.6); most were positively associated with sepsis risk, while PLT# appeared to be protective. NLR was the most predictive component (AUC: 0.825), while RET-HE was the least (AUC: 0.703).

The scoring system demonstrated strong overall performance (AUC: 0.906, 0.879, and 0.813, in training, testing and validation, respectively) (Table 5.4; Figure 5.7). Sensitivity and DOR also remained high, while specificity and PPV declined from training to

Table 5.6: Structure of a scoring system for late-onset sepsis.

Parameter	Optimal Threshold (1x AUC)	Threshold for 85% Specificity (2x AUC)	Threshold for 95% Specificity (4x AUC)	AUC
RET-HE (pg)	29.6	26.0	23.1	0.703
NEUT# (x 10^9/L)	5.46	6.39	8.68	0.796
IG# (x 10^9/L)	0.19	0.30	0.70	0.745
NLR (%)	1.46	1.80	2.52	0.825
NEUT-RI (FI)	46.1	48.0	51.3	0.812
NEUT-TW (au)	37.3	40.0	55.1	0.811
PLT# (x 10^9/L)	257	206	112	0.731
IPF (%)	7.5	17.8	25.5	0.740

Abbreviations: AUC – area under the curve; IG# – immature granulocyte count; IPF – immature platelet fraction; NEUT# – neutrophil count; NEUT-RI – neutrophil reactive index; NEUT-TW – neutrophil distribution width; NLR – neutrophil-to-lymphocyte ratio; PLT# – platelet count; RET-HE – reticulocyte haemoglobin equivalent.

validation. NPV was higher in validation (89% versus 74% in training and 71% in testing). Threshold adjustment had limited effects on performance (Table A4.6, p. 269).

k-Nearest Neighbour

The kNN model did not generate interpretable parameters, as classification is based on proximity to existing data. The value of k was set to the square root of the number of observations in the training dataset ($k = 20$). Performance was weaker than all other models, with AUCs of 0.836 (training), 0.716 (testing), and 0.751 (validation; Table 5.4; Figure 5.8). Specificity and PPV fell sharply in the external dataset. Threshold tuning did not improve performance (Table A4.7, p. 270), and overall diagnostic utility remained limited.

5.3.3 Comparison of Diagnostic Performance

Most models achieved strong internal performance (AUCs > 0.85), but generalisability varied (Table 5.4; Figure 5.9; Figure 5.10). In all models, performance declined between the training, testing, and validation datasets. Notably, three models (logistic regression, random forest, and the scoring system) retained an AUC above 0.8 in external validation, indicating more robust diagnostic performance.

The random forest model outperformed all other models in terms of sensitivity (95.1%, 86.8%, and 91.0%, respectively) and DOR (368.92, 16.57, and 14.85, respectively). However, specificity and PPV declined substantially in external validation. The logistic regression model and the scoring system also demonstrated consistent performance, with high NPVs supporting their use as rule-out tools. In contrast, both decision trees were prone to overfitting, and the kNN model consistently underperformed.

To explore the limits of model performance, the most extreme outliers from the training and testing datasets were reviewed (Table 5.7). Among the 12 cases classified as *unlikely sepsis*, eight were flagged by at least five models, likely reflecting infections missed by the reference classifications or post-infection leucocytosis. Conversely, all 12 *probable*

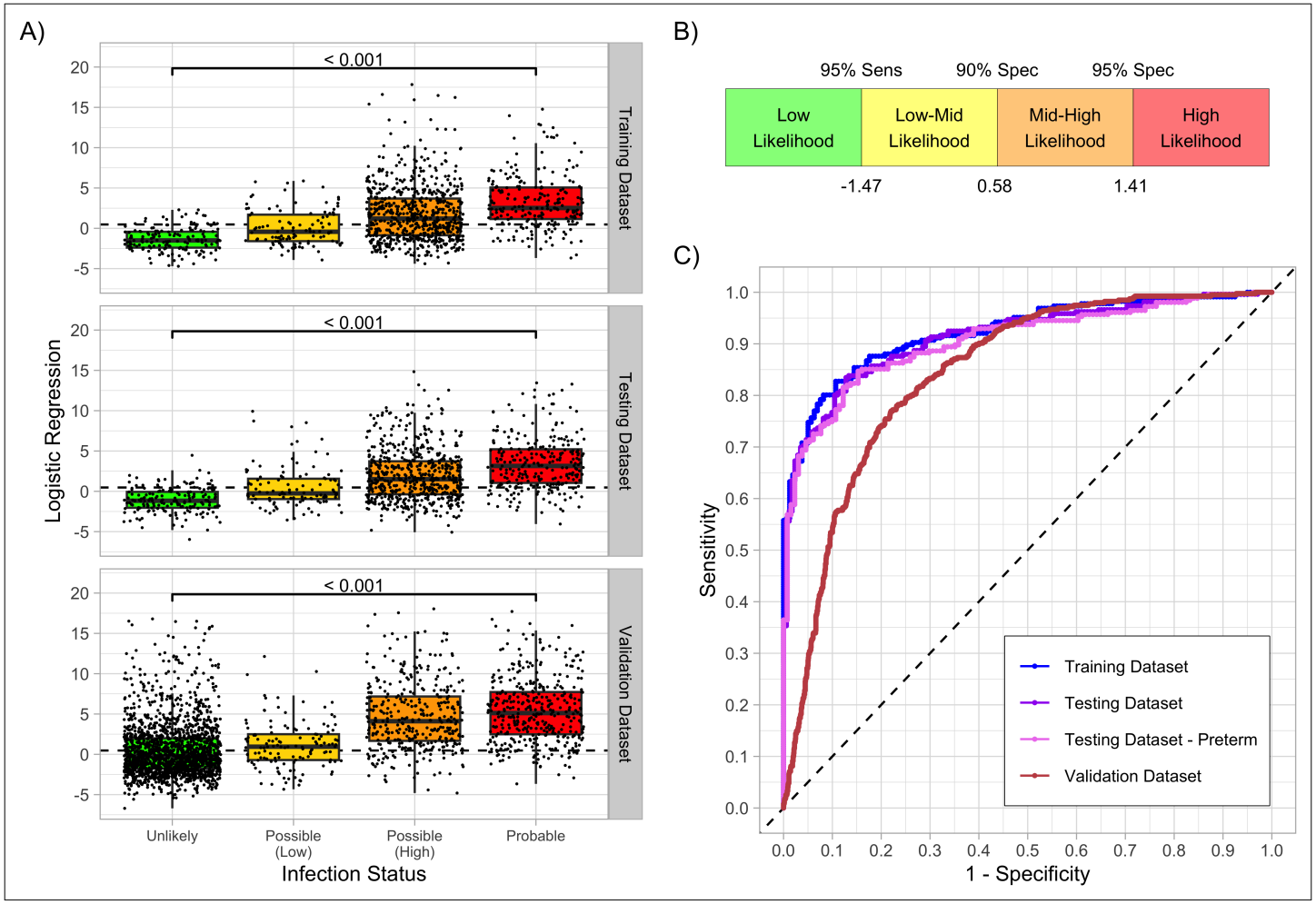


Figure 5.3: Diagnostic performance of a logistic regression model for late-onset sepsis.

A) Box plot of values against clinical suspicion of sepsis, with optimal threshold (dashed line). **B)** Clinically relevant thresholds calculated from ROC curve analysis. **C)** ROC curves for three datasets, plus a subset of the testing dataset as a comparator for the validation cohort.

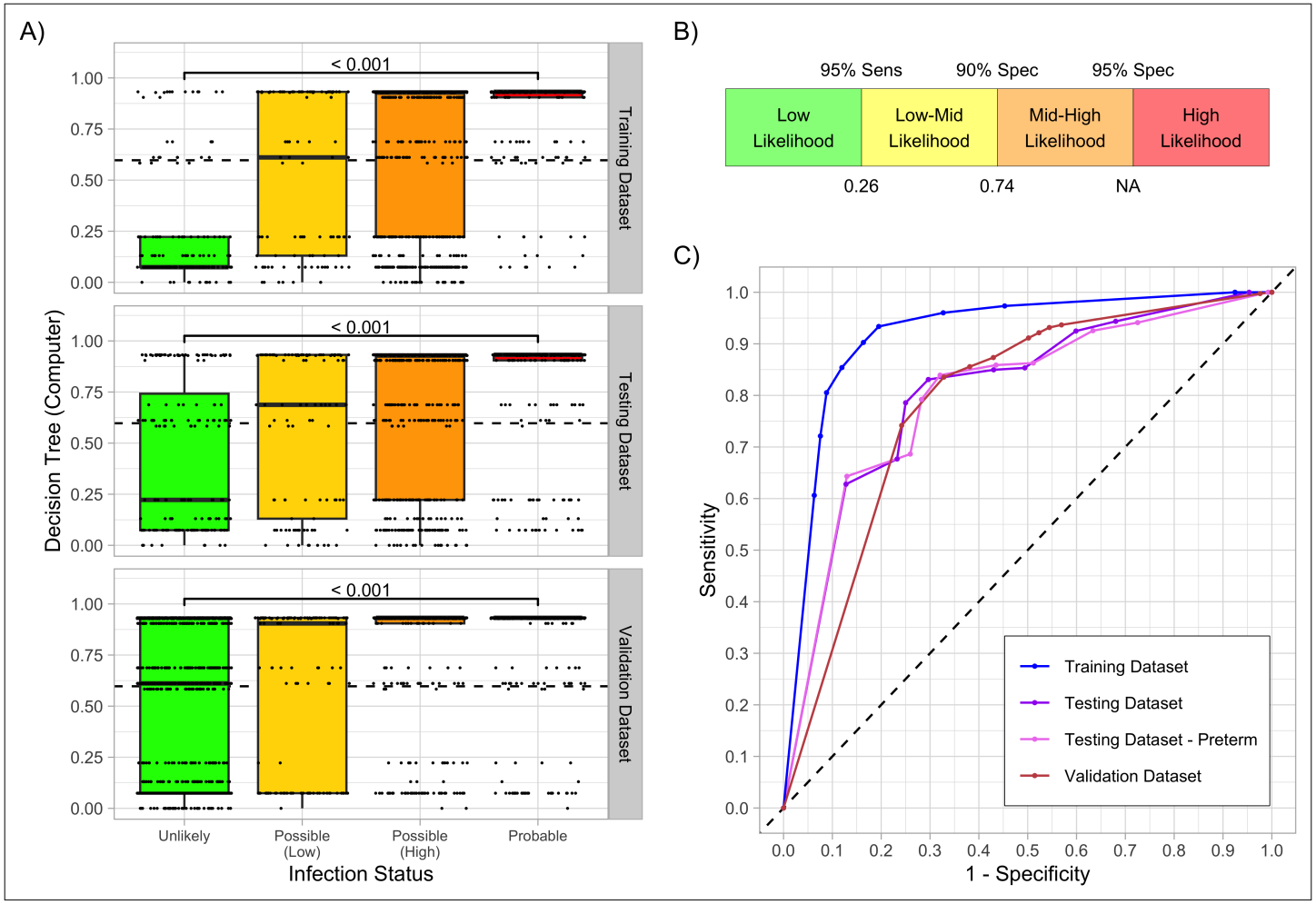


Figure 5.4: Diagnostic performance of a computer-generated decision tree model for late-onset sepsis.

A) Box plot of values against clinical suspicion of sepsis, with optimal threshold (dashed line). **B)** Clinically relevant thresholds calculated from ROC curve analysis. **C)** ROC curves for three datasets, plus a subset of the testing dataset as a comparator for the validation cohort.

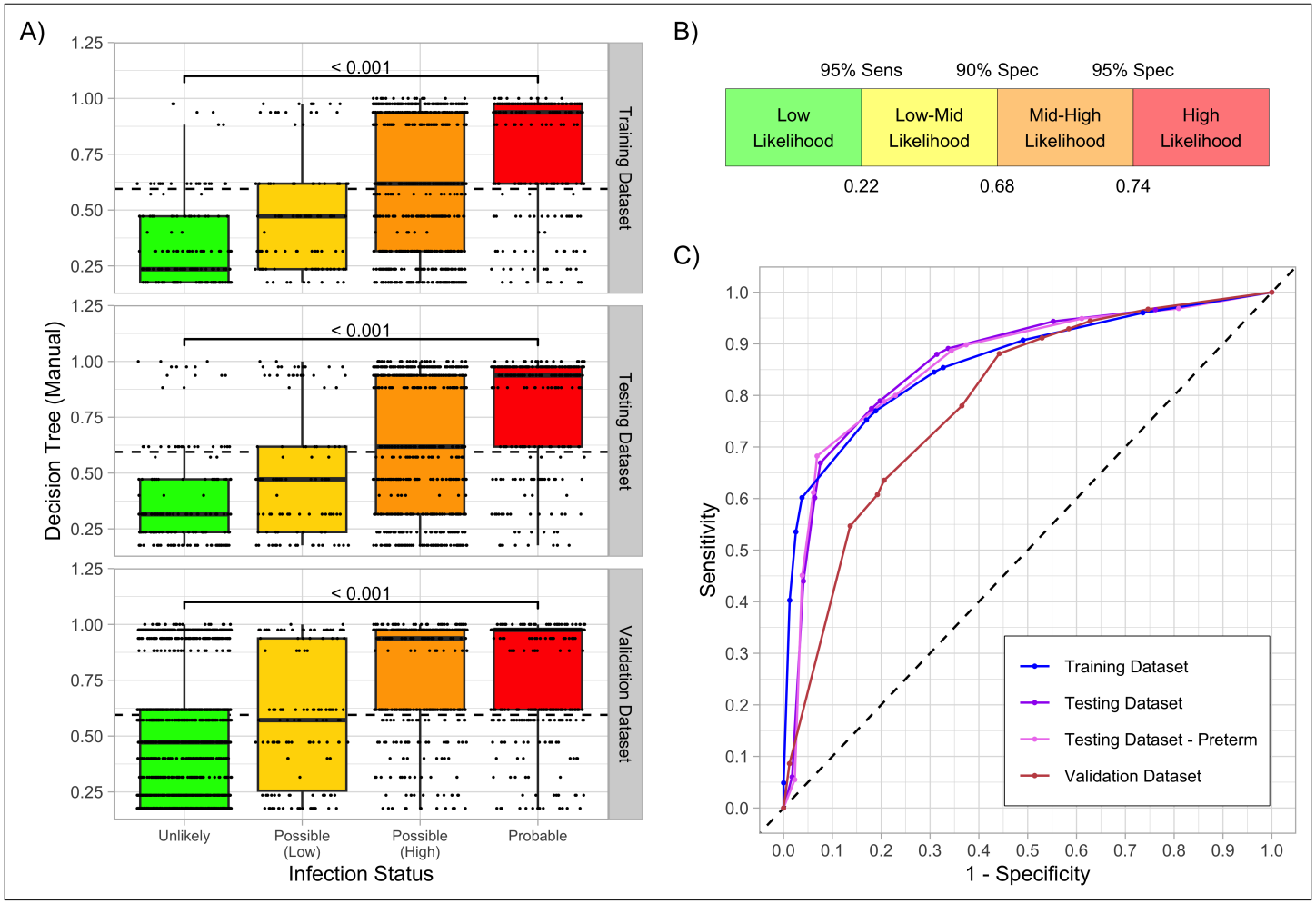


Figure 5.5: Diagnostic performance of a manually constructed decision tree model for late-onset sepsis.

A) Box plot of values against clinical suspicion of sepsis, with optimal threshold (dashed line). **B)** Clinically relevant thresholds calculated from ROC curve analysis. **C)** ROC curves for three datasets, plus a subset of the testing dataset as a comparator for the validation cohort.

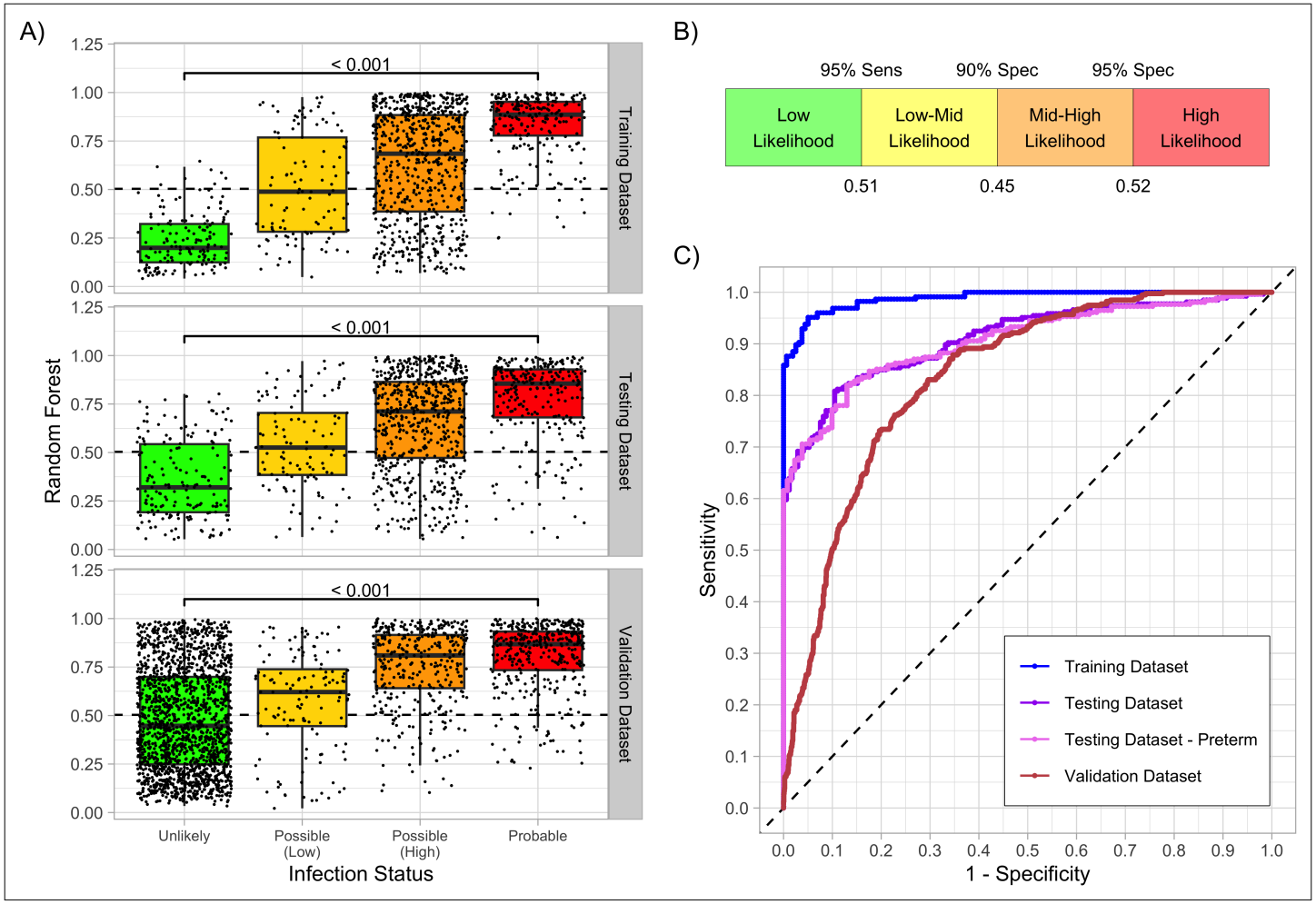


Figure 5.6: Diagnostic performance of a random forest model for late-onset sepsis.

A) Box plot of values against clinical suspicion of sepsis, with optimal threshold (dashed line). **B)** Clinically relevant thresholds calculated from ROC curve analysis. **C)** ROC curves for three datasets, plus a subset of the testing dataset as a comparator for the validation cohort.

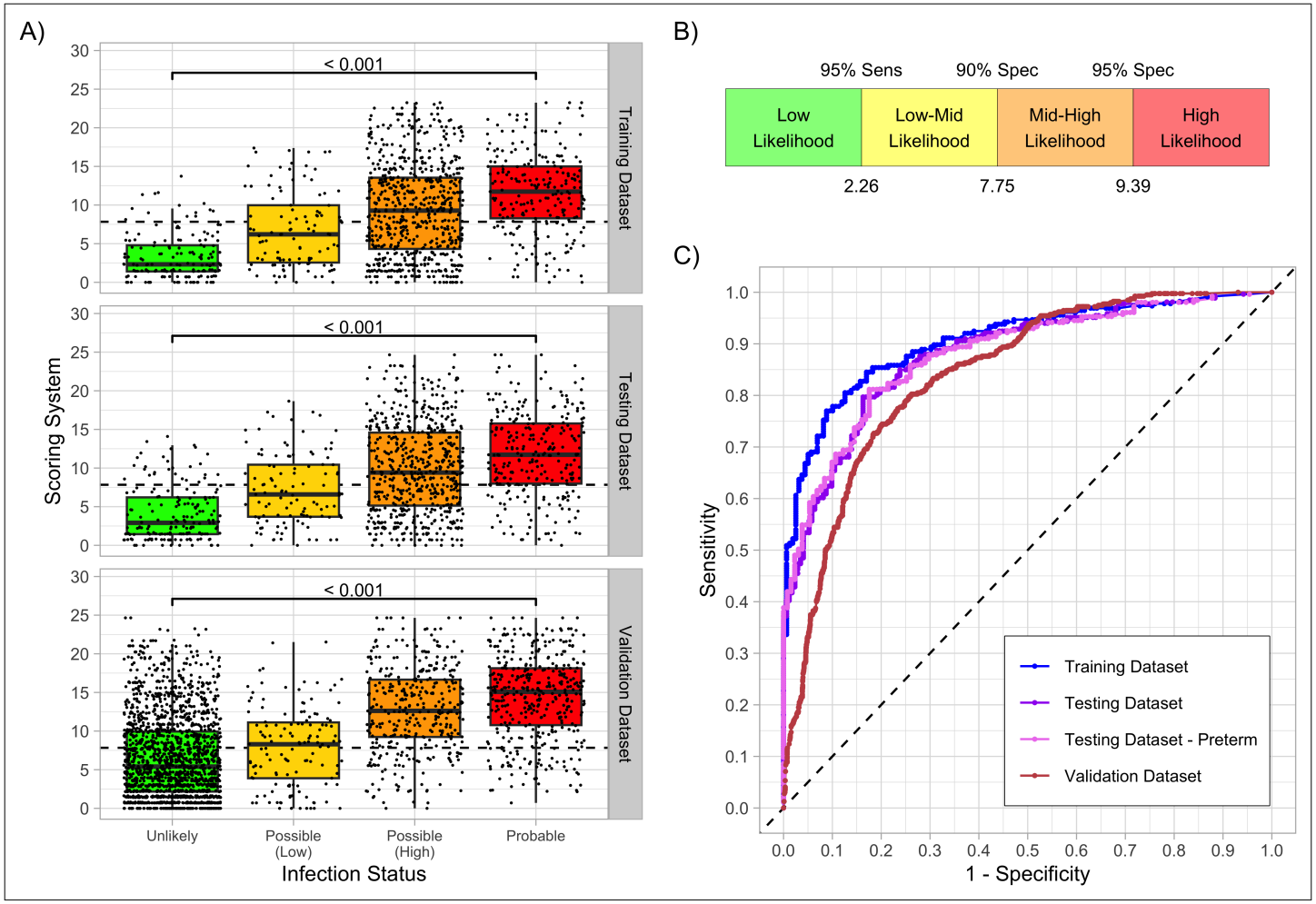


Figure 5.7: Diagnostic performance of a scoring system for late-onset sepsis.

A) Box plot of values against clinical suspicion of sepsis, with optimal threshold (dashed line). **B)** Clinically relevant thresholds calculated from ROC curve analysis. **C)** ROC curves for three datasets, plus a subset of the testing dataset as a comparator for the validation cohort.

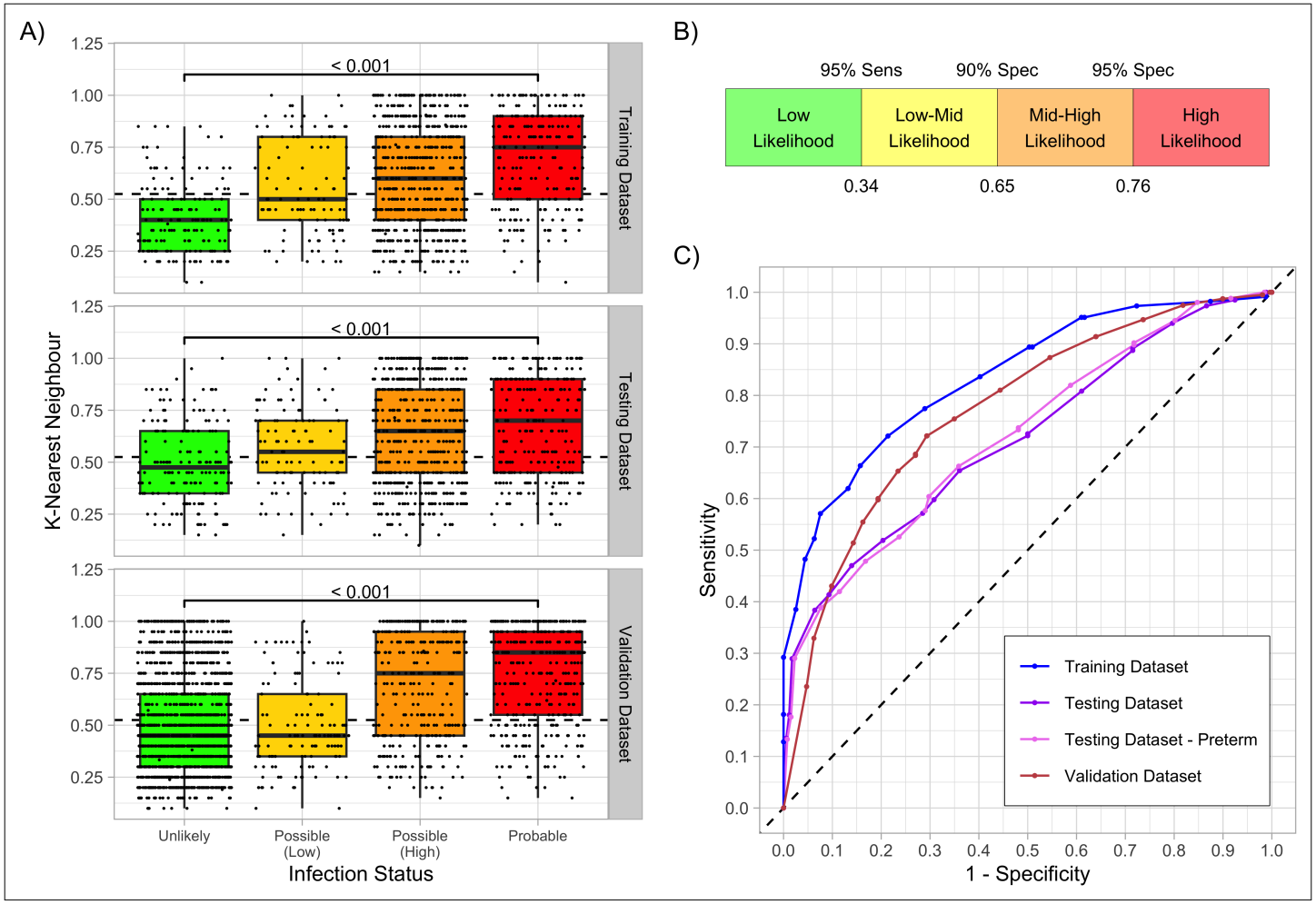


Figure 5.8: Diagnostic performance of a k-nearest neighbour model for late-onset sepsis.

A) Box plot of values against clinical suspicion of sepsis, with optimal threshold (dashed line). **B)** Clinically relevant thresholds calculated from ROC curve analysis. **C)** ROC curves for three datasets, plus a subset of the testing dataset as a comparator for the validation cohort.

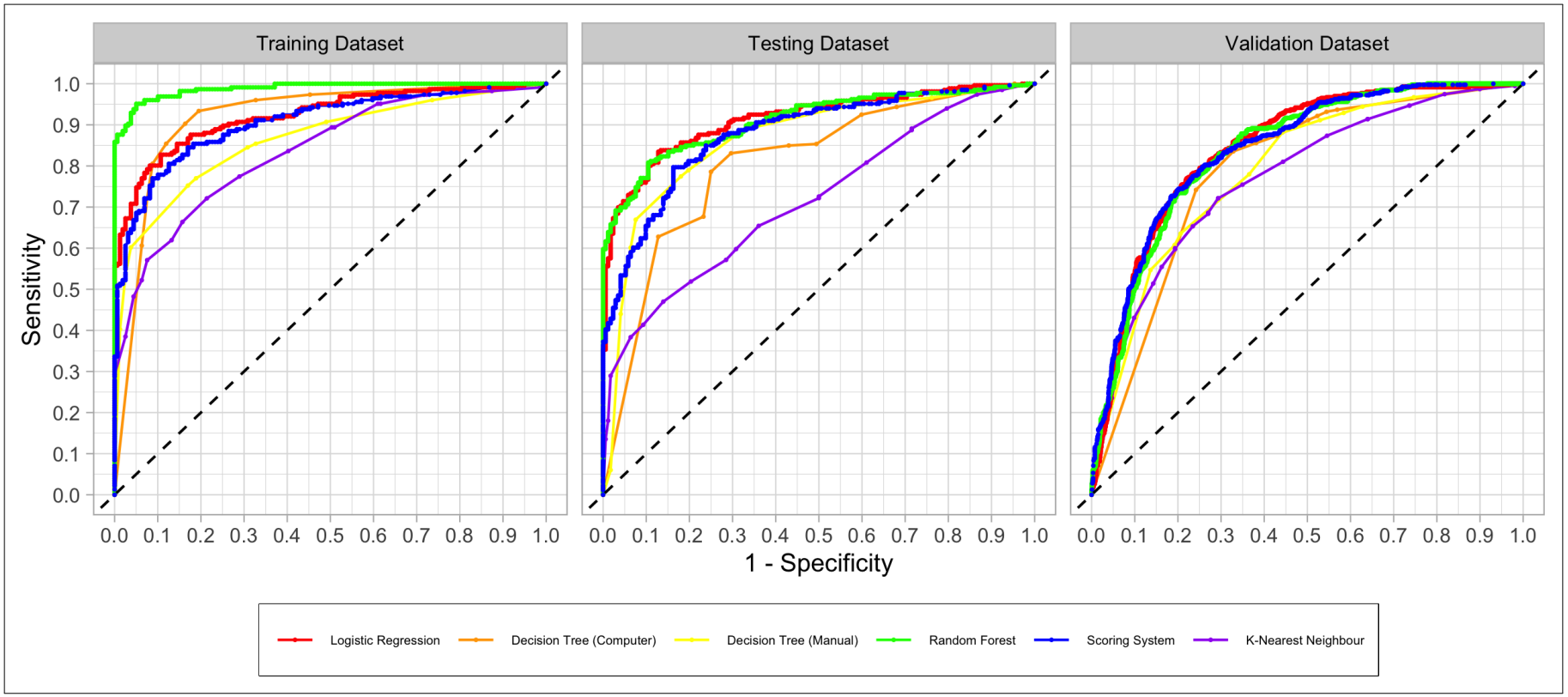


Figure 5.9: The diagnostic accuracy of six diagnostic models for late-onset sepsis.

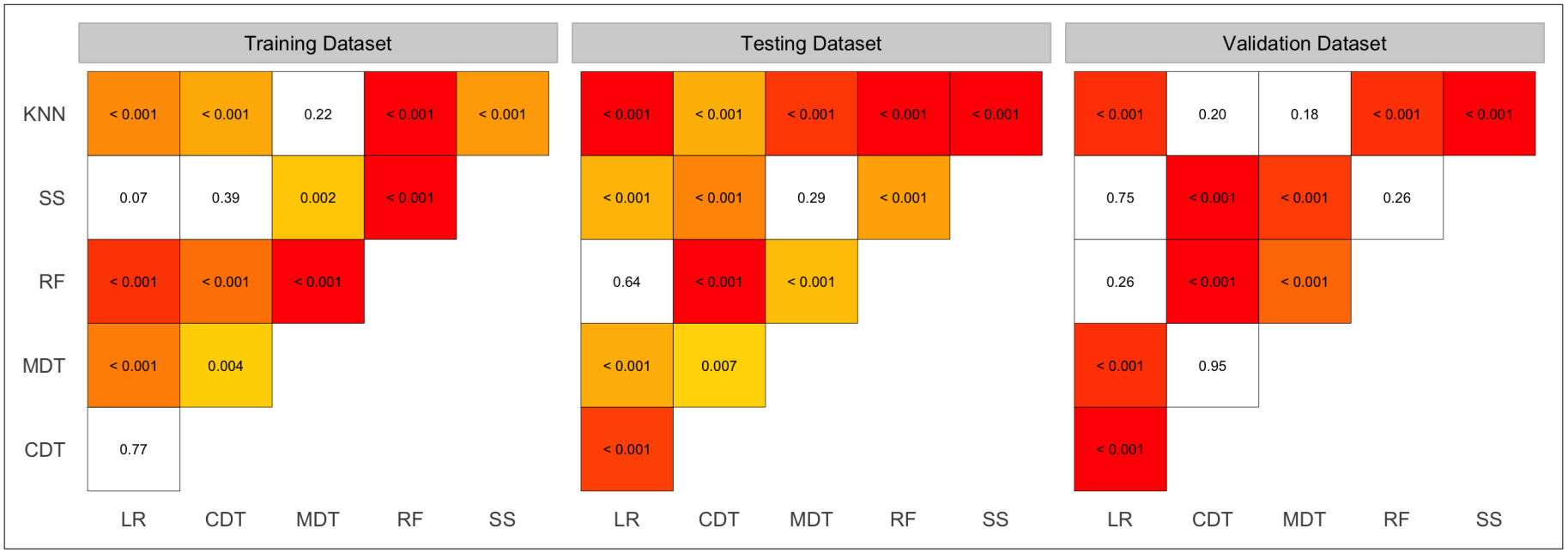


Figure 5.10: Heatmap of DeLong Test p-values for the comparison of receiver operating characteristic curve analysis of six diagnostic models for late-onset sepsis.

Abbreviations: CDT – computer-generated decision tree; KNN – k-nearest neighbour; LR – logistic regression; MDT – manually constructed decision tree; RF – random forest; SS – scoring system.

Table 5.7: Review of the extreme outliers for each diagnostic model for late-onset sepsis. The most extreme values from the *unlikely* and *probable sepsis* groups were identified for each model and dataset and reviewed to understand why they were misclassified.

Model	Dataset	Group	ID	Age	Value	Alignment with Other Models	Conclusion
Logistic Regression	Training Dataset	*Unlikely Sepsis*	B97329	70	2.29	"False positive" in two models	Respiratory culture positive for *E.coli* six days before – post-infection leucocytosis?
		Probable Sepsis	C05498	26	-3.68	"False negative" in all models, extreme outlier in five models	Respiratory culture positive for *Klebsiella oxytoca* without clinical deterioration – contamination?
	Testing Dataset	*Unlikely Sepsis*	C32089	31	4.47	"False positive" in all models, extreme outlier in two models	Diagnosed with periventricular/intraventricular haemorrhage the day before – possible non-infective immune activation?
		Probable Sepsis	C41906	20	-4.06	"False negative" in all models, extreme outlier in three models	Blood culture positive for CoNS two days before then negative on the day – contamination?
Computer Generated Decision Tree	Training Dataset	*Unlikely Sepsis*	B94058	7	0.93	"False positive" in three models	Suspicion of maternal and neonatal infection at birth, treated with three days of antibiotics for EOS – post-infection leucocytosis?
		Probable Sepsis	B85864	21	0.07	"False negative" in all models, extreme outlier in five models	Respiratory culture positive for *Staphylococcus aureus* with no clinical deterioration – contamination?
	Testing Dataset	*Unlikely Sepsis*	B86286	32	0.93	"False positive" in three models	Respiratory culture positive for *Klebsiella aerogenes* six days before with no positive cultures after – post-infection leucocytosis?
		Probable Sepsis	B74313	94	0.07	"False negative" in all models, extreme outlier in five models	Respiratory culture positive for *Staphylococcus aureus* with no clinical deterioration – contamination?

Continued on next page

Continued from last page

Model	Dataset	Group	ID	Age	Value	Alignment with Other Models	Conclusion
Manually Constructed Decison Tree	Training Dataset	*Unlikely Sepsis*	C06806	31	0.98	"False positive" in five models, extreme outlier in four models	Clinical deterioration for two days before, antibiotics started five days after – possible under-recognised infection?
		Probable Sepsis	B85864	21	0.18	"False negative" in all models, extreme outlier in five models	Respiratory culture positive for *Staphylococcus aureus* with no clinical deterioration – contamination?
	Testing Dataset	*Unlikely Sepsis*	C05658	64	1.00	"False positive" in four models	Inguinal herniotomy one day before – misclassified result that should have been excluded
		Probable Sepsis	B74313	94	0.18	"False negative" in all models, extreme outlier in five models	Respiratory culture positive for *Staphylococcus aureus* with no clinical deterioration – contamination?
Random Forest	Training Dataset	*Unlikely Sepsis*	C53241	6	0.65	False positive" in five models, extreme outlier in three models	Suspicion of maternal infection at birth, treated with three days of antibiotics – post-EOS leucocytosis?
		Probable Sepsis	C51561	19	0.25	"False negative" in five models, extreme outlier in four models	Isolated blood culture positive for CoNS with no clinical deterioration – contamination?
	Testing Dataset	*Unlikely Sepsis*	C32089	70	0.80	"False positive" in all models, extreme outlier in three models	Respiratory culture positive for *Enterobacter* nine days before, clinical deteroiration, raised CRP – post-infection leucocytosis?
		Probable Sepsis	C41906	20	0.06	"False negative" in all models, extreme outlier in three models	Isolated blood culture positive for CoNS with no clinical deterioration – contamination?

Continued on next page

Continued from last page

Model	Dataset	Group	ID	Age	Value	Alignment with Other Models	Conclusion
Scoring System	Training Dataset	*Unlikely Sepsis*	C30306	14	13.73	"False positive" in five models, extreme outlier in three models	Six days of antibiotics, raised CRP, clinical deterioration despite negative blood culture six days before – missed infection?
		Probable Sepsis	B85864	21	0.00	"False negative" in all models, extreme outlier in five models	Respiratory culture positive for *Staphylococcus aureus* with no clinical deterioration – contamination?
	Testing Dataset	*Unlikely Sepsis*	C32089	70	14.09	"False positive" in all models, extreme outlier in three models	Respiratory culture positive for *Enterobacter* nine days before, clinical deteroiration, raised CRP – post-infection leucocytosis?
		Probable Sepsis	B74313	94	0.00	"False negative" in all models, extreme outlier in five models	Respiratory culture positive for *Staphylococcus aureus* with no clinical deterioration – contamination?
K-Nearest Neighbour	Training Dataset	*Unlikely Sepsis*	C30306	14	0.85	False positive in five models, extreme outlier in three models	Raised CRP, and signs of clinical deterioration despite negative blood culture six days before – missed infection?
		Probable Sepsis	B85864	21	0.10	"False negative" in all models, extreme outlier in five models	Respiratory culture positive for *Staphylococcus aureus* with no clinical deterioration – contamination?
	Testing Dataset	*Unlikely Sepsis*	C38738	121	1.00	"False positive" in all models, extreme outlier in three models	Eight days of antibiotics before, necrotising enterocolitis suspected, all cultures negative – missed infection?
		Probable Sepsis	B74313	94	0.20	"False negative" in all models, extreme outlier in five models	Respiratory culture positive for *Staphylococcus aureus* with no clinical deterioration – contamination?

Abbreviations: CoNS – coagulase negative Staphylococcus.

sepsis outliers showed features suggestive of culture contamination, based on pathogen type and the absence of clinical deterioration.

Comparison with Individual Parameters

All models, except the kNN model, outperformed individual CBC parameters (e.g., WBC#, NLR, and NEUT-RI), indicating that multivariable approaches offer substantial improvements over single-marker strategies (Figures 5.11 – 5.17). Most models also outperformed CRP; however, comparisons should be interpreted cautiously, as CRP levels greater than 10 mg/L were an exclusion criterion for the *unlikely sepsis* group.

Comparison with Alternative Models

Exploratory models that included GA and postnatal age did not improve performance in any model (Table A4.8, p. 271). Both the alternative random forest and kNN models included both parameters (by design), and the alternative computer-generated decision tree and scoring system retained GA after refinement, suggesting perceived utility. Overall, neither variable consistently enhanced model performance.

5.3.4 Longitudinal Case Studies

Three representative case studies, one from each dataset, were selected to illustrate the longitudinal performance of the diagnostic models in comparison to CRP. The kNN model was excluded due to its consistently poor performance across datasets.

Figure 5.18 presents a neonate from the training dataset with *S. marcescens* bacteraemia on postnatal days 27 and 29. CRP began to rise on day 23, peaked on day 28, and returned to baseline by day 36 (Figure 5.18F). All diagnostic models crossed their respective thresholds by day 23, four days before the first positive blood culture, and remained elevated throughout the infection period.

Figure 5.19 shows a neonate from the testing dataset with a prolonged respiratory infection. Despite microbiological evidence of disease (*K. aerogenes* on postnatal days 26 and 35, and *S. viridans/Staphylococci* on day 45), blood cultures remained negative. CRP exceeded 10 mg/L only once, on day 32 (Figure 5.19F). All models exceeded

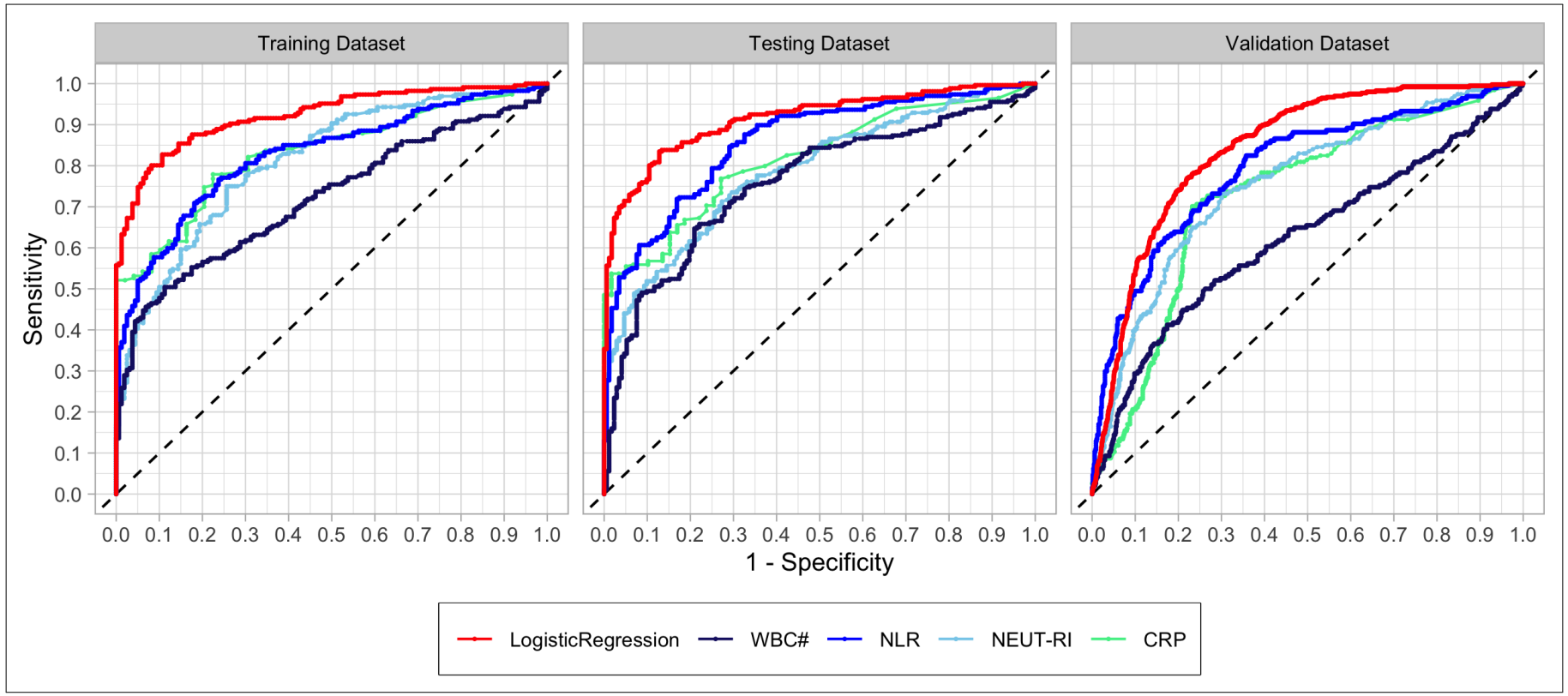


Figure 5.11: The diagnostic performance of a logistic regression model for late-onset sepsis compared to traditional biomarkers.

Abbreviations: CRP – C-reactive protein; NLR – neutrophil-to-lymphocyte ratio; NEUT-RI – neutrophil reactive index; WBC# – white blood cell count.

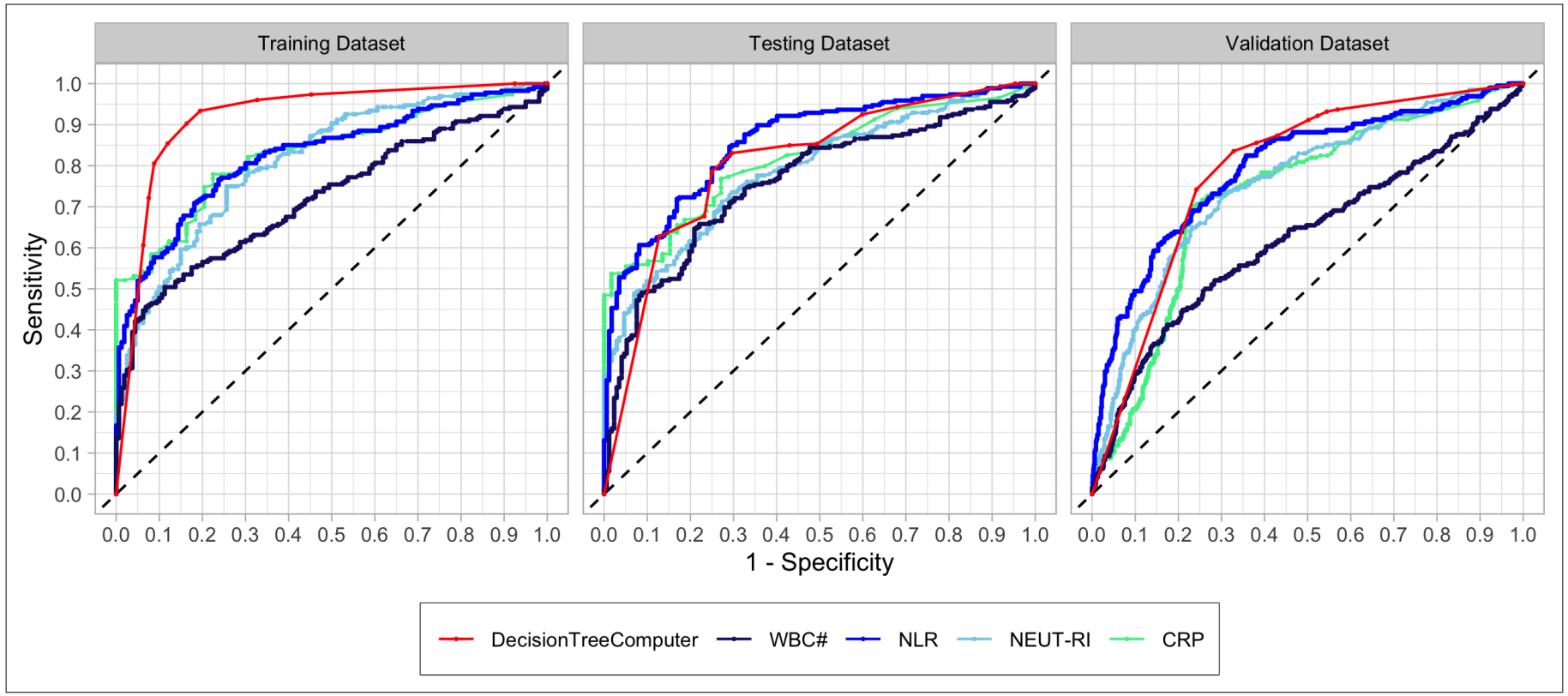


Figure 5.12: The diagnostic performance of a computer-generated decision tree model for late-onset sepsis compared to traditional biomarkers.

Abbreviations: CRP – C-reactive protein; NLR – neutrophil-to-lymphocyte ratio; NEUT-RI – neutrophil reactive index; WBC# – white blood cell count.

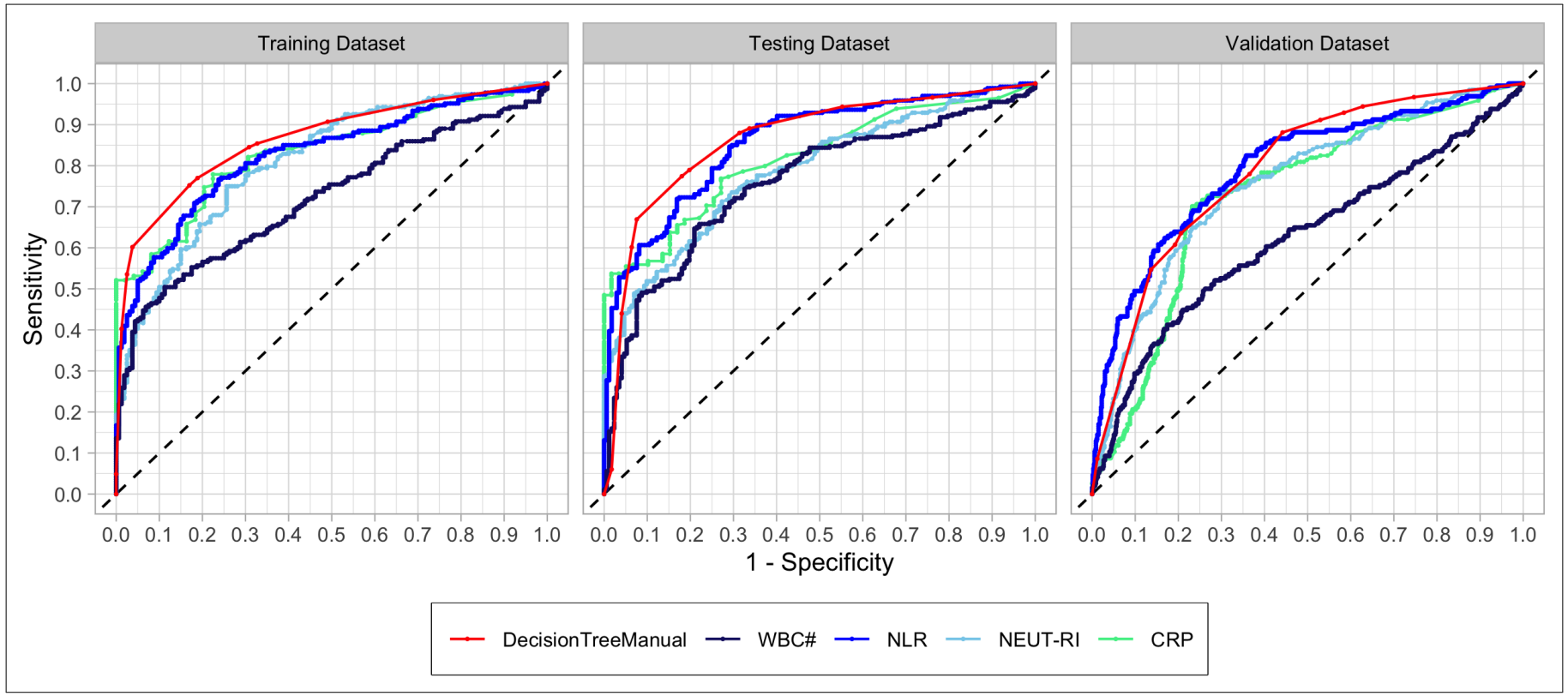


Figure 5.13: The diagnostic performance of a manually constructed decision tree model for late-onset sepsis compared to traditional biomarkers.

Abbreviations: CRP – C-reactive protein; NLR – neutrophil-to-lymphocyte ratio; NEUT-RI – neutrophil reactive index; WBC# – white blood cell count.

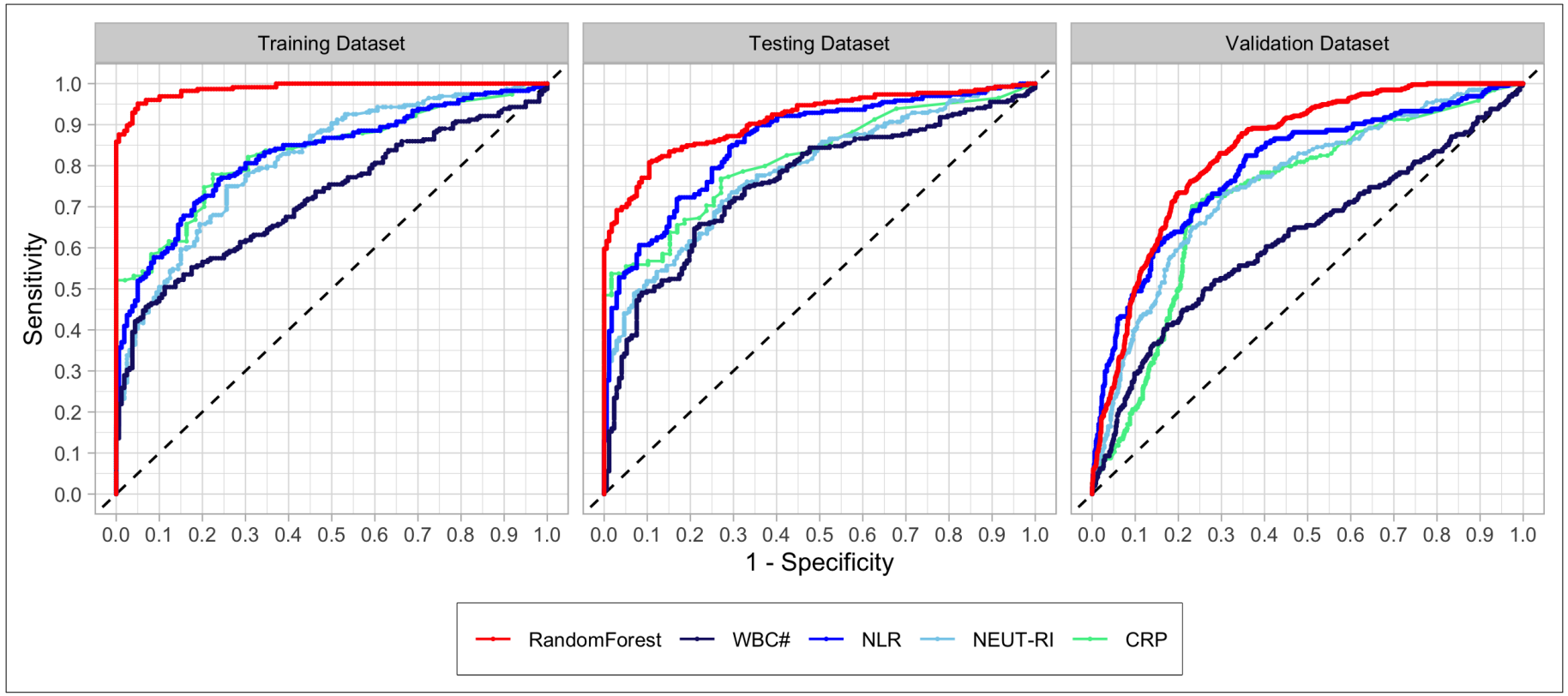


Figure 5.14: The diagnostic performance of a random forest model for late-onset sepsis compared to traditional biomarkers.

Abbreviations: CRP – C-reactive protein; NLR – neutrophil-to-lymphocyte ratio; NEUT-RI – neutrophil reactive index; WBC# – white blood cell count.

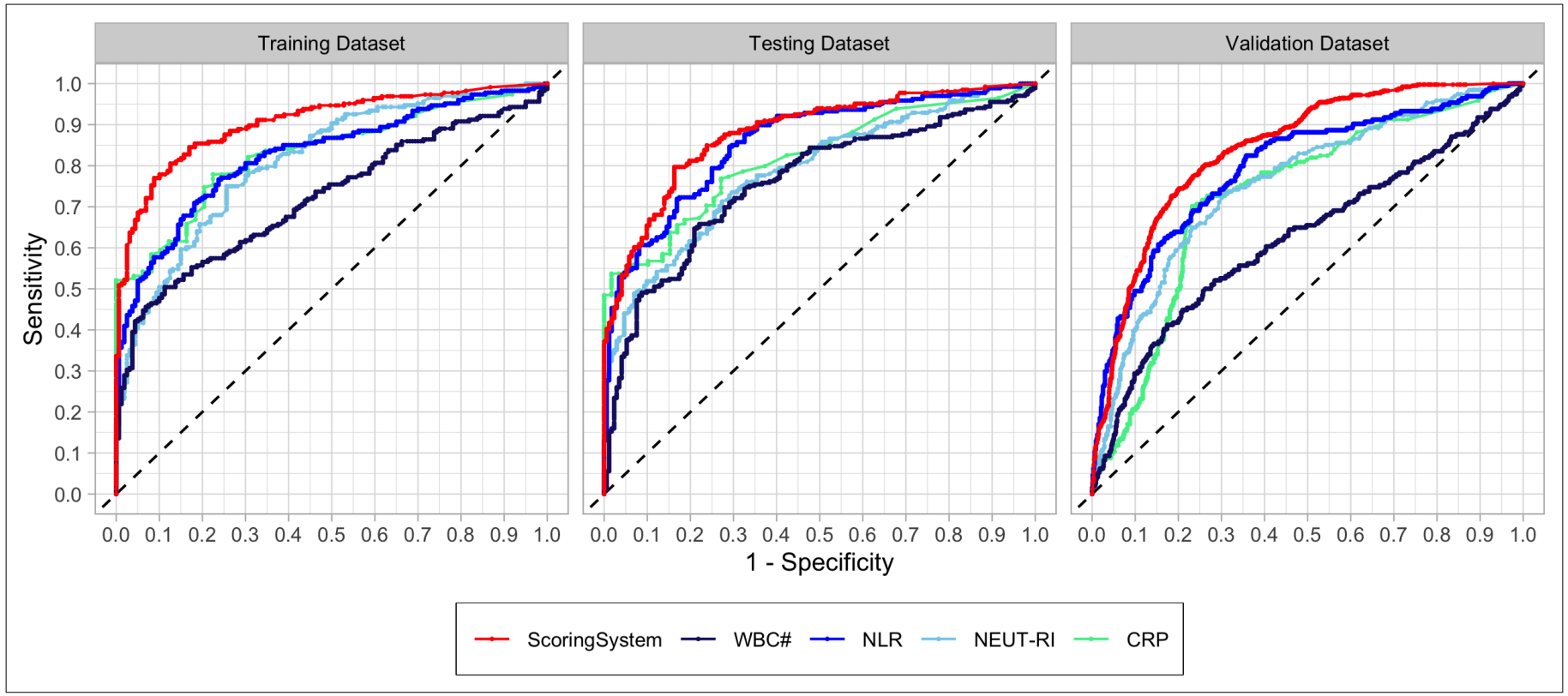


Figure 5.15: The diagnostic performance of a scoring system for late-onset sepsis compared to traditional biomarkers.

Abbreviations: CRP – C-reactive protein; NLR – neutrophil-to-lymphocyte ratio; NEUT-RI – neutrophil reactive index; WBC# – white blood cell count.

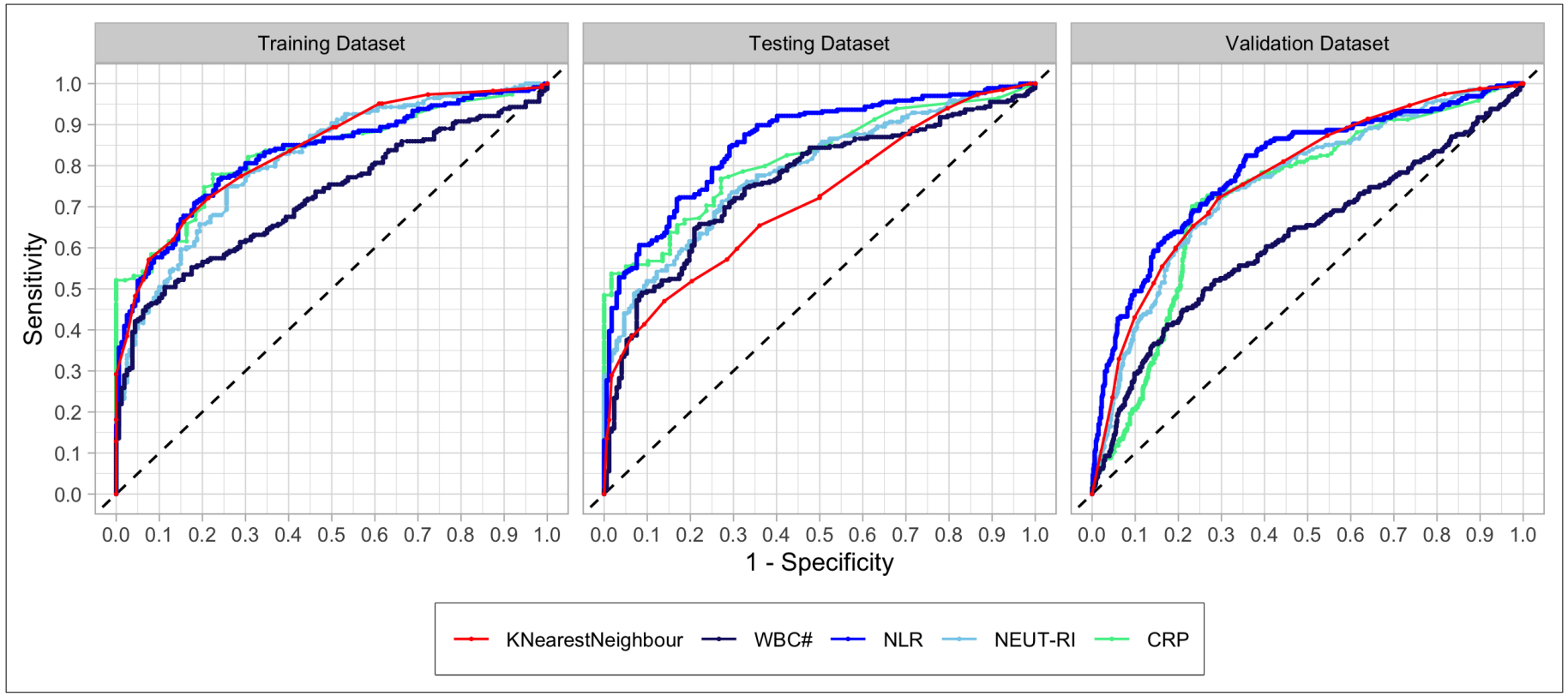


Figure 5.16: The diagnostic performance of a k-nearest neighbour model for late-onset sepsis compared to traditional biomarkers.

Abbreviations: CRP – C-reactive protein; NLR – neutrophil-to-lymphocyte ratio; NEUT-RI – neutrophil reactive index; WBC# – white blood cell count.

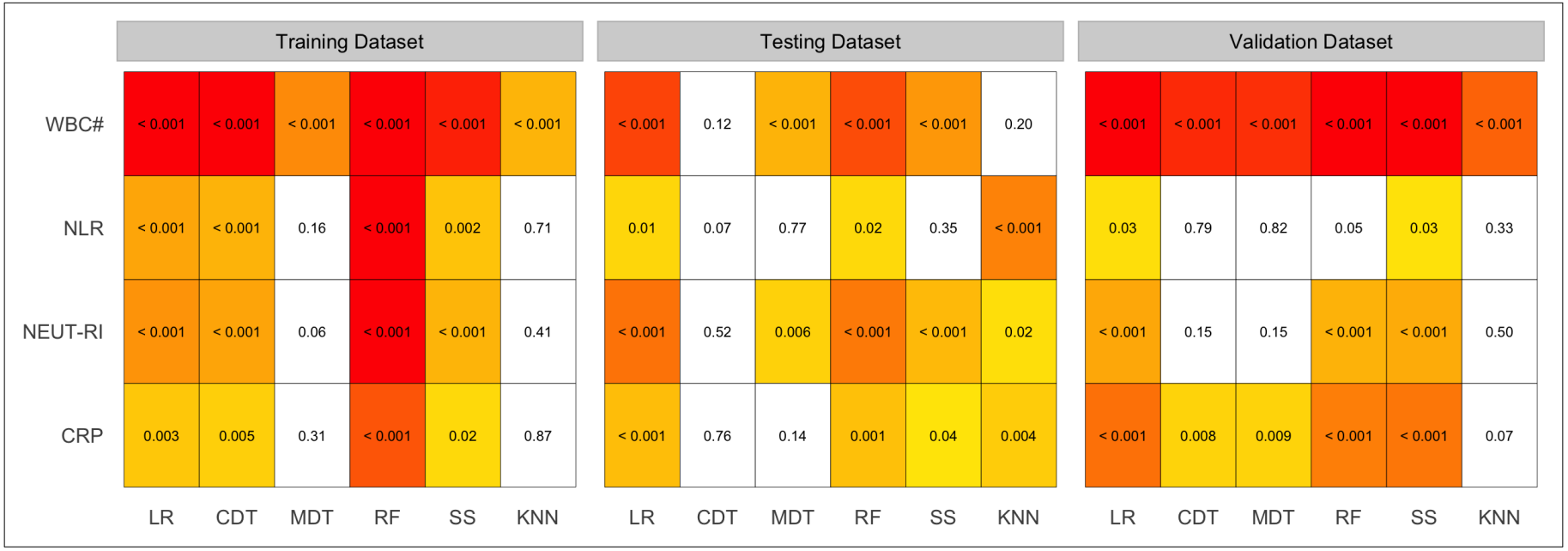


Figure 5.17: Heatmap of DeLong Test p-values for the comparison of receiver operating characteristic curve analysis of six diagnostic models for late-onset sepsis with traditional haematological biomarkers.

Abbreviations: CDT – computer-generated decision tree model; CRP – C-reactive protein; KNN – k-nearest neighbour model; LR – logistic regression model; MDT – manually constructed decision tree model; NLR – neutrophil-to-lymphocyte ratio; NEUT-RI – neutrophil reactive index; RF – random forest model; SS – scoring system model; WBC# – white blood cell count.

their respective thresholds by day 24 and remained elevated throughout the infection period. The computer-generated decision tree (Figure 5.19B), random forest model (Figure 5.19D), and scoring system (Figure 5.19E) indicated high likelihoods of infection from the first measurement on day 21.

Figure 5.20 presents a neonate from the external validation dataset with positive blood cultures for *K. pneumonia* on postnatal days 90 and 98. All models rose sharply between days 88 and 90, aligning with the first positive culture, and remained elevated throughout. The logistic regression model (Figure 5.20A), the manually constructed decision tree (Figure 5.20C), the random forest model (Figure 5.20D), and the scoring system (Figure 5.20E) all remained above their respective thresholds until day 110.

Together, these case studies highlight the potential for using longitudinal models to monitor neonates throughout their stay in the NICU. The diagnostic models often identified infective episodes earlier than CRP or microbiology cultures, and remained elevated throughout the clinical course, even when culture results were negative or CRP values fluctuated.

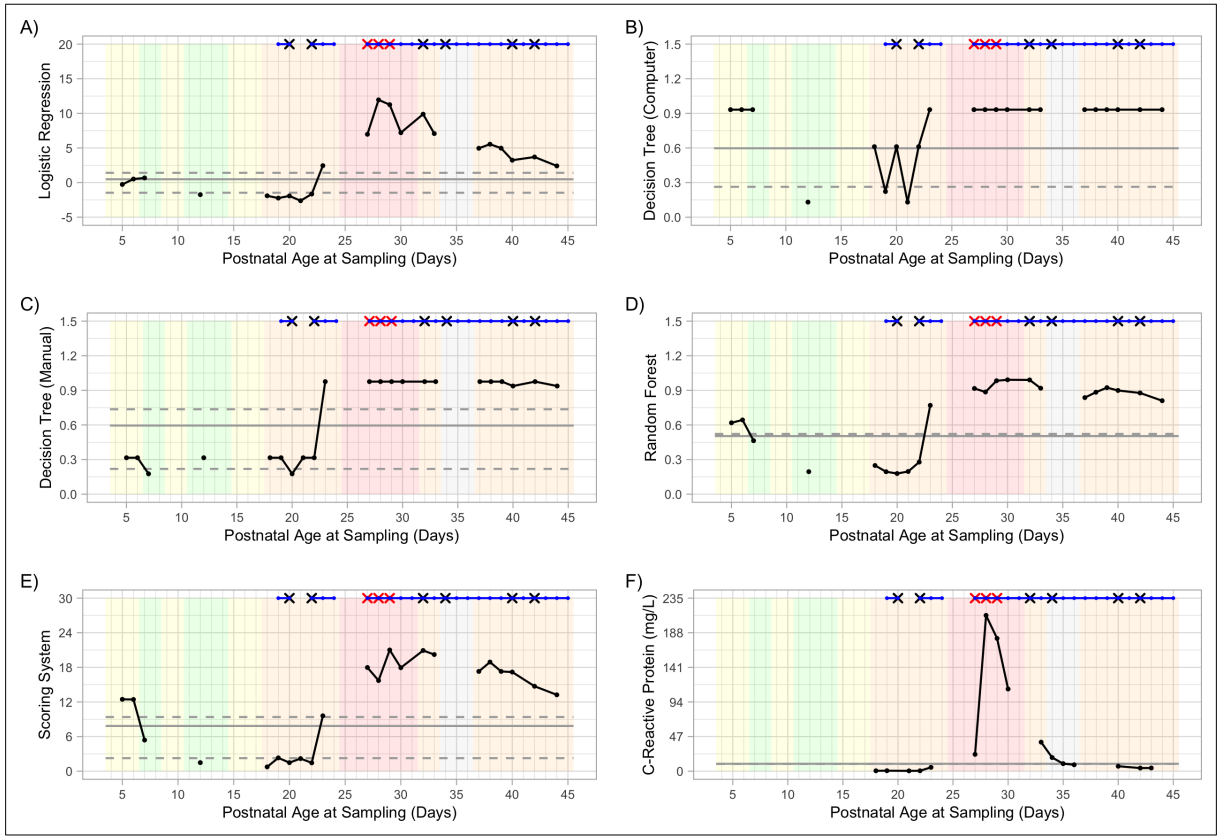


Figure 5.18: Longitudinal case study of a neonate born at 30 weeks' gestation and admitted to the John Radcliffe Hospital Neonatal Intensive Care Unit. Blood cultures were positive for *Serratia marcescens* on postnatal days 27, 28, and 29.

A) Logistic regression model. **B)** Computer-generated decision tree. **C)** Manually constructed decision tree. **D)** Random forest model. **E)** Scoring system. **F)** C-reactive protein (CRP; mg/L). Background colours indicate infection status (green: *unlikely*; yellow: *possible (low)*; orange: *possible (high)*; red: *probable*; grey: excluded). Blue lines denote antibiotic administration. Blood culture results are marked with red (positive) or black (negative) crosses. Solid grey lines indicate optimal thresholds (10 mg/L for CRP) and dashed grey lines mark thresholds for 95% sensitivity and 95% specificity.

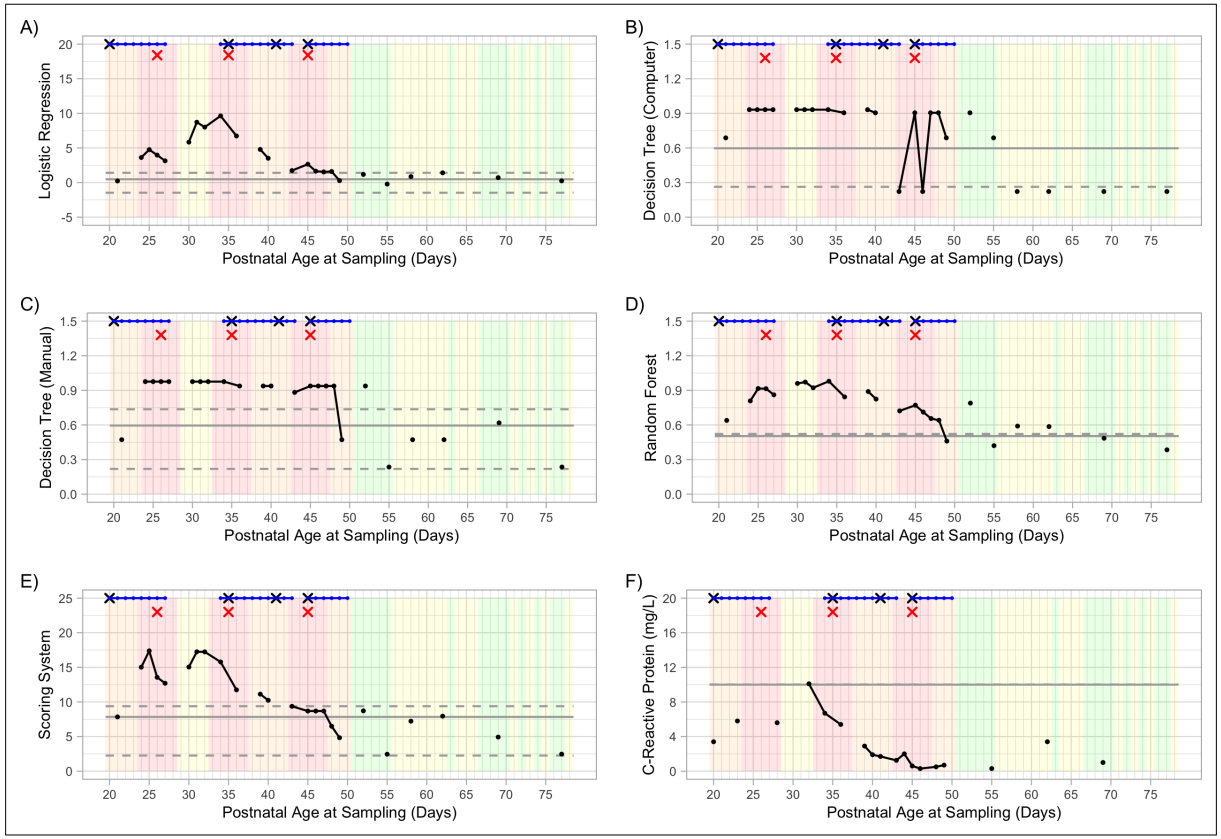


Figure 5.19: Longitudinal case study of a neonate born at 27 weeks' gestation and admitted to the John Radcliffe Hospital Neonatal Intensive Care Unit. Respiratory cultures were positive for *Klebsiella aerogenes* on postnatal days 26 and 35, and for *Streptococcus viridans* on day 45.

A) Logistic regression model. **B)** Computer-generated decision tree. **C)** Manually constructed decision tree. **D)** Random forest model. **E)** Scoring system. **F)** C-reactive protein (CRP; mg/L). Background colours indicate infection status (green: *unlikely*; yellow: *possible (low)*; orange: *possible (high)*; red: *probable*; grey: excluded). Blue lines denote antibiotic administration. Negative blood culture results are marked with black crosses. Positive respiratory cultures are marked with red crosses. Solid grey lines indicate optimal thresholds (10 mg/L for CRP) and dashed grey lines mark thresholds for 95% sensitivity and 95% specificity.

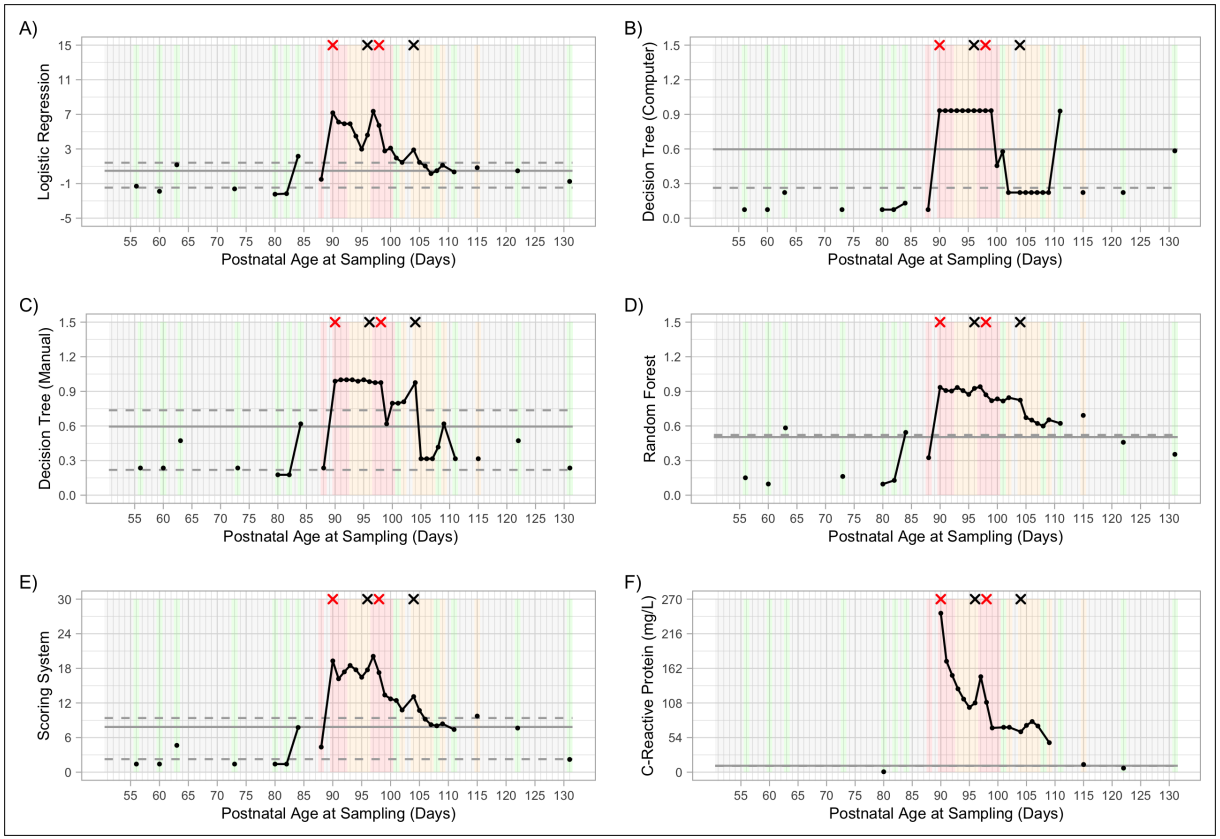


Figure 5.20: Longitudinal case study of a neonate born at 26 weeks' gestation and admitted to the John Radcliffe Hospital Neonatal Intensive Care Unit. Blood cultures were positive for *Klebsiella pneumoniae* on postnatal days 90 and 98.

A) Logistic regression model. **B)** Computer-generated decision tree. **C)** Manually constructed decision tree. **D)** Random forest model. **E)** Scoring system. **F)** C-reactive protein (CRP; mg/L). Background colours indicate infection status (green: *unlikely*; yellow: *possible (low)*; orange: *possible (high)*; red: *probable*; grey: data unavailable). Blood culture results are marked with red (positive) or black (negative) crosses. Solid grey lines indicate optimal thresholds (10 mg/L for CRP) and dashed grey lines mark thresholds for 95% sensitivity and 95% specificity.

5.4 Discussion

5.4.1 Summary of Findings

This chapter describes the development of six diagnostic models for LOS using multiple CBC parameters: a logistic regression model, a random forest model, a scoring system, a computer-generated decision tree, a manually constructed decision tree, and a kNN model. Except for the kNN model, all models demonstrated strong internal performance. External validation confirmed the generalisability of the logistic regression model, random forest model, and scoring system. Although sensitivity and specificity declined with validation, the NPV remained high in all datasets and models, suggesting that all models would function well as rule-out tools for LOS.

The random forest model achieved the highest performance, maintaining strong AUCs across all datasets (0.988 in training, 0.911 in testing, and 0.852 in validation). The logistic regression model and scoring system also demonstrated consistent performance (AUCs > 0.8), although both exhibited reduced specificity during external validation. In contrast, the decision trees lacked generalisability, and the kNN model underperformed across all datasets.

The initial set of 18 CBC parameters was selected based on statistical significance (Chapter 4), biological plausibility, and efforts to reduce redundancy. After refinement, NEUT# and NEUT-RI were retained in all six models, while NLR, PLT#, and IPF were included in five, highlighting their robustness within multivariable frameworks. In contrast, NRBC#, RET#, LYMPH#, ITR, and NEUT-GI were only retained in models that did not apply feature selection (random forest and kNN), suggesting more limited predictive value. Further refinement by excluding less predictive parameters may enhance generalisability.

Comparison with Individual Parameters

Although widely used in clinical practice, the diagnostic performance of the WBC# was limited, as shown in Chapter 4. In the present chapter, all six models outperformed WBC# in both the training and validation datasets, and all except the kNN model did so

in the testing dataset. NLR and NEUT-RI emerged as the strongest individual parameters in Chapter 4; however, both were significantly outperformed by the logistic regression and random forest models (DeLong $p < 0.01$). These multivariable models can be derived from standard CBC data, offering enhanced diagnostic accuracy without the need for additional testing, thereby reducing reliance on weaker individual markers.

CRP was also included for comparison; however, its interpretation is limited by the study design since a CRP value greater than 10 mg/L was an exclusion criterion for the *unlikely sepsis* group. Nonetheless, the logistic regression, random forest, and scoring system models consistently outperformed CRP. While CRP remains clinically useful, its delayed response (12 to 24 hours after the onset of infection) limits its effectiveness as an early diagnostic tool [18]. In contrast, the case studies presented here illustrate that several of the developed models identified infection earlier and more consistently than CRP.

Comparison with Alternative Models

Incorporating GA and postnatal age did not improve diagnostic performance. This likely reflects the fact that many CBC parameters are age-sensitive, as shown in Chapter 3, making the inclusion of explicit age variables redundant. For example, MONO#, NEUT-RI, and IPF, each of which varies with age, were replaced by GA in the alternative decision tree, suggesting that when age-sensitive parameters are already incorporated, the explicit inclusion of age variables confers limited additional value.

5.4.2 Comparison with Existing Literature

Existing Clinical Scores

In comparison with previously published scores, the models developed here demonstrated stronger diagnostic performance and generalisability. For example, the Töllner score shows varied diagnostic utility: one study reported a sensitivity and specificity exceeding 90% [145], while another reported substantially lower values [146], even underperforming the kNN model developed in this chapter. This inconsistency likely reflects a reliance on both objective (e.g., WBC#, blood pH) and subjective (e.g., skin colour, respiratory

distress) criteria. In contrast, the models developed here are based solely on objective CBC data, eliminating inter-observer variability and improving reproducibility.

Similarly, the Rodwell HSS and Sepscore, while based on the CBC, demonstrated lower AUCs, sensitivities, and specificities than any of the six models developed here [122, 123]. These older scoring systems rely on conventional markers and fixed diagnostic cut-offs. In contrast, the models developed here incorporate novel markers of immune activation and more adaptable decision thresholds. Even the scoring system developed here incorporates thresholds aligned with clinically meaningful specificities, enhancing both diagnostic accuracy and clinical utility.

Comparison with Machine Learning Models

A recent systematic review of machine learning models for neonatal sepsis reported pooled metrics of AUC 0.94, sensitivity 87%, and specificity 89% (Table 5.1) [155]. Only four of the 19 reviewed studies were directly comparable to the present study – that is, they used CBC data and culture-proven sepsis definitions, and reported full performance metrics [157–160]. Among the 21 models from these four studies, only six achieved AUCs above 0.8, and only one (an artificial neural network built on 25 variables [159]) exceeded 85% specificity.

None of the reviewed models outperformed the logistic regression, random forest, or scoring system models developed here, all of which underwent both internal and external validation. These findings highlight not only the strong diagnostic performance of the current models but also their superior generalisability compared to much of the existing literature. Furthermore, most previously published models are trained and validated on blood culture-positive results, meaning that their diagnostic accuracy in real-world situations of suspected sepsis may be significantly lower. In the current study, longitudinal case studies illustrated that several models could offer utility even before culture positivity, allowing for earlier interventions in those with developing sepsis and avoiding unnecessary treatment in those without.

5.4.3 Implications of Findings

Chapter 4 established that several CBC parameters possess moderate to strong individual diagnostic value. This chapter demonstrates that combining CBC parameters into multivariable models significantly improves accuracy. The logistic regression, random forest, and scoring system models consistently outperformed traditional clinical scores, individual CBC parameters, and many existing published machine learning models.

Each model offers distinct clinical advantages. The scoring system and decision trees are highly interpretable with clear relative contributions for each included parameter, which may improve clinician trust. The logistic regression model is also relatively transparent, although its interpretation is slightly more complex. In contrast, the random forest and kNN models act as "black boxes" where the contribution of each parameter to the final output is unclear, which may limit clinical adoption.

Technical feasibility also varies across models. The random forest and kNN models would require integration into hospital IT systems and thus may be less suited to low-resource or point-of-care settings. Conversely, the logistic regression model, scoring system, and decision trees can be implemented directly into haematological analysers, enabling the real-time generation of results alongside the routine CBC results and allowing direct implementation into point-of-care settings.

Three longitudinal case studies illustrated how these models could support ongoing monitoring and early detection of neonatal sepsis. In each instance, model outputs rose before or alongside clinical deterioration, often before CRP increased or cultures became positive. These findings, coupled with the external validation in a temporally and geographically distinct cohort, reinforce the practical value of these models for daily monitoring in NICU settings.

Crucially, all models can be calculated automatically from a routine CBC result without the need for additional sampling and therefore benefit from the same advantages as the CBC over other diagnostic tools – namely, minimal blood volumes, low costs, and short

turnaround times. These features support broad implementation, even in resource-limited settings. The low blood volume requirement is particularly advantageous for fragile NICU patients, allowing models to be used as a routine part of infection monitoring.

5.4.4 Strengths and Limitations

Many strengths align with those discussed in Chapter 3 and Chapter 4. Both the internal and external datasets were derived from large, real-world NICU cohorts. The external validation dataset was geographically and temporally distinct from the training dataset, providing a robust test of each model's generalisability. The addition of longitudinal case studies from each dataset demonstrates practical utility and further supports clinical relevance.

All models were constructed using parameters with established statistical and biological relevance to neonatal sepsis (Chapter 4). Where applicable, models were refined to minimise overfitting and reduce complexity. For example, the initial random forest model overfitted the training dataset (AUC: 1; data not shown), which was mitigated in the final model by limiting the number of trees. Similarly, reducing the number of included parameters improved performance in the logistic regression model, decision trees, and scoring system.

Cases and controls were defined using strict inclusion and exclusion criteria, allowing the models to differentiate clearly between hospitalised but clinically well neonates and those with cultured-confirmed sepsis. In clinical practice, the ambiguous groups (*possible low* and *possible high*) are most relevant as they represent culture-negative neonates with unexplained signs of clinical deterioration. Across all six models, outputs increased proportionally with clinical suspicion, suggesting a strong capacity to identify unwell neonates, even when CRP remained low or cultures were negative.

This study also has several limitations. As a retrospective analysis, it is susceptible to potential bias related to data quality and completeness. Despite efforts to minimise misclassification, missing clinical data may have affected diagnostic categorisation.

Interestingly, some *unlikely sepsis* cases with high model scores likely represented missed infections or post-infective leucocytosis. In contrast, low-scoring *probable sepsis* cases were often associated with culture contaminants, such as a single positive culture for CoNS without signs of clinical deterioration.

Eighteen CBC parameters were included based on statistical significance and biological plausibility. However, less prominent markers excluded from this selection may have performed better in multivariable combinations. An exhaustive evaluation of all possible parameter combinations was not feasible due to time constraints and the smaller size of the datasets; however, this could be a promising area for future research. Similarly, while parameter elimination optimised the logistic regression model, decision trees, and scoring system, the true optimal set of parameters may not have been fully captured within the tested iterations.

Finally, while the inclusion of GA and postnatal age did not improve performance, other clinically relevant non-CBC variables, such as weight z-score, respiratory or heart rate, or biochemical markers, were not evaluated. Although their inclusion would complicate implementation, future studies could explore whether combining these features with CBC parameters could offer improved performance.

5.5 Chapter Summary

In this chapter, I developed and validated six diagnostic models for LOS using routine CBC parameters and a range of analytic techniques. All models underwent internal testing, and several demonstrated strong and generalisable performance in external validation. While the random forest model achieved the highest overall accuracy, the logistic regression model and scoring system offered greater interpretability and would require minimal technical infrastructure, making them well-suited for clinical implementation.

Each model can be automatically calculated from the standard CBC output, offering the potential for real-time monitoring and assessment of infection risk. Their low cost, minimal blood volume requirement, and compatibility with existing laboratory systems make them attractive tools for early detection of LOS in NICU settings.

The next chapter will apply the methodological framework established in this thesis (reference interval construction, individual parameter analysis, and multivariable diagnostic modelling) to two clinically significant sample types: CBC results from neonates within the first three postnatal days, and CBC results derived from umbilical cord blood samples. Although the models in this chapter were developed to target LOS, evaluating their utility in these earlier contexts may reveal broader diagnostic applicability.

Chapter 6

The Application of Analytical Methods to Other Clinically Relevant Samples for the Diagnosis of Early-Onset Neonatal Sepsis

6.1 Introduction

Some of the results presented in this chapter have been submitted to a peer-reviewed journal [Reference to be added upon publication].

EOS remains a significant cause of neonatal morbidity and mortality, with diagnosis often complicated by non-specific clinical signs and delayed microbiological confirmation (as discussed in Chapter 1). This chapter focuses on the interpretation and utility of CBC results within the first three days of life, including samples drawn at birth from the umbilical cord, to explore opportunities for earlier and less invasive diagnosis of EOS.

6.1.1 CBC Results From the First Three Days of Life

Evaluation of the neonatal CBC during the first three days of life remains challenging due to the absence of a robust, healthy control group. As detailed in Chapter 3, blood samples are rarely obtained from healthy neonates without a medical indication; therefore, most published research is derived from hospitalised but clinically stable newborns. These cohorts are often poorly defined, introducing heterogeneity that complicates interpretation (as discussed in Chapter 2).

Establishing the Baseline CBC Profile

Establishing reference ranges for the neonatal CBC has historically been limited by the assumption that haematological parameters fluctuate widely after birth. Significant shifts in the WBC# were first reported in the 1930s [161], with later studies confirming similar trends across the differential counts [140, 162]. For instance, NEUT# values

exceeding 25 x 10^9/L have been reported within the first 24 hours, while published reference ranges outside this early newborn period rarely report upper limits beyond 8 x 10^9/L [127, 140]. These changes are believed to reflect the early immune response to bacterial gut colonisation [127]. Similar fluctuations in NRBC# and RET# values in the first few weeks of life reflect the physiological transition from fetal to adult haemoglobin [129, 130].

As discussed in Chapter 3, multiple clinical variables influence the neonatal CBC. For example, sex and GA have both been shown to significantly affect the neonatal haematological profile in the immediate postnatal period [129, 132]. Another important but often overlooked variable is the MOD. Neonates born via Caesarean section (CS) often experience reduced stress, altered gut microbiota, and lower levels of systemic inflammation markers such as CRP than those born by spontaneous vaginal delivery (SVD) [163–165]. These differences may contribute to systematic variation in CBC results and must be considered when establishing baseline values.

Diagnostic Utility of the CBC for EOS

As discussed in Chapter 1, EOS is defined as a systemic bloodstream infection that presents within the first 72 hours of life [15]. It is typically caused by the vertical transmission of bacteria from the maternal genital tract; hence, risk factors include preterm or prolonged membrane rupture and maternal colonisation with GBS [11].

Despite its widespread use in sepsis workups, the diagnostic accuracy of the CBC for EOS is poorly established. Four studies identified in the systematic review and meta-analysis (Chapter 2) specifically assessed the performance of individual CBC parameters for EOS (Table 6.1) [95, 100, 101, 107]. Among these, only one study reported an AUC greater than 0.8 for ITR and PLR [101]. Most studies reported poor sensitivity and specificity for traditional markers, including WBC#, NEUT#, NLR, and PLT#.

Notably, there is limited evidence regarding the diagnostic potential of novel immune activation markers, despite their increasing use in the diagnosis of adult infections.

Table 6.1: Diagnostic accuracy of the complete blood count for early-onset sepsis from previously published studies.

Parameter	Study	AUC	Threshold	Sensitivity (%)	Specificity (%)
WBC# (x 10^9/L)	Manucha et al., 2002 [107]	–	10.00	57.0	76.0
NEUT# (x 10^9/L)	Manucha et al., 2002 [107]	–	8.00	76.0	67.0
NLR	Mahmoud et al., 2022 [95]	0.790	0.10[a]	67.0	99.0
	Zhang 2021 [100]	0.788	3.17[a]	77.0	78.0
ITR	Arcagok et al., 2019 [101]	0.955	0.22	54.3	98.1
	Manucha et al., 2002 [107]	–	0.14	63.0	72.0
PLT# (x 10^9/L)	Manucha et al., 2002 [107]	–	150	62.0	71.0
PLR	Arcagok et al., 2019 [101]	0.847	57.70[a]	91.3	97.6
	Mahmoud et al., 2022 [95]	0.780	7.00[a]	70.0	73.0
	Zhang 2021 [100]	0.699	90.85[a]	64.9	80.0

AUC values $>$ 0.8 and sensitivities/specificities $>$ 85% are highlighted in green, suggesting strong diagnostic performance. [a]*Study was not included in meta-analyses of diagnostic performance because values showed high variance. Abbreviations: ITR – immature-to-total neutrophil ratio; NEUT# – neutrophil count; NLR – neutrophil-to-lymphocyte ratio; PLR – platelet-to-lymphocyte ratio; PLT# – platelet count; WBC# – white blood cell count.*

Parameters such as NEUT-RI and NEUT-GI reflect the internal structure of circulating immune cells, so they show changes in response to infection much earlier than traditional counts, which only rise after a period of immune activation [72]. Therefore, characterising these parameters in a control cohort would identify abnormal values that may signal infection before clinical deterioration.

6.1.2 CBC Results From Umbilical Cord Blood Samples

However, repeated sampling, especially in preterm neonates, can induce iatrogenic anaemia [2]. One promising alternative to peripheral sampling is the use of umbilical cord blood, which provides an immediately available and non-invasive reflection of the neonate's haematological status at birth [166]. Unlike peripheral blood draws, which can be painful, technically challenging, and limited by the neonate's small circulating blood volume, umbilical cord blood can be collected without causing distress or harm. In addition, cord blood is typically discarded after delivery or used solely for blood gas analysis [167], representing an underutilised resource for early diagnostic testing.

Incorporating umbilical cord CBC analysis into the early sepsis workup could reduce the need for additional blood draws and minimise iatrogenic anaemia, particularly in preterm infants. Critically, umbilical cord blood samples offer an opportunity to characterise the neonatal haematological profile before postnatal environmental exposures, clinical deterioration, or non-inflammatory WBC spikes in response to bacterial colonisation of the gut [113], therefore facilitating earlier decision-making in the management of at-risk infants. Moreover, the use of umbilical cord blood overcomes the ethical and logistical barriers of sampling healthy term neonates for constructing reference ranges.

Establishing the Baseline Profile

Several studies have described baseline values for umbilical cord CBC results, as summarised in Table 6.2. However, most present only descriptive statistics (e.g., mean $\pm$ SD) rather than formal reference ranges, as recommended by the CLSI guidelines [137]. This lack of standardisation limits the applicability of existing data in clinical decision-making, as

Table 6.2: Umbilical cord complete blood count reference ranges from previously published studies.

Parameter	Sabnis et al., 2023 [172]	Muddathir et al., 2021 [168]	Manna et al., 2021 [173]	Angelo et al., 2021 [169]	Zitouni et al., 2020 [171]
RBC# (x 10^{12}/L)	2.45 – 6.27	4.14 ± 0.96[ab]	4.30 ± 0.60	3.55 – 5.52[b]	4.16 ± 0.65[c]
HGB (g/L)	81 – 214	145 ± 36[b]	152 ± 14	124 – 197	148 ± 23[c]
HCT (%)	29.0 – 67.0	44.9 ± 12.4[b]	46.7 ± 4.8	37.9 – 56.3	45.4 ± 7.2[c]
WBC# (x 10^9/L)	2.56 – 21.19[b]	11.74 ± 7.74[b]	16.20 ± 4.70	6.60 – 19.40[b]	13.64 ± 4.25
NEUT# (x 10^9/L)	3.40 – 8.70[b]	–	–	2.70 – 12.90[b]	6.62 ± 2.35
LYMPH# (x 10^9/L)	2.58 – 5.97[b]	–	–	1.90 – 8.30	5.15 ± 1.95[c]
MONO# (x 10^9/L)	0.09 – 1.57[b]	–	–	–	1.28 ± 0.68
EOSI# (x 10^9/L)	0.05 – 0.59	–	–	–	0.30 ± 0.18[c]
BASO# (x 10^9/L)	0.00 – 0.08	–	–	–	0.34 ± 0.38
PLT# (x 10^9/L)	17 – 479	257 ± 128[b]	227 ± 73	146 – 438	222 ± 95[c]

Statistically significant differences are denoted for the following clinical factors: [a]sex; [b]mode of delivery; and [c]gestational age at birth. Abbreviations: BASO# – basophil count; EOSI# – eosinophil count; HCT – haematocrit; HGB – haemoglobin content; LYMPH# – lymphocyte count; MONO# – monocyte count; NEUT# – neutrophil count; PLT# – platelet count; RBC# – red blood cell count; WBC# – white blood cell count.

formal reference ranges provide clinically actionable thresholds for interpreting individual results.

As with peripheral CBCs, previous research has shown variability in parameters of the umbilical cord CBC with key clinical factors. Male neonates typically present with higher RBC# values than females (4.14 and 4.13 x 10^{12}/L, respectively) [168], while MOD significantly influences RBC#, WBC#, and PLT# [168, 169]. These physiological differences are thought to reflect varying degrees of birth stress, oxygenation, and immune activation. GA remains a major confounder, with consistent effects on RBC#, HGB, and HCT values [170, 171].

Despite these findings, no published studies have established reference ranges for newer reticulocyte characteristics or immune cell activation markers, such as RET-HE, IRF, and NEUT-RI. These parameters may provide earlier or more specific insights into neonatal immune activation and bone marrow activity. However, their diagnostic potential in this context remains unknown.

Diagnostic Utility of the CBC for EOS

The diagnostic potential of the umbilical cord CBC for EOS remains underexplored. Most studies have focused on umbilical cord blood cultures, demonstrating high correlation with peripheral blood cultures in confirmed EOS cases [174, 175]. Only a few studies have investigated haematological parameters, with one reporting elevated ITR values in neonates with EOS compared to healthy controls (0.29 versus 0.19, $p = 0.001$) [176]. However, no study has systematically assessed the diagnostic performance of modern CBC parameters in umbilical cord blood samples.

Given their ease of collection, physiological relevance, and potential to inform early intervention, there is a strong rationale to study the diagnostic value of umbilical cord blood CBCs. These advantages underscore the need to evaluate the diagnostic performance, feasibility, and clinical utility of umbilical cord blood-derived CBC parameters in the early management of EOS.

6.1.3 Aims of the Chapter

This chapter aimed to conduct exploratory analyses of CBC results from both peripheral samples taken within the first three days of life and umbilical cord blood collected at delivery, using methods developed in earlier chapters. For each dataset, there were three key objectives: (1) establish reference ranges in line with CLSI guidelines; (2) evaluate the influence of key clinical variables on each CBC parameter, including sex, MOD, and GA; and (3) investigate the diagnostic potential of each CBC parameter for EOS.

Given the limited sample sizes in both cohorts, these analyses were exploratory in nature and lacked sufficient power for definitive conclusions. Nevertheless, this work provides foundational insights that may inform larger prospective studies and support the clinical integration of CBC-based diagnostics for EOS. Importantly, it also offers one of the first systematic evaluations of novel immune and reticulocyte indices in both peripheral and umbilical cord blood samples in neonates during the first three days of life.

6.2 Methods

6.2.1 Data Collection

CBC Results From the First Three Days of Life

The dataset used in this chapter is the same as described in Chapter 3. Ethical approval was obtained from the Oxford Research Ethics Committee (REC 08/H0606/139). An opt-out consent process was implemented, whereby parents were informed via patient information leaflets and consent was assumed unless explicitly declined. Data were anonymised before analyses to ensure compliance with ethical guidelines.

CBC results were obtained from neonatal samples analysed using a Sysmex XN-2000 (Kobe, Japan) in the John Radcliffe (Oxford, UK) haematology laboratory and an XN-450 in the NICU. These results were linked with clinical data extracted from EPR, Badgernet, and observation charts. CBC results obtained within two days after a surgical intervention or blood transfusion were excluded due to their potentially confounding effects on haematological parameters.

Where multiple CBC results were recorded for the same neonate on the same day, daily means were calculated for each parameter to minimise the influence of repeated sampling. Results were categorised based on clinical and laboratory signs of neonatal sepsis (Figure 6.1) – the key difference from the previous classifications was the addition of maternal and neonatal infection (defined by ICD-10 codes) as an exclusion criterion in the *unlikely sepsis* group. The dataset was filtered for results from the first three days of life.

CBC Results From Umbilical Cord Blood Samples

Ethical approval was obtained from the Mid and South Buckinghamshire Research Ethics Committee (REC 07/H0607/74). Written informed consent was obtained from pregnant women admitted to the John Radcliffe Maternity Unit (Oxford, UK) before delivery and sample collection. Data were anonymised before analyses to ensure compliance with ethical guidelines.

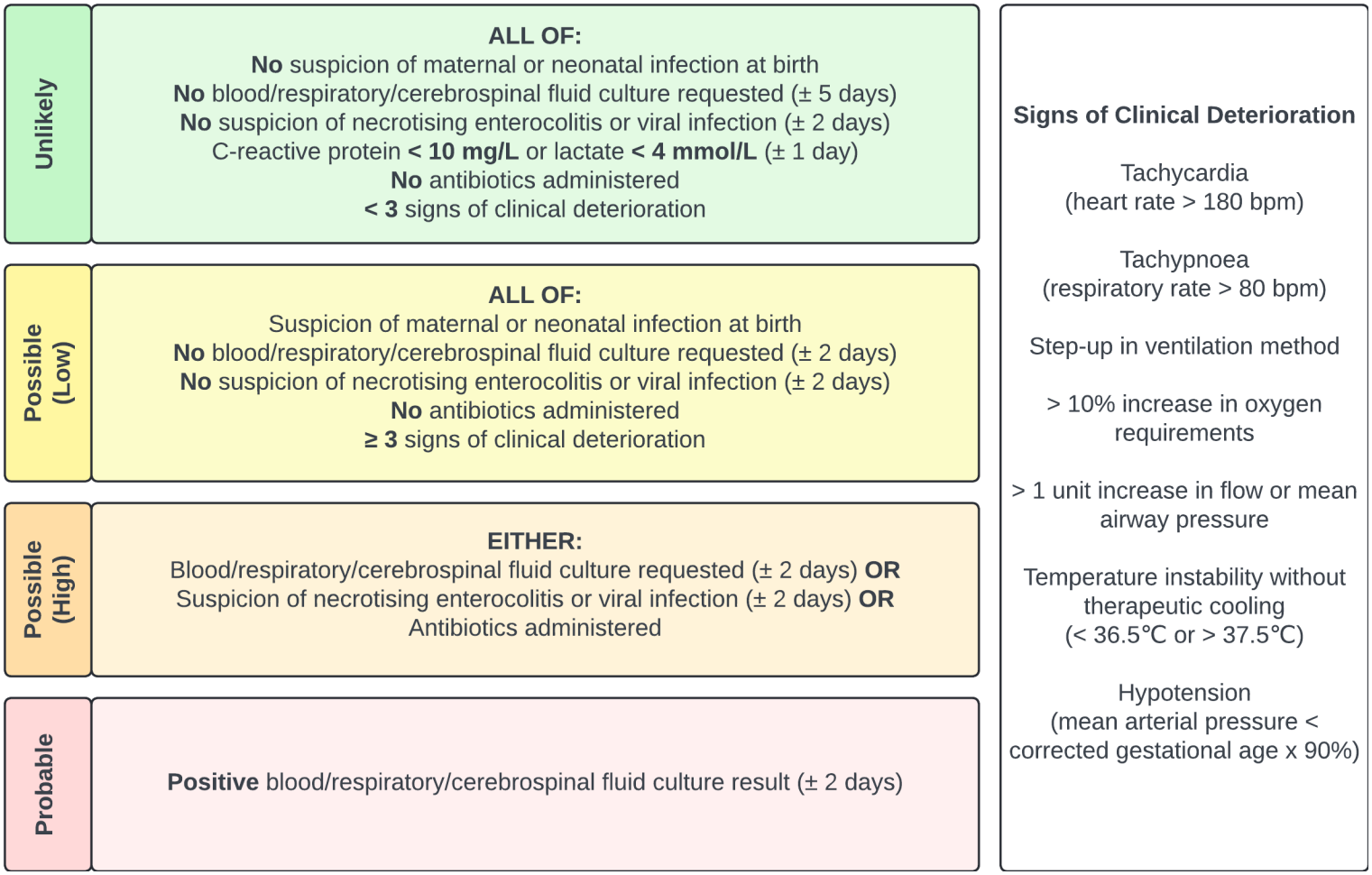


Figure 6.1: Criteria for the categorisation of complete blood count results for early-onset sepsis, based on clinical and laboratory signs of sepsis.

Immediately following delivery, paired umbilical cord venous and arterial blood samples were collected in EDTA tubes and analysed in duplicate within four hours using the XN-450 haematology analyser in the NICU. Matched maternal and neonatal clinical data were extracted from EPR.

CBC results were excluded from the initial analyses if there were maternal or neonatal concerns at birth (e.g. pneumothorax, congenital anomalies, or hydronephrosis) or if the newborn required NICU admission. Where paired arterial and venous samples were available, differences were assessed using paired Student's t-tests or Wilcoxon signed-rank tests. Due to better availability and minimal variation, only CBC results from the umbilical vein were retained for further analysis.

6.2.2 Evaluation of the Baseline Profile

Reference ranges were established for each dataset using the same statistical approach as described in Chapter 3, following the CLSI guidelines (EP28-A3C) [137]. The distribution of each CBC parameter was assessed using the Shapiro-Wilk test. Non-normally distributed parameters were transformed using an ordered quantile transformation and re-tested for normality. For normally distributed data, reference ranges were calculated as the mean $\pm$ 1.96 $\times$ SD, with 95% CIs for the bounds calculated as 1.96 $\times$ SD divided by the square root of the sample size. For non-parametric data, the 2.5^{th} and 97.5^{th} centiles were used, with CIs calculated using bootstrap methods.

CBC Results From the First Three Days of Life

The dataset from the first three days of life was filtered to include only CBC results categorised as *unlikely sepsis*. The distribution of values for each parameter was compared to that of the primary study cohort (i.e., CBC results classified as *unlikely sepsis* taken at least four days after birth) using the Kolmogorov-Smirnov test to evaluate whether the early postnatal period differed from the later neonatal reference profile. The effects of sex, postnatal age, GA, and MOD were assessed using a multivariate ANOVA:

$$value \sim sex + postnatal\ age + GA + MOD$$

CBC Results From Umbilical Cord Blood Samples

In the umbilical cord CBC dataset, the distribution of values for each parameter was compared to the early postnatal dataset (CBC results within the first three days of life, categorised as *unlikely sepsis*) using the Kolmogorov-Smirnov test, to determine whether umbilical cord CBC results are representative of the early postnatal haematological profile. The effects of neonatal sex, GA, and MOD were assessed using a multivariate ANOVA with the following model:

$$value \sim sex + GA + MOD$$

6.2.3 Diagnostic Accuracy of Individual CBC Parameters

The diagnostic potential of individual CBC parameters for EOS was evaluated in both datasets. For each parameter, comparisons were made using a Student's t-test or Wilcoxon signed-rank test, depending on the underlying distribution. Diagnostic accuracy was evaluated using ROC curve analysis, and the AUC, optimal threshold, sensitivity and specificity were calculated. Due to small sample sizes, ROC analyses were considered exploratory, and results were interpreted with caution.

CBC Results From the First Three Days of Life

The dataset was filtered to include only CBC results categorised as *unlikely* or *probable sepsis*. AUCs were compared to those calculated for LOS (as detailed in Chapter 4) using DeLong's test, where applicable. For parameters whose direction of the effect differed between EOS and LOS, direct comparison using DeLong's test could not be performed. Sensitivity and specificity were recalculated by applying LOS-specific optimal thresholds to determine generalisability between datasets.

CBC Results From Umbilical Cord Blood Samples

Two analyses were conducted comparing baseline CBC results with results from neonates (1) admitted to the NICU, and (2) with suspected sepsis. To assess whether diagnostic accuracy for EOS differed by sample source, AUCs for each CBC parameter were compared between peripheral and cord blood samples using DeLong's test, where applicable.

6.2.4 Diagnostic Accuracy of Previously Developed Models

CBC Results From the First Three Days of Life

To evaluate the potential utility of the previously developed diagnostic models (Chapter 5) for EOS, the models were applied to the dataset from the first few days of life. CBC results were re-categorised as *unlikely* or *probable sepsis* based on model-generated values and comparison with either the pre-defined or dataset-specific optimal thresholds. For each model, the AUC, optimal threshold, sensitivity, and specificity were calculated using ROC curve analysis.

To explore whether parameter selection differed between EOS and LOS, a logistic regression model was developed on the current dataset. The original list of 18 parameters was filtered to include only parameters which showed statistical significance in the current dataset. Model refinement was conducted using a stepwise selection approach. Given the limited sample size, this model was developed to examine which CBC parameters were retained and how they aligned with those in the previous chapter (Chapter 5). The AUC, optimal threshold, sensitivity, and specificity were calculated for descriptive purposes. All analyses were conducted using RStudio (R version 4.3.0) [86].

6.3 Results

Figures for the analysis of the baseline profile and diagnostic performance of each CBC parameter are available online at `https://4fsqon-emily-hyde.shinyapps.io/ThesisFigures/` [138] under the following tabs: *Newborn CBC Baseline Profile*; *Newborn CBC for EOS Diagnosis*; *Umbilical Cord CBC Baseline Profile*; *Umbilical Cord CBC for EOS Diagnosis*. Representative figures for the analysis of WBC# are presented throughout (Figures 6.2 – 6.5).

6.3.1 Sample Characteristics

Due to the limited sample sizes in both datasets, all analyses were exploratory and lacked sufficient robustness to draw solid conclusions.

CBC Results From the First Three Days of Life

During the study period, 875 CBC results from 487 neonates were collected during the first three days of life. Of these, 63 CBC results were categorised as *unlikely sepsis* and 27 as *probable sepsis* (Table 6.3). There were statistically significant differences between the *unlikely* and *probable sepsis* groups concerning sex, GA, postnatal age, and MOD (all p-values < 0.05). CBC results categorised as *probable sepsis* were from more preterm neonates than those categorised as *unlikely sepsis*.

CBC Results From Umbilical Cord Samples

Umbilical cord blood samples were collected from 114 neonates during the study period. A total of 98 neonates were classified as "healthy" and comprised the baseline cohort (Table 6.4). Of these, 92 neonates had matched venous and arterial samples available for analysis.

There were no significant differences in sex distribution between the baseline cohort and either the NICU-admitted group ($p = 0.98$) or the suspected sepsis group ($p = 0.48$). Both the NICU-admitted and suspected sepsis groups included significantly more preterm neonates compared to the baseline group ($p < 0.01$). Additionally, neonates with

Table 6.3: Sample characteristics for complete blood count results used for the evaluation of the baseline profile and diagnostic utility for early-onset sepsis in the first three days of life in neonates admitted to the Neonatal Intensive Care Unit.

		All Values	*Unlikely Sepsis*	*Probable Sepsis*	P-Value
CBC Results		875	63	27	–
Neonates		487	48	15	–
Sex (Male)		273 (56.1%)	15 (31.3%)	10 (66.7%)	0.03[a]
GA (Weeks)		36.0 (30.0 – 39.0)	37.1 (34.5 – 39.0)	31.7 (24.3 – 38.0)	0.02[b]
GA	Term ($\geq$ 37 weeks)	202 (41.5%)	26 (54.2%)	5 (33.3%)	0.27[a]
	Preterm ($\leq$ 36 weeks)	285 (58.5%)	22 (45.8%)	10 (66.7%)	
Postnatal Age (Days)		1 (0 – 2)	1 (0 – 1)	1 (1 – 3)	0.008[b]
MOD	Elective CS	68 (14.0%)	18 (37.5%)	1 (6.7%)	0.01[b]
	Emergency CS	218 (44.8%)	12 (25.0%)	2 (13.3%)	
	SVD	193 (39.6%)	17 (35.4%)	12 (80.0%)	

Values are presented as counts (%) or median (interquartile range). P-values are calculated from [a] Pearson's Chi-squared test or [b] Wilcoxon rank sum test. Significant p-values (< 0.05) are highlighted in red. Abbreviations: CBC – complete blood count; CS – Caesarean section; GA – gestational age at birth; MOD – mode of delivery; SVD – spontaneous vaginal delivery.

Table 6.4: Sample characteristics for umbilical cord complete blood count results used for the evaluation of the baseline profile and diagnostic utility for early-onset sepsis.

		All Samples	Vessel Comparison	Baseline Profile	NICU Admission	Suspected Sepsis	NICU Admission P-Value	Sepsis P-Value
Neonates		114	92	98	8	12	–	–
Sex (Male)		50 (43.9%)	39 (42.4%)	42 (42.9%)	4 (50.0%)	7 (58.3%)	0.98[a]	0.48[a]
GA (Weeks)		39.0 (37.1 – 39.6)	39.1 (37.7 – 39.8)	39.1 (37.4 – 39.6)	35.4 (30.6 – 37.6)	36.9 (33.6 – 38.0)	0.003[b]	0.004[b]
GA	Term ($\geq$ 37 weeks)	94 (82.5%)	83 (90.2%)	86 (87.8%)	3 (37.5%)	6 (50.0%)	0.001[a]	0.004[a]
GA	Preterm ($\leq$ 36 weeks)	20 (17.5%)	9 (9.8%)	12 (12.2%)	5 (62.5%)	6 (50.0%)		
MOD	CS Not in Labour	60 (52.6%)	52 (56.5%)	55 (56.1%)	2 (25.0%)	2 (16.7%)	0.06[a]	0.02[a]
MOD	CS in Labour	12 (10.5%)	9 (9.8%)	10 (10.2%)	0 (0.0%)	1 (8.3%)		
MOD	SVD	42 (36.8%)	31 (33.7%)	33 (33.7%)	6 (75.0%)	9 (75.0%)		

Values are presented as counts (%) or median (interquartile range). P-values are calculated from [a] Pearson's Chi-squared test or [b] Wilcoxon rank sum test. Significant p-values (< 0.05) are highlighted in red. Abbreviations: CS – Caesarean section; GA – gestational age at birth; MOD – mode of delivery; NICU – Neonatal Intensive Care Unit; SVD – spontaneous vaginal delivery.

suspected sepsis were more likely to have had a SVD compared to the baseline cohort (75% versus 34%; p = 0.02).

6.3.2 Evaluation of the Baseline Profile

CBC Results From the First Three Days of Life

Values from the first three days of life significantly differed from CBC results thereafter in all but five parameters (MCHC, LYMPH#, ITR, NEUT-RI, and RE-MONO), suggesting that these parameters are stable across the neonatal period (Table 6.5; Figure 6.2). Similarly, 33 parameters were statistically different between umbilical cord blood CBC results and peripheral CBC results, suggesting that the umbilical cord baseline profile is not fully representative of the early postnatal period.

Parametric reference ranges were calculated for all parameters except AS-LYMPH, HFLC, and RE-MONO, which could not be normalised due to a high number of zero values and thus had non-parametric reference ranges defined by the 2.5^{th} and 97.5^{th} percentiles.

Twelve parameters were significantly different between males and females ($p < 0.05$), and four showed potential clinical relevance (pes > 0.10). For example, NEUT-GI was higher in females than in males. Fourteen parameters exhibited significant age-related changes, of which 13 showed potentially clinically relevant trends. Specifically, IG# and ITR were significantly higher in CBC results obtained on the day of birth compared to those taken between days one and three of life.

Only 12 parameters were statistically significantly different between term and preterm neonates, with nine of these also being clinically relevant. Finally, four parameters were both statistically and clinically different between the different MODs. WBC# (Figure 6.2) and NEUT# were both lower in preterm neonates compared to term neonates, and higher in neonates delivered by SVD than those delivered by elective CS.

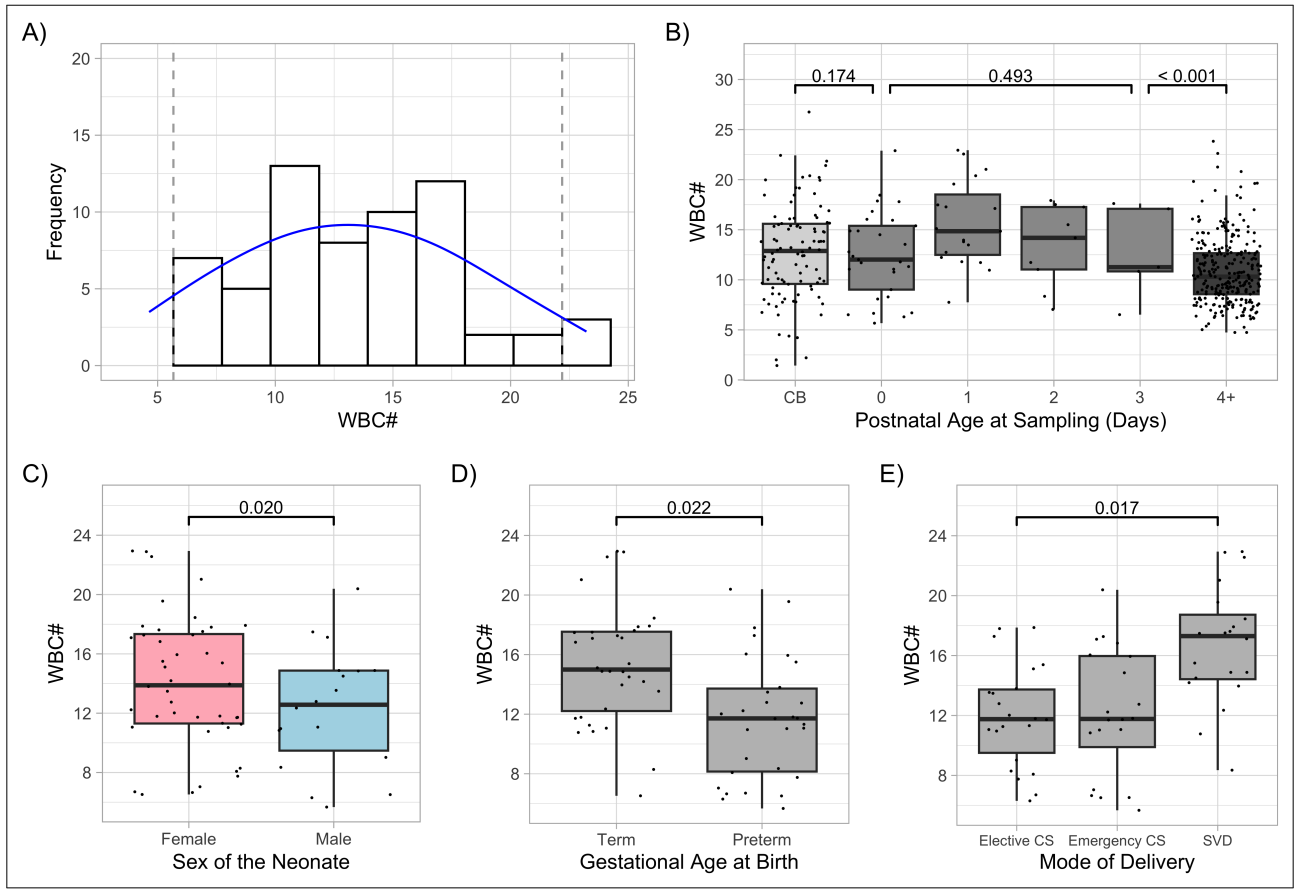


Figure 6.2: Analysis of the baseline white blood cell count (WBC#; x 10^9/L) in the first three days of life in clinically stable neonates admitted to a Neonatal Intensive Care Unit.

A) Histogram with reference ranges (grey dotted lines) and distribution curve (blue line). **B)** Box plot for postnatal age at sampling, with comparisons to values from umbilical cord blood (CB) samples and values from day four of life onwards (4+). **C)** Box plot by sex. **D)** Box plot by gestational age at birth. **E)** Box plot by mode of delivery.

Abbreviations: CB – umbilical cord blood; CS – Caesarean section; SVD – spontaneous vaginal delivery.

Table 6.5: Evaluation of the baseline profile for complete blood count parameters in the first three days of life.

	KS P-Value		Reference Ranges			Sex		Age		GA		MOD	
Parameter	UC vs Newborn	Newborn vs Neonatal	Lower	Average	Upper	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect
Red blood cell parameters													
RBC#	<0.001	<0.001	3.85 (3.84 – 4.14)	5.14 (4.98 – 5.30)	6.43 (6.14 – 6.72)	0.20	0.03	0.83	0.00	0.43	0.01	0.61	0.02
HGB	<0.001	<0.001	143 (143 – 153)	190 (184 – 196)	237 (227 – 239)	0.92	0.00	0.99	0.00	0.51	0.01	0.28	0.05
HCT	<0.001	<0.001	40.0 (39.6 – 42.6)	52.0 (50.5 – 53.5)	64.0 (61.4 – 64.1)	0.87	0.00	0.38	0.01	0.43	0.01	0.23	0.05
MCV	0.003	<0.001	89.0 (86.3 – 91.7)	101.2 (99.6 – 102.7)	113.3 (110.6 – 116.1)	0.02	0.10	0.005	0.13	0.99	0.00	0.10	0.08
MCH	0.02	<0.001	32.2 (31.1 – 33.2)	36.9 (36.3 – 37.4)	41.5 (40.5 – 42.2)	0.010	0.11	0.38	0.01	0.88	0.00	0.21	0.06
MCHC	<0.001	0.51	341 (336 – 346)	364 (361 – 367)	387 (382 – 388)	0.39	0.01	<0.001	0.21	0.84	0.00	0.69	0.01
RDW-SD	0.03	<0.001	45.8 (43.8 – 49.0)	60.1 (58.3 – 61.9)	74.4 (71.2 – 77.1)	0.04	0.08	0.30	0.02	0.70	0.00	0.34	0.04
RDW-CV	0.003	0.02	13.8 (13.2 – 14.5)	16.8 (16.4 – 17.2)	19.8 (19.2 – 20.3)	0.38	0.01	0.58	0.01	0.46	0.01	0.95	0.00
NRBC#	<0.001	<0.001	0.01 (0.00 – 0.02)	0.24 (0.13 – 0.32)	1.86 (1.07 – 2.81)	0.16	0.04	0.007	0.12	0.52	0.01	0.48	0.03
NRBC%	<0.001	0.002	0.1 (0.0 – 0.1)	1.9 (1.2 – 2.7)	15.3 (11.1 – 17.5)	0.51	0.01	0.01	0.11	0.96	0.00	0.20	0.06

Continued on next page

Continued from last page

	KS P-Value		Reference Ranges			Sex		Age		GA		MOD	
Parameter	UC vs Newborn	Newborn vs Neonatal	Lower	Average	Upper	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect
RET#	<0.001	<0.001	0.121 (0.091 – 0.152)	0.257 (0.240 – 0.274)	0.393 (0.362 – 0.423)	0.65	0.00	0.54	0.01	0.83	0.00	0.53	0.02
RET%	0.003	<0.001	1.6 (1.4 – 3.5)	5.0 (4.7 – 5.2)	8.5 (6.8 – 8.7)	0.25	0.02	0.75	0.00	0.30	0.02	0.20	0.06
IRF	<0.001	<0.001	24.7 (21.5 – 27.9)	39.0 (37.2 – 40.8)	53.3 (50.1 – 56.5)	0.14	0.04	0.06	0.06	0.84	0.00	0.54	0.02
RBC-HE	<0.001	<0.001	30.5 (29.8 – 31.3)	34.1 (33.6 – 34.5)	37.6 (36.8 – 37.6)	0.11	0.04	0.91	0.00	0.52	0.01	0.47	0.03
RET-HE	0.15	<0.001	29.4 (28.8 – 30.4)	34.0 (33.4 – 34.6)	38.6 (37.6 – 39.6)	0.02	0.09	0.005	0.13	0.45	0.01	0.57	0.02
DELTA-HE	0.04	<0.001	-4.4 (-6.5 – -3.4)	0.0 (-0.4 – 0.2)	2.4 (2.0 – 2.6)	0.42	0.01	<0.001	0.23	0.68	0.00	0.06	0.10
HYPO-HE	0.11	<0.001	0.3 (0.3 – 0.4)	0.6 (0.5 – 0.7)	2.1 (1.3 – 2.4)	0.50	0.01	0.61	0.01	0.19	0.03	0.12	0.07
HYPER-HE	<0.001	<0.001	1.0 (0.9 – 1.0)	3.2 (2.4 – 3.8)	8.6 (8.0 – 14.7)	0.07	0.06	0.89	0.00	0.58	0.01	0.39	0.03
MICRO-RBC	<0.001	<0.001	0.8 (0.7 – 1.2)	1.7 (1.5 – 1.8)	3.6 (2.8 – 5.9)	0.07	0.06	0.002	0.16	0.13	0.04	0.04	0.11
MACRO-RBC	0.003	<0.001	4.7 (4.4 – 6.0)	11.4 (9.8 – 12.6)	28.0 (24.8 – 30.6)	0.02	0.10	0.006	0.13	0.98	0.00	0.18	0.06

Continued on next page

Continued from last page

	KS P-Value		Reference Ranges			Sex		Age		GA		MOD	
Parameter	UC vs Newborn	Newborn vs Neonatal	Lower	Average	Upper	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect
White blood cell parameters													
WBC#	0.17	<0.001	5.67 (5.67 – 6.88)	13.58 (12.48 – 14.67)	22.19 (20.27 – 22.94)	0.02	0.09	0.49	0.01	0.02	0.09	0.02	0.14
NEUT#	0.10	<0.001	0.62 (0.62 – 1.85)	7.50 (6.58 – 8.42)	14.77 (13.15 – 16.39)	0.02	0.09	0.65	0.00	0.002	0.16	0.02	0.12
NEUT%	<0.001	<0.001	16.3 (9.6 – 28.2)	55.9 (52.5 – 59.1)	69.8 (68.6 – 73.1)	0.08	0.05	0.77	0.00	<0.001	0.21	0.15	0.07
LYMPH#	0.18	0.79	2.03 (1.90 – 2.52)	4.20 (3.80 – 4.45)	7.44 (6.00 – 9.05)	0.16	0.04	0.07	0.06	0.46	0.01	0.20	0.06
LYMPH%	0.17	<0.001	17.7 (15.3 – 20.1)	30.2 (28.7 – 34.1)	64.6 (56.2 – 76.7)	0.24	0.03	0.23	0.03	0.007	0.12	0.11	0.08
MONO#	0.50	<0.001	0.37 (0.32 – 0.55)	1.25 (1.19 – 1.41)	2.72 (2.53 – 3.14)	0.07	0.06	0.13	0.04	0.80	0.00	0.11	0.08
MONO%	0.03	<0.001	5.0 (4.3 – 5.8)	9.9 (9.5 – 10.7)	18.8 (14.7 – 26.6)	0.87	0.00	0.01	0.11	0.11	0.05	0.73	0.01
EOSI#	<0.001	<0.001	0.01 (0.00 – 0.04)	0.23 (0.19 – 0.27)	0.95 (0.73 – 1.77)	0.003	0.15	0.06	0.06	0.63	0.00	0.09	0.08
EOSI%	<0.001	<0.001	0.1 (0.0 – 0.3)	2.1 (1.4 – 2.5)	7.2 (6.5 – 10.2)	<0.001	0.19	0.05	0.07	0.70	0.00	0.20	0.06
BASO#	0.004	0.008	0.03 (0.03 – 0.04)	0.10 (0.10 – 0.11)	0.34 (0.27 – 0.49)	0.19	0.03	0.23	0.03	0.13	0.04	0.17	0.06

Continued on next page

Continued from last page

	KS P-Value		Reference Ranges			Sex		Age		GA		MOD	
Parameter	UC vs Newborn	Newborn vs Neonatal	Lower	Average	Upper	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect
BASO%	<0.001	0.01	0.3 (0.3 – 0.4)	0.7 (0.7 – 0.8)	2.0 (1.7 – 2.8)	0.78	0.00	0.13	0.04	0.85	0.00	0.56	0.02
IG#	<0.001	<0.001	0.04 (0.02 – 0.05)	0.21 (0.17 – 0.25)	1.43 (1.29 – 1.64)	0.42	0.01	<0.001	0.19	<0.001	0.30	0.56	0.02
IG%	<0.001	<0.001	0.5 (0.4 – 0.6)	1.6 (1.4 – 2.1)	7.8 (7.3 – 9.3)	0.94	0.00	<0.001	0.26	<0.001	0.31	0.55	0.02
NLR	0.03	<0.001	0.12 (0.12 – 0.35)	1.88 (1.63 – 2.13)	3.85 (3.41 – 4.28)	0.25	0.02	0.50	0.01	0.005	0.13	0.08	0.09
ITR	<0.001	0.12	0.01 (0.01 – 0.01)	0.03 (0.02 – 0.04)	0.14 (0.12 – 0.16)	0.48	0.01	<0.001	0.25	0.002	0.16	0.33	0.04
NEUT-RI	<0.001	0.14	34.7 (33.1 – 36.3)	42.0 (41.1 – 42.9)	49.4 (47.7 – 49.9)	0.99	0.00	0.76	0.00	0.28	0.02	0.27	0.05
NEUT-GI	0.005	0.02	140.2 (138.1 – 142.2)	149.3 (148.2 – 150.5)	158.5 (156.4 – 160.5)	0.004	0.14	0.10	0.05	<0.001	0.18	0.47	0.03
NEUT-TW[a]	0.35	0.02	25.1 (23.0 – 27.0)	33.8 (32.0 – 35.0)	51.7 (48.5 – 52.0)	0.71	0.00	0.003	0.15	0.004	0.14	0.76	0.01
RE-LYMPH	0.34	<0.001	0.1 (0.0 – 0.2)	1.1 (0.8 – 1.7)	10.7 (10.5 – 12.2)	0.03	0.08	0.43	0.01	0.03	0.08	0.40	0.03
AS-LYMPH[a]	<0.001	<0.001	0.0 (0.0 – 0.0)	0.0 (0.0 – 0.0)	0.9 (0.7 – 0.9)	0.12	0.04	0.24	0.02	0.03	0.08	0.03	0.12
HFLC[a]	<0.001	<0.001	0.0 (0.0 – 0.0)	0.0 (0.0 – 0.0)	0.0 (0.0 – 0.0)	0.03	0.08	0.62	0.00	0.07	0.06	0.16	0.06

Continued on next page

Continued from last page

	KS P-Value		Reference Ranges			Sex		Age		GA		MOD	
Parameter	UC vs Newborn	Newborn vs Neonatal	Lower	Average	Upper	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect
RE-MONO[a]	0.81	0.48	0.0 (0.0 – 0.0)	0.2 (0.2 – 0.3)	2.4 (1.4 – 6.1)	0.82	0.00	0.05	0.07	0.78	0.00	0.41	0.03
Platelet parameters													
PLT#	0.79	<0.001	97 (26 – 118)	248 (237 – 275)	385 (338 – 566)	0.27	0.02	0.84	0.00	0.69	0.00	0.14	0.07
PDW	<0.001	<0.001	9.2 (9.1 – 9.7)	10.9 (10.7 – 11.5)	15.2 (14.0 – 15.5)	0.95	0.00	0.60	0.01	0.82	0.00	0.63	0.02
MPV	0.02	<0.001	9.0 (8.8 – 9.3)	10.4 (10.2 – 10.6)	11.8 (11.5 – 12.1)	0.29	0.02	0.09	0.05	0.12	0.04	0.09	0.09
P-LCR	0.005	<0.001	16.7 (16.1 – 19.0)	27.2 (25.8 – 28.5)	37.7 (35.3 – 39.4)	0.56	0.01	0.08	0.05	0.27	0.02	0.11	0.08
PCT	0.87	<0.001	0.1 (0.1 – 0.2)	0.3 (0.2 – 0.3)	0.4 (0.4 – 0.4)	0.20	0.03	0.76	0.00	0.62	0.01	0.19	0.06
IPF	<0.001	<0.001	2.6 (2.1 – 3.2)	5.3 (4.8 – 5.7)	12.7 (10.4 – 47.9)	0.80	0.00	0.30	0.02	0.99	0.00	0.56	0.02
PLR	<0.001	0.02	0.028 (0.007 – 0.032)	0.056 (0.053 – 0.064)	0.109 (0.094 – 0.152)	0.57	0.01	0.49	0.01	0.82	0.00	0.29	0.04

Effect refers to partial eta-squared (pes) values from multivariate analyses of variance. All values were calculated parametrically unless transformation failed to normalise values ([a]). Average refers to the mean or median, depending on the underlying distribution. Significant p-values ($p < 0.05$) and effect sizes ($pes > 0.10$) are highlighted in red. Abbreviations: GA – gestational age at birth; KS – Kolmogorov-Smirnov; MOD – mode of delivery; UC – umbilical cord. For parameter abbreviations, see List of Complete Blood Count Parameter Abbreviations (p. xxi).

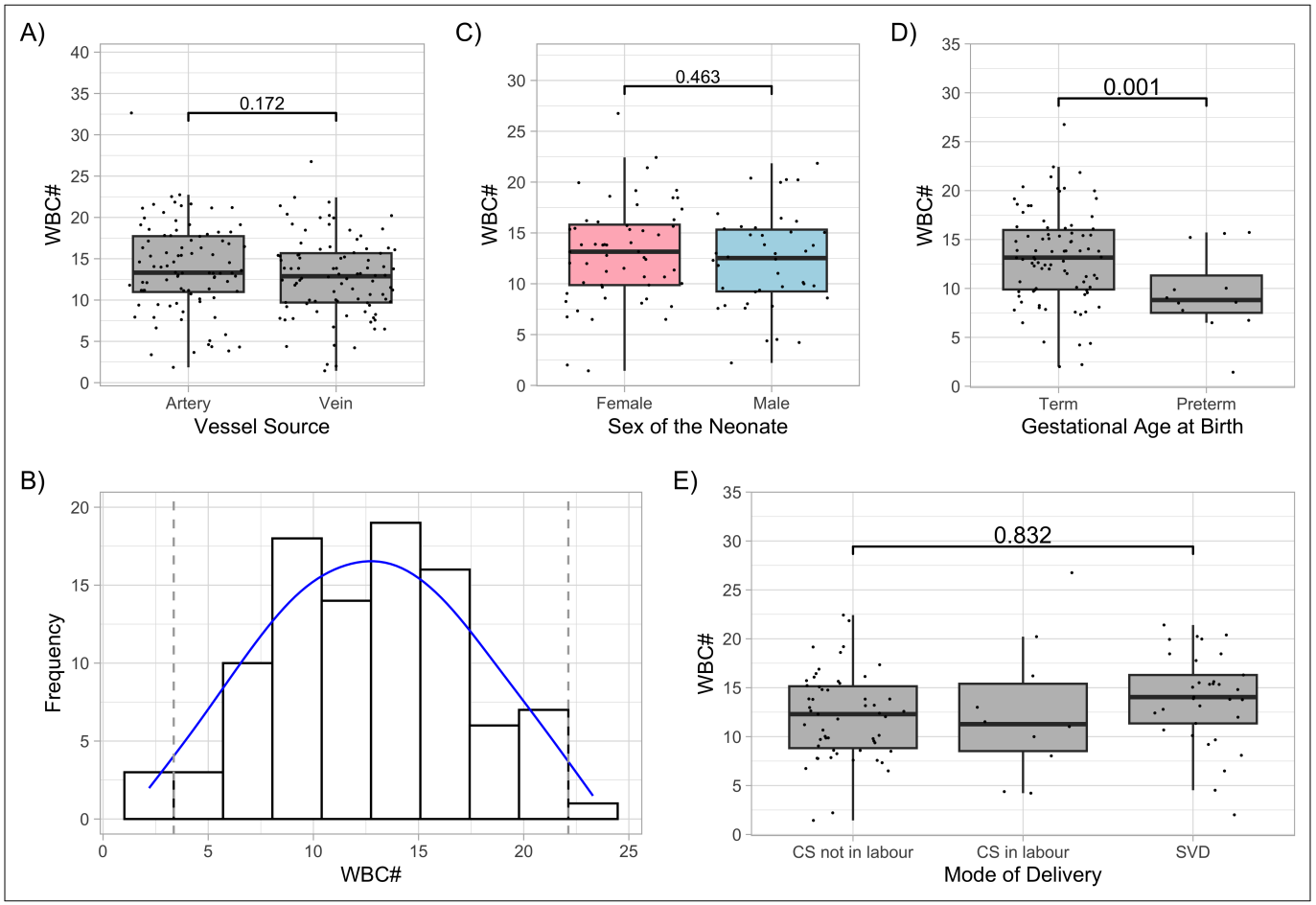


Figure 6.3: Analysis of the baseline umbilical cord white blood cell count (WBC#; x 10^9/L).

A) Box plot by umbilical vessel source. **B)** Histogram with reference ranges (grey dotted lines) and distribution curve (blue line). **C)** Box plot by sex. **D)** Box plot by gestational age at birth. **E)** Box plot by mode of delivery.

Abbreviations: CS – Caesarean section; SVD – spontaneous vaginal delivery.

Table 6.6: Evaluation of the baseline profile for umbilical cord complete blood count parameters.

	Vessel		Reference Ranges			Sex		GA		MOD	
Parameter	P-Value	Effect	Lower	Average	Upper	P-Value	Effect	P-Value	Effect	P-Value	Effect
Red blood cell parameters											
RBC#	0.16	0.21	3.16 (2.96 – 3.37)	4.32 (4.20 – 4.43)	5.47 (5.27 – 5.67)	0.60	0.00	0.59	0.00	0.02	0.08
HGB	0.19	0.19	113 (105 – 120)	155 (151 – 159)	198 (190 – 205)	0.67	0.00	0.46	0.01	0.007	0.10
HCT	0.53	0.09	32.7 (30.5 – 34.8)	45.1 (43.9 – 46.4)	57.6 (55.4 – 59.8)	0.78	0.00	0.69	0.00	0.004	0.11
MCV	0.20	-0.19	93.7 (91.8 – 95.6)	104.7 (103.5 – 105.8)	115.6 (113.7 – 116.6)	0.38	0.01	0.02	0.05	0.16	0.04
MCH	0.61	0.08	32.4 (31.8 – 33.0)	36.0 (35.6 – 36.3)	39.5 (38.9 – 40.0)	0.67	0.00	0.002	0.10	0.53	0.01
MCHC	0.02	0.36	324 (321 – 328)	344 (342 – 346)	364 (360 – 367)	0.41	0.01	0.20	0.02	0.30	0.03
RDW-SD	0.40	-0.12	53.4 (50.2 – 54.9)	61.7 (60.7 – 63.1)	78.0 (75.5 – 79.9)	0.85	0.00	0.68	0.00	0.19	0.04
RDW-CV	0.75	0.05	14.5 (14.5 – 14.7)	16.0 (15.8 – 16.3)	19.2 (18.7 – 20.0)	0.65	0.00	0.08	0.03	0.53	0.01
NRBC#	0.58	0.08	0.09 (0.08 – 0.13)	0.62 (0.52 – 0.69)	2.18 (2.01 – 2.28)	0.96	0.00	0.34	0.01	0.14	0.04
NRBC%	0.91	-0.02	0.7 (0.5 – 1.1)	4.8 (4.4 – 5.2)	13.7 (12.3 – 30.8)	0.94	0.00	0.45	0.01	0.08	0.05
RET#	0.66	0.06	0.107 (0.102 – 0.122)	0.192 (0.183 – 0.200)	0.276 (0.261 – 0.291)	0.05	0.04	0.87	0.00	0.06	0.06
RET%	0.71	-0.05	2.7 (2.5 – 3.0)	4.5 (4.3 – 4.6)	6.2 (5.9 – 6.5)	0.008	0.07	0.84	0.00	0.75	0.01
IRF	0.44	-0.11	26.2 (24.7 – 27.8)	35.0 (34.1 – 35.9)	43.8 (42.3 – 45.4)	0.002	0.09	0.41	0.01	0.04	0.07
RBC-HE	0.40	-0.13	30.8 (28.7 – 31.8)	35.6 (35.1 – 36.0)	38.6 (38.1 – 38.8)	0.09	0.03	0.003	0.09	0.005	0.11
RET-HE	0.20	-0.19	27.9 (25.3 – 30.3)	34.5 (33.9 – 34.9)	39.8 (39.3 – 40.4)	0.17	0.02	0.05	0.04	0.03	0.07
DELTA-HE	0.42	-0.12	-4.4 (-5.1 – -3.8)	-0.9 (-1.3 – -0.6)	2.6 (2.0 – 2.7)	0.76	0.00	0.53	0.00	0.39	0.02
HYPO-HE	0.03	0.32	0.3 (0.3 – 0.4)	0.7 (0.6 – 0.8)	2.7 (1.6 – 4.5)	0.38	0.01	0.45	0.01	0.02	0.08

Continued on next page

Continued from last page

	Vessel		Reference Ranges			Sex		GA		MOD	
Parameter	P-Value	Effect	Lower	Average	Upper	P-Value	Effect	P-Value	Effect	P-Value	Effect
HYPER-HE	0.70	-0.06	1.3 (1.1 – 1.7)	6.0 (5.1 – 6.9)	14.2 (13.1 – 15.1)	0.12	0.03	0.003	0.09	0.02	0.08
MICRO-RBC	0.10	0.24	0.6 (0.6 – 0.7)	1.2 (1.1 – 1.3)	2.6 (2.1 – 4.0)	0.29	0.01	0.74	0.00	0.62	0.01
MACRO-RBC	0.23	-0.18	6.0 (5.6 – 7.5)	15.4 (14.1 – 16.2)	32.2 (29.9 – 35.2)	0.38	0.01	0.05	0.04	0.12	0.05
White blood cell parameters											
WBC#	0.17	0.20	3.36 (1.71 – 5.00)	12.74 (11.79 – 13.69)	22.12 (20.48 – 23.77)	0.46	0.01	<0.001	0.11	0.83	0.00
NEUT#	0.56	0.09	0.98 (0.77 – 1.95)	6.51 (5.95 – 7.07)	12.05 (11.08 – 13.02)	0.22	0.02	<0.001	0.19	0.96	0.00
NEUT%	0.21	-0.18	28.8 (25.5 – 32.1)	47.7 (45.8 – 49.7)	66.7 (63.4 – 69.8)	0.16	0.02	<0.001	0.18	0.18	0.04
LYMPH#	0.08	0.26	0.79 (0.39 – 1.42)	4.38 (4.01 – 4.74)	7.97 (7.34 – 8.60)	0.75	0.00	0.27	0.01	0.50	0.02
LYMPH%	0.20	0.19	15.5 (12.6 – 18.5)	32.4 (30.7 – 34.1)	49.2 (46.3 – 52.2)	0.62	0.00	<0.001	0.12	0.39	0.02
MONO#	0.22	0.18	0.34 (0.27 – 0.49)	1.24 (1.13 – 1.31)	2.77 (2.40 – 2.89)	0.91	0.00	0.04	0.05	0.63	0.01
MONO%	0.96	0.01	5.7 (5.2 – 6.4)	9.6 (9.2 – 10.0)	13.4 (12.7 – 14.1)	0.39	0.01	0.21	0.02	0.10	0.05
EOSI#	0.89	0.02	0.03 (0.02 – 0.10)	0.38 (0.32 – 0.44)	1.28 (1.08 – 1.33)	0.005	0.08	0.25	0.01	0.36	0.02
EOSI%	0.52	-0.10	0.8 (0.2 – 0.9)	3.0 (2.8 – 3.3)	7.7 (7.0 – 8.9)	<0.001	0.12	0.22	0.02	0.45	0.02
BASO#	0.21	0.19	0.03 (0.02 – 0.04)	0.13 (0.11 – 0.15)	0.45 (0.39 – 0.55)	0.72	0.00	0.68	0.00	0.08	0.05
BASO%	0.53	0.09	0.5 (0.4 – 0.5)	1.0 (0.9 – 1.2)	3.7 (2.9 – 4.3)	0.44	0.01	<0.001	0.13	0.02	0.08
IG#	0.48	0.10	0.04 (0.03 – 0.09)	0.42 (0.32 – 0.50)	1.60 (1.27 – 2.01)	0.15	0.02	<0.001	0.14	0.25	0.03
IG%	0.98	0.00	0.9 (0.8 – 1.1)	3.1 (2.7 – 3.8)	8.5 (7.5 – 9.0)	0.13	0.03	0.006	0.08	0.10	0.05
NLR	0.16	-0.21	0.60 (0.42 – 0.69)	1.60 (1.46 – 1.74)	4.70 (2.94 – 5.70)	0.33	0.01	<0.001	0.16	0.41	0.02
ITR	0.59	0.08	0.02 (0.02 – 0.02)	0.07 (0.06 – 0.08)	0.16 (0.15 – 0.18)	0.28	0.01	0.12	0.03	0.07	0.06

Continued on next page

Continued from last page

	Vessel		Reference Ranges			Sex		GA		MOD	
Parameter	P-Value	Effect	Lower	Average	Upper	P-Value	Effect	P-Value	Effect	P-Value	Effect
NEUT-RI	0.54	0.09	35.3 (34.6 – 36.1)	39.8 (39.3 – 40.2)	44.2 (43.4 – 45.0)	0.13	0.02	0.003	0.09	0.72	0.01
NEUT-GI	0.03	-0.33	136.8 (135.1 – 138.6)	146.6 (145.6 – 147.6)	156.4 (154.7 – 158.1)	0.96	0.00	0.05	0.04	0.96	0.00
NEUT-TW	0.65	0.07	24.8 (23.6 – 25.8)	33.2 (32.1 – 34.9)	57.3 (50.2 – 58.7)	0.38	0.01	0.26	0.01	0.48	0.02
RE-LYMPH	0.01	0.38	0.0 (0.0 – 0.1)	1.5 (1.1 – 1.9)	6.6 (5.3 – 7.7)	0.81	0.00	0.57	0.00	0.08	0.05
AS-LYMPH[a]	0.57	0.04	0.0 (0.0 – 0.0)	0.0 (0.0 – 0.0)	0.0 (0.0 – 0.1)	0.18	0.02	0.34	0.01	0.44	0.02
HFLC[a]	0.06	0.14	0.0 (0.0 – 0.0)	0.0 (0.0 – 0.0)	0.0 (0.0 – 0.0)	0.63	0.00	0.64	0.00	0.10	0.05
RE-MONO	0.18	0.20	0.0 (0.0 – 0.0)	0.2 (0.2 – 0.3)	1.9 (1.4 – 2.4)	0.11	0.03	0.46	0.01	0.22	0.03
Platelet parameters											
PLT#	0.37	-0.13	23 (16 – 41)	263 (246 – 278)	446 (400 – 453)	0.58	0.00	<0.001	0.11	0.58	0.01
PDW	0.003	0.47	8.3 (8.1 – 8.4)	10.1 (10.0 – 10.3)	17.1 (13.4 – 20.2)	0.99	0.00	0.26	0.02	0.03	0.08
MPV	0.004	0.45	8.8 (8.3 – 8.9)	10.0 (9.8 – 10.1)	11.8 (11.4 – 13.4)	0.92	0.00	0.03	0.05	0.05	0.07
P-LCR	0.002	0.48	14.5 (11.9 – 15.7)	23.5 (22.3 – 24.4)	41.6 (34.9 – 47.7)	0.91	0.00	0.04	0.05	0.03	0.08
PCT	0.39	-0.13	0.1 (0.0 – 0.1)	0.3 (0.2 – 0.3)	0.4 (0.4 – 0.5)	0.84	0.00	<0.001	0.13	0.93	0.00
IPF	<0.001	0.50	1.8 (1.4 – 2.0)	3.4 (3.2 – 3.7)	31.8 (16.5 – 35.0)	0.28	0.01	0.23	0.02	0.18	0.04
PLR	0.006	0.41	0.006 (0.004 – 0.008)	0.017 (0.016 – 0.019)	0.080 (0.038 – 0.145)	0.91	0.00	0.14	0.02	0.60	0.01

Vessel Effect refers to Cohen's d or Wilcoxon r effect size. Sex/GA/MOD Effect refers to partial eta-squared (pes) values from multivariable analyses of variance. All values were calculated parametrically unless transformation failed to normalise values ([a]). Average refers to the mean or median, depending on the underlying distribution. Significant p-values ($p < 0.05$) and effect sizes (absolute Cohen's $d > 0.5$; absolute Wilcoxon $d > 0.3$; $pes > 0.1$) are highlighted in red. Abbreviations: GA – gestational age at birth; KS – Kolmogorov-Smirnov; MOD – mode of delivery. For parameter abbreviations, see List of Complete Blood Count Parameter Abbreviations (p. xxi).

CBC Results From Umbilical Cord Samples

Statistically significant differences were observed between venous and arterial samples in nine parameters, although none demonstrated a large effect size (Table 6.6; Figure 6.3). Parametric reference ranges were calculated for all parameters except AS-LYMPH and HFLC, which could not be normalised and thus had non-parametric reference ranges defined by the 2.5^{th} and 97.5^{th} centiles.

In a multivariate analysis, sex significantly impacted four parameters with large effect sizes in only EOSI% values, which were higher in males than in females. GA affected 20 parameters, with ten showing large effect sizes. NEUT#, IG#, and PLT# all increased with GA ($p \leq 0.001$). MOD influenced 11 parameters, with two of these having large effect sizes.

6.3.3 Diagnostic Accuracy of Individual CBC Parameters

CBC Results From the First Three Days of Life

Of the 49 CBC parameters evaluated in the peripheral CBC dataset, 29 parameters showed statistically significant differences between *unlikely* and *probable sepsis* groups, including NRBC#, NEUT-RI, and PLT# (Table 6.7; Figure 6.4). All except IG# and ITR showed large effect sizes (absolute Cohen's $d > 0.50$ or Wilcoxon $r > 0.30$).

Five parameters achieved an AUC greater than 0.8, indicating strong diagnostic accuracy. NRBC# achieved the highest specificity (95%), with an optimal threshold of 1.18 x 10^9/L and 82% sensitivity. NRBC% also performed well, with an optimal threshold of 6.5% achieving 82% sensitivity and 89% specificity.

However, applying the pre-defined optimal thresholds from Chapter 4 (Table 4.2, p. 97) yielded poor performance for most parameters. IRF could not be evaluated as the pre-defined threshold was lower than all observed values. NRBC# and NRBC% both achieved a high sensitivity (89%), while RET-HE had the highest specificity (95%).

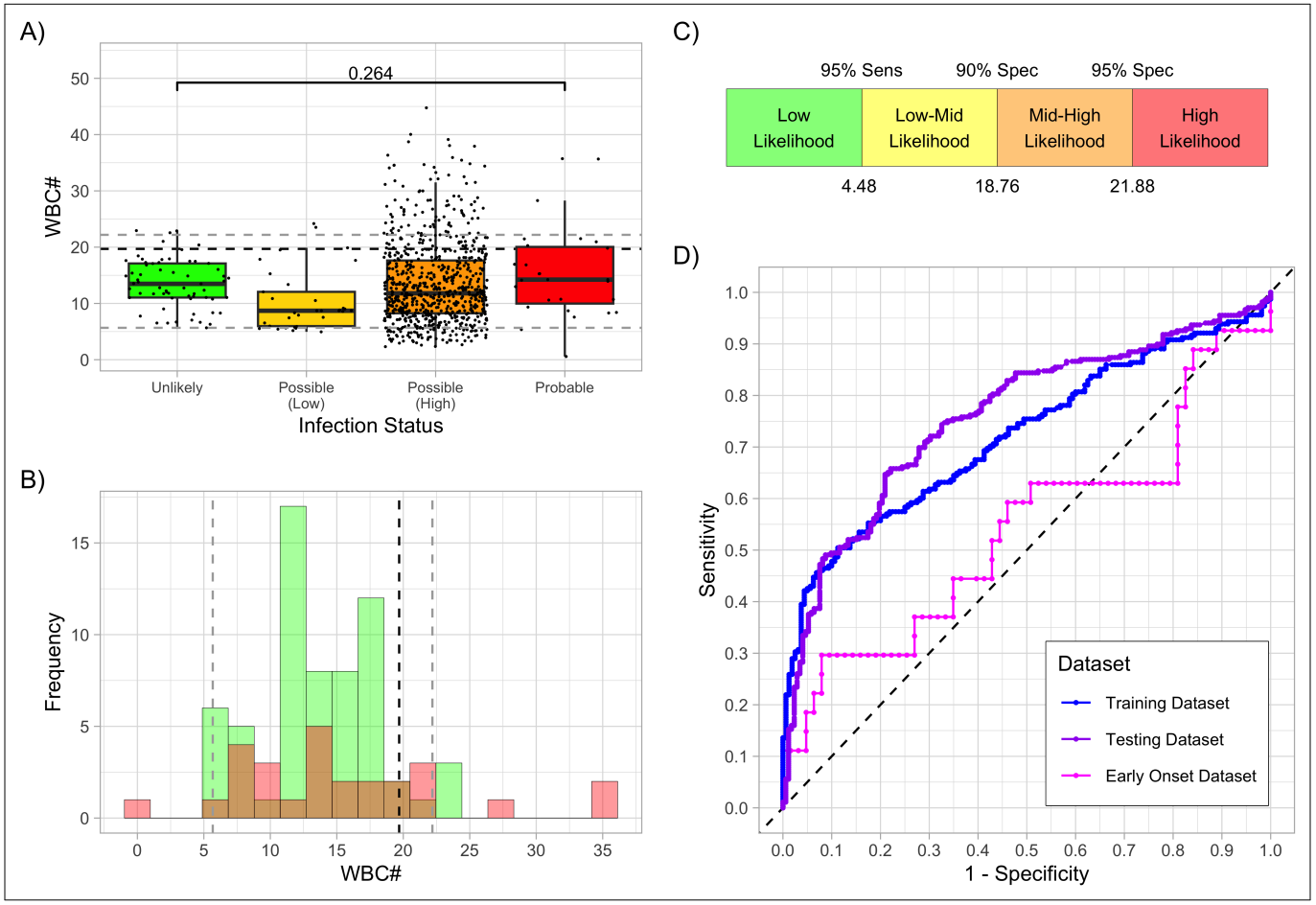


Figure 6.4: Evaluation of the diagnostic accuracy of white blood cell count (WBC#; x 10^9/L) for early-onset sepsis.

A) Box plot of values with clinical suspicion of sepsis. **B)** Histogram of *unlikely* (green) and *probable sepsis* (red) values with the dataset-specific optimal threshold (dashed black) and reference ranges (dashed grey). **C)** Clinically relevant thresholds calculated from ROC curve analysis. **D)** ROC curves for three datasets, including late-onset sepsis training and testing datasets.

Table 6.7: Diagnostic potential of complete blood count parameters in the first few days of life for early-onset sepsis.

						Dataset-Specific Threshold			Pre-Defined Threshold (Chpt 4)		
Parameter	*Unlikely* Average	*Probable* Average	P-Value	Effect	AUC (95% CI)	Threshold	Sens (%)	Spec (%)	Threshold	Sens (%)	Spec (%)
Red blood cell parameters											
RBC#	5.13	4.29	<0.001	1.11	0.767 (0.659 – 0.876)	4.46	55.6	88.9	3.23	88.9	0.0
HGB	189	158	<0.001	1.15	0.779 (0.667 – 0.890)	160	55.6	90.5	104	96.3	0.0
HCT	51.9	44.6	<0.001	0.95	0.752 (0.629 – 0.874)	44.7	55.6	88.9	28.2	96.3	0.0
MCV	101.0	104.5	0.03	–0.52	0.638 (0.507 – 0.769)	103.4	63.0	66.7	96.4	14.8	76.2
MCH	36.7	36.8	0.81	–0.05	0.532 (0.403 – 0.661)	37.6	51.9	66.7	36.8	40.7	42.9
MCHC	362	354	<0.001	0.90	0.734 (0.632 – 0.835)	365	96.3	47.6	368	96.3	38.1
RDW-SD	60.0	64.9	0.02	–0.60	0.634 (0.501 – 0.766)	66.7	46.2	82.5	52.7	96.2	15.9
RDW-CV	16.6	17.0	0.22	–0.30	0.448 (0.302 – 0.594)	16.5	50.0	65.1	16.9	46.2	55.6
NRBC#	0.16	2.45	<0.001	–1.63	0.887 (0.793 – 0.980)	1.18	81.5	95.2	0.17	88.9	46.0
NRBC%	1.5	15.0	<0.001	–1.55	0.873 (0.776 – 0.969)	6.5	81.5	88.9	0.6	88.9	27.0
RET#	0.255	0.261	0.79	–0.07	0.494 (0.351 – 0.638)	0.299	37.0	76.2	0.130	0.0	96.8
RET%	5.0	5.3	0.08	–0.45	0.620 (0.471 – 0.769)	6.9	40.7	95.2	4.3	29.6	71.4
IRF	38.9	36.6	0.27	0.27	0.578 (0.435 – 0.722)	32.4	37.0	87.3	16.7	–	–
RBC-HE	34.0	33.3	0.07	0.43	0.631 (0.509 – 0.752)	34.5	81.5	47.6	31.0	11.1	95.2

Continued on next page

Continued from last page

						Dataset-Specific Threshold			Pre-Defined Threshold (Chpt 4)		
Parameter	*Unlikely* Average	*Probable* Average	P-Value	Effect	AUC (95% CI)	Threshold	Sens (%)	Spec (%)	Threshold	Sens (%)	Spec (%)
RET-HE	33.9	29.2	<0.001	1.33	0.830 (0.727 – 0.933)	32.0	70.4	82.5	29.6	51.9	95.2
DELTA-HE	0.1	–3.4	<0.001	1.21	0.807 (0.696 – 0.919)	–2.1	74.1	81.0	–2.4	66.7	84.1
HYPO-HE	0.6	0.9	<0.001	–0.90	0.749 (0.630 – 0.869)	0.8	74.1	71.4	1.4	25.9	96.8
HYPER-HE	3.0	2.4	0.20	0.30	0.558 (0.434 – 0.682)	4.2	85.2	33.3	0.8	96.3	0.0
MICRO-RBC	1.8	1.8	0.62	–0.12	0.536 (0.395 – 0.676)	1.9	48.2	66.7	4.3	3.7	98.4
MACRO-RBC	11.4	14.5	0.06	–0.46	0.633 (0.500 – 0.767)	13.2	70.4	65.1	4.3	100.0	1.6
White blood cell parameters											
WBC#	13.56	15.48	0.26	–0.29	0.539 (0.395 – 0.683)	19.69	29.6	92.1	15.26	44.4	65.1
NEUT#	7.51	9.03	0.18	–0.33	0.578 (0.438 – 0.717)	11.35	37.0	88.9	5.46	70.4	30.2
NEUT%	55.2	59.6	0.19	–0.32	0.571 (0.431 – 0.712)	60.5	51.9	73.0	45.5	74.1	28.6
LYMPH#	4.12	3.56	0.03	0.55	0.670 (0.528 – 0.812)	3.36	55.6	85.7	3.84	55.6	57.1
LYMPH%	31.5	23.1	<0.001	0.91	0.735 (0.607 – 0.863)	22.2	55.6	92.1	30.5	74.1	47.6
MONO#	1.24	1.61	0.06	–0.49	0.648 (0.503 – 0.794)	1.83	51.9	85.7	2.70	33.3	96.8
MONO%	9.9	12.0	0.05	–0.52	0.677 (0.534 – 0.820)	13.6	51.9	88.9	21.0	18.5	98.4
EOSI#	0.23	0.13	0.01	0.61	0.688 (0.563 – 0.814)	0.14	59.3	74.6	0.20	66.7	60.3
EOSI%	1.9	0.9	0.001	0.74	0.695 (0.585 – 0.806)	2.3	85.2	49.2	2.1	81.5	50.8

Continued on next page

Continued from last page

						Dataset-Specific Threshold			Pre-Defined Threshold (Chpt 4)		
Parameter	*Unlikely* Average	*Probable* Average	P-Value	Effect	AUC (95% CI)	Threshold	Sens (%)	Spec (%)	Threshold	Sens (%)	Spec (%)
BASO#	0.11	0.11	0.96	0.01	0.503 (0.362 – 0.645)	0.13	37.0	79.4	0.06	81.5	12.7
BASO%	0.7	0.7	0.86	0.04	0.502 (0.370 – 0.634)	0.7	70.4	42.9	0.7	48.1	57.1
IG#	0.18	0.32	0.05	–0.47	0.604 (0.473 – 0.735)	0.46	44.4	76.2	0.19	51.9	44.4
IG%	1.5	2.8	0.02	–0.56	0.634 (0.511 – 0.757)	0.9	100.0	22.2	1.4	77.8	41.3
NLR	1.84	2.65	0.004	–0.74	0.691 (0.557 – 0.825)	2.36	63.0	76.2	1.46	77.8	31.7
ITR	0.03	0.06	0.04	–0.50	0.618 (0.495 – 0.741)	0.03	81.5	44.4	0.06	37.0	63.5
NEUT-RI	42.0	49.7	0.002	–0.92	0.769 (0.641 – 0.897)	46.8	63.0	92.1	46.1	63.0	88.9
NEUT-GI	149.5	145.9	0.02	0.59	0.639 (0.500 – 0.778)	142.9	40.7	93.7	144.2	40.7	90.5
NEUT-TW[a]	33.7	49.0	<0.001	0.52	0.830 (0.711 – 0.949)	36.0	88.9	71.4	37.3	81.5	76.2
RE-LYMPH	1.3	1.6	0.65	–0.11	0.539 (0.391 – 0.686)	2.0	55.6	65.1	14.2	7.4	100.0
AS-LYMPH[a]	0.0	0.2	0.11	0.17	0.596 (0.478 – 0.713)	0.1	55.6	65.1	1.0	3.7	100.0
HFLC[a]	0.0	0.0	0.25	0.12	0.573 (0.444 – 0.701)	0.0	18.5	100.0	0.0	3.7	100.0
RE-MONO[a]	0.2	0.7	<0.001	0.35	0.722 (0.599 – 0.845)	0.3	74.1	66.7	0.9	44.4	87.3
Platelet parameters											
PLT#	251	185	<0.001	1.02	0.763 (0.662 – 0.865)	229	74.1	71.4	257	88.9	47.6
PDW	10.9	12.2	0.02	–0.60	0.655 (0.529 – 0.782)	11.6	62.5	66.1	13.7	16.7	91.9

Continued on next page

Continued from last page

						Dataset-Specific Threshold			Pre-Defined Threshold (Chpt 4)		
Parameter	*Unlikely* Average	*Probable* Average	P-Value	Effect	AUC (95% CI)	Threshold	Sens (%)	Spec (%)	Threshold	Sens (%)	Spec (%)
MPV	10.4	10.8	0.06	–0.50	0.620 (0.481 – 0.758)	10.5	66.7	59.7	11.5	20.8	91.9
P-LCR	27.2	30.5	0.02	–0.60	0.654 (0.529 – 0.778)	27.9	70.8	56.5	35.2	16.7	91.9
PCT	0.3	0.2	0.001	0.82	0.720 (0.606 – 0.835)	0.3	87.5	45.2	0.3	95.8	17.7
IPF	5.4	8.4	<0.001	–0.98	0.763 (0.653 – 0.873)	7.8	63.0	77.8	7.5	63.0	76.2
PLR	0.055	0.057	0.79	–0.07	0.499 (0.356 – 0.642)	0.045	33.3	77.8	0.059	59.3	44.4

Effect refers to Cohen's d or Wilcoxon r effect size. All values were calculated parametrically unless transformation failed to normalise values ([a]). Average refers to the mean or median, depending on the underlying distribution. Significant p-values (< 0.05) and effect sizes (absolute Cohen's $d > 0.5$ and Wilcoxon $r > 0.3$) are highlighted in red. AUC values > 0.8 and sensitivities/specificities $> 85\%$ are highlighted in green, suggesting strong diagnostic performance. Abbreviations: AUC – area under the curve; CI – confidence interval; Sens – sensitivity; Spec – specificity. For parameter abbreviations, see List of Complete Blood Count Parameter Abbreviations (p. xxi).

CBC Results From Umbilical Cord Samples

Nine CBC parameters were significantly different between the baseline umbilical cord blood dataset and neonates admitted to the NICU (Table 6.8; Figure 6.5). Among these, RET% and IRF demonstrated the highest diagnostic accuracy, with AUCs greater than 0.8. RET% had an optimal threshold of 5.3%, yielding 88% sensitivity and 82% specificity. RET# and BASO% also showed good diagnostic potential for NICU admission, with sensitivities of 75% and specificities of 77% and 90%, respectively.

In the comparison of the baseline cohort to neonates with suspected sepsis, only six parameters were significantly different, including RET#, RET%, and IRF. No parameters achieved an AUC greater than 0.8. IRF showed the best performance, with an AUC of 0.781 and an optimal threshold of 41.4, achieving 58% sensitivity and 95% specificity. These findings suggest a potential role for RET% and IRF as predictors of neonatal compromise in cord blood, although validation in larger cohorts is needed.

6.3.4 Diagnostic Accuracy of Previously Developed Models

CBC Results From the First Three Days of Life

In an exploratory analysis of the performance of the diagnostic models developed for LOS in Chapter 5 in this cohort, results varied (Table 6.9). Among the models tested (logistic regression, computer-generated and manually constructed decision trees, random forest, scoring system, and kNN), the scoring system and kNN models achieved the highest AUCs (0.864 and 0.870, respectively) and a high specificity (97%) (Figures 6.6 and 6.7). No model achieved a sensitivity greater than 85%.

Twelve parameters were retained for inclusion in a logistic regression model for EOS, based on statistical significance. After stepwise refinement, four parameters were retained: NRBC#, NLR, NEUT-RI, and PLT#. The final model achieved an AUC of 0.932, with 89% sensitivity and 92% specificity (Figure 6.8). Although this model is underpowered and not suitable for clinical use, it may help guide parameter selection in future studies.

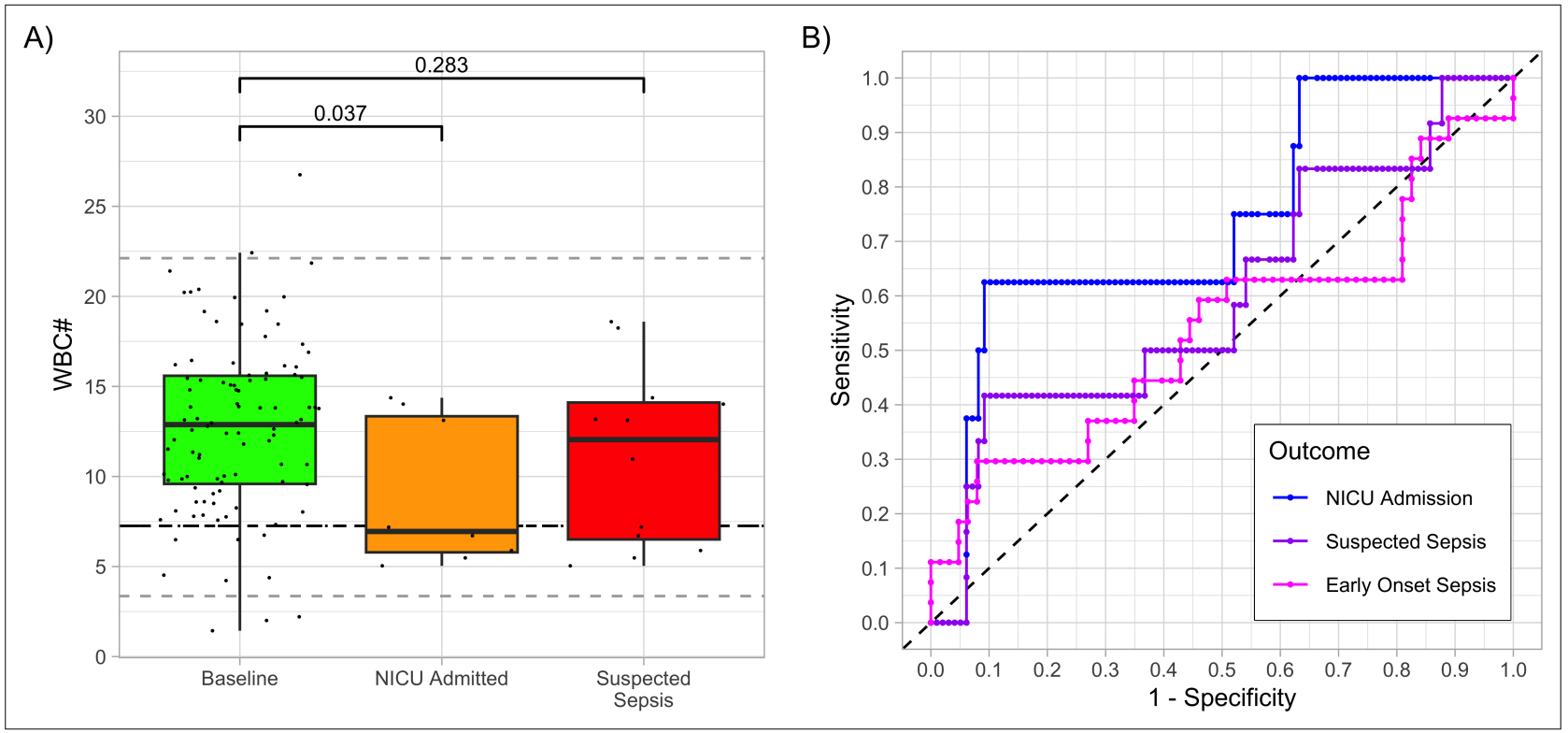


Figure 6.5: Evaluation of the diagnostic accuracy of white blood cell count (WBC#; x 10^9/L) for admission to the Neonatal Intensive Care Unit and for suspected early-onset sepsis.

A) Box plot of values with admission to the Neonatal Intensive Care Unit (NICU) and suspected early-onset sepsis. **B)** ROC curves, including ROC curve for the diagnostic accuracy of WBC# in the first three days of life for culture-confirmed early-onset sepsis.

Table 6.8: Diagnostic potential of umbilical cord complete blood count parameters for Neonatal Intensive Care Unit admission and suspected early-onset sepsis.

	NICU Admission						Suspected EOS					
Parameter	P-Value	Effect	AUC (95% CI)	Threshold	Sens (%)	Spec (%)	P-Value	Effect	AUC (95% CI)	Threshold	Sens (%)	Spec (%)
Red blood cell parameters												
RBC#	0.17	0.57	0.631 (0.390 – 0.871)	3.57	50.0	88.8	0.29	0.34	0.444 (0.259 – 0.630)	4.37	58.3	56.1
HGB	0.29	0.39	0.594 (0.393 – 0.796)	169	100.0	24.5	0.47	0.21	0.557 (0.379 – 0.736)	141	41.7	75.5
HCT	0.53	0.23	0.467 (0.260 – 0.674)	46.1	62.5	59.2	0.80	0.08	0.497 (0.324 – 0.670)	46.1	58.3	59.2
MCV	0.22	– 0.54	0.615 (0.406 – 0.825)	120.5	25.0	100.0	0.35	–0.34	0.580 (0.388 – 0.771)	117.2	25.0	100.0
MCH	0.22	– 0.54	0.616 (0.406 – 0.826)	40.2	25.0	100.0	0.41	–0.29	0.574 (0.379 – 0.769)	36.7	50.0	68.4
MCHC	0.23	0.53	0.634 (0.397 – 0.871)	327	37.5	95.9	0.09	0.57	0.653 (0.475 – 0.831)	341	66.7	65.3
RDW-SD	0.24	– 0.51	0.622 (0.400 – 0.845)	64.2	62.5	64.3	0.39	–0.30	0.577 (0.382 – 0.771)	72.4	25.0	93.9
RDW-CV	0.66	–0.16	0.550 (0.327 – 0.774)	17.2	50.0	79.6	0.76	–0.09	0.474 (0.302 – 0.647)	18.3	100.0	12.2
NRBC#	0.39	–0.40	0.655 (0.386 – 0.924)	0.98	62.5	77.6	0.20	–0.48	0.663 (0.459 – 0.866)	0.88	66.7	71.4
NRBC%	0.20	– 0.60	0.675 (0.398 – 0.951)	11.9	50.0	93.9	0.14	– 0.55	0.651 (0.441 – 0.860)	11.9	41.7	93.9
RET#	0.02	– 1.20	0.793 (0.602 – 0.984)	0.223	75.0	76.5	0.04	– 0.77	0.685 (0.494 – 0.876)	0.278	41.7	98.0
RET%	0.01	– 1.31	0.845 (0.635 – 1.000)	5.3	87.5	81.6	0.04	– 0.79	0.713 (0.499 – 0.928)	5.7	58.3	91.8
IRF	0.004	– 1.44	0.849 (0.720 – 0.979)	41.4	62.5	94.9	0.005	– 1.11	0.781 (0.626 – 0.937)	41.4	58.3	94.9
RBC-HE	0.36	–0.41	0.582 (0.319 – 0.846)	36.8	50.0	77.6	0.61	–0.18	0.523 (0.308 – 0.738)	37.4	33.3	86.7
RET-HE	0.22	– 0.51	0.629 (0.408 – 0.850)	35.5	62.5	65.3	0.29	–0.36	0.604 (0.415 – 0.793)	36.9	41.7	82.7
DELTA-HE	0.42	–0.32	0.605 (0.397 – 0.814)	–1.2	87.5	44.9	0.26	–0.36	0.623 (0.447 – 0.799)	0.1	58.3	70.4
HYPO-HE	0.02	– 0.82	0.721 (0.574 – 0.868)	0.6	100.0	42.9	0.04	– 0.64	0.681 (0.531 – 0.831)	0.6	91.7	42.9

Continued on next page

Continued from last page

	NICU Admission						Suspected EOS					
Parameter	P-Value	Effect	AUC (95% CI)	Threshold	Sens (%)	Spec (%)	P-Value	Effect	AUC (95% CI)	Threshold	Sens (%)	Spec (%)
HYPER-HE	0.42	–0.37	0.579 (0.306 – 0.852)	10.3	50.0	82.7	0.66	–0.16	0.524 (0.302 – 0.746)	10.3	41.7	82.7
MICRO-RBC	0.36	–0.35	0.643 (0.444 – 0.843)	1.3	87.5	57.1	0.34	–0.31	0.612 (0.438 – 0.787)	1.3	75.0	57.1
MACRO-RBC	0.22	– 0.55	0.622 (0.409 – 0.836)	13.2	87.5	38.8	0.31	–0.36	0.584 (0.392 – 0.777)	36.5	25.0	100.0
White blood cell parameters												
WBC#	0.04	0.85	0.733 (0.538 – 0.929)	7.26	62.5	90.8	0.28	0.35	0.602 (0.414 – 0.790)	7.26	41.7	90.8
NEUT#	0.02	1.02	0.768 (0.597 – 0.940)	3.74	62.5	80.6	0.11	0.52	0.634 (0.464 – 0.804)	7.37	91.7	37.8
NEUT%	0.02	1.11	0.776 (0.612 – 0.939)	45.4	87.5	66.3	0.07	0.67	0.673 (0.495 – 0.852)	45.4	66.7	66.3
LYMPH#	0.31	0.37	0.610 (0.400 – 0.820)	3.16	50.0	78.4	0.85	0.06	0.548 (0.360 – 0.735)	3.16	41.7	78.4
LYMPH%	0.46	–0.29	0.606 (0.386 – 0.825)	42.6	37.5	87.6	0.76	–0.10	0.547 (0.361 – 0.732)	28.5	83.3	36.1
MONO#	0.09	0.75	0.693 (0.484 – 0.901)	0.42	37.5	95.9	0.37	0.30	0.566 (0.374 – 0.757)	0.42	25.0	95.9
MONO%	0.34	0.46	0.638 (0.390 – 0.886)	9.7	75.0	54.1	0.37	0.34	0.597 (0.418 – 0.775)	9.7	75.0	54.1
EOSI#	0.64	0.20	0.562 (0.309 – 0.816)	0.17	37.5	87.8	0.75	0.10	0.536 (0.345 – 0.726)	0.30	50.0	64.3
EOSI%	0.93	0.04	0.566 (0.299 – 0.834)	1.8	50.0	81.6	0.79	0.09	0.571 (0.370 – 0.772)	1.8	41.7	81.6
BASO#	0.03	– 1.08	0.771 (0.576 – 0.966)	0.47	50.0	99.0	0.003	– 1.11	0.783 (0.647 – 0.918)	0.16	83.3	64.3
BASO%	0.03	– 1.11	0.793 (0.555 – 1.000)	2.1	75.0	89.8	0.005	– 1.11	0.793 (0.632 – 0.953)	1.3	83.3	66.3
IG#	0.27	0.48	0.659 (0.402 – 0.916)	0.16	62.5	80.6	0.42	0.28	0.588 (0.372 – 0.803)	0.16	50.0	80.6
IG%	0.43	0.36	0.631 (0.379 – 0.882)	2.5	75.0	62.2	0.31	0.36	0.608 (0.412 – 0.804)	2.5	66.7	62.2
NLR	0.11	0.74	0.713 (0.477 – 0.948)	1.08	75.0	77.3	0.22	0.44	0.630 (0.427 – 0.833)	1.08	58.3	77.3
ITR	0.68	–0.17	0.482 (0.257 – 0.707)	0.06	62.5	56.1	0.86	–0.06	0.489 (0.290 – 0.688)	0.06	58.3	56.1

Continued on next page

Continued from last page

	NICU Admission						Suspected EOS					
Parameter	P-Value	Effect	AUC (95% CI)	Threshold	Sens (%)	Spec (%)	P-Value	Effect	AUC (95% CI)	Threshold	Sens (%)	Spec (%)
NEUT-RI	0.82	0.10	0.548 (0.243 – 0.854)	37.9	50.0	85.7	0.82	–0.08	0.527 (0.300 – 0.754)	44.4	25.0	95.9
NEUT-GI	0.59	0.19	0.570 (0.359 – 0.782)	144.8	62.5	67.4	0.85	0.06	0.516 (0.335 – 0.696)	144.8	50.0	67.4
NEUT-TW	0.19	–0.50	0.639 (0.442 – 0.837)	34.4	75.0	56.1	0.08	– 0.53	0.649 (0.495 – 0.802)	34.4	75.0	56.1
RE-LYMPH	0.20	– 0.62	0.675 (0.404 – 0.947)	5.5	50.0	95.9	0.25	–0.43	0.614 (0.386 – 0.842)	5.5	41.7	95.9
AS-LYMPH[a]	0.80	0.03	0.495 (0.485 – 0.505)	0.0	99.0	5.0	0.75	0.03	0.495 (0.485 – 0.505)	0.0	99.0	5.0
HFLC[a]	0.55	0.06	0.555 (0.327 – 0.784)	0.0	37.5	92.8	0.44	0.07	0.559 (0.380 – 0.739)	0.0	33.3	92.8
RE-MONO	0.36	–0.43	0.608 (0.323 – 0.893)	0.7	50.0	88.8	0.25	–0.42	0.608 (0.402 – 0.813)	0.7	41.7	87.8
Platelet parameters												
PLT#	0.56	0.27	0.552 (0.250 – 0.854)	132	50.0	87.8	0.91	–0.04	0.562 (0.336 – 0.789)	303	50.0	76.5
PDW	0.70	–0.20	0.593 (0.251 – 0.936)	11.3	66.7	79.1	0.84	–0.08	0.534 (0.306 – 0.761)	10.8	60.0	65.9
MPV	0.88	–0.08	0.540 (0.219 – 0.862)	10.9	33.3	89.0	0.78	0.11	0.475 (0.245 – 0.705)	10.0	60.0	51.7
P-LCR	0.80	–0.13	0.576 (0.244 – 0.908)	26.3	66.7	72.5	0.87	0.06	0.496 (0.261 – 0.731)	25.5	50.0	70.3
PCT	0.94	–0.04	0.548 (0.231 – 0.864)	0.3	66.7	68.1	0.46	–0.27	0.619 (0.412 – 0.826)	0.3	70.0	68.1
IPF	0.65	–0.20	0.594 (0.309 – 0.879)	7.8	50.0	86.7	0.90	0.05	0.512 (0.297 – 0.727)	2.3	25.0	90.8
PLR	0.61	–0.23	0.463 (0.183 – 0.743)	0.012	37.5	77.6	0.79	–0.10	0.502 (0.289 – 0.714)	0.015	50.0	64.3

Effect refers to Cohen's d or Wilcoxon r effect size. All values were calculated parametrically unless transformation failed to normalise values ([a]). Average refers to the mean or median, depending on the underlying distribution. Significant p-values (< 0.05) and effect sizes (absolute Cohen's $d > 0.5$ and Wilcoxon $r > 0.3$) are highlighted in red. AUC values > 0.8 and sensitivities/specificities $> 85\%$ are highlighted in green, suggesting strong diagnostic performance. Abbreviations: AUC – area under the curve; CI – confidence interval; Sens – sensitivity; Spec – specificity. For parameter abbreviations, see List of Complete Blood Count Parameter Abbreviations (p. xxi).

Table 6.9: Diagnostic potential of diagnostic models developed for late-onset sepsis for early-onset sepsis

						Dataset-Specific Threshold			Pre-Defined Threshold (Chapter 5)		
Model	*Unlikely* Average	*Probable* Average	P-Value	Effect	AUC (95% CI)	Threshold	Sens (%)	Spec (%)	Threshold	Sens (%)	Spec (%)
Logistic Regression	–0.80	2.60	<0.001	–1.25	0.845 (0.740 – 0.950)	0.55	81.5	84.1	0.47	81.5	79.4
Computer-Generated Decision Tree[a]	0.69	0.93	<0.001	0.36	0.723 (0.600 – 0.846)	0.93	74.1	77.8	0.60	81.5	34.9
Manually Constructed Decision Tree[a]	0.47	0.62	<0.001	0.42	0.764 (0.657 – 0.871)	0.59	59.3	85.7	0.59	59.3	85.7
Random Forest	0.41	0.73	<0.001	–1.27	0.824 (0.724 – 0.924)	0.64	74.1	84.1	0.50	85.2	68.3
Scoring System	4.63	11.36	<0.001	–1.55	0.864 (0.772 – 0.955)	10.54	70.4	96.8	7.84	74.1	73.0
k-Nearest Neighbour	0.46	0.74	<0.001	–1.62	0.870 (0.790 – 0.951)	0.73	59.3	96.8	0.53	85.2	61.9
EOS-Specific Logistic Regression	–2.97	6.02	<0.001	–1.77	0.932 (0.859 – 1.000)	–1.02	88.9	92.1			

Effect refers to Cohen's d or Wilcoxon r effect size. All values were calculated parametrically unless transformation failed to normalise values ([a]). Average refers to the mean or median, depending on the underlying distribution. Significant p-values (< 0.05) and effect sizes (absolute Cohen's $d > 0.5$ and Wilcoxon $r > 0.3$) are highlighted in red. AUC values > 0.8 and sensitivities/specificities $> 85\%$ are highlighted in green, suggesting strong diagnostic performance. Abbreviations: AUC – area under the curve; CI – confidence interval; EOS – early-onset sepsis; Sens – sensitivity; Spec – specificity. For parameter abbreviations, see List of Complete Blood Count Parameter Abbreviations (p. xxi).

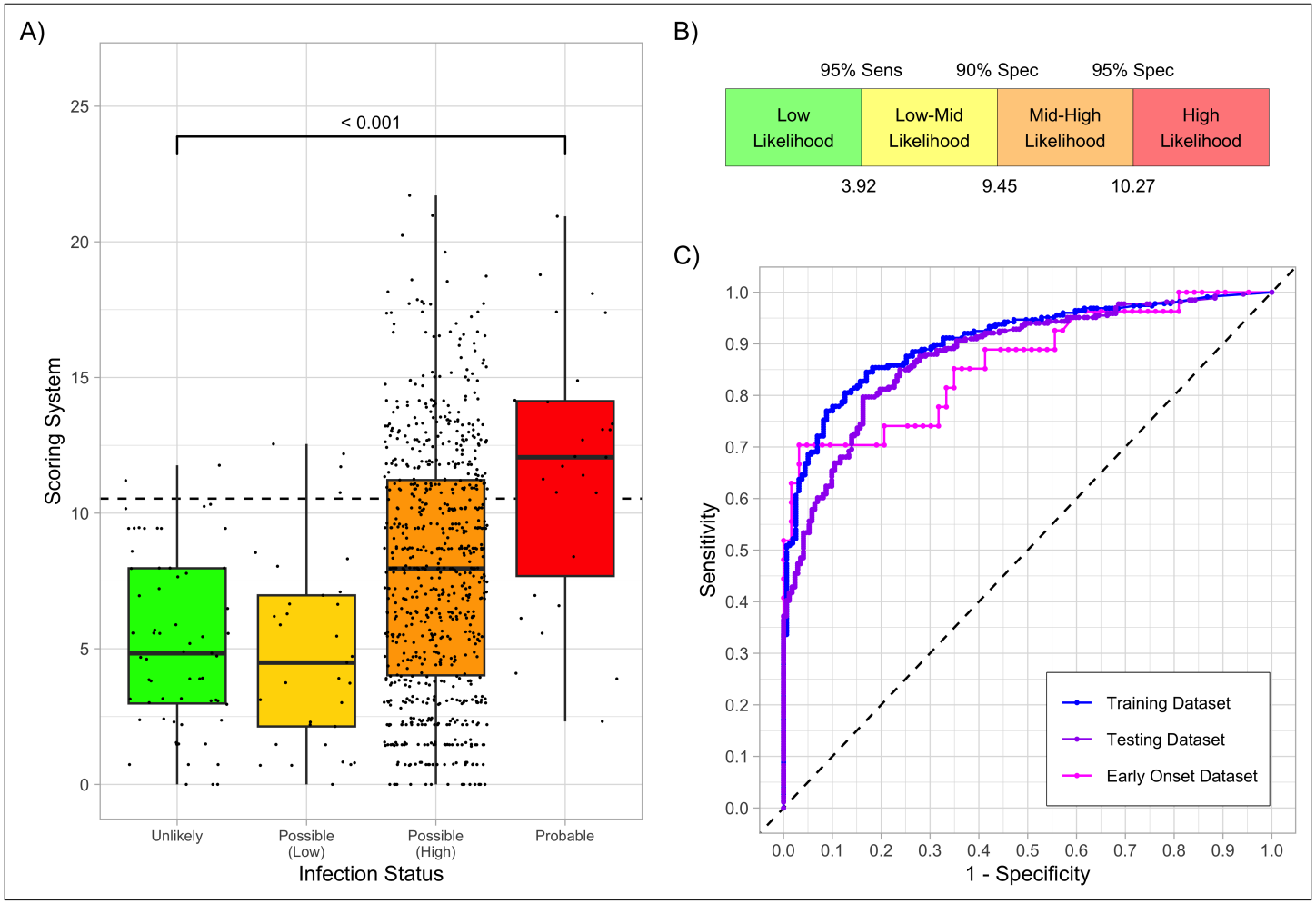


Figure 6.6: Diagnostic performance of a scoring system for late-onset sepsis in a cohort of neonates with early-onset sepsis.

A) Box plot of values with level of suspicion of sepsis. **B)** Optimal thresholds, according to ROC curve analysis of the late-onset sepsis training dataset. **C)** ROC curves for three datasets, including late-onset sepsis training and testing datasets.

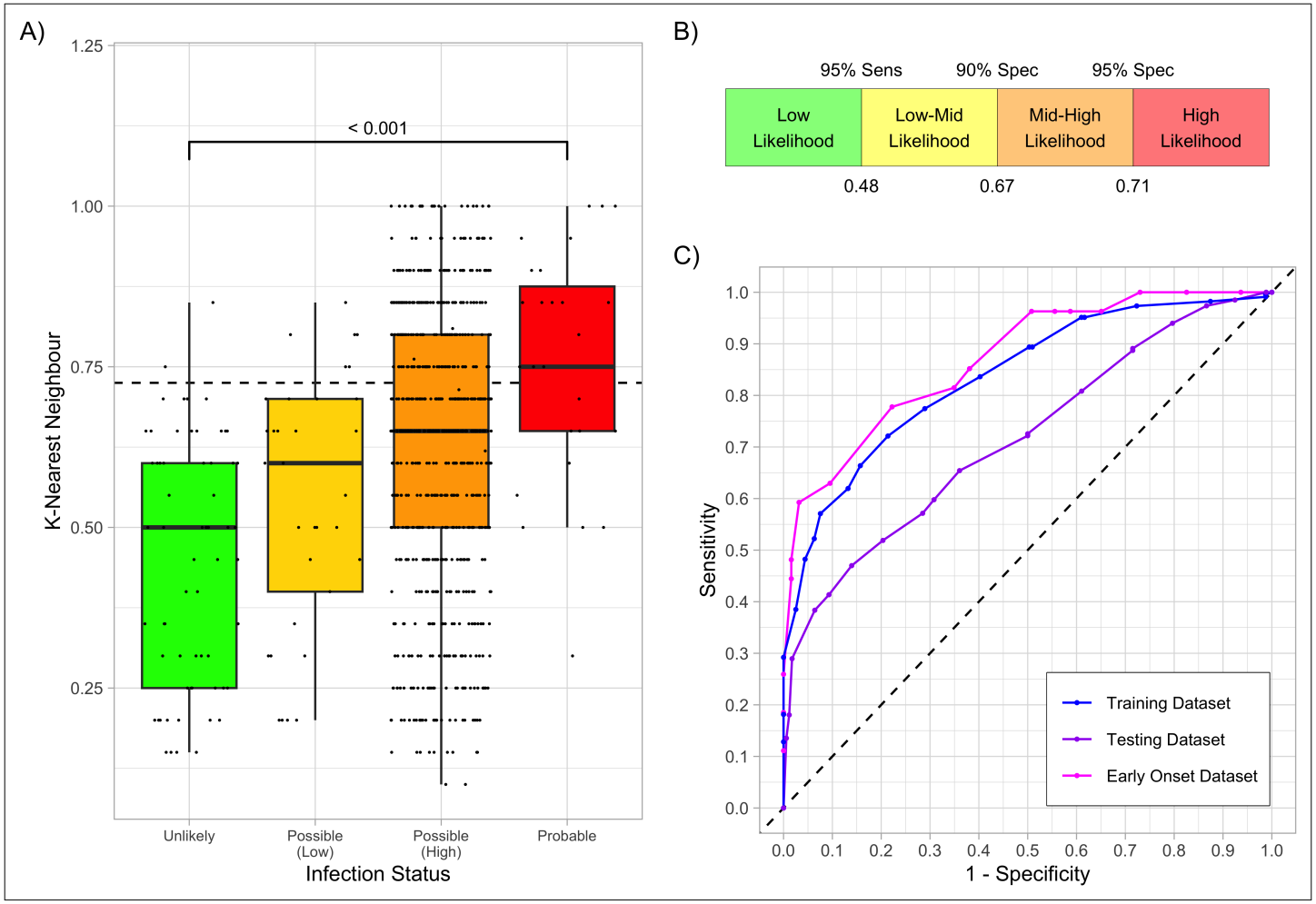


Figure 6.7: Diagnostic performance of a k-nearest neighbour model for late-onset sepsis in a cohort of neonates with early-onset sepsis. **A)** Box plot of values with level of suspicion of sepsis. **B)** Optimal thresholds, according to ROC curve analysis of the late-onset sepsis training dataset. **C)** ROC curves for three datasets, including late-onset sepsis training and testing datasets.

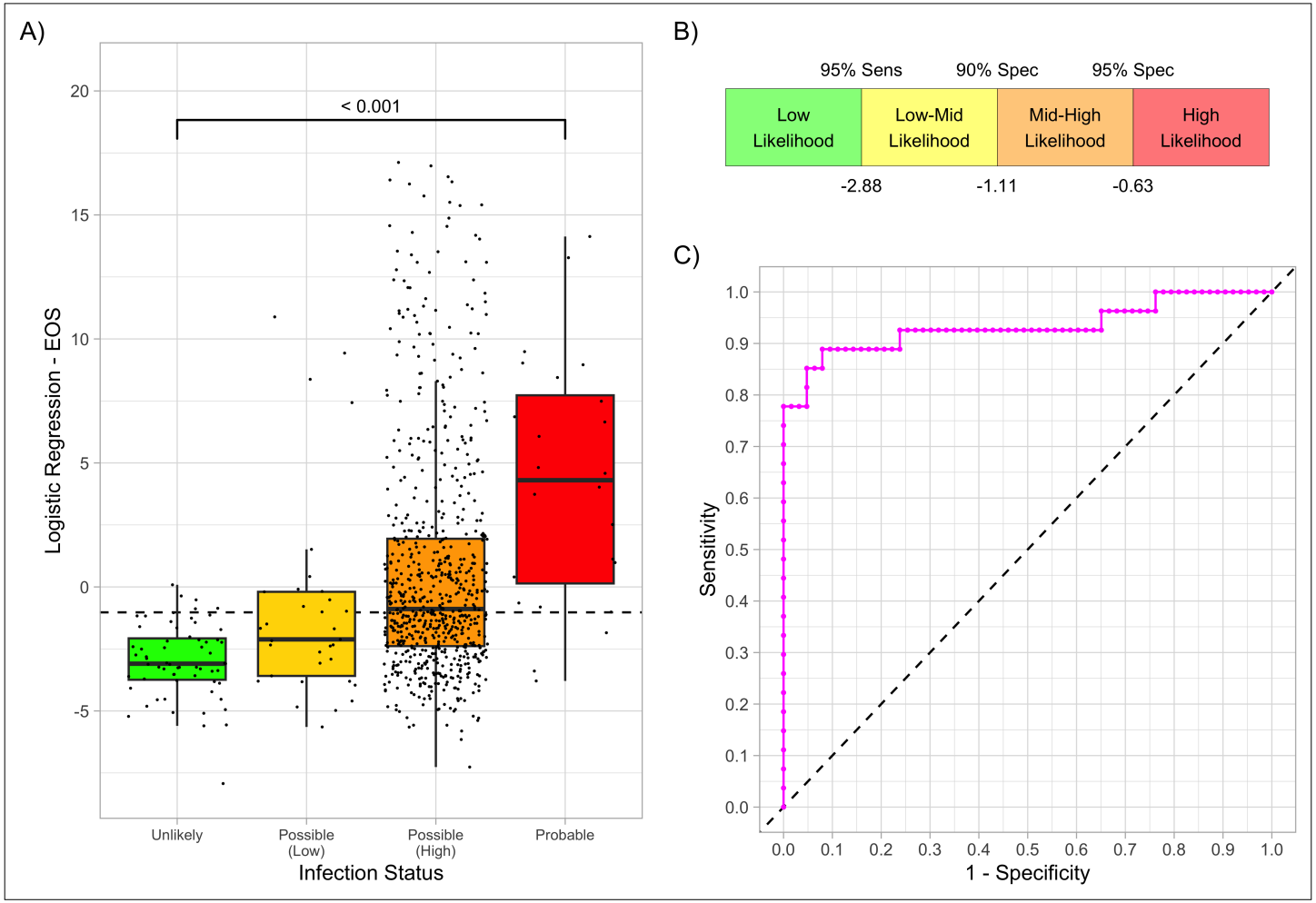


Figure 6.8: Diagnostic performance of a logistic regression model for early-onset sepsis. **A)** Box plot of values with level of suspicion of sepsis. **B)** Optimal thresholds, according to ROC curve analysis. **C)** ROC curve.

6.4 Discussion

6.4.1 Summary of Findings

This chapter applied analytical methods developed in previous chapters to two clinically relevant datasets: CBC results from the first three days of life and from umbilical cord blood samples. The chapter aimed to (1) characterise baseline haematological profiles in each dataset, (2) examine the influence of clinical factors on the baseline profile, and (3) evaluate the diagnostic accuracy of individual parameters and previously developed models for EOS.

In both datasets, sex, GA, and MOD significantly influenced multiple CBC parameters. In the newborn dataset, postnatal age also had a marked effect. When comparing both datasets, EOSI% was the only parameter to show statistically significant sex differences with a clinically meaningful effect size (pes $>$ 0.10). Five parameters, including NEUT#, IG#, and NLR, showed clinically relevant differences by GA, suggestive of a non-specific bone marrow push in term neonates. No single parameter was significantly affected by MOD in either dataset, underscoring potential context-specific variation.

Several parameters showed diagnostic potential in detecting EOS, although most were dataset-specific. This may be due to the differences in diagnostic criteria between the two datasets: a positive blood culture in the newborn dataset and a clinical suspicion of sepsis in the umbilical cord dataset. In the former, NRBC# demonstrated the highest diagnostic accuracy (AUC: 0.887), while IRF showed the strongest performance in the umbilical cord dataset (AUC: 0.781). Notably, IRF was the only parameter that showed consistent diagnostic potential across both cohorts.

Exploratory analyses of previously developed diagnostic models (as detailed in Chapter 5) demonstrated moderate to high performance for EOS in the newborn dataset. The scoring system and kNN model achieved the highest diagnostic accuracies (AUC: 0.864 and 0.870, respectively), both with high specificity (97%). A new logistic regression model developed within this chapter retained four parameters after stepwise refinement

(NRBC#, NLR, NEUT-RI, and PLT#) and had the highest overall diagnostic performance (AUC: 0.932). Together, these findings suggest that there is a potential for developing CBC-based diagnostic models for EOS, and may inform future model development efforts.

6.4.2 Comparison with Existing Literature

CBC Results From the First Three Days of Life

Age-related changes in red cell parameters have been widely documented in neonates, including reductions in HGB, NRBC#, RET#, IRF, and RET-HE during the early postnatal period [129, 130]. The current study replicates these trends, particularly in RET-HE and NRBC#, with IRF also showing a non-significant decline over time. These findings reinforce the notion that the early neonatal period is haematologically distinct from the later neonatal period, and age-specific reference ranges are essential for use in clinical care.

GA-related trends were partly consistent with previous research. For example, increases in EOSI# and MONO# with increasing GA have been previously reported [177], which is consistent with the non-significant trends observed in this study. Furthermore, prior studies have identified differences with MOD in multiple parameters, including HGB, HCT, RBC#, WBC#, and PLT# [178]. In alignment with this, WBC# values in this research were significantly higher in neonates born via SVD compared to CS.

In terms of diagnostic accuracy, previous studies identified ITR and PLR as promising markers for EOS [101]. However, the current findings found limited support for their utility, with AUCs of 0.618 and 0.499, respectively. In contrast, NRBC# and RET-HE demonstrated the highest diagnostic performance in this dataset (AUC: 0.887 and 0.830, respectively). As markers of erythropoietic activity, their elevation in EOS cases may reflect suppressed erythropoiesis, a phenomenon previously documented in the context of adult sepsis [179].

CBC Results From Umbilical Cord Samples

Comparison with existing literature on umbilical cord CBC results is challenging due to the limited sample size and scarcity of comparable studies, particularly with diagnostic outcomes for EOS. Nevertheless, the current findings partially aligned with existing reference values. For example, the median RBC# in the current study (4.32 x 10^{12}/L) closely aligned with previously published reference values (4.14, 4.30, and 4.16 x 10^{12}/L) [168, 171, 173]. Similarly, PLT# values (263 x 10^{9}/L) also aligned with earlier studies (257, 227, and 222 x 10^{9}/L) [168, 171, 173].

Sex-based differences have previously been reported only for RBC#, with males typically presenting higher counts [168]. However, this study found no sex-based differences in RBC# ($p = 0.60$). As in prior studies [168], HCT was higher in neonates born by SVD than by CS. GA affected PLT# with increasing maturity linked to higher values, aligning with prior reports [171].

Evidence on the diagnostic accuracy of umbilical cord CBC parameters remains sparse. One study found that ITR was significantly higher in EOS cases compared to controls (0.29 vs 0.19, $p = 0.001$), reporting a high diagnostic accuracy (AUC: 0.859) [176]. In contrast, none of the parameters evaluated in the current study achieved an AUC above 0.8. IRF demonstrated the best performance (AUC: 0.781), but this fell short of thresholds typically considered clinically actionable.

6.4.3 Implications of Findings

The findings from this chapter highlight the diagnostic potential of the CBC as a rapid, cost-effective, and minimally invasive tool for EOS detection in neonates. While blood cultures remain the diagnostic gold standard, their low sensitivity and slow turnaround time reduce their clinical utility. By contrast, the CBC can be rapidly processed and requires only small blood volumes, making it a promising tool for early risk stratification.

A novel contribution of this study is the evaluation of routinely discarded umbilical cord blood samples as a potential source for diagnosis. If validated in larger cohorts, this

approach could enable EOS risk stratification at birth without the need for additional invasive sampling. Such implementation could help reduce unnecessary antibiotic exposure and support earlier clinical decision-making.

These findings also highlight the importance of using reference ranges in neonatal haematology that account for the significant variability seen across sex, GA, postnatal age, and MOD. This suggests that fixed diagnostic cut-offs, particularly those adapted from older paediatric or adult populations, may be inappropriate in neonates. Future research should integrate these factors to improve the accuracy of baseline analyses.

6.4.4 Strengths and Limitations

The major strength of this chapter lies in its use of real-world clinical data. The neonatal cohort represents a subset of a broader hospital-based dataset, while the umbilical cord blood samples reflect the general population of neonates born at the John Radcliffe Hospital (Oxford, UK). This enhances the clinical relevance of findings.

The study also incorporated both conventional and extended CBC parameters, including immune activation indices such as NEUT-RI, RE-LYMPH, and IRF, offering a more nuanced understanding of neonatal haematological responses. Reference ranges were calculated following CLSI guidelines, with 95% CIs presented throughout. The influences of key clinical factors (sex, GA, and MOD) were carefully evaluated, providing insight into potential confounders and sources of biological variation.

Despite these strengths, several limitations must be acknowledged. The main limitation is the small sample size, which did not meet the CLSI-recommended minimum of 120 observations for establishing a reference range. As a result, all analyses were underpowered to detect subtle effects. To minimise this, emphasis was placed on effect sizes and CIs to support interpretation. Furthermore, diagnostic models were only evaluated in the neonatal dataset, as differences in EOS definitions and sample size in the cord blood dataset limited applicability.

As in the previous chapters, clinical information for both datasets was collected retrospectively; therefore, it is possible that misclassifications occurred due to missing clinical data. The criteria applied to the umbilical cord blood dataset were also less robust than those applied to the neonatal dataset, as culture results were not readily available in this cohort.

In addition, detailed information on maternal factors, such as antenatal corticosteroid exposure, was not available to allow meaningful assessment of their impact on neonatal CBC parameters in the first three days of life. The timing, dosage, and indication for maternal corticosteroid administration relative to delivery and blood sampling are likely to be critical determinants of any effects, and retrospective data were insufficient to capture this complexity reliably. Prospective evaluation of maternal exposures would therefore be an important component of future research.

Together, these limitations highlight the need for larger, prospective studies to validate these findings and further explore the diagnostic potential of the CBC for EOS diagnosis.

6.5 Chapter Summary

This chapter applied statistical analysis techniques developed in earlier chapters to two clinically relevant datasets – umbilical cord blood CBC results and neonatal CBC results from the first three days of life. The aim was to characterise the baseline haematological profile of each cohort and evaluate their potential for EOS diagnosis. Key clinical factors, including MOD, GA, and postnatal age, significantly influenced multiple CBC parameters, underscoring the importance of population-specific reference ranges for care in the immediate postnatal period.

In terms of diagnostic performance, NRBC# demonstrated moderate diagnostic accuracy for detecting culture-positive EOS in neonatal samples (AUC > 0.8). Previously developed diagnostic models (Chapter 5) also showed moderate performance in this cohort. In the umbilical cord dataset, IRF exhibited moderate predictive value for NICU admission; however, no single parameter reliably identified clinical EOS in this group.

The next chapter will synthesise the overall findings of this thesis within a broader clinical context and propose directions for future research.

Chapter 7

Discussion

7.1 Summary of Chapters

This thesis aimed to investigate the diagnostic utility of the CBC for neonatal sepsis, with a particular focus on LOS, defined as a systemic bloodstream infection occurring after three days of life. The primary hypothesis was that specific parameters of the routine neonatal CBC, either individually or in combination, could distinguish between neonates with and without sepsis, and that multivariable diagnostic models incorporating these parameters would achieve sufficient accuracy to support clinical decision-making.

Chapter 2 presented a meta-analysis of existing literature on the diagnostic performance of CBC parameters for neonatal sepsis. Seven parameters were evaluated, with ITR and MNV demonstrating the highest diagnostic accuracy. However, MNV was only assessed in three studies, which showed substantial heterogeneity, limiting the robustness of this finding. The chapter also highlighted considerable variability in study design, particularly in how sepsis and control groups were defined, which resulted in overlapping patient characteristics across studies. Despite these limitations, the review established a strong rationale for further investigation of the CBC as a diagnostic tool in neonatal care.

Chapter 3 addressed the methodological issues identified in the review by establishing a cohort of neonates admitted to the NICU but otherwise clinically well, providing a more appropriate reference population for neonatal sepsis research. In this cohort, I determined reference ranges for each CBC parameter and assessed the individual and interactive effects of sex, postnatal age, and GA. I also evaluated the impact of corticosteroid exposure, given its frequent use in labour and NICU settings, and its known effects on adult CBC values. The findings of this chapter underscore the need for subgroup-specific

reference ranges and, to my knowledge, marks the first time that novel immune activation parameters and corticosteroid effects have been characterised in the neonatal CBC.

Chapter 4 assessed the diagnostic accuracy of individual CBC parameters for LOS by comparing neonates with culture-confirmed sepsis to the control cohort defined in Chapter 3. Multiple parameters, including NEUT-RI and NLR, showed strong diagnostic potential. These findings expanded on the systematic review (Chapter 2) by incorporating underreported but routinely generated immune cell activation parameters, such as NEUT-RI, NEUT-GI, RE-LYMPH, and RE-MONO. Results were validated in an independent testing subset of the same cohort, improving the robustness of the conclusions.

Chapter 5 built on these findings by constructing six multivariable diagnostic models using the most promising parameters identified in Chapter 4. These models were developed on a training cohort and validated using both an internal testing dataset and an external validation cohort. The scoring system, logistic regression model, and random forest model achieved high diagnostic accuracy, with AUCs exceeding 0.8 across all datasets. To demonstrate clinical utility, I applied these models longitudinally in three case studies (one from each dataset), showing that each model detected infection earlier than CRP or blood cultures and maintained elevated scores throughout the infection period.

Chapter 6 characterised CBC results collected within the first three days of life and from umbilical cord blood samples. In each dataset, I examined the baseline CBC profile, determined the individual effects of sex, GA, and MOD, and assessed the diagnostic accuracy of individual parameters for EOS. I also explored the feasibility of developing multivariable models for EOS, although small sample sizes limited this effort. Nevertheless, these analyses provide a valuable foundation for future work on early sepsis diagnostics in neonates.

Taken together, the work presented in this thesis provides strong evidence in support of the primary hypothesis. Across multiple complementary approaches – from the systematic

review, to establishment of reference ranges, to evaluation of individual diagnostic parameters, and ultimately to the development and validation of multivariable diagnostic models – I have demonstrated that the routine neonatal CBC provides clinically meaningful information for the diagnosis of neonatal sepsis. Several individual parameters, including both established and novel indices, showed diagnostic potential, while multivariable models consistently achieved high diagnostic accuracy and outperformed CRP in longitudinal case studies. Although further validation in larger, multicentre cohorts will be required before clinical implementation, these findings confirm that the CBC, a widely available and inexpensive test requiring minimal blood volumes, can form the basis of reliable diagnostic tools for neonatal sepsis.

7.2 Clinical Implications of Findings

Current NICE guidelines recommend a positive blood culture as the diagnostic gold standard and suggest serial CRP measurements to monitor disease progression and recovery [11]. However, cultures require up to 48 hours for results and have limited sensitivity, particularly in cases of intermittent or low-level bacteraemia [2, 19]. While CRP is valuable for monitoring, it does not reliably rise until 10 – 12 hours after the onset of infection, making it unsuitable for early detection [2, 19].

My findings suggest that the CBC, which is already performed routinely in neonatal care, could fill this diagnostic gap. By utilising both traditional indices and newer immune activation parameters, CBC-based models can provide rapid and reliable risk assessment at the bedside. These models performed consistently across datasets, supporting their potential to be generalisable beyond the current study population.

In practice, this could mean earlier identification of sepsis, fewer missed infections, and a reduction in unnecessary antibiotic exposure. More broadly, the results highlight a need to reconsider current diagnostic pathways and to test CBC-derived models in multicentre clinical trials. If validated further, they could provide a low-cost, widely available tool to support faster, evidence-based decisions in neonatal sepsis care.

7.3 Marketability of CBC-Based Diagnostic Tools

The research presented in this thesis was partially funded by Sysmex, a manufacturer and distributor of haematology analysers used throughout the UK, including at the John Radcliffe Hospital. As part of this partnership, I developed a business plan to explore the commercial potential of marketing point-of-care analysers and the CBC-based multivariable diagnostic models developed in Chapter 5.

Four potential use cases were evaluated based on national NICU admission statistics, the estimated incidence of culture-confirmed sepsis, and summary data from the study cohort analysed in this thesis (Table 7.1). These scenarios ranged from a conservative approach (monitoring only culture-confirmed sepsis cases) to a comprehensive approach (continuous daily monitoring of all neonates admitted to the NICU). The projected test volume across the UK in these scenarios ranged from approximately 9,000 to 800,000 CBC results per year, equating to 129 to 11,000 CBC results per NICU per year.

Assuming that each NICU would be required to purchase a Sysmex XN-450 analyser as a point-of-care device, along with the necessary licences for the full CBC panel (including WBC and reticulocyte parameters), initial setup costs for each NICU would total approximately £7,300, generating £525,000 in total upfront revenue across all 72 UK NICUs. Reagent consumption would drive recurring revenue. Assuming a cost per test ranging from £0.41 (in high-volume settings) to £10.12 (low-volume use), annual reagent revenue would range from £93,000 to £330,000, depending on uptake and test frequency.

The diagnostic model itself could be commercialised in two ways: as an additional software licence, priced at £200 per NICU and generating an extra £14,400 in initial revenue; or as an added cost per test, for example a cost of £0.50 per CBC result would generate an annual recurring revenue of £4,500 to £400,000, depending on the use case scenario. These prices are exemplary, based on the current prices of similar Sysmex licences or NHS tests.

Table 7.1: Estimated revenue for the sale of the complete blood count as a diagnostic test for neonatal sepsis

	Monitoring Confirmed Sepsis	Monitoring Suspected Sepsis	Monitoring All Admissions	Continuous Monitoring
Estimation of CBC Measurements				
Patients	4,640 [1]	50,604[a]	100,762 [14]	100,762 [14]
CBC Results per Patient	2[b]	2[b]	2[b]	8[c]
Annual CBC Results	9,280	101,207	201,524	806,096
Annual CBC Results per NICU[d]	129	1,406	2,799	11,196
Daily CBC Results per NICU[d]	0.35	3.85	7.67	30.67
Reccurring Revenue for Reagents[e]				
Recurring Revenue per NICU[d]	£1,304.41	£1,506.02	£1,827.02	£4,631.83
Actual Cost per Test	£10.12	£1.07	£0.65	£0.41
Annual Recurring Revenue	£93,917.73	£108,433.60	£131,545.25	£333,491.60
Reccurring Revenue for Generation of Model Output				
Cost per Test	£0.50	£0.50	£0.50	£0.50
Annual Recurring Revenue	£4,640.00	£50,603.62	£100,762.00	£403,048.00
Annual Revenue per NICU[d]	£64.44	£702.83	£1,399.47	£5,597.89
Total Reccurring Revenue (Reagent + Model Output Cost)				
Cost per Test	£10.62	£1.57	£1.15	£0.91
Annual Recurring Revenue	£98,557.73	£159,037.22	£232,307.25	£736,539.60
Annual Revenue per NICU[d]	£1,368.86	£2,208.85	£3,226.49	£10,229.72

Estimates from study cohort: [a] 50% of admitted neonates have one episode of suspected sepsis; [b] median number of CBC results per neonate; [c] median duration of stay. [d] Based on 72 NICUs in the United Kingdom [181]. [e] Calculated by reagent consumption (Supplementary Table A5.1). Abbreviations: CBC – complete blood count; NICU – Neonatal Intensive Care Unit.

By comparison, the average NHS cost of a CRP test is approximately £1.50 [180]. Under the current costing model, excluding profit margins, all but the most conservative scenarios offer a CBC-based alternative at a lower per-test cost than CRP. Furthermore, the CBC offers broader diagnostic utility beyond sepsis, including assessments of anaemia. Therefore, integrating point-of-care CBC analysers into NICUs and embedding CBC-based diagnostics into standard neonatal care pathways could be a cost-effective strategy for healthcare providers, while representing a scalable commercial opportunity for Sysmex.

7.4 Limitations and Suggestions for Future Work

Despite efforts to conduct a robust study, several limitations could not be fully overcome. The main limitation is the retrospective nature of all the evaluated cohorts, which introduced missing data and likely led to misclassification of some cases. While every effort was made to include only neonates with complete records, and those with significant missing data were excluded, an inherent drawback of retrospective research is that it is not always clear which data are missing. For instance, data on antibiotic and corticosteroid administration were extracted from Badgernet, which relies on manual input. It is likely that not all instances of drug administration were accurately recorded, leading to misclassification of some CBC results as *unlikely* or *possible (low) sepsis*, thereby introducing bias into the results.

Similarly, clinical deterioration was assessed using hourly recorded vital sign data. Although efforts were made to account for potential errors, such as excluding isolated outliers, nuanced fluctuations within an hour would not have been captured. In addition, ventilation settings and changes were also used to assess clinical deterioration, but these are reactive and clinician-dependent indicators. Different staff members may respond to an unstable neonate in different ways, meaning that such data are subject to variability in clinical judgement rather than objective measures of patient status.

Due to time and resource constraints, it was not feasible to review each neonate's case manually to verify CBC classifications, culture results, or signs of deterioration.

Some positive cultures likely resulted from contamination rather than actual infection. Indeed, extreme outliers in the diagnostic model outputs often reflected such misclassified cases, suggesting that the models were capable of identifying culture contamination. This limitation was better addressed in the external validation cohort, where all positive cultures were systematically reviewed to exclude contaminants before classification, thereby reducing the risk of misclassification compared to the original cohort. Consequently, the *probable sepsis* group in this cohort was more accurately defined, lending greater confidence to the conclusions drawn from this analysis.

A prospective study, in which neonates are actively followed, and CBC classifications are made in consensus with multiple clinicians, would address many of these limitations. A clinical trial comparing standard clinical care with an intervention arm incorporating CBC-based evaluation, or multivariable diagnostic models, would allow for a direct evaluation of their impact on diagnostic accuracy, antibiotic stewardship, and clinical outcomes.

Although the sample size of the primary study cohort was sufficient for the analyses in Chapters 3 – 5, a larger dataset would have enabled more detailed analyses. For example, the reference range analyses (Chapter 3) showed that sex, postnatal age, and GA all significantly influenced many CBC parameters. Ideally, subgroup-specific reference ranges would be generated for clinical implementation. However, there were insufficient cases in each subgroup to calculate these ranges according to CLSI guidelines [137].

A larger sample size would also allow for investigation of additional clinical variables (such as birth weight for GA or sex) and their impact on the baseline haematological profile, evaluation of whether specific CBC parameters have differing diagnostic utility across neonatal subgroups, and the construction of more complex diagnostic models, including artificial intelligence and machine learning techniques. For example, although the random forest model developed in Chapter 5 outperformed simpler models in some respects, it was prone to overfitting, a common issue when training machine learning models on smaller datasets. A larger and more diverse dataset would allow for more reliable training, validation, and generalisation of such models.

A further consideration for clinical implementation is that this work is based on CBC parameters generated using Sysmex analysers, including several immune activation indices that are not universally available across all haematology platforms. While the underlying biological processes reflected by these parameters are generalisable, numerical outputs and parameter definitions differ between manufacturers. As such, the application of these diagnostic models beyond Sysmex platforms would require further validation using alternative analysers, or re-expression of the models using biologically equivalent parameters available across platforms.

7.5 Concluding Remarks

Neonatal sepsis remains a global diagnostic challenge, with current tools often failing to provide timely and accurate diagnosis, delaying targeted intervention. This thesis explored the underappreciated potential of the CBC, a rapid, low-cost test, to support early diagnosis and monitoring of neonatal sepsis.

Using a range of analytical techniques to characterise the baseline haematological profile, evaluate the diagnostic accuracy of individual parameters, and develop multivariable diagnostic models, I demonstrated that the CBC could provide meaningful diagnostic value beyond its current applications. These models were externally validated, demonstrating promising diagnostic performance and offering a potential route to early detection.

While further prospective validation is needed, the findings presented here support the CBC as a valuable tool for improving neonatal outcomes. Integrating such diagnostics into routine clinical workflows may reduce diagnostic uncertainty, optimise antibiotic stewardship, and improve care for this vulnerable population.

Bibliography

[1] Li, J., Xiang, L., Chen, X., Li, S., Sun, Q., et al. Global, regional, and national burden of neonatal sepsis and other neonatal infections, 1990–2019: Findings from the Global Burden of Disease Study 2019. *European Journal of Pediatrics*, 182(5): 2335–2343, 2023. doi: 10.1007/s00431-023-04911-7.

[2] Celik, I.H., Hanna, M., Canpolat, F.E., and Pammi, M. Diagnosis of neonatal sepsis: The past, present and future. *Pediatric Research*, 91(2):337–350, 2022. doi: 10.1038/s41390-021-01696-z.

[3] Ronchetti, M.P., Bersani, I., Piersigilli, F., and Auriti, C. Neonatal sepsis. *Archives of Paediatrics and Developmental Pathology*, 1(3):1015, 2017. doi: 10.47739/2641-774X/1015.

[4] Miselli, F., Crestani, S., Maugeri, M., Passini, E., Spaggiari, V., et al. Late-onset sepsis mortality among preterm infants: Beyond time to first antibiotics. *Microorganisms*, 11(2):396, 2023. doi: 10.3390/microorganisms11020396.

[5] Fleischmann, C., Reichert, F., Cassini, A., Horner, R., Harder, T., et al. Global incidence and mortality of neonatal sepsis: A systematic review and meta-analysis. *Archives of Disease in Childhood*, 106(8):745–752, 2021. doi: 10.1136/archdischild-2020-320217.

[6] Vergnano, S., Menson, E., Kennea, N., Embleton, N., Russell, A.B., et al. Neonatal infections in England: The NeonIN surveillance network. *Archives of Disease in Childhood: Fetal and Neonatal Edition*, 96(1):F9–F14, 2011. doi: 10.1136/adc.2009.178798.

[7] Raturi, A. and Chandran, S. Neonatal sepsis: Aetiology, pathophysiology, diagnostic advances and management strategies. *Clinical Medicine Insights: Pediatrics*, 18: 1–11, 2024. doi: 10.1177/11795565241281337.

[8] Russell, N., Barday, M., Okomo, U., Dramowski, A., Sharland, M., et al. Early- versus late-onset sepsis in neonates – time to shift the paradigm? *Clinical Microbiology and Infection*, 30(1):38–43, 2024. doi: 10.1016/j.cmi.2023.07.023.

[9] Simonsen, K.A., Anderson-Berry, A.L., Delair, S.F., and Davies, H.D. Early-onset neonatal sepsis. *Clinical Microbiology Reviews*, 27(1):21–47, 2014. doi: 10.1128/CMR.00031-13.

[10] Fox, G., Hoque, N., and Watts, T. Infection. In Fox, G., Hoque, N., and Watts, T., editors, *Oxford Handbook of Neonatology*, pages 323–370. Oxford University Press, Oxford, United Kingdom, 2nd edition, 2017. ISBN 978-0-19-177312-9 978-0-19-870395-2. doi: 10.1093/med/9780198703952.003.0012.

[11] NICE Guideline Updates Team. Neonatal infection: Antibiotics for prevention and treatment. Report NG195, National Institute for Health and Care Excellence (NICE), London, United Kingdom, 2024. `https://www.nice.org.uk/guidance/ng195`. Accessed on: 2025-02-07.

[12] Kuzniewicz, M.W., Puopolo, K.M., Fischer, A., Walsh, E.M., Li, S., et al. A quantitative, risk-based approach to the management of neonatal early-onset sepsis. *JAMA Pediatrics*, 171(4):365–371, 2017. doi: 10.1001/jamapediatrics.2016.4678.

[13] Pettinger, K.J., Mayers, K., McKechnie, L., and Phillips, B. Sensitivity of the Kaiser Permanente early-onset sepsis calculator: A systematic review and meta-analysis. *eClinicalMedicine*, 19:100227, 2020. doi: 10.1016/j.eclinm.2019.11.020.

[14] Bliss. Statistics about neonatal care, 2024. `https://www.bliss.org.uk/research-campaigns/neonatal-care-statistics/statistics-about-neonatal-care`. Accessed on: 2025-08-18.

[15] Bedford Russell, A.R. Neonatal sepsis. *Paediatrics and Child Health*, 25(6):271–275, 2015. doi: 10.1016/j.paed.2015.02.005.

[16] Zea-Vera, A. and Ochoa, T.J. Challenges in the diagnosis and management of neonatal sepsis. *Journal of Tropical Pediatrics*, 61:1–13, 2015. doi: 10.1093/tropej/fmu079.

[17] Tang, Y.H., Jeng, M.J., Wang, H.H., Tsao, P.C., Chen, W.Y., et al. Risk factors and predictive markers for early and late-onset neonatal bacteremic sepsis in preterm and term infants. *Journal of the Chinese Medical Association*, 85(4):507–513, 2022. doi: 10.1097/JCMA.0000000000000681.

[18] Sharma, D., Farahbakhsh, N., Shastri, S., and Sharma, P. Biomarkers for diagnosis of neonatal sepsis: A literature review. *The Journal of Maternal-Fetal & Neonatal Medicine*, 31(12):1646–1659, 2018. doi: 10.1080/14767058.2017.1322060.

[19] Boscarino, G., Migliorino, R., Carbone, G., Davino, G., Dell'Orto, V.G., et al. Biomarkers of neonatal sepsis: Where we are and where we are going. *Antibiotics*, 12:1233, 2023. doi: 10.3390/antibiotics12081233.

[20] O'Connell, A.E., Rivers, A.E., and Slayton, W.B. The development of the human immune system. In De Alarcón, P.A., Werner, E.J., Christensen, R.D., and Sola-Visner, M.C., editors, *Neonatal Hematology: Pathogenesis, Diagnosis, and Management of Hematologic Problems*, pages 25–42. Cambridge University Press, Cambridge, United Kingdom, 3rd edition, 2021. ISBN 978-1-108-77358-4 978-1-108-48898-3. doi: 10.1017/9781108773584.005.

[21] Bakhuizen, S.E., de Haan, T.R., Teune, M.J., van Wassenaer-Leemhuis, A.G., van der Heyden, J.L., et al. Meta-analysis shows that infants who have suffered neonatal sepsis face an increased risk of mortality and severe complications. *Acta Paediatrica*, 103(12):1211–1218, 2014. doi: 10.1111/apa.12764.

[22] Mukhopadhyay, S., Puopolo, K.M., Hansen, N.I., Lorch, S.A., DeMauro, S.B., et al. Neurodevelopmental outcomes following neonatal late-onset sepsis and blood culture-negative conditions. *Archives of Disease in Childhood: Fetal and Neonatal Edition*, 106(5):467–473, 2021. doi: 10.1136/archdischild-2020-320664.

[23] Xu, F., Luo, J., and Li, W. Comparison of the diagnostic accuracy of resistin and CRP levels for sepsis in neonates and children: A systematic review and meta-analysis. *Frontiers in Pediatrics*, 13:1555671, 2025. doi: 10.3389/fped.2025.1555671.

[24] Liu, Y., Zhao, L., and Wu, Z. Accuracy of C-reactive protein test for neonatal septicemia: A diagnostic meta-analysis. *Medical Science Monitor*, 25:4076–4081, 2019. doi: 10.12659/MSM.916968.

[25] Ruan, L., Chen, G.Y., Liu, Z., Zhao, Y., Xu, G.Y., et al. The combination of procalcitonin and C-reactive protein or presepsin alone improves the accuracy of diagnosis of neonatal sepsis: A meta-analysis and systematic review. *Critical Care*, 22(1):316, 2018. doi: 10.1186/s13054-018-2236-1.

[26] Xu, L., Li, Q., Mo, Z., and You, P. Diagnostic value of C-reactive protein in neonatal sepsis: A meta-analysis. *European Journal of Inflammation*, 14(2):100–108, 2016. doi: 10.1177/1721727X16646787.

[27] Yuan, H., Huang, J., Lv, B., Yan, W., Hu, G., et al. Diagnosis value of the serum amyloid a test in neonatal sepsis: A meta-analysis. *BioMed Research International*, 2013:520294, 2013. doi: 10.1155/2013/520294.

[28] Eschborn, S. and Weitkamp, J.H. Procalcitonin versus C-reactive protein: Review of kinetics and performance for diagnosis of neonatal sepsis. *Journal of Perinatology*, 39(7):893–903, 2019. doi: 10.1038/s41372-019-0363-4.

[29] Hedegaard, S.S., Wisborg, K., and Hvas, A.M. Diagnostic utility of biomarkers for neonatal sepsis – a systematic review. *Infectious Diseases*, 47(3):117–124, 2015. doi: 10.3109/00365548.2014.971053.

[30] Yu, Z., Liu, J., Sun, Q., Qiu, Y., Han, S., et al. The accuracy of the procalcitonin test for the diagnosis of neonatal sepsis: A meta-analysis. *Scandinavian Journal of Infectious Diseases*, 42(10):723–733, 2010. doi: 10.3109/00365548.2010.489906.

[31] Vouloumanou, E.K., Plessa, E., Karageorgopoulos, D.E., Mantadakis, E., and Falagas, M.E. Serum procalcitonin as a diagnostic marker for neonatal sepsis: A systematic review and meta-analysis. *Intensive Care Medicine*, 37(5):747–762, 2011. doi: 10.1007/s00134-011-2174-8.

[32] Xing, W., Wang, Y., Liu, J., Pei, J., and Yu, C. Role of interleukins in the detection of neonatal sepsis: A network meta-analysis. *Frontiers in Pediatrics*, 11:1267777, 2023. doi: 10.3389/fped.2023.1267777.

[33] Shahkar, L., Keshtkar, A., Mirfazeli, A., Ahani, A., and Roshandel, G. The role of IL-6 for predicting neonatal sepsis: A systematic review and meta-analysis. *Iranian Journal of Pediatrics*, 21(4):411–417, 2011. `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3446138/`. Accessed on: 2025-07-29.

[34] Eichberger, J. and Resch, B. Reliability of interleukin-6 alone and in combination for diagnosis of early onset neonatal sepsis: Systematic review. *Frontiers in Pediatrics*, 10:840778, 2022. doi: 10.3389/fped.2022.840778.

[35] Boskabadi, H. and Zakerihamidi, M. Evaluate the diagnosis of neonatal sepsis by measuring interleukins: A systematic review. *Pediatrics & Neonatology*, 59(4): 329–338, 2018. doi: 10.1016/j.pedneo.2017.10.004.

[36] Zhou, M., Cheng, S., Yu, J., and Lu, Q. Interleukin-8 for diagnosis of neonatal sepsis: A meta-analysis. *PLOS ONE*, 10(5):e0127170, 2015. doi: 10.1371/journal.pone.0127170.

[37] Bellos, I., Fitrou, G., Pergialiotis, V., Thomakos, N., Perrea, D.N., et al. The diagnostic accuracy of presepsin in neonatal sepsis: A meta-analysis. *European Journal of Pediatrics*, 177(5):625–632, 2018. doi: 10.1007/s00431-018-3114-1.

[38] Bellos, I., Fitrou, G., Daskalakis, G., Thomakos, N., Papantoniou, N., et al. Soluble TREM-1 as a predictive factor of neonatal sepsis: A meta-analysis. *Inflammation Research*, 67(7):571–578, 2018. doi: 10.1007/s00011-018-1149-4.

[39] Shi, J., Tang, J., and Chen, D. Meta-analysis of diagnostic accuracy of neutrophil CD64 for neonatal sepsis. *Italian Journal of Pediatrics*, 42:57, 2016. doi: 10.1186/s13052-016-0268-1.

[40] Sweeney, T.E., Wynn, J.L., Cernada, M., Serna, E., Wong, H.R., et al. Validation of the sepsis MetaScore for diagnosis of neonatal sepsis. *Journal of the Pediatric Infectious Diseases Society*, 7(2):129–135, 2018. doi: 10.1093/jpids/pix021.

[41] Rashid, A., Al-Obeidat, F., Kanthimathinathan, H.K., Benakatti, G., Hafez, W., et al. Advancing sepsis clinical research: Harnessing transcriptomics for an omics-based strategy - a comprehensive scoping review. *Informatics in Medicine Unlocked*, 44:101419, 2024. doi: 10.1016/j.imu.2023.101419.

[42] Lee, C.Y., Zven, S.E., Sathya, S.A., Abukhalaf, D., Sahoo, S., et al. Using molecular methods to diagnose, classify, and treat neonatal sepsis: A scoping review. *Frontiers in Pediatrics*, 13, 2025. doi: 10.3389/fped.2025.1625449.

[43] Bai, Y., Zhao, N., Zhang, Z., Jia, Y., Zhang, G., et al. Identification and validation of a novel four-gene diagnostic model for neonatal early-onset sepsis with bacterial infection. *European Journal of Pediatrics*, 182(3):977–985, 2023. doi: 10.1007/s00431-022-04753-9.

[44] Cernada, M., Pinilla-González, A., Kuligowski, J., Morales, J.M., Lorente-Pozo, S., et al. Transcriptome profiles discriminate between Gram-positive and Gram-negative sepsis in preterm neonates. *Pediatric Research*, 91(3):637–645, 2022. doi: 10.1038/s41390-021-01444-3.

[45] Qi, Y., Chen, X., Wu, N., Ma, C., Cui, X., et al. Identification of risk factors for sepsis-associated mortality by gene expression profiling analysis. *Molecular Medicine Reports*, 17(4):5350–5355, 2018. doi: 10.3892/mmr.2018.8491.

[46] Ng, S., Strunk, T., Lee, A.H., Gill, E.E., Falsafi, R., et al. Whole blood transcriptional responses of very preterm infants during late-onset sepsis. *PLOS ONE*, 15(6):e0233841, 2020. doi: 10.1371/journal.pone.0233841.

[47] Mukhopadhyay, S., Sengupta, S., and Puopolo, K.M. Challenges and opportunities for antibiotic stewardship among preterm infants. *Archives of Disease in Childhood: Fetal and Neonatal Edition*, 104(3):F327–F332, 2019. doi: 10.1136/archdischild-2018-315412.

[48] Tapiainen, T., Koivusaari, P., Brinkac, L., Lorenzi, H.A., Salo, J., et al. Impact of intrapartum and postnatal antibiotics on the gut microbiome and emergence of antimicrobial resistance in infants. *Scientific Reports*, 9(1):10635, 2019. doi: 10.1038/s41598-019-46964-5.

[49] Esmaeilizand, R., Shah, P.S., Seshia, M., Yee, W., Yoon, E.W., et al. Antibiotic exposure and development of necrotizing enterocolitis in very preterm neonates. *Paediatrics & Child Health*, 23(4):e56–e61, 2018. doi: 10.1093/pch/pxx169.

[50] Zhang, Z., Wang, J., Wang, H., Li, Y., Jia, Y., et al. Association of infant antibiotic exposure and risk of childhood asthma: A meta-analysis. *The World Allergy Organization Journal*, 14(11):100607, 2021. doi: 10.1016/j.waojou.2021.100607.

[51] Canova, C., Ludvigsson, J.F., Di Domenicantonio, R., Zanier, L., Barbiellini Amidei, C., et al. Perinatal and antibiotic exposures and the risk of developing childhood-onset inflammatory bowel disease: A nested case-control study based on a population-based birth cohort. *International Journal of Environmental Research and Public Health*, 17(7):2409, 2020. doi: 10.3390/ijerph17072409.

[52] Li, G., Bielicki, J.A., Ahmed, A.S.M.N.U., Islam, M.S., Berezin, E.N., et al. Towards understanding global patterns of antimicrobial use and resistance in neonatal sepsis: Insights from the NeoAMR network. *Archives of Disease in Childhood*, 105(1): 26–31, 2020. doi: 10.1136/archdischild-2019-316816.

[53] Howie, S.R.C. Blood sample volumes in child health research: Review of safe limits. *Bulletin of the World Health Organization*, 89(1):46–53, 2011. doi: 10.2471/BLT.10.080010.

[54] Hay, D., King, A., and Desborough, M. *Haematology*. Lecture Notes. Wiley-Blackwell, 11th edition, 2022. ISBN 978-1-119-84121-0.

[55] Hoffbrand, A.V., Chowdary, P., Collins, G., and Loke, J. *Hoffbrand's Essential Haematology*. Wiley -Blackwell, Hoboken (New Jersey), ninth edition edition, 2024. ISBN 978-1-394-16817-0 978-1-394-16815-6.

[56] Sysmex Europe. Haematopoiesis, 2025. `https://www.sysmex.co.uk/education/knowledge-centre/haematopoiesis/`. Accessed on: 2025-07-30.

[57] Papayannopoulou, T. and Migliaccio, A.R. Biology of erythropoiesis, erythroid differentiation, and maturation. In Hoffman, R., Benz, E.J., Silberstein, L.E., Heslop, H.E., Weitz, J.I., et al., editors, *Hematology*, pages 297–320.e14. Elsevier, 7th edition, 2018. ISBN 978-0-323-35762-3. doi: 10.1016/B978-0-323-35762-3.00026-3.

[58] Pikora, K., Krętowska-Grunwald, A., Krawczuk-Rybak, M., and Sawicka-Żukowska, M. Diagnostic value and prognostic significance of nucleated red blood cells (NRBCs) in selected medical conditions. *Cells*, 12(14):1817, 2023. doi: 10.3390/cells12141817.

[59] Basha, S., Surendran, N., and Pichichero, M. Immune responses in neonates. *Expert Review of Clinical Immunology*, 10(9):1171–1184, 2014. doi: 10.1586/1744666X.2014.942288.

[60] Camacho-Gonzalez, A., Spearman, P.W., and Stoll, B.J. Neonatal infectious diseases: Evaluation of neonatal sepsis. *Pediatric Clinics of North America*, 60(2):367–389, 2013. doi: 10.1016/j.pcl.2012.12.003.

[61] Khanna-Gupta, A. and Berliner, N. Granulocytopoiesis and monocytopoiesis. In Hoffman, R., Benz, E.J., Silberstein, L.E., Heslop, H.E., Weitz, J.I., et al., editors, *Hematology*, pages 321–333.e1. Elsevier, 7th edition, 2018. ISBN 978-0-323-35762-3. doi: 10.1016/B978-0-323-35762-3.00027-5.

[62] van Pelt, J., Klatte, S., Hwandih, T., Barcaru, A., Riphagen, I.J., et al. Reference intervals for Sysmex XN hematological parameters as assessed in the Dutch Lifelines cohort. *Clinical Chemistry and Laboratory Medicine*, 60(6):907–920, 2022. doi: 10.1515/cclm-2022-0094.

[63] Cantor, A.B. Thrombocytopoiesis. In Hoffman, R., Benz, E.J., Silberstein, L.E., Heslop, H.E., Weitz, J.I., et al., editors, *Hematology*, pages 334–349. Elsevier, 7th edition, 2018. ISBN 978-0-323-35762-3. doi: 10.1016/B978-0-323-35762-3.00028-7.

[64] Sysmex Europe. DC sheath flow detection method, 2025. `https://www.sysmex-europe.com/academy/knowledge-centre/technologies/dc-sheath-flow-detection-method/`. Accessed on: 2025-07-30.

[65] Sysmex Europe. SLS detection method, 2025. `https://www.sysmex-europe.com/academy/knowledge-centre/technologies/sls-detection-method/`. Accessed on: 2025-07-30.

[66] Sysmex Europe. Fluorescence flow cytometry, 2025. `https://www.sysmex-europe.com/academy/knowledge-centre/technologies/fluorescence-flow-cytometry/`. Accessed on: 2025-07-30.

[67] De La Salle, B. Pre- and postanalytical errors in haematology. *International Journal of Laboratory Hematology*, 41(S1):170–176, 2019. doi: 10.1111/ijlh.13007.

[68] Gunawardena, D., Jayaweera, S., Madhubhashini, G., Lokumarakkala, D.D., and Senanayake, S.J. Reliability of parameters of complete blood count with different storage conditions. *Journal of Clinical Laboratory Analysis*, 31(2):e22042, 2017. doi: 10.1002/jcla.22042.

[69] Gulati, G., Uppal, G., and Gong, J. Unreliable automated complete blood count results: Causes, recognition, and resolution. *Annals of Laboratory Medicine*, 42(5):515–530, 2022. doi: 10.3343/alm.2022.42.5.515.

[70] Vis, J.Y. and Huisman, A. Verification and quality control of routine hematology analyzers. *International Journal of Laboratory Hematology*, 38(S1):100–109, 2016. doi: 10.1111/ijlh.12503.

[71] Rice, L. and Jung, M. Neutrophilic leukocytosis, neutropenia, monocytosis, and monocytopenia. In Hoffman, R., Benz, E.J., Silberstein, L.E., Heslop, H.E., Weitz, J.I., et al., editors, *Hematology*, pages 675–681. Elsevier, 7th edition, 2018. ISBN 978-0-323-35762-3. doi: 10.1016/B978-0-323-35762-3.00048-2.

[72] Agnello, L., Giglio, R.V., Bivona, G., Scazzone, C., Gambino, C.M., et al. The value of a complete blood count (CBC) for sepsis diagnosis and prognosis. *Diagnostics*, 11(10):1881, 2021. doi: 10.3390/diagnostics11101881.

[73] Nierhaus, A., Linssen, J., Wichmann, D., Braune, S., and Kluge, S. Use of a weighted, automated analysis of the differential blood count to differentiate sepsis from non-infectious systemic inflammation: The intensive care infection score (ICIS). *Inflammation & Allergy-Drug Targets*, 11(2):109–115, 2012. doi: 10.2174/187152812800392841.

[74] Weimann, K., Zimmermann, M., Spies, C.D., Wernecke, K.D., Vicherek, O., et al. Intensive care infection score - a new approach to distinguish between infectious and noninfectious processes in intensive care and medicosurgical patients. *The Journal of International Medical Research*, 43(3):435–451, 2015. doi: 10.1177/0300060514557711.

[75] Van Der Geest, P.J., Mohseni, M., Linssen, J., Duran, S., De Jonge, R., et al. The intensive care infection score - a novel marker for the prediction of infection and its severity. *Critical Care*, 20(1):180, 2016. doi: 10.1186/s13054-016-1366-6.

[76] Hyde, E., Anthony, M., Kennedy, S., and Vatish, M. Role of complete blood count in the diagnosis of culture-proven neonatal sepsis: A systematic review and meta-analysis. *Archives of Disease in Childhood*, 110, 2025. doi: 10.1136/archdischild-2025-328523.

[77] NICE Guideline Updates Team. Evidence review for investigations before starting treatment for late-onset neonatal infection: Neonatal infection: Antibiotics for prevention and treatment. Report Evidence Review G, National Institute for Health and Care Excellence (NICE), London, United Kingdom, 2021. https://www.nice.org.uk/guidance/ng195/evidence. Accessed on: 2025-02-07.

[78] European Medicines Agency (EMA). Report on the expert meeting on neonatal and paediatric sepsis. Report EMA/477725/2010, European Medicines Agency, London, United Kingdom, 2010. https://www.ema.europa.eu/en/documents/report/report-expert-meeting-neonatal-and-paediatric-sepsis_en.pdf.

[79] World Health Organization, Child Health and Development (CHD), Guidelines Review Committee, and Maternal, Newborn, Child & Adolescent Health & Ageing (MCA). *Pocket Book of Hospital Care for Children: Guidelines for the Management of Common Childhood Illnesses*. World Health Organization, Geneva, Switzerland, 2nd edition, 2013. ISBN 978-92-4-154837-3.

[80] Deeks, J.J., Bossuyt, P.M., Leeflang, M.M., and Takwoingi, Y., editors. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy*. John Wiley & Sons, Chichester, United Kingdom, 1st edition, 2023. ISBN 978-1-119-75616-3.

[81] Hyde, E., Saker, J., Anthony, M., and Vatish, M. The accuracy of haematological parameters for the diagnosis of neonatal sepsis. Report CRD42023476510, PROSPERO, 2023. https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42023476510. Accessed on: 2024-11-27.

[82] Veritas Health Innovation. Covidence systematic review software, 2024. https://www.covidence.org/. Accessed on: 2024-10-28.

[83] Whiting, P.F., Rutjes, A.W., Westwood, M.E., Mallett, S., Deeks, J.J., et al. QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine*, 155(8):529–536, 2011. doi: 10.7326/0003-4819-155-8-201110180-00009.

[84] Critical Appraisal Skills Programme. CASP diagnostic test study checklist, 2018. https://casp-uk.net/casp-tools-checklists/diagnostic-study-checklist/. Accessed on: 2022-01-20.

[85] Wells, G., Shea, B., O'Connell, D., Robertson, J., Peterson, J., et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomized studies in meta-analysis, 2021. https://www.ohri.ca/programs/clinical_epidemiology/oxford.asp. Accessed on: 2022-02-21.

[86] RStudio Team. RStudio: Integrated Development Environment for R. RStudio, PBC, 2021. http://www.rstudio.com/.

[87] Abiramalatha, T., Santhanam, S., Mammen, J., Rebekah, G., Shabeer, M., et al. Utility of neutrophil volume conductivity scatter (VCS) parameter changes as sepsis screen in neonates. *Journal of Perinatology*, 36(9):733–738, 2016. doi: 10.1038/jp.2016.69.

[88] Aboud, M.I., Waise, M.M.A., and Shakerdi, L.A. Procalcitonin as a marker of neonatal sepsis in intensive care units. *Iranian Journal of Medical Sciences*, 35(3): 205–210, 2010. https://ijms.sums.ac.ir/article_39783.html.

[89] Buyukeren, M., Celik, H.T., Yigit, S., Portakal, O., Sancak, B., et al. The role of the delta neutrophil index in determining the etiology of neonatal sepsis. *Turkish Journal of Biochemistry*, 46(1):73–79, 2021. doi: 10.1515/TJB-2020-0030.

[90] Celik, H.T., Portakal, O., Yigit, S., Hascelik, G., Korkmaz, A., et al. Efficacy of new leukocyte parameters versus serum C-reactive protein, procalcitonin, and interleukin-6 in the diagnosis of neonatal sepsis. *Pediatrics International*, 58(2): 119–125, 2016. doi: 10.1111/ped.12754.

[91] Celik, I.H., Arifoglu, I., Arslan, Z., Aksu, G., Bas, A.Y., et al. The value of delta neutrophil index in neonatal sepsis diagnosis, follow-up and mortality prediction. *Early Human Development*, 131:6–9, 2019. doi: 10.1016/j.earlhumdev.2019.02.003.

[92] Elawady, S., Botros, S.K., Sorour, A.E., Ghany, E.A., Elbatran, G., et al. Neutrophil CD64 as a diagnostic marker of sepsis in neonates. *Journal of Investigative Medicine*, 62(3):644–649, 2014. doi: 10.2310/JIM.0000000000000060.

[93] Kumar, P., Anjula, C., Bhardwaj, P., Chauhan, L., and Karol, M. D-dimer: A useful marker in neonatal sepsis. *Journal of Clinical Neonatology*, 4(2):101–103, 2015. doi: 10.4103/2249-4847.154552.

[94] Leibovitch, L., Zohar, H., Gavri-Beker, A., Goshen, A., and Strauss, T. Mean platelet volume in preterm infants as a predictor of late-onset neonatal sepsis: A retrospective comparative study. *BMJ Paediatrics Open*, 8(1):e002698, 2024. doi: 10.1136/bmjpo-2024-002698.

[95] Mahmoud, N.M.S., Abdelhakeem, M., Mohamed, H.H.A., and Baheeg, G. Validity of platelet to lymphocyte ratio and neutrophil to lymphocyte ratio in diagnosing early-onset neonatal sepsis in full-term newborns. *Journal of Comprehensive Pediatrics*, 13(1):e115378, 2022. doi: 10.5812/compreped.115378.

[96] Nesargi, P., Niranjan, H., Bandiya, P., and Benakappa, N. Neutrophil volume, conductivity and scatter (VCS) as a screening tool in neonatal sepsis. *Scientific Reports*, 10(1):4457, 2020. doi: 10.1038/s41598-020-61434-z.

[97] Schrama, A., de Beaufort, A., Poorthuis, B., Berger, H., and Walther, F. Secretory phospholipase A2 in newborn infants with sepsis. *Journal of Perinatology*, 28(4): 291–296, 2008. doi: 10.1038/sj.jp.7211929.

[98] Tosson, A.M., Koptan, D.M., Kamal, M., and Elhady, M.A. Assessment of serum interleukin-27 and mean platelet volume in late-onset neonatal sepsis. *American Journal of Perinatology*, 41(9):1232–1237, 2024. doi: 10.1055/s-0042-1748165.

[99] Vardar, G. and Ozek, E. Carboxyhemoglobin levels in preterm neonatal late-onset sepsis: To predict or not to predict. *Mediterranean Journal of Hematology and Infectious Diseases*, 15(1):e2023017, 2023. doi: 10.4084/mjhid.2023.017.

[100] Zhang, S., Luan, X., Zhang, W., and Jin, Z. Platelet-to-lymphocyte and neutrophil-to-lymphocyte ratio as predictive biomarkers for early-onset neonatal sepsis. *Journal of the College of Physicians and Surgeons*, 31(7):821–824, 2021. doi: 10.29271/jcpsp.2021.07.821.

[101] Arcagok, B.C. and Karabulut, B. Platelet to lymphocyte ratio in neonates: A predictor of early onset neonatal sepsis. *Mediterranean Journal of Hematology and Infectious Diseases*, 11(1):e2019055, 2019. doi: 10.4084/MJHID.2019.055.

[102] Guney Varal, I. and Dogan, P. Can neutrophil-lymphocyte ratio be a predictor of late-onset sepsis in preterm infants? *Annals of Medical Research*, 27(1):23–28, 2020. doi: 10.5455/annalsmedres.2020.01.023.

[103] Hanaganahalli, S.B., Sreeram, S., Bompada, M., Kuppannagari, S.K., Suresh, P.K., et al. Is MPV a predictive marker for neonatal sepsis? A pilot study. *Journal of Pediatric Hematology/Oncology*, 40(7):548–552, 2018. doi: 10.1097/MPH.0000000000001272.

[104] Madani, S.H., Amiri, S., Khazaei, S., Khadem Erfan, M.B., Rostami-Far, Z., et al. Platelet indices as useful indicators of neonatal sepsis. *Journal of Evolution of Medical and Dental Sciences*, 8(20):1612–1617, 2019. doi: 10.14260/jemds/2019/357.

[105] Panda, S.K., Nayak, M.K., Rath, S., and Das, P. The utility of the neutrophil-lymphocyte ratio as an early diagnostic marker in neonatal sepsis. *Cureus*, 13(1):e12891, 2021. doi: 10.7759/cureus.12891.

[106] Guney Varal, I., Dogan, P., Acar Celik, E., and Güler Kazancı, E. Mean platelet volume and mean platelet volume/platelet count ratio are predictors of late-onset sepsis in preterm infants: A case-control study. *Fetal and Pediatric Pathology*, 42(1):44–54, 2023. doi: 10.1080/15513815.2022.2064574.

[107] Manucha, V., Rusia, U., Sikka, M., Faridi, M., and Madan, N. Utility of haematological parameters and C-reactive protein in the detection of neonatal

sepsis. *Journal of Paediatrics and Child Health*, 38(5):459–464, 2002. doi: 10.1046/j.1440-1754.2002.00018.x.

[108] Panda, S.K., Nayak, M.K., Thangaraj, J., Das, P., and Pugalia, R. Platelet parameters as a diagnostic marker in early diagnosis of neonatal sepsis - seeking newer answers for older problems. *Journal of Family Medicine and Primary Care*, 11(5):1748–54, 2022. doi: 10.4103/jfmpc.jfmpc_1271_21.

[109] Yang, A.P., Liu, J., Yue, L.H., Wang, H.Q., Yang, W.J., et al. Neutrophil CD64 combined with PCT, CRP and WBC improves the sensitivity for the early diagnosis of neonatal sepsis. *Clinical Chemistry and Laboratory Medicine*, 54(2):345–351, 2016. doi: 10.1515/cclm-2015-0277.

[110] Zaki, M.E.S. and El Sayed, H. Evaluation of microbiologic and hematologic parameters and E-selectin as early predictors for outcome of neonatal sepsis. *Archives of Pathology and Laboratory Medicine*, 133(8):1291–1296, 2009. doi: 10.5858/133.8.1291.

[111] Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372:n71, 2021. doi: 10.1136/bmj.n71.

[112] Honda, T., Uehara, T., Matsumoto, G., Arai, S., and Sugano, M. Neutrophil left shift and white blood cell count as markers of bacterial infection. *Clinica Chimica Acta*, 457:46–53, 2016. doi: 10.1016/j.cca.2016.03.017.

[113] Christensen, R.D., Henry, E., Jopling, J., and Wiedmeier, S.E. The CBC: Reference ranges for neonates. *Seminars in Perinatology*, 33(1):3–11, 2009. doi: 10.1053/j.semperi.2008.10.010.

[114] Jia, W.Y. and Zhang, J.J. Effects of glucocorticoids on leukocytes: Genomic and non-genomic mechanisms. *World Journal of Clinical Cases*, 10(21):7187–7194, 2022. doi: 10.12998/wjcc.v10.i21.7187.

[115] Doyle, L.W., Davis, P.G., Morley, C.J., McPhee, A., Carlin, J.B., et al. Low-dose dexamethasone facilitates extubation among chronically ventilator-dependent infants: A multicenter, international, randomized, controlled trial. *Pediatrics*, 117 (1):75–83, 2006. doi: 10.1542/peds.2004-2843.

[116] Roberts, I., Stanworth, S., and Murray, N.A. Thrombocytopenia in the neonate. *Blood Reviews*, 22(4):173–186, 2008. doi: 10.1016/j.blre.2008.03.004.

[117] Haque, K.N. Definitions of bloodstream infection in the newborn. *Pediatric Critical Care Medicine*, 6(3 Suppl):S45–S49, 2005. doi: 10.1097/01.PCC.0000161946.73305.0A.

[118] Chen, J., Yasrebinia, S., Ghaedi, A., Khanzadeh, M., Quintin, S., et al. Meta-analysis of the role of neutrophil to lymphocyte ratio in neonatal sepsis. *BMC Infectious Diseases*, 23:837, 2023. doi: 10.1186/s12879-023-08800-0.

[119] Xin, Y., Shao, Y., Mu, W., Li, H., Zhou, Y., et al. Accuracy of the neutrophil-to-lymphocyte ratio for the diagnosis of neonatal sepsis: A systematic review and meta-analysis. *BMJ Open*, 12:e060391, 2022. doi: 10.1136/bmjopen-2021-060391.

[120] Toro-Huamanchumo, C.J., Cabanillas-Ramirez, C., Quispe-Vicuña, C., Caballero-Alvarado, J.A., León-Figueroa, D.A., et al. Mean platelet volume in neonatal sepsis: Meta-analysis of observational studies. *Children*, 9(12):1821, 2022. doi: 10.3390/children9121821.

[121] Narasimha, A. and Harendra Kumar, M. Significance of hematological scoring system (HSS) in early diagnosis of neonatal sepsis. *Indian Journal of Hematology and Blood Transfusion*, 27(1):14–17, 2011. doi: 10.1007/s12288-010-0050-2.

[122] Rodwell, R.L., Leslie, A.L., and Tudehope, D.I. Early diagnosis of neonatal sepsis using a hematologic scoring system. *The Journal of Pediatrics*, 112(5):761–767, 1988. doi: 10.1016/S0022-3476(88)80699-1.

[123] Padhy, S., Thandaveshwara, D., Doreswamy, S.M., and Krishnamurthy, V. Sepscore - an improved armament in the diagnosis of neonatal sepsis. *Journal of Clinical and Diagnostic Research*, 17(4):5–7, 2023. doi: 10.7860/JCDR/2023/59408.17710.

[124] Huang, B., Wang, R., Masino, A.J., and Obstfeld, A.E. Aiding clinical assessment of neonatal sepsis using hematological analyzer data with machine learning techniques. *International Journal of Laboratory Hematology*, 43(6):1341–1356, 2021. doi: 10.1111/ijlh.13549.

[125] Nassin, M.L., Lapping-Carr, G., and de Jong, J.L.O. Anemia in the neonate: The differential diagnosis and treatment. *Pediatric Annals*, 44(7):e159–e163, 2015. doi: 10.3928/00904481-20150710-08.

[126] Hyde, E., Saker, J., Kennedy, S., Anthony, M., and Vatish, M. Establishing reference ranges and evaluating clinical factors of the complete blood count in neonates admitted to a neonatal intensive care unit. *International Journal of Laboratory Hematology*, 48(1):81–92, 2025. doi: 10.1111/ijlh.70014.

[127] Manroe, B.L., Weinberg, A.G., Rosenfeld, C.R., and Browne, R. The neonatal blood count in health and disease. I. Reference values for neutrophilic cells. *The Journal of Pediatrics*, 95(1):89–98, 1979. doi: 10.1016/s0022-3476(79)80096-7.

[128] Kalbermatter, C., Fernandez Trigo, N., Christensen, S., and Ganal-Vonarburg, S.C. Maternal microbiota, early life colonization and breast milk drive immune development in the newborn. *Frontiers in Immunology*, 12:683022, 2021. doi: 10.3389/fimmu.2021.683022.

[129] Christensen, R.D., Henry, E., Andres, R.L., and Bennett, S.T. Reference ranges for blood concentrations of nucleated red blood cells in neonates. *Neonatology*, 99 (4):289–294, 2011. doi: 10.1159/000320148.

[130] Christensen, R.D., Henry, E., Bennett, S.T., and Yaish, H.M. Reference intervals for reticulocyte parameters of infants during their first 90 days after birth. *Journal of Perinatology*, 36(1):61–66, 2016. doi: 10.1038/jp.2015.140.

[131] Christensen, R.D., Baer, V.L., Gordon, P.V., Henry, E., Whitaker, C., et al. Reference ranges for lymphocyte counts of neonates: Associations between abnormal counts and outcomes. *Pediatrics*, 129(5):e1165–1172, 2012. doi: 10.1542/peds.2011-2661.

[132] Schmutz, N., Henry, E., Jopling, J., and Christensen, R.D. Expected ranges for blood neutrophil concentrations of neonates: The Manroe and Mouzinho charts revisited. *Journal of Perinatology*, 28(4):275–281, 2008. doi: 10.1038/sj.jp.7211916.

[133] Jefferies, AL. Postnatal corticosteroids to treat or prevent chronic lung disease in preterm infants. *Paediatrics & Child Health*, 17(10):573–574, 2012. doi: 10.1093/pch/17.10.573.

[134] Barden, A., Phillips, M., Hill, L.M., Fletcher, E.M., Mas, E., et al. Antiemetic doses of dexamethasone and their effects on immune cell populations and plasma mediators of inflammation resolution in healthy volunteers. *Prostaglandins, Leukotrienes and Essential Fatty Acids*, 139:31–39, 2018. doi: 10.1016/j.plefa.2018.11.004.

[135] Roberts, I.A.G. and Murray, N.A. Haematology. In Rennie, J.M., editor, *Rennie & Roberton's Textbook of Neonatology*, pages 755–790. Elsevier Health Sciences UK, Edinburgh, Scotland, 5th edition, 2012. ISBN 978-0-7020-3479-4.

[136] Fox, G., Hoque, N., and Watts, T. Normal values, therapeutic drug levels, and useful formulae. In Fox, G., Watts, T., and Hoque, N., editors, *Oxford Handbook of Neonatology*, pages 531–542. Oxford University Press, Oxford, United Kingdom, 2nd edition, 2017. ISBN 978-0-19-870395-2. doi: 10.1093/med/9780198703952.003.0021.

[137] Clinical and Laboratory Standards Institute (CLSI). Defining, establishing, and verifying reference intervals in the clinical laboratory: Approved guideline. Report

EP28-A3C, Clinical and Laboratory Standards Institute (CLSI), 2010. https://clsi.org/shop/standards/ep28/. Accessed on: 2025-08-20.

[138] Hyde, E. Characterisation of the blood cell response in neonatal sepsis, 2025. https://4fsqon-emily-hyde.shinyapps.io/ThesisFigures/. Accessed on: 2025-07-15.

[139] Zhang, X., Zhang, Y., Xu, Y., Liu, J., Fu, M., et al. Age- and sex-specific reference intervals for complete blood count parameters in capillary blood for Chinese neonates and infants: A prospective study. *Clinica Chimica Acta*, 538: 104–112, 2023. doi: 10.1016/j.cca.2022.11.016.

[140] Christensen, R.D. Reference intervals in neonatal hematology. In Werner, E.J., Sola-Visner, M.C., de Alarcón, P.A., and Christensen, R.D., editors, *Neonatal Hematology: Pathogenesis, Diagnosis, and Management of Hematologic Problems*, pages 440–469. Cambridge University Press, Cambridge, United Kingdom, 3rd edition, 2021. ISBN 978-1-108-48898-3.

[141] Shannon, O. The role of platelets in sepsis. *Research and Practice in Thrombosis and Haemostasis*, 5(1):27–37, 2021. doi: 10.1002/rth2.12465.

[142] Rees, C.A., Lim, J., Westbrook, A.L., El Helou, R., Schmid, A., et al. Systematic review and meta-analysis of the diagnostic value of four biomarkers in detecting neonatal sepsis in low- and middle-income countries. *BMJ Paediatrics Open*, 7(1): e001627, 2023. doi: 10.1136/bmjpo-2022-001627.

[143] Puopolo, K.M., Draper, D., Wi, S., Newman, T.B., Zupancic, J., et al. Estimating the probability of neonatal early-onset infection on the basis of maternal risk factors. *Pediatrics*, 128(5):e1155, 2011. doi: 10.1542/peds.2010-3464.

[144] Töllner, U. Early diagnosis of septicemia in the newborn: Clinical studies and sepsis score. *European Journal of Pediatrics*, 138(4):331–337, 1982. doi: 10.1007/BF00442511.

[145] Putri, V.R., Lubis, B.M., and Lubis, I.N.D. Effectiveness of Tollner's score and procalcitonin for diagnosing neonatal sepsis. *International Journal of Innovative Science and Research Technology*, 7(7):947–950, 2022. doi: 10.5281/zenodo.6974466.

[146] Özdemir, Ö.M., Erdal, B., and Turgut, M. Evaluation of EMA, Töllner and Rodwell scores in the diagnosis of neonatal sepsis. *Pamukkale Medical Journal*, 17:746–754, 2024. doi: 10.31362/patd.1457265.

[147] Mahieu, L.M., De Muynck, A.O., De Dooy, J.J., Laroche, S.M., and Van Acker, K.J. Prediction of nosocomial sepsis in neonates by means of a computer-weighted bedside scoring system (NOSEP score). *Critical Care Medicine*, 28(6):2026–2033, 2000. doi: 10.1097/00003246-200006000-00058.

[148] Mahieu, L.M., De Dooy, J.J., Cossey, V.R., Goossens, L.L., Vrancken, S.L., et al. Internal and external validation of the NOSEP prediction score for nosocomial sepsis in neonates. *Critical Care Medicine*, 30(7):1459–1466, 2002. doi: 10.1097/00003246-200207000-00010.

[149] Singh, S.A., Dutta, S., and Narang, A. Predictive clinical scores for diagnosis of late onset neonatal septicemia. *Journal of Tropical Pediatrics*, 49(4):235–239, 2003. doi: 10.1093/tropej/49.4.235.

[150] Kudawla, M., Dutta, S., and Narang, A. Validation of a clinical score for the diagnosis of late onset neonatal septicemia in babies weighing 1000-2500 g. *Journal of Tropical Pediatrics*, 54(1):66–69, 2008. doi: 10.1093/tropej/fmm065.

[151] Rosenberg, R.E., Ahmed, A.N.U., Saha, S.K., Chowdhury, M.A., Ahmed, S., et al. Nosocomial sepsis risk score for preterm infants in low-resource settings. *Journal of Tropical Pediatrics*, 56(2):82–89, 2009. doi: 10.1093/tropej/fmp061.

[152] Okascharoen, C., Sirinavin, S., Thakkinstian, A., Kitayaporn, D., and Supapanachart, S. A bedside prediction-scoring model for late-onset neonatal sepsis. *Journal of Perinatology*, 25(12):778–783, 2005. doi: 10.1038/sj.jp.7211404.

[153] Okascharoen, C., Hui, C., Cairnie, J., Morris, A.M., and Kirpalani, H. External validation of bedside prediction score for diagnosis of late-onset neonatal sepsis. *Journal of Perinatology*, 27(8):496–501, 2007. doi: 10.1038/sj.jp.7211767.

[154] Husada, D., Chanthavanich, P., Chotigeat, U., Sunttarattiwong, P., Sirivichayakul, C., et al. Predictive model for bacterial late-onset neonatal sepsis in a tertiary care hospital in Thailand. *BMC Infectious Diseases*, 20(1):1–11, 2020. doi: 10.1186/s12879-020-4875-5.

[155] Kainth, D., Prakash, S., and Sankar, M.J. Diagnostic performance of machine learning-based models in neonatal sepsis: A systematic review. *Pediatric Infectious Disease Journal*, 43(9):889–901, 2024. doi: 10.1097/INF.0000000000004409.

[156] Mani, S., Ozdas, A., Aliferis, C., Varol, H.A., Chen, Q., et al. Medical decision support using machine learning for early detection of late-onset neonatal sepsis. *Journal of the American Medical Informatics Association: JAMIA*, 21(2):326–336, 2014. doi: 10.1136/amiajnl-2013-001854.

[157] Masino, A.J., Harris, M.C., Forsyth, D., Ostapenko, S., Srinivasan, L., et al. Machine learning models for early sepsis recognition in the neonatal intensive care unit using readily available electronic health record data. *PLOS ONE*, 14(2): e0212665, 2019. doi: 10.1371/journal.pone.0212665.

[158] Thakur, J., Pahuja, S.K., and Pahuja, R. Performance comparison of prediction models for neonatal sepsis using logistic regression, multiple discriminant analysis and artificial neural network. *Biomedical Physics & Engineering Express*, 5(3): 035013, 2019. doi: 10.1088/2057-1976/aaf677.

[159] Helguera-Repetto, A.C., Soto-Ramírez, M.D., Villavicencio-Carrisoza, O., Yong-Mendoza, S., Yong-Mendoza, A., et al. Neonatal sepsis diagnosis decision-making based on artificial neural networks. *Frontiers in Pediatrics*, 8:525, 2020. doi: 10.3389/fped.2020.00525.

[160] Stocker, M., Daunhawer, I., van Herk, W., el Helou, S., Dutta, S., et al. Machine learning used to compare the diagnostic accuracy of risk factors, clinical signs and biomarkers and to develop a new prediction model for neonatal early-onset sepsis. *The Pediatric Infectious Disease Journal*, 41(3):248, 2022. doi: 10.1097/INF.0000000000003344.

[161] Kato, K. Leucocytes in infancy and childhood. *The Journal of Pediatrics*, 7(1): 7–15, 1935. doi: 10.1016/S0022-3476(35)80143-1.

[162] Xanthou, M. Leucocyte blood picture in healthy full-term and premature babies during neonatal period. *Archives of Disease in Childhood*, 45(240):242–249, 1970. doi: 10.1136/adc.45.240.242.

[163] Prior, T. and Kumar, S. Mode of delivery has an independent impact on neonatal condition at birth. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 181:135–139, 2014. doi: 10.1016/j.ejogrb.2014.07.041.

[164] Kozlovski, T., Yochpaz, S., Shachar, I., Friedman, N., and Marom, R. Does the delivery mode affect post-birth neonatal serum C-reactive protein levels? A causal effect analysis. *Journal of Paediatrics and Child Health*, 58(8):1330–1336, 2022. doi: 10.1111/jpc.15975.

[165] Lai, C., Huang, L., Wang, Y., Huang, C., Luo, Y., et al. Effect of different delivery modes on intestinal microbiota and immune function of neonates. *Scientific Reports*, 14(1):17452, 2024. doi: 10.1038/s41598-024-68599-x.

[166] Carroll, P.D., Nankervis, C.A., Iams, J., and Kelleher, K. Umbilical cord blood as a replacement source for admission complete blood count in premature infants. *Journal of Perinatology*, 32(2):97–102, 2012. doi: 10.1038/jp.2011.60.

[167] Victory, R., Penava, D., Da Silva, O., Natale, R., and Richardson, B. Umbilical cord pH and base excess values in relation to adverse outcome events for infants delivering at term. *American Journal of Obstetrics and Gynecology*, 191(6): 2021–2028, 2004. doi: 10.1016/j.ajog.2004.04.026.

[168] Muddathir, A.R.M., Abdelgadir, R.E., Awdoon, A.Y., Allah, E.I.A., and Waggiallah, H.A. Hematological parameters reference range in Sudanese neonatal cord blood in normal and caesarian delivery. *Entomology and Applied Science Letters*, 8(4-2021): 59–63, 2021. doi: 10.51847/s9NYGrb3zH.

[169] Angelo, A., Derbie, G., Demtse, A., and Tsegaye, A. Umbilical cord blood hematological parameters reference interval for newborns from Addis Ababa, Ethiopia. *BMC Pediatrics*, 21:275, 2021. doi: 10.1186/s12887-021-02722-z.

[170] Abd El-Gaber Ali, A.E.N. and Mahmoud, A.M. Correlation between mode of delivery and newborn's cord blood composition. *Open Journal of Obstetrics and Gynecology*, 9(3):395–403, 2019. doi: 10.4236/ojog.2019.93040.

[171] Zitouni, S., Bouatrous, E., Laabidi, O., Boudrigua, I., Chaouachi, D., et al. Tunisian newborn's cord blood: Reference values of complete blood count and hemoglobin fractions. *American Journal of Perinatology*, 39:1241–1247, 2020. doi: 10.1055/s-0040-1722199.

[172] Sabnis, K., Ghanghurde, S., Shukla, A., Sukheja, D., and Rojekar, M.V. An Indian perspective for umbilical cord blood haematological parameters reference interval. *BMC Pediatrics*, 23(1):287, 2023. doi: 10.1186/s12887-023-04090-2.

[173] Manna, S., Dhar, I., Naskar, T.K., and Law, S. Cord blood haematological parameters reference range difference in urban and rural population of West Bengal as compared to global scenario. *Journal of Biomedical Research & Environmental Sciences*, 2(3):206–212, 2021. doi: 10.37871/JBRES1212.

[174] Meena, J., Charles, M.V.P., Ali, A., Ramakrishnan, S., Gosh, S., et al. Utility of cord blood culture in early onset neonatal sepsis. *The Australasian Medical Journal*, 8(8):263–267, 2015. doi: 10.4066/AMJ.2015.2460.

[175] Bensouda, B., Mandel, R., Altit, G., and Ali, N. Umbilical cord blood culture for early onset sepsis in preterm infants. *Pediatric Research*, 96(7):1765–1770, 2024. ISSN 1530-0447. doi: 10.1038/s41390-024-03183-7.

[176] Shakhawat Alam, M., Mannan, M., Kumer Dey, S., Raihanur Rahman, A.Z.M., and Shahidullah, M. Umbilical cord blood for the screening of early onset neonatal sepsis among those at risk of infection. *Scholars Journal of Applied Medical Sciences*, 9(3):452–461, 2021. doi: 10.36347/sjams.2021.v09i03.032.

[177] Christensen, R.D., Jensen, J., Maheshwari, A., and Henry, E. Reference ranges for blood concentrations of eosinophils and monocytes during the neonatal period defined from over 63 000 records in a multihospital health-care system. *Journal of Perinatology*, 30(8):540–545, 2010. doi: 10.1038/jp.2009.196.

[178] El Gendy, F., Allam, A., Allam, M., and Allam, R. Haematological parameters of newborns delivered vaginally versus caesarean section. *Menoufia Medical Journal*, 29(2):259–264, 2016. doi: 10.4103/1110-2098.192429.

[179] Gomes, A.C. and Gomes, M.S. Hematopoietic niches, erythropoiesis and anemia of chronic infection. *Experimental Hematology*, 44(2):85–91, 2016. doi: 10.1016/j.exphem.2015.11.007.

[180] Mughal, Z., Narayanan, A., Gupta, V., and Reay-Jones, N. Clinical need-directed blood tests: A step in saving the NHS? *Annals of Clinical Biochemistry: International Journal of Laboratory Medicine*, 53(5):568–574, 2016. doi: 10.1177/0004563215617782.

[181] Infant Journal. Neonatal unit guide, 2025. `https://www.infantjournal.co.uk/nicu_list.html`. Accessed on: 2025-08-28.

Appendix A1

The Diagnosis Accuracy of the Complete Blood Count for Neonatal Sepsis: A Systematic Review and Meta-Analysis

---

A1.1 Search Strategies

MEDLINE and Embase via Ovid

1. Exp Neonatal Sepsis/
2. (Neonatal adj2 infection).mp.
3. 1 OR 2
4. Exp Leukocytes/
5. Exp Blood cell count/
6. Exp Neutrophils/
7. Exp Lymphocytes/
8. Exp Basophils/
9. Exp Monocytes/
10. Exp Macrophages/
11. Exp Eosinophils/
12. 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11
13. Exp "Sensitivity and Specificity"/
14. (Sensitivity).mp.
15. (Specificity).mp.
16. (Accuracy).mp.
17. 13 OR 14 OR 15 OR 16
18. 3 AND 12 AND 17

Cochrane

1. (MH “Neonatal Sepsis”)
2. Neonatal infection
3. 1 OR 2
4. (MH “Leukocytes”)
5. (MH “Blood cell count")
6. (MH “Neutrophils")
7. (MH “Lymphocytes")
8. (MH “Basophils")
9. (MH “Monocytes")
10. (MH “Macrophages")
11. (MH “Eosinophils")
12. 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11
13. (MH “Sensitivity and Specificity”)
14. Sensitivity
15. Specificity
16. Accuracy
17. 13 OR 14 OR 15 OR 16
18. 3 AND 12 AND 17

CINAHL

1. MeSH descriptor: [Neonatal Sepsis] explode all trees
2. (Neonatal infection)
3. 1 OR 2
4. MeSH descriptor: [Leukocytes] explode all trees
5. MeSH descriptor: [Blood cell count] explode all trees
6. MeSH descriptor: [Neutrophils] explode all trees

7. MeSH descriptor: [Lymphocytes] explode all trees
8. MeSH descriptor: [Basophils] explode all trees
9. MeSH descriptor: [Monocytes] explode all trees
10. MeSH descriptor: [Macrophages] explode all trees
11. MeSH descriptor: [Eosinophils] explode all trees
12. 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11
13. MeSH descriptor: [Sensitivity and Specificity] explode all trees
14. (Sensitivity)
15. (Specificity)
16. (Accuracy)
17. 13 OR 14 OR 15 OR 16
18. 3 AND 12 AND 17

A1.2 Level of Applicability

Table A1.1: Level of applicability of the studies included in the systematic review and meta-analysis, adapted from QUADAS-2 [83], the CASP diagnostic checklist [84], and the Newcastle-Ottawa scale [85].

Study No	Is there evidence of a sample size/power calculation?	Were patients enrolled consecutively or randomly?	Were the exclusion criteria adequate and clearly described?	Is the case definition adequate and clearly described?	Is the control definition adequate and clearly described?	Were controls from the same population as cases?	Were the reference and index tests conducted at the same time?	Was the method for performing the index test described?	Total
#1	Y	Y	Y	Y	Y	Y	Y	Y	8.0
#2	N	U	Y	Y	Y	Y	Y	N	5.5
#3	Y	U	Y	Y	Y	Y	U	N	6.0
#4	N	Y	N	Y	Y	Y	Y	Y	6.0
#5	N	Y	Y	Y	Y	Y	Y	Y	7.0
#6	N	Y	Y	Y	Y	Y	U	Y	6.5
#7	N	Y	N	Y	Y	Y	U	Y	5.5
#8	N	Y	Y	Y	U	Y	Y	Y	6.5
#9	Y	Y	Y	Y	N	Y	Y	Y	7.0
#10	Y	Y	Y	Y	Y	Y	Y	Y	8.0
#11	N	Y	Y	Y	Y	Y	U	N	5.5

Continued on next page

Continued from last page

Study No	Is there evidence of a sample size/power calculation?	Were patients enrolled consecutively or randomly?	Were the exclusion criteria adequate and clearly described?	Is the case definition adequate and clearly described?	Is the control definition adequate and clearly described?	Were controls from the same population as cases?	Were the reference and index tests conducted at the same time?	Was the method for performing the index test described?	Total
#12	N	Y	Y	Y	Y	Y	Y	N	6.0
#13	Y	U	N	N	Y	Y	U	Y	5.0
#14	Y	Y	Y	Y	Y	U	U	Y	7.0
#15	N	U	N	N	Y	U	Y	Y	4.0
#16	N	Y	Y	Y	Y	Y	Y	Y	7.0
#17	N	Y	N	Y	Y	Y	Y	Y	6.0
#18	N	Y	Y	Y	Y	Y	Y	Y	7.0
#19	N	Y	Y	Y	Y	Y	Y	N	6.0
#20	N	U	Y	Y	Y	Y	Y	Y	6.5
#21	N	Y	Y	Y	Y	Y	U	N	5.5
#22	N	Y	N	Y	Y	Y	Y	N	5.0
#23	N	Y	Y	Y	Y	Y	U	Y	6.5
#24	N	U	Y	Y	Y	Y	Y	N	5.5

A1.3 Study Values

Table A1.2: Diagnostic accuracy values for parameters of the complete blood count for neonatal sepsis, from studies included in the systematic review and meta-analysis.

Study No	Cases[a]	Controls[a]	Optimal Threshold	Sens (%)	Spec (%)	PPV (%)	NPV (%)	LR^+	LR^-	TP	TN	FP	FN	AUC	Youden Index
White blood cell count (x10^9/L)															
#1	–	–	< 5.00	91.4	63.0	13.3	99.2	2.50	0.10	35	227	133	3	–	0.54
#2	11.00	7.50	< 4.00 > 12.00	72.0	63.0	66.0	69.0	–	–	18	14	8	7	0.670	0.35
#5	15.85	12.76	> 19.60	33.3	95.7	63.2	78.8	–	–	13	106	5	27	0.590	0.29
#7	15.55	8.50	> 10.50[c]	64.0	92.0	71.9	88.9	–	–	16	23	2	9	0.710	0.56
#11	16.54	10.89	–	69.3	90.7	82.8	82.9	–	–	70	87	9	31	–	0.60
#15	11.50	19.70	< 5.00	24.0	97.0	63.0	89.0	–	–	5	39	1	16	–	0.21
#15	–	–	< 10.00	57.0	76.0	28.0	92.0	–	–	12	30	10	9	–	0.33
#15	–	–	> 20.00	10.0	86.0	10.0	85.0	–	–	2	34	6	19	–	-0.04
#22	11.19	9.02	> 8.99	63.6	70.7	70.0	64.4	2.17	0.51	38	42	18	22	0.671	0.34
#23	13.80	9.90	< 5.00 > 20.00	48.0	77.0	67.0	62.0	2.10	0.68	28	23	7	30	–	0.25

Continued on next page

Continued from last page

Study No	Cases[a]	Controls[a]	Optimal Threshold	Sens (%)	Spec (%)	PPV (%)	NPV (%)	LR^+	LR^-	TP	TN	FP	FN	AUC	Youden Index
Neutrophil count (x10^9/L)															
#1	–	–	< 1.50	85.7	61.5	7.5	99.2	2.20	0.20	33	221	139	5	–	0.47
#5	7.76	6.51	> 10.08	37.5	90.3	50.0	78.4	–	–	15	100	11	25	0.542	0.28
#7	15.30	2.94	> 4.13[c]	76.0	80.0	76.9	79.2	–	–	19	20	5	6	0.816	0.56
#15	6.60	15.40	< 8.00	76.0	67.0	28.0	94.0	–	–	16	27	13	5	–	0.43
#23	6.90	4.90	< 2.00 > 7.50	55.0	75.0	67.0	64.0	2.10	0.61	32	23	8	26	–	0.30
Immature-to-total neutrophil ratio															
#1	–	–	> 0.20	81.8	60.8	3.8	99.4	2.10	0.30	31	219	141	7	–	0.43
#3	0.33	0.10	> 0.22	54.3	98.1	96.6	68.2	12.50	–	36	90	2	31	0.955	0.52
#5	0.25	0.11	> 0.19	62.5	92.5	79.4	88.9	–	–	25	103	8	15	0.815	0.55
#7	–	–	> 0.20[c]	80.0	100.0	88.3	100.0	–	–	20	25	0	5	0.902	0.80
#15	0.18	0.07	> 0.14	63.0	72.0	26.0	93.0	–	–	13	29	11	8	–	0.35
#15	–	–	> 0.16	47.0	82.0	29.0	91.0	–	–	10	33	7	11	–	0.29
#19	0.33	0.09	> 0.20	75.0	76.0	–	–	–	–	18	42	13	6	0.903	0.51
#20	0.13	0.07	> 0.13	53.3	95.6	92.3	67.2	–	–	24	43	2	21	0.780	0.49
#23	0.30	0.12	> 0.20	76.0	87.0	85.0	79.0	5.80	0.28	44	26	4	14	0.800	0.63

Continued on next page

Continued from last page

Study No	Cases[a]	Controls[a]	Optimal Threshold	Sens (%)	Spec (%)	PPV (%)	NPV (%)	LR^+	LR^-	TP	TN	FP	FN	AUC	Youden Index
Neutrophil-to-lymphocyte ratio															
#8	3.20	1.40	> 1.57	68.0	82.0	–	–	–	–	52	33	7	24	0.723	0.50
#14[b]	0.80	0.08	> 0.10	67.0	99.0	98.0	75.0	–	–	54	79	1	26	0.790	0.66
#17	3.88	2.34	> 1.70	68.3	46.2	0.0	64.9	–	–	28	24	28	13	0.623	0.15
#21	2.10	0.70	> 1.37	63.3	79.7	75.0	69.4	–	–	49	61	16	28	0.764	0.43
#24	4.16	2.35	> 3.17[c]	77.0	78.0	85.0	75.0	–	–	57	39	11	17	0.788	0.55
Delta neutrophil index															
#4	1.50	0.10	> 0.65	96.2	97.4	–	–	–	–	74	128	3	3	–	0.94
#6	20.80	1.40	> 4.60	85.0	80.0	87.0	77.0	–	–	94	70	17	17	0.880	0.65
Mean neutrophil volume (au)															
#1	162	147	> 151	71.3	71.9	62.9	79.0	2.50	0.40	27	259	101	11	0.783	0.43
#5	153	145	> 160	37.5	94.6	71.4	80.8	–	–	15	105	6	25	0.634	0.32
#16	180	150	> 158[c]	97.0	96.0	92.0	98.0	–	–	81	154	6	3	0.990	0.93
Mean neutrophil conductivity (au)															
#5	138	142	< 145	78.4	46.8	35.2	86.7	–	–	31	52	59	9	0.626	0.25
#16	142	137	> 138	63.0	57.0	43.0	74.0	1.46	0.64	53	91	69	31	0.590	0.20

Continued on next page

Continued from last page

Study No	Cases[a]	Controls[a]	Optimal Threshold	Sens (%)	Spec (%)	PPV (%)	NPV (%)	LR^+	LR^-	TP	TN	FP	FN	AUC	Youden Index
Mean neutrophil scatter (au)															
#5	134	138	< 142	86.5	37.6	32.7	87.2	–	–	35	42	69	5	0.645	0.24
#16	120	126	< 120	47.0	32.0	26.0	53.0	0.69	1.65	39	51	109	45	0.280	-0.21
Platelet count (x10^9/L)															
#1	–	–	< 150	63.2	65.6	26.7	90.0	1.80	0.60	24	236	124	14	–	0.29
#7	214	289	< 165	92.0	68.0	89.5	74.2	–	–	23	17	8	2	0.893	0.60
#9	109	400	< 278[c]	85.9	90.9	–	–	–	–	55	40	4	9	0.920	0.77
#15	135	251	< 150	62.0	71.0	26.0	92.0	–	–	13	28	12	8	–	0.33
#18	187	285	< 150	41.5	96.2	89.5	67.6	–	–	18	52	2	25	0.797	0.38
#23	220	242	< 150	41.0	87.0	50.0	52.0	1.10	0.99	24	26	4	34	–	0.28
Mean platelet volume (fL)															
#9	9.7	8.9	> 9.2	62.5	72.7	–	–	–	–	40	32	12	24	0.679	0.35
#10	9.6	8.6	> 9.5	56.0	85.0	48.3	88.5	–	–	36	60	11	28	0.734	0.41
#12	9.3	8.9	> 8.5	81.7	51.2	–	–	–	–	51	32	31	12	0.600	0.33
#13	10.3	9.3	> 10.0	65.3	75.0	31.8	27.8	2.60	0.47	13	15	5	7	0.730	0.40
#18	10.0	9.2	> 9.0	63.4	53.8	52.0	65.1	–	–	27	29	25	16	0.641	0.17
#20	12.0	9.4	> 11.6[c]	77.8	97.8	97.2	81.5	–	–	35	44	1	10	0.890	0.76

Continued on next page

Continued from last page

Study No	Cases[a]	Controls[a]	Optimal Threshold	Sens (%)	Spec (%)	PPV (%)	NPV (%)	LR^+	LR^-	TP	TN	FP	FN	AUC	Youden Index
Platelet-to-lymphocyte ratio															
#3	62.40	15.30	> 57.70[c]	91.3	97.6	97.4	91.8	9.2	–	61	90	2	6	0.847	0.89
#14	15.00	5.90	> 7.00	70.0	73.0	72.0	71.0	–	–	56	58	22	24	0.780	0.43
#24	99.57	71.53	> 90.85	64.9	80.0	79.0	76.9	–	–	48	40	10	26	0.699	0.45

[a] *Mean or median.* [b] *Excluded from meta-analysis.* [c] *Optimal threshold according to Youden Index. AUC values > 0.8 and sensitivities/specificities $>$ 85% are highlighted in green, suggesting strong diagnostic performance. Abbreviations: AUC – area under the curve; FN – false negative; FP – false positive; LR^+ – LR^+; LR^- – LR^-; NPV – negative predictive value; PPV – positive predictive value; Sens – sensitivity; Spec – specificity; TN – true negative; TP – true positive.*

Appendix A2

Establishing a Baseline Profile for the Neonatal Complete Blood Count in a Cohort of Neonates Admitted to a Neonatal Intensive Care Unit

---

A2.1 Validation of the XN-450 Point-of-Care Analyser

Rationale

Before commencing data collection, a Sysmex (Kobe, Japan) XN-450 haematology analyser was installed in the John Radcliffe Hospital (Oxford, UK) NICU for use as a point-of-care device. To ensure that CBC results obtained from the XN-450 were comparable with those generated by the laboratory-based Sysmex XN-2000 analysers, a formal validation study was performed. This validation was necessary to support both routine clinical care in the NICU and the integrity of research data collected for this study.

Methods

Fifty adult blood samples were analysed on both the XN-2000 and XN-450 within 30 minutes to minimise the risk of sample degradation. Twenty-five routinely reported CBC parameters were compared between the analysers. Depending on the underlying distribution of each parameter, comparisons were made using paired Student's t-tests or Wilcoxon rank sum tests.

Correlations were assessed using Spearman's rank correlation coefficients, and agreement between analysers was further evaluated through linear regression analyses. Bland-Altman plots, box plots, and scatter plots were generated to provide a visual assessment of differences. The same analyses were subsequently replicated on 325 neonatal CBC results.

To evaluate the precision of the XN-450, intra-run variability was assessed by analysing ten consecutive quality control samples, while inter-run variability was examined across 52 quality control samples obtained over a three-month period. Variability was evaluated using chi-squared tests; results were interpreted with reference to the stability of measurement distributions across runs.

Results

The evaluation of WBC is presented in Figure A2.1 as an example of these analyses. Figures were produced for all 25 parameters, but are not shown here.

In the adult validation set, only three parameters (MCV, MCHC, and IG%) showed a statistically significant difference between analysers (Table A2.1). Of these, only the difference in MCHC values reached a magnitude considered clinically relevant. In the neonatal dataset, 18 parameters differed significantly between analysers, of which five demonstrated large effect sizes.

Despite these differences, all CBC parameters showed statistically significant correlations between the XN-450 and XN-2000, indicating strong overall agreement between analyser outputs. Visual inspection of Bland–Altman plots and other graphical analyses confirmed that the differences observed were small and unlikely to affect clinical interpretation.

Both inter- and intra-run variability were minimal across all 25 parameters (all p-values $>$ 0.6), indicating stable performance of the XN-450 throughout the study period.

Conclusion

The validation confirmed that the Sysmex XN-450 haematology analyser produced results that were highly comparable to the laboratory-based XN-2000, with only minor differences of limited clinical significance. The analyser also demonstrated stable performance with minimal intra- and inter-run variability. These findings support the reliability of the XN-450 for point-of-care use in the NICU and justify its use for both routine clinical care and research data collection in this study.

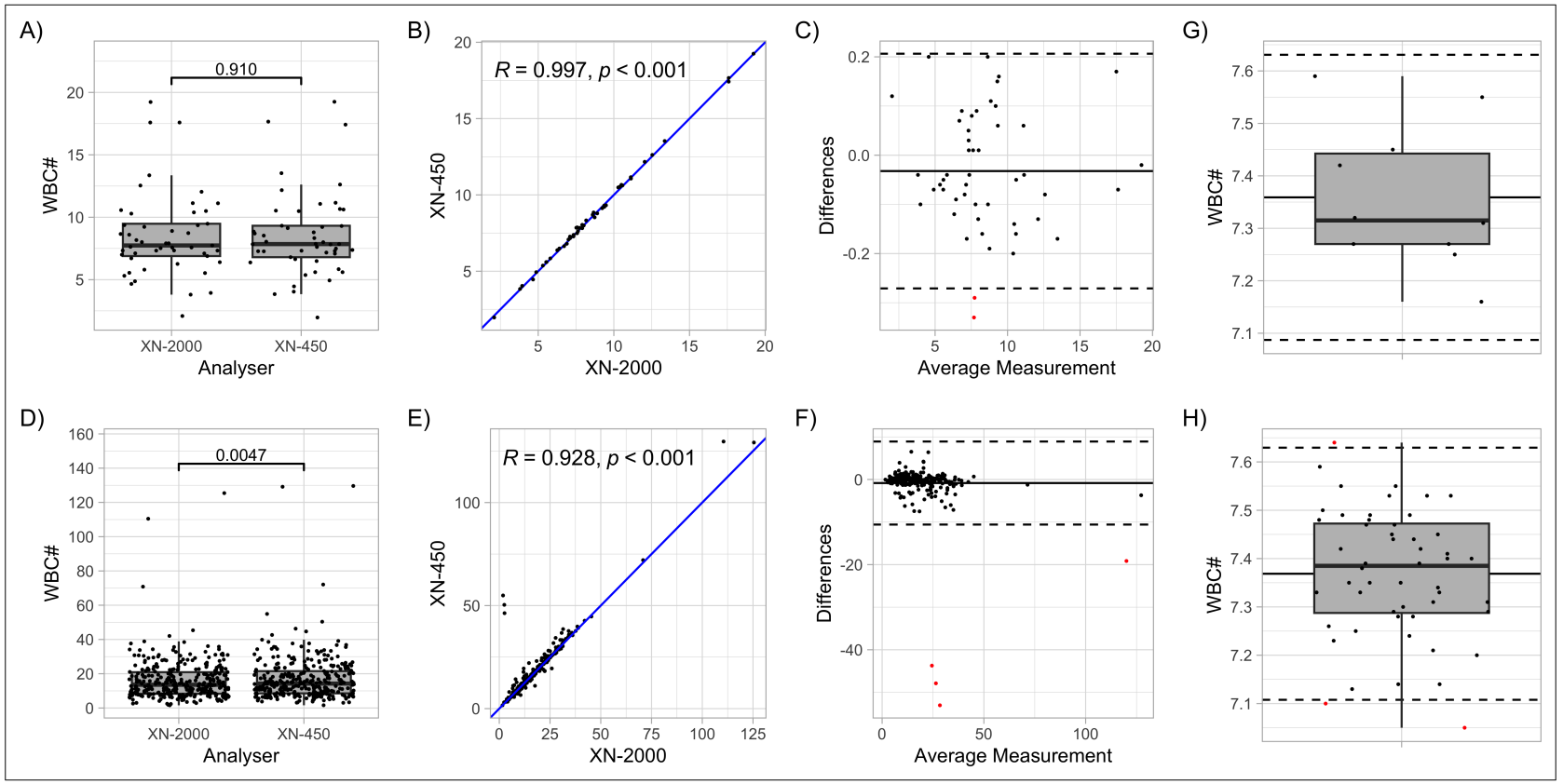


Figure A2.1: Validation of white blood cell count (WBC#; x 10^9/L) on a Sysmex XN-450 haematology analyser against an XN-2000 analyser in the haematology laboratory, for use as a point-of-care device in a Neonatal Intensive Care Unit.

A/D) Box plots of values from (A) adult and (D) neonatal blood samples. **B/E)** Scatter plots of values from (B) adult and (E) neonatal blood samples, with Spearman's correlation coefficients and a line of best fit (blue). **C/F)** Bland-Altman plots comparing values from (C) adult and (F) neonatal blood samples, with a solid line representing the mean bias and dashed lines showing the 95% limits of agreement. **G)** Box plot of ten consecutive quality control measurements, illustrating intra-run variability. **H)** Box plot of quality control measurements obtained over a three-month period illustrating inter-run variability.

Table A2.1: Validation of a Sysmex XN-450 haematology analyser for use as a point-of-care device in a Neonatal Intensive Care Unit, including comparisons of adult and neonatal complete blood count results, and inter- and intra-run quality control results.

	Adult CBC Results					Neonatal CBC Results					Consecutive QC ChiSq P-Value	Cumulative QC ChiSq P-Value
Parameter	P-Value	Effect	Corr P-Value	Corr R	Coeff	P-Value	Effect	Corr P-Value	Corr R	Coeff		
Red blood cell parameters												
RBC#	0.87	-0.02	<0.001	0.996	0.99	0.75	-0.02	<0.001	0.940	0.81	1.00	1.00
HGB	0.92	0.01	<0.001	0.997	1.01	0.76	0.02	<0.001	0.951	0.90	1.00	1.00
HCT	0.11	-0.23	<0.001	0.990	0.96	0.19	-0.07	<0.001	0.949	0.88	1.00	1.00
MCV	0.02	-0.35	<0.001	0.991	0.95	<0.001	-0.24	<0.001	0.963	0.99	1.00	1.00
MCH	0.82	0.03	<0.001	0.989	0.98	0.10	0.09	<0.001	0.958	0.98	1.00	1.00
MCHC	<0.001	0.99	<0.001	0.894	1.00	<0.001	0.34	<0.001	0.802	0.87	1.00	1.00
RDW-SD	0.07	-0.27	<0.001	0.990	0.91	<0.001	-0.46	<0.001	0.966	0.97	1.00	1.00
RDW-CV	0.89	0.02	<0.001	0.995	0.98	0.33	-0.05	<0.001	0.961	0.94	1.00	1.00
White blood cell parameters												
WBC#	0.91	-0.02	<0.001	0.997	1.00	0.005	-0.16	<0.001	0.928	0.85	1.00	1.00
NEUT#	0.72	-0.05	<0.001	0.999	0.99	0.005	-0.16	<0.001	0.994	0.91	1.00	1.00
NEUT%	0.75	-0.05	<0.001	0.995	1.02	<0.001	0.43	<0.001	0.920	0.88	1.00	1.00
LYMPH#	0.81	0.03	<0.001	0.992	1.01	0.90	0.01	<0.001	0.734	0.17	1.00	1.00
LYMPH%	0.76	0.04	<0.001	0.994	1.02	<0.001	0.37	<0.001	0.813	0.77	1.00	1.00

Continued on next page

Continued from last page

	Adult CBC Results					Neonatal CBC Results					Consecutive QC ChiSq P-Value	Cumulative QC ChiSq P-Value
Parameter	P-Value	Effect	Corr P-Value	Corr R	Coeff	P-Value	Effect	Corr P-Value	Corr R	Coeff		
MONO#	0.51	0.10	<0.001	0.894	1.01	0.01	-0.14	<0.001	0.918	0.80	1.00	1.00
MONO%	0.34	0.14	<0.001	0.866	0.95	0.01	0.14	<0.001	0.747	0.64	1.00	1.00
EOSI#	0.94	0.01	<0.001	0.970	0.98	0.33	0.05	<0.001	0.938	1.12	1.00	1.00
EOSI%	0.97	0.00	<0.001	0.976	0.97	<0.001	0.39	<0.001	0.917	1.05	1.00	1.00
BASO#	0.39	-0.12	<0.001	0.648	0.74	<0.001	-0.76	<0.001	0.512	0.02	1.00	1.00
BASO%	0.28	-0.16	<0.001	0.544	0.56	<0.001	0.62	<0.001	0.405	0.08	1.00	1.00
IG#	0.05	0.29	<0.001	0.807	1.11	<0.001	0.22	<0.001	0.948	0.90	1.00	1.00
IG%	0.02	0.35	<0.001	0.756	1.03	<0.001	0.35	<0.001	0.905	0.95	1.00	1.00
Platelet parameters												
PLT#	0.19	0.19	<0.001	0.949	1.01	0.03	0.12	<0.001	0.977	1.03	0.81	0.66
MPV	0.55	-0.09	<0.001	0.969	0.94	<0.001	0.57	<0.001	0.846	0.84	1.00	1.00
P-LCR	0.67	-0.06	<0.001	0.969	0.96	<0.001	0.81	<0.001	0.911	0.93	1.00	1.00
PCT	0.96	0.01	<0.001	0.980	1.02	<0.001	0.43	<0.001	0.954	1.03	1.00	1.00

Effect refers to Cohen's d or Wilcoxon r effect size. Significant p-values (< 0.05) and effect sizes (absolute Cohen's $d > 0.50$ or Wilcoxon $r > 0.3$) are highlighted in red. Abbreviations: CBC – complete blood count; ChiSq – chi-square test; Coeff – linear regression coefficient; Corr – spearman correlation; QC – quality control. For parameter abbreviations, see List of Complete Blood Count Parameter Abbreviations (p. xxi).

A2.2 Clinical Observation Charts

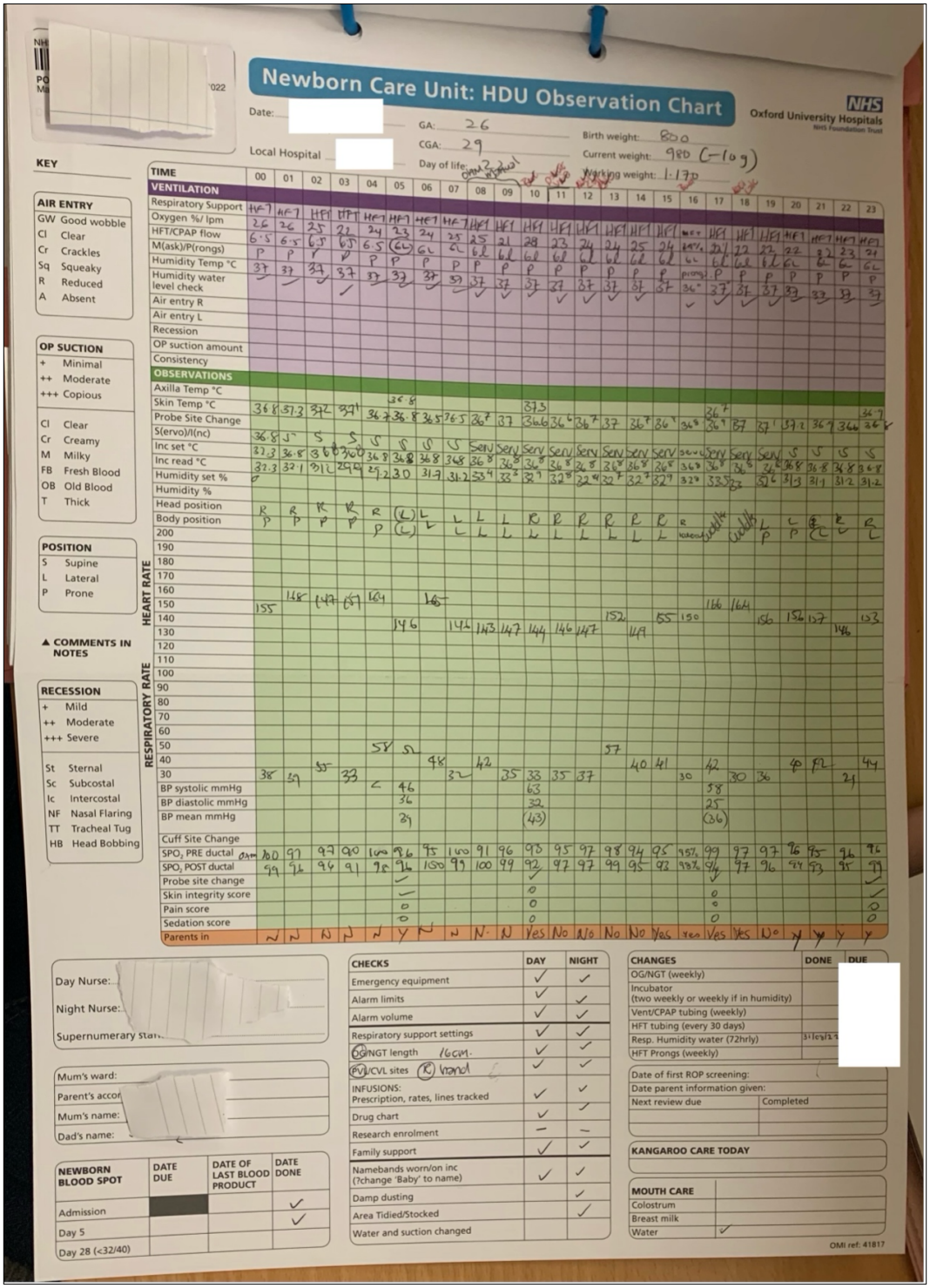


Figure A2.2: Example of daily observation chart from paper clinical records. Summarised clinical data were extracted from these records for each day of a neonate's stay in the Neonatal Intensive Care Unit, including the highest and lowest heart rate, the highest respiratory rate, and the lowest blood pressure.

A2.3 List of Surgical Interventions

Abdominal and Gastrointestinal

Small/large bowel resection, duodenal/jejunal/colonic atresia repair, gastroschisis repair, Ladd's procedure, exploratory laparotomy, relaparotomy, adhesiolysis, ileostomy, jejunostomy, drainage of abdominal abscess, stoma formation/closure/refashioning, appendicectomy, closure of bowel/stomach perforation, PEG insertion, liver/colon biopsy

Thoracic, Spinal, and Neurosurgical

Thoracoscopy/thoracotomy (with or without biopsy/lobectomy), diaphragmatic hernia repair (primary/prosthetic), plication of diaphragm, tracheostomy, tracheostomy tube replacement, needle thoracocentesis, spinal meningocele closure, ventricular shunt insertion, endoscopic third ventricle exteriorisation

Genital and Urinary

Cystoscopy, urethroscopy, urinary diversion, resection of posterior urethral valves, bladder exstrophy closure

Miscellaneous

Wound debridement, abscess drainage, percutaneous muscle biopsy, sacrococcygeal teratoma excision, anal dilation, amputation of an accessory finger, EUA of the rectum, Hickman line insertion, oesophageal/tracheoesophageal repairs

Appendix A3

Evaluation of the Diagnostic Accuracy of the Complete Blood Count for Late-Onset Neonatal Sepsis

---

A3.1 Evaluation of the Training Dataset

Table A3.1: Evaluation of the diagnostic accuracy of the neonatal complete blood count for late-onset sepsis using receiver operating characteristic curve analysis of the training dataset.

Parameter	*Unlikely* Average	*Probable* Average	P-Value	Effect	AUC	Cut-Off	Sens (%)	Spec (%)	TP	TN	FP	FN	PPV (%)	NPV (%)	LR^+	LR^-	DOR	Youden Index
Red blood cell parameters																		
RBC#	3.44	3.62	0.02	-0.25	0.588	3.23	77.6	45.6	177	73	87	51	67.0	58.9	1.43	0.49	2.91	0.23
HGB[a]	109.89	117.00	0.02	0.12	0.569	103.89	80.7	44.4	184	71	89	44	67.4	61.7	1.45	0.43	3.34	0.25
HCT[a]	30.60	32.90	0.02	0.12	0.572	28.20	87.7	37.5	200	60	100	28	66.7	68.2	1.40	0.33	4.29	0.25
MCV	91.89	90.70	0.07	0.19	0.552	96.35	75.4	35.0	172	56	104	56	62.3	50.0	1.16	0.70	1.65	0.10
MCH	33.30	32.86	0.09	0.18	0.545	36.83	92.1	21.9	210	35	125	18	62.7	66.0	1.18	0.36	3.27	0.14
MCHC	360.06	358.97	0.34	0.10	0.530	368.25	76.8	36.3	175	58	102	53	63.2	52.3	1.20	0.64	1.88	0.13
RDW-SD	54.96	58.84	< 0.001	-0.46	0.624	52.68	80.3	40.0	183	64	96	45	65.6	58.7	1.34	0.49	2.71	0.20
RDW-CV	16.50	18.50	< 0.001	-0.74	0.710	16.90	75.4	58.8	172	94	66	56	72.3	62.7	1.83	0.42	4.37	0.34
NRBC#[a]	0.06	0.30	< 0.001	0.33	0.694	0.17	61.0	69.4	139	111	49	89	73.9	55.5	1.99	0.56	3.54	0.30
NRBC%[a]	0.60	1.88	< 0.001	0.26	0.650	0.60	73.3	51.3	167	82	78	61	68.2	57.3	1.50	0.52	2.88	0.25
RET#	0.10	0.08	0.005	0.30	0.586	0.13	80.2	39.6	182	63	96	45	65.5	58.3	1.33	0.50	2.65	0.20
RET%	3.04	2.32	< 0.001	0.37	0.602	4.31	87.7	34.0	199	54	105	28	65.5	65.9	1.33	0.36	3.66	0.22
IRF	33.06	30.49	0.09	0.18	0.542	16.74	20.7	91.2	47	145	14	180	77.0	44.6	2.35	0.87	2.70	0.12
RBC-HE	30.61	29.72	< 0.001	0.40	0.626	31.00	74.9	50.9	170	81	78	57	68.5	58.7	1.53	0.49	3.10	0.26
RET-HE	29.46	26.87	< 0.001	0.72	0.703	29.58	77.5	56.6	176	90	69	51	71.8	63.8	1.79	0.40	4.50	0.34
DELTA-HE	-1.65	-2.86	< 0.001	0.39	0.615	-2.43	56.4	64.8	128	103	56	99	69.6	51.0	1.60	0.67	2.38	0.21

Continued on next page

Continued from last page

Parameter	*Unlikely* Average	*Probable* Average	P-Value	Effect	AUC	Cut-Off	Sens (%)	Spec (%)	TP	TN	FP	FN	PPV (%)	NPV (%)	LR^+	LR^-	DOR	Youden Index
HYPO-HE	1.62	2.30	< 0.001	-0.42	0.619	1.43	76.7	45.3	174	72	87	53	66.7	57.6	1.40	0.52	2.72	0.22
HYPER-HE	0.90	0.91	0.92	-0.01	0.521	0.77	64.3	46.5	146	74	85	81	63.2	47.7	1.20	0.77	1.57	0.11
MICRO-RBC	3.60	4.66	< 0.001	-0.40	0.630	4.27	62.3	66.9	142	107	53	86	72.8	55.4	1.88	0.56	3.33	0.29
MACRO-RBC	4.81	5.53	0.009	-0.28	0.582	4.32	73.7	46.9	168	75	85	60	66.4	55.6	1.39	0.56	2.47	0.21
White blood cell parameters																		
WBC#	10.44	14.43	< 0.001	-0.79	0.721	15.26	50.4	88.8	115	142	18	113	86.5	55.7	4.48	0.56	8.03	0.39
NEUT#[a]	3.74	8.35	< 0.001	0.50	0.796	5.46	70.6	78.8	161	126	34	67	82.6	65.3	3.32	0.37	8.90	0.49
NEUT%	36.00	54.16	< 0.001	-1.17	0.806	45.45	72.4	77.5	165	124	36	63	82.1	66.3	3.22	0.36	9.02	0.50
LYMPH#	4.17	3.63	< 0.001	0.59	0.659	3.84	62.1	64.4	141	103	57	86	71.2	54.5	1.74	0.59	2.96	0.26
LYMPH%	43.24	23.91	< 0.001	1.34	0.836	30.49	71.8	85.0	163	136	24	64	87.2	68.0	4.79	0.33	14.44	0.57
MONO#	1.59	2.30	< 0.001	-0.67	0.687	2.70	43.9	90.0	100	144	16	128	86.2	52.9	4.39	0.62	7.03	0.34
MONO%	15.32	15.54	0.40	-0.09	0.523	20.99	18.9	91.3	43	146	14	185	75.4	44.1	2.16	0.89	2.42	0.10
EOSI#[a]	0.40	0.29	0.01	0.13	0.576	0.20	42.1	85.0	96	136	24	132	80.0	50.7	2.81	0.68	4.12	0.27
EOSI%	4.13	2.49	< 0.001	0.51	0.658	2.09	50.0	84.4	114	135	25	114	82.0	54.2	3.20	0.59	5.40	0.34
BASO#[a]	0.08	0.11	< 0.001	0.24	0.640	0.06	85.5	40.0	195	64	96	33	67.0	66.0	1.43	0.36	3.94	0.26
BASO%	0.60	0.66	0.002	-0.32	0.591	0.74	34.2	80.6	77	129	31	151	71.3	46.1	1.77	0.82	2.16	0.15
IG#	0.09	0.27	< 0.001	-0.92	0.745	0.19	64.0	76.3	146	122	38	82	79.3	59.8	2.70	0.47	5.72	0.40
IG%	1.00	2.01	< 0.001	-0.80	0.723	1.39	69.3	68.8	158	110	50	70	76.0	61.1	2.22	0.45	4.97	0.38
NLR	0.98	2.40	< 0.001	-1.28	0.825	1.46	76.7	76.3	174	122	38	53	82.1	69.7	3.23	0.31	10.54	0.53

Continued on next page

Continued from last page

Parameter	*Unlikely* Average	*Probable* Average	P-Value	Effect	AUC	Cut-Off	Sens (%)	Spec (%)	TP	TN	FP	FN	PPV (%)	NPV (%)	LR^+	LR^-	DOR	Youden Index
ITR	0.03	0.04	< 0.001	-0.37	0.611	0.06	36.4	83.8	83	134	26	145	76.1	48.0	2.24	0.76	2.95	0.20
NEUT-RI	43.23	49.65	< 0.001	-1.25	0.812	46.12	75.0	74.4	171	119	41	57	80.7	67.6	2.93	0.34	8.71	0.49
NEUT-GI	147.73	145.93	< 0.001	0.34	0.598	144.15	40.4	76.9	92	123	37	136	71.3	47.5	1.75	0.78	2.25	0.17
NEUT-TW	31.00	42.37	< 0.001	-1.17	0.811	37.34	68.4	81.9	156	131	29	72	84.3	64.5	3.78	0.39	9.79	0.50
RE-LYMPH	9.89	13.40	< 0.001	-0.61	0.675	14.20	49.6	77.5	113	124	36	115	75.8	51.9	2.20	0.65	3.38	0.27
AS-LYMPH[a]	0.42	0.90	< 0.001	0.26	0.649	0.97	32.5	93.8	74	150	10	154	88.1	49.3	5.19	0.72	7.21	0.26
HFLC[a]	0.00	0.00	< 0.001	0.33	0.696	0.00	41.0	91.9	95	144	16	132	85.6	52.2	5.05	0.64	7.85	0.33
RE-MONO[a]	0.34	0.78	< 0.001	0.28	0.667	0.89	47.4	82.5	108	132	28	120	79.4	52.4	2.71	0.64	4.24	0.30
Platelet parameters																		
PLT#	370.04	211.45	< 0.001	0.87	0.731	256.76	65.4	75.6	149	121	39	79	79.3	60.5	2.68	0.46	5.85	0.41
PDW	13.30	15.59	< 0.001	-0.55	0.656	13.73	71.0	57.0	98	81	61	40	61.6	66.9	1.65	0.51	3.25	0.28
MPV	11.49	12.22	< 0.001	-0.47	0.640	11.53	69.6	54.9	96	78	64	42	60.0	65.0	1.54	0.55	2.79	0.25
P-LCR	35.46	41.60	< 0.001	-0.51	0.646	35.15	73.9	50.7	102	72	70	36	59.3	66.7	1.50	0.51	2.91	0.25
PCT	0.44	0.35	< 0.001	0.52	0.633	0.32	44.9	80.3	62	114	28	76	68.9	60.0	2.28	0.69	3.32	0.25
IPF	7.52	14.38	< 0.001	-0.91	0.740	7.51	84.6	52.8	192	84	75	35	71.9	70.6	1.79	0.29	6.14	0.37
PLR	0.08	0.06	< 0.001	0.40	0.610	0.06	52.0	70.0	118	112	48	109	71.1	50.7	1.73	0.69	2.53	0.22

Effect refers to Cohen's d or Wilcoxon r effect size. All values were calculated parametrically unless transformation failed to normalise values ([a]). Average refers to the mean or median, depending on the underlying distribution. Significant p-values (< 0.05) and effect sizes (absolute Cohen's $d > 0.50$ and Wilcoxon $r > 0.3$) are highlighted in red. AUC values > 0.8 and sensitivities/specificities $> 85\%$ are highlighted in green, suggesting strong diagnostic performance. Abbreviations: AUC – area under the curve; DOR – diagnostic odds ratio; FN – false negative; FP – false positive; LR^+ – positive likelihood ratio; LR^- – negative likelihood ratio; NPV – negative predictive value; PPV – positive predictive value; Sens – sensitivity; Spec – specificity; TN – true negative; TP – true positive. For parameter abbreviations, see List of Complete Blood Count Parameter Abbreviations (p. xxi).

A3.2 Evaluation of the Testing Dataset

Table A3.2: Validation of the diagnostic accuracy of the neonatal complete blood count for late-onset sepsis using receiver operating characteristic curve analysis of the testing dataset.

Parameter	*Unlikely* Average	*Probable* Average	P-Value	Effect	AUC	Sens (%)	Spec (%)	TP	TN	FP	FN	PPV (%)	NPV (%)	LR^+	LR^-	DOR	Youden Index
Red blood cell parameters																	
RBC#	3.60	3.59	0.70	0.04	0.502	73.0	33.7	197	58	114	73	63.3	44.3	1.10	0.80	1.37	0.07
HGB[a]	117.66	114.87	0.10	0.08	0.547	73.7	34.9	199	60	112	71	64.0	45.8	1.13	0.75	1.50	0.09
HCT[a]	32.91	32.09	0.09	0.08	0.547	78.9	30.8	213	53	119	57	64.2	48.2	1.14	0.69	1.66	0.10
MCV	92.03	87.97	< 0.001	0.60	0.666	88.1	39.5	238	68	104	32	69.6	68.0	1.46	0.30	4.86	0.28
MCH[a]	33.30	31.38	< 0.001	0.25	0.649	96.3	27.3	260	47	125	10	67.5	82.5	1.33	0.14	9.78	0.24
MCHC	359.92	357.75	0.05	0.19	0.562	84.1	21.5	227	37	135	43	62.7	46.3	1.07	0.74	1.45	0.06
RDW-SD	55.94	55.91	0.97	0.00	0.523	70.7	28.5	191	49	123	79	60.8	38.3	0.99	1.03	0.96	-0.01
RDW-CV	16.70	17.78	< 0.001	-0.51	0.635	67.4	51.2	182	88	84	88	68.4	50.0	1.38	0.64	2.17	0.19
NRBC#[a]	0.04	0.27	< 0.001	0.31	0.686	57.2	70.9	154	122	50	115	75.5	51.5	1.97	0.60	3.27	0.28
NRBC%[a]	0.40	1.50	< 0.001	0.24	0.642	66.4	58.1	178	100	72	90	71.2	52.6	1.59	0.58	2.75	0.25
RET#	0.11	0.08	< 0.001	0.58	0.653	75.7	38.4	203	66	106	65	65.7	50.4	1.23	0.63	1.94	0.14
RET%	3.12	2.07	< 0.001	0.54	0.642	88.1	32.0	236	55	117	32	66.9	63.2	1.29	0.37	3.47	0.20
IRF	32.50	27.48	< 0.001	0.34	0.597	25.7	89.0	69	153	19	199	78.4	43.5	2.33	0.83	2.79	0.15
RBC-HE	30.24	29.05	< 0.001	0.70	0.695	88.1	43.6	236	75	97	32	70.9	70.1	1.56	0.27	5.70	0.32
RET-HE	29.91	26.73	< 0.001	0.89	0.737	78.4	55.8	210	96	76	58	73.4	62.3	1.77	0.39	4.57	0.34
DELTA-HE	-1.30	-2.51	< 0.001	0.37	0.604	51.9	65.1	139	112	60	129	69.8	46.5	1.49	0.74	2.01	0.17

Continued on next page

Continued from last page

Parameter	*Unlikely* Average	*Probable* Average	P-Value	Effect	AUC	Sens (%)	Spec (%)	TP	TN	FP	FN	PPV (%)	NPV (%)	LR^+	LR^-	DOR	Youden Index
HYPO-HE	1.80	2.30	< 0.001	-0.33	0.592	73.1	38.4	196	66	106	72	64.9	47.8	1.19	0.70	1.69	0.12
HYPER-HE[a]	1.04	0.60	< 0.001	0.27	0.661	36.6	39.0	98	67	105	170	48.3	28.3	0.60	1.63	0.37	-0.24
MICRO-RBC	4.13	5.42	< 0.001	-0.66	0.682	74.1	55.2	200	95	77	70	72.2	57.6	1.65	0.47	3.53	0.29
MACRO-RBC	5.50	4.10	< 0.001	0.49	0.628	44.8	36.6	121	63	109	149	52.6	29.7	0.71	1.51	0.47	-0.19
White blood cell parameters																	
WBC#[a]	9.89	15.54	< 0.001	0.44	0.758	50.9	87.2	137	150	22	132	86.2	53.2	3.98	0.56	7.08	0.38
NEUT#[a]	3.53	9.49	< 0.001	0.56	0.830	77.0	75.6	208	130	42	62	83.2	67.7	3.15	0.30	10.38	0.53
NEUT%	35.78	58.69	< 0.001	-1.42	0.845	79.1	71.5	212	123	49	56	81.2	68.7	2.78	0.29	9.50	0.51
LYMPH#	3.94	3.45	< 0.001	0.65	0.679	70.0	54.7	187	94	78	80	70.6	54.0	1.54	0.55	2.82	0.25
LYMPH%	40.97	22.08	< 0.001	1.49	0.861	75.3	77.3	201	133	39	66	83.8	66.8	3.32	0.32	10.39	0.53
MONO#	1.64	2.09	< 0.001	-0.52	0.654	34.7	91.9	93	158	14	175	86.9	47.4	4.26	0.71	6.00	0.27
MONO%	15.85	14.80	0.01	0.24	0.571	13.1	84.9	35	146	26	233	57.4	38.5	0.86	1.02	0.84	-0.02
EOSI#	0.38	0.29	0.02	0.22	0.561	41.4	77.3	111	133	39	157	74.0	45.9	1.83	0.76	2.41	0.19
EOSI%[a]	3.82	1.94	< 0.001	0.22	0.630	52.2	77.3	140	133	39	128	78.2	51.0	2.30	0.62	3.73	0.30
BASO#[a]	0.08	0.11	< 0.001	0.19	0.611	82.5	36.6	222	63	109	47	67.1	57.3	1.30	0.48	2.73	0.19
BASO%	0.65	0.66	0.38	-0.09	0.522	26.9	74.4	72	128	44	196	62.1	39.5	1.05	0.98	1.07	0.01
IG#[a]	0.08	0.26	< 0.001	0.36	0.713	57.8	72.7	155	125	47	113	76.7	52.5	2.12	0.58	3.65	0.31
IG%	1.02	1.88	< 0.001	-0.56	0.662	58.2	66.3	156	114	58	112	72.9	50.4	1.73	0.63	2.74	0.24
NLR	1.05	2.62	< 0.001	-1.47	0.857	85.0	70.3	227	121	51	40	81.7	75.2	2.87	0.21	13.46	0.55

Continued on next page

Continued from last page

Parameter	*Unlikely* Average	*Probable* Average	P-Value	Effect	AUC	Sens (%)	Spec (%)	TP	TN	FP	FN	PPV (%)	NPV (%)	LR^+	LR^-	DOR	Youden Index
ITR	0.03	0.03	0.24	-0.11	0.539	31.3	77.3	84	133	39	184	68.3	42.0	1.38	0.89	1.56	0.09
NEUT-RI	42.94	48.90	< 0.001	-1.13	0.788	67.5	74.4	181	128	44	87	80.4	59.5	2.64	0.44	6.05	0.42
NEUT-GI	147.44	144.07	< 0.001	0.58	0.658	51.5	76.2	138	131	41	130	77.1	50.2	2.16	0.64	3.39	0.28
NEUT-TW	31.50	43.05	< 0.001	-1.00	0.762	64.6	72.1	173	124	48	95	78.3	56.6	2.31	0.49	4.70	0.37
RE-LYMPH	9.48	14.08	< 0.001	-0.75	0.710	50.4	79.1	135	136	36	133	78.9	50.6	2.41	0.63	3.83	0.29
AS-LYMPH[a]	0.50	0.90	< 0.001	0.27	0.659	34.0	93.0	91	160	12	177	88.3	47.5	4.87	0.71	6.85	0.27
HFLC[a]	0.00	0.00	< 0.001	0.31	0.684	43.1	88.4	115	152	20	152	85.2	50.0	3.70	0.64	5.75	0.31
RE-MONO[a]	0.22	0.77	< 0.001	0.34	0.704	45.1	81.4	121	140	32	147	79.1	48.8	2.43	0.67	3.60	0.27
Platelet parameters																	
PLT#	305.88	216.82	< 0.001	0.56	0.647	60.4	58.7	163	101	71	107	69.7	48.6	1.46	0.67	2.17	0.19
PDW	14.04	15.01	0.02	-0.26	0.598	66.5	51.7	119	75	70	60	63.0	55.6	1.38	0.65	2.13	0.18
MPV	11.80	12.09	0.02	-0.27	0.594	72.6	45.5	130	66	79	49	62.2	57.4	1.33	0.60	2.22	0.18
P-LCR	38.09	40.32	0.006	-0.31	0.604	75.4	40.0	135	58	87	44	60.8	56.9	1.26	0.61	2.05	0.15
PCT	0.40	0.32	< 0.001	0.44	0.610	49.7	66.2	89	96	49	90	64.5	51.6	1.47	0.76	1.94	0.16
IPF	9.32	13.66	< 0.001	-0.62	0.678	83.6	39.0	224	67	105	44	68.1	60.4	1.37	0.42	3.25	0.23
PLR	0.07	0.07	0.20	0.12	0.529	43.8	62.8	117	108	64	150	64.6	41.9	1.18	0.89	1.32	0.07

Effect refers to Cohen's d or Wilcoxon r effect size. All values were calculated parametrically unless transformation failed to normalise values ([a]). Average refers to the mean or median, depending on the underlying distribution. Significant p-values (< 0.05) and effect sizes (absolute Cohen's $d > 0.50$ and Wilcoxon $r > 0.3$) are highlighted in red. AUC values > 0.8 and sensitivities/specificities $> 85\%$ are highlighted in green, suggesting strong diagnostic performance. Abbreviations: AUC – area under the curve; DOR – diagnostic odds ratio; FN – false negative; FP – false positive; LR^+ – positive likelihood ratio; LR^- – negative likelihood ratio; NPV – negative predictive value; PPV – positive predictive value; Sens – sensitivity; Spec – specificity; TN – true negative; TP – true positive. For parameter abbreviations, see List of Complete Blood Count Parameter Abbreviations (p. xxi).

Appendix A4

Development and Validation of Diagnostic Algorithms for Late-Onset Neonatal Sepsis

---

A4.1 Logistic Regression Model

Table A4.1: Diagnostic accuracy of a logistic regression model for late-onset sepsis.

	Training Dataset	Testing Dataset	Validation Dataset	Testing Dataset	Validation Dataset
P-Value	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Effect Size	0.72	0.70	0.50	0.70	0.50
AUC	0.923 (0.897 – 0.949)	0.914 (0.888 – 0.940)	0.818 (0.802 – 0.834)	0.914 (0.888 – 0.940)	0.818 (0.802 – 0.834)
Threshold	0.47	0.47	0.47	0.60	1.14
TP	187	223	767	222	705
TN	142	147	1,315	150	1,478
FP	17	25	752	22	589
FN	39	43	120	44	182
Sensitivity (%)	82.7 (77.2 – 87.4)	83.8 (78.8 – 88.0)	86.5 (84.0 – 88.7)	83.5 (78.4 – 87.7)	79.5 (76.7 – 82.1)
Specificity (%)	89.3 (83.4 – 93.6)	85.5 (79.3 – 90.4)	63.6 (61.5 – 65.7)	87.2 (81.3 – 91.8)	71.5 (69.5 – 73.4)
PPV (%)	91.7 (87.0 – 95.1)	89.9 (85.5 – 93.4)	50.5 (47.9 – 53.0)	91.0 (86.7 – 94.3)	54.5 (51.7 – 57.2)
NPV (%)	78.5 (71.7 – 84.2)	77.4 (70.8 – 83.1)	91.6 (90.1 – 93.0)	77.3 (70.8 – 83.0)	89.0 (87.4 – 90.5)
LR^+	7.74 (4.92 – 12.18)	5.77 (4.00 – 8.32)	2.38 (2.23 – 2.53)	6.52 (4.40 – 9.67)	2.79 (2.59 – 3.01)
LR^-	0.19 (0.14 – 0.26)	0.19 (0.14 – 0.25)	0.21 (0.18 – 0.25)	0.19 (0.14 – 0.25)	0.29 (0.25 – 0.33)
DOR	40.05 (21.76 – 73.71)	30.49 (17.86 – 52.07)	11.18 (9.04 – 13.82)	34.40 (19.81 – 59.75)	9.72 (8.05 – 11.74)
Youden Index	0.72 (0.61 – 0.81)	0.69 (0.58 – 0.78)	0.50 (0.46 – 0.54)	0.71 (0.60 – 0.80)	0.51 (0.46 – 0.56)
DeLong's P-Value	–	0.63	< 0.001	0.63	< 0.001

Significant p-values (< 0.05) and effect sizes (absolute Wilcoxon $r > 0.30$) are highlighted in red. AUC values > 0.8 and sensitivities/specificities $> 85\%$ are highlighted in green, suggesting strong diagnostic performance. Dataset-specific thresholds are highlighted in yellow. Abbreviations: AUC – area under the curve; DOR – diagnostic odds ratio; FN – false negative; FP – false positive; LR^+ – positive likelihood ratio; LR^- – negative likelihood ratio; NPV – negative predictive value; PPV – positive predictive value; TN – true negative; TP – true positive.

A4.2 Computer-Generated Decision Tree

Table A4.2: Diagnostic accuracy of a computer-generated decision tree model for late-onset sepsis.

	Training Dataset	Testing Dataset	Validation Dataset	Testing Dataset	Validation Dataset
P-Value	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Effect Size	0.74	0.54	0.47	0.54	0.47
AUC	0.918 (0.888 – 0.949)	0.807 (0.766 – 0.848)	0.786 (0.770 – 0.802)	0.807 (0.766 – 0.848)	0.786 (0.770 – 0.802)
Threshold	0.60	0.60	0.60	0.80	0.80
TP	204	226	790	209	729
TN	133	98	1,096	129	1,348
FP	26	74	971	43	719
FN	22	40	97	57	158
Sensitivity (%)	90.3 (85.6 – 93.8)	85.0 (80.1 – 89.0)	89.1 (86.8 – 91.0)	78.6 (73.1 – 83.3)	82.2 (79.5 – 84.7)
Specificity (%)	83.7 (77.0 – 89.0)	57.0 (49.2 – 64.5)	53.0 (50.8 – 55.2)	75.0 (67.8 – 81.3)	65.2 (63.1 – 67.3)
PPV	88.7 (83.9 – 92.5)	75.3 (70.1 – 80.1)	44.9 (42.5 – 47.2)	82.9 (77.7 – 87.4)	50.3 (47.7 – 53.0)
NPV	85.8 (79.3 – 90.9)	71.0 (62.7 – 78.4)	91.9 (90.2 – 93.4)	69.4 (62.2 – 75.9)	89.5 (87.8 – 91.0)
LR^+	5.52 (3.87 – 7.87)	1.97 (1.65 – 2.36)	1.90 (1.80 – 2.00)	3.14 (2.41 – 4.10)	2.36 (2.21 – 2.53)
LR^-	0.12 (0.08 – 0.17)	0.26 (0.19 – 0.36)	0.21 (0.17 – 0.25)	0.29 (0.22 – 0.37)	0.27 (0.24 – 0.32)
DOR	47.43 (25.82 – 87.15)	7.48 (4.76 – 11.76)	9.19 (7.32 – 11.55)	11.00 (7.00 – 17.30)	8.65 (7.12 – 10.51)
Youden Index	0.74 (0.63 – 0.83)	0.42 (0.29 – 0.54)	0.42 (0.38 – 0.46)	0.54 (0.41 – 0.65)	0.47 (0.43 – 0.52)
DeLong's P-Value	–	< 0.001	< 0.001	< 0.001	< 0.001

Significant p-values (< 0.05) and effect sizes (absolute Wilcoxon $r > 0.30$) are highlighted in red. AUC values > 0.8 and sensitivities/specificities $> 85\%$ are highlighted in green, suggesting strong diagnostic performance. Dataset-specific thresholds are highlighted in yellow. Abbreviations: AUC – area under the curve; DOR – diagnostic odds ratio; FN – false negative; FP – false positive; LR^+ – positive likelihood ratio; LR^- – negative likelihood ratio; NPV – negative predictive value; PPV – positive predictive value; TN – true negative; TP – true positive.

A4.3 Manually Constructed Decision Tree

Table A4.3: Diagnostic accuracy of a manually constructed decision tree model for late-onset sepsis.

	Training Dataset	Testing Dataset	Validation Dataset	Validation Dataset
P-Value	< 0.001	< 0.001	< 0.001	< 0.001
Effect Size	0.63	0.62	0.46	0.46
AUC	0.864 (0.829 – 0.900)	0.864 (0.828 – 0.900)	0.788 (0.771 – 0.805)	0.788 (0.771 – 0.805)
Threshold	0.59	0.59	0.59	0.75
TP	170	206	684	565
TN	132	141	1,375	1,690
FP	27	31	692	377
FN	56	60	203	322
Sensitivity (%)	75.2 (69.1 – 80.7)	77.4 (71.9 – 82.3)	77.1 (74.2 – 79.8)	63.7 (60.4 – 66.9)
Specificity (%)	83.0 (76.3 – 88.5)	82.0 (75.4 – 87.4)	66.5 (64.4 – 68.6)	81.8 (80.0 – 83.4)
PPV (%)	86.3 (80.7 – 90.8)	86.9 (82.0 – 90.9)	49.7 (47.0 – 52.4)	60.0 (56.8 – 63.1)
NPV (%)	70.2 (63.1 – 76.6)	70.1 (63.3 – 76.4)	87.1 (85.4 – 88.7)	84.0 (82.3 – 85.6)
LR^+ (%)	4.43 (3.12 – 6.30)	4.30 (3.10 – 5.95)	2.30 (2.15 – 2.47)	3.49 (3.15 – 3.87)
LR^- (%)	0.30 (0.24 – 0.38)	0.28 (0.22 – 0.35)	0.34 (0.30 – 0.39)	0.44 (0.41 – 0.49)
DOR	14.84 (8.89 – 24.77)	15.62 (9.63 – 25.33)	6.70 (5.58 – 8.03)	7.87 (6.59 – 9.39)
Youden Index	0.58 (0.45 – 0.69)	0.59 (0.47 – 0.70)	0.44 (0.39 – 0.48)	0.45 (0.40 – 0.50)
DeLong's P-Value	–	1.00	< 0.001	< 0.001

Significant p-values (< 0.05) and effect sizes (absolute Wilcoxon $r > 0.30$) are highlighted in red. AUC values > 0.8 and sensitivities/specificities $> 85\%$ are highlighted in green, suggesting strong diagnostic performance. Dataset-specific thresholds are highlighted in yellow. Abbreviations: AUC – area under the curve; DOR – diagnostic odds ratio; FN – false negative; FP – false positive; LR^+ – positive likelihood ratio; LR^- – negative likelihood ratio; NPV – negative predictive value; PPV – positive predictive value; TN – true negative; TP – true positive.

A4.4 Random Forest Model

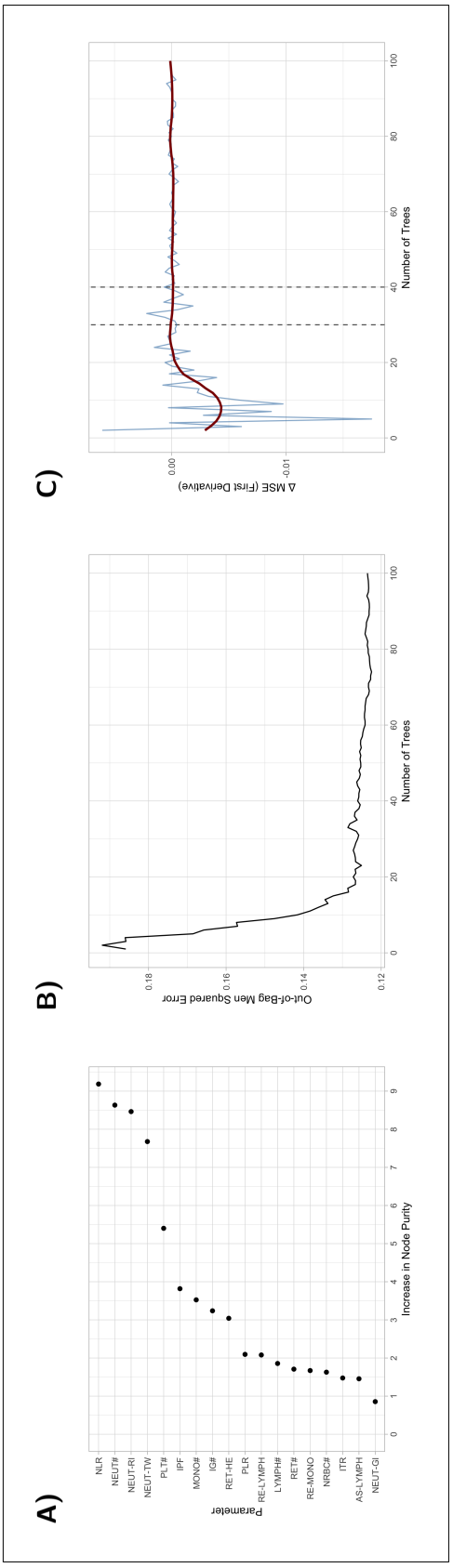


Figure A4.1: Development of a random forest diagnostic model for late-onset sepsis.
A) Variable importance. **B)** Out-of-bag MSE. **C)** Change in ΔMSE. *Abbreviations: AS-LYMPH – antibody-synthesising lymphocyte count; IG# – immature granulocyte count; IPF – immature platelet fraction; ITR – immature-to-total neutrophil ratio; LYMPH# – lymphocyte count; MONO# – monocyte count; MSE – mean squared error; NEUT-GI – neutrophil granularity index; NEUT-RI – neutrophil reactive index; NEUT-TW – neutrophil distribution width; NLR – neutrophil-to-lymphocyte ratio; NRBC# – nucleated red blood cell count; PLR – platelet-to-lymphocyte ratio; PLT# – platelet count; RET# – reticulocyte count; RET-HE – reticulocyte haemoglobin equivalent; RE-LYMPH – reactive lymphocyte count; RE-MONO – reactive monocyte count.*

Table A4.4: Diagnostic accuracy of a random forest model for late-onset sepsis.

	Training Dataset	Testing Dataset	Validation Dataset	Testing Dataset	Validation Dataset
P-Value	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Effect Size	0.83	0.69	0.56	0.69	0.56
AUC	0.988 (0.981 – 0.996)	0.911 (0.885 – 0.937)	0.852 (0.838 – 0.867)	0.911 (0.885 – 0.937)	0.852 (0.838 – 0.867)
Threshold	0.50	0.50	0.50	0.65	0.70
TP	215	231	807	215	686
TN	151	123	1231	154	1631
FP	8	49	836	18	436
FN	11	35	80	51	201
Sensitivity (%)	95.1 (91.5 – 97.5)	86.8 (82.2 – 90.7)	91.0 (88.9 – 92.8)	80.8 (75.6 – 85.4)	77.5 (74.4 – 80.1)
Specificity (%)	95.0 (90.3 – 97.8)	71.5 (64.1 – 78.1)	59.6 (57.4 – 61.7)	89.5 (84.0 – 93.7)	78.9 (77.1 – 80.6)
PPV (%)	96.4 (93.1 – 98.4)	82.5 (77.5 – 86.8)	49.1 (46.7 – 51.6)	92.3 (88.1 – 95.4)	61.1 (58.2 – 64.0)
NPV (%)	93.2 (88.2 – 96.6)	77.8 (70.6 – 84.1)	93.9 (92.5 – 95.1)	75.1 (68.6 – 80.9)	89.0 (87.5 – 90.4)
LR^+	18.91 (9.62 – 37.17)	3.05 (2.39 – 3.88)	2.25 (2.13 – 2.38)	7.72 (4.97 – 12.00)	3.67 (3.35 – 4.01)
LR^-	0.05 (0.03 – 0.09)	0.18 (0.13 – 0.25)	0.15 (0.12 – 0.19)	0.21 (0.17 – 0.28)	0.29 (0.25 – 0.32)
DOR	368.92 (144.95 – 938.96)	16.57 (10.19 – 26.93)	14.85 (11.61 – 19.00)	36.07 (20.28 – 64.15)	12.77 (10.56 – 15.43)
Youden Index	0.90 (0.82 – 0.95)	0.58 (0.46 – 0.69)	0.51 (0.46 – 0.54)	0.70 (0.60 – 0.79)	0.56 (0.52 – 0.61)
DeLong's P-Value	–	< 0.001	< 0.001	< 0.001	< 0.001

Significant p-values (< 0.05) and effect sizes (absolute Wilcoxon $r > 0.30$) are highlighted in red. AUC values > 0.8 and sensitivities/specificities $> 85\%$ are highlighted in green, suggesting strong diagnostic performance. Dataset-specific thresholds are highlighted in yellow. Abbreviations: AUC – area under the curve; DOR – diagnostic odds ratio; FN – false negative; FP – false positive; LR^+ – positive likelihood ratio; LR^- – negative likelihood ratio; NPV – negative predictive value; PPV – positive predictive value; TN – true negative; TP – true positive.

A4.5 Scoring System

Table A4.5: Development of a diagnostic scoring system for late-onset sepsis.

Model	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
AUC	0.899	0.899	0.898	0.896	0.901	0.905	0.901	0.902	0.901	0.900	0.906	0.902	0.893	0.892	0.898	0.886	0.877	0.803
NLR	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
NEUT-RI	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	
NEUT-TW	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y		
NEUT#	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y			
IG#	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y				
IPF	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y					
PLT#	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y						
RET-HE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y							
NRBC#	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y								
MONO#	Y	Y	Y	Y	Y	Y	Y	Y	Y									
RE-LYMPH	Y	Y	Y	Y	Y	Y	Y	Y										
RE-MONO	Y	Y	Y	Y	Y	Y	Y											
LYMPH#	Y	Y	Y	Y	Y	Y												
AS-LYMPH	Y	Y	Y	Y	Y													
ITR	Y	Y	Y	Y														
PLR	Y	Y	Y															
NEUT-GI	Y	Y																
RET#	Y																	

Final model is highlighted in green. Abbreviations: AUC – area under the curve; AS-LYMPH – antibody-synthesising lymphocyte count; IG# – immature granulocyte count; IPF – immature platelet fraction; ITR – immature-to-total neutrophil ratio; LYMPH# – lymphocyte count; MONO# – monocyte count; NEUT-GI – neutrophil granularity index; NEUT-RI – neutrophil reactive index; NEUT-TW – neutrophil distribution width; NLR – neutrophil-to-lymphocyte ratio; NRBC# – nucleated red blood cell count; PLR – platelet-to-lymphocyte ratio; PLT# – platelet count; RET# – reticulocyte count; RET-HE – reticulocyte haemoglobin equivalent; RE-LYMPH – reactive lymphocyte count; RE-MONO – reactive monocyte count.

Table A4.6: Diagnostic accuracy of a scoring system for late-onset sepsis.

	Training Dataset	Testing Dataset	Validation Dataset	Testing Dataset	Validation Dataset
P-Value	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Effect Size	0.69	0.64	0.50	0.64	0.50
AUC	0.906 (0.876 – 0.935)	0.879 (0.848 – 0.911)	0.813 (0.797 – 0.830)	0.879 (0.848 – 0.911)	0.813 (0.797 – 0.830)
Threshold	7.84	7.84	7.84	7.74	7.75
TP	174	207	718	212	725
TN	145	144	1417	144	1408
FP	14	28	650	28	659
FN	52	59	169	54	162
Sensitivity (%)	77.0 (70.9 – 82.3)	77.8 (72.3 – 82.7)	80.9 (78.2 – 83.5)	79.7 (74.4 – 84.4)	81.7 (79.0 – 84.2)
Specificity (%)	91.2 (85.7 – 95.1)	83.7 (77.3 – 88.9)	68.6 (66.5 – 70.6)	83.7 (77.3 – 88.9)	68.1 (66.1 – 70.1)
PPV (%)	92.6 (87.8 – 95.9)	88.1 (83.2 – 91.9)	52.5 (49.8 – 55.2)	88.3 (83.6 – 92.1)	52.4 (49.7 – 55.0)
NPV (%)	73.6 (66.9 – 79.6)	70.9 (64.2 – 77.1)	89.3 (87.7 – 90.8)	72.7 (66.0 – 78.8)	89.7 (88.1 – 91.1)
LR^+	8.74 (5.28 – 14.49)	4.78 (3.39 – 6.75)	2.57 (2.40 – 2.76)	4.90 (3.47 – 6.91)	2.56 (2.39 – 2.75)
LR^-	0.25 (0.20 – 0.32)	0.26 (0.21 – 0.33)	0.28 (0.24 – 0.32)	0.24 (0.19 – 0.31)	0.27 (0.23 – 0.31)
DOR	34.66 (18.46 – 65.07)	18.04 (10.97 – 29.68)	9.26 (7.65 – 11.22)	20.19 (12.21 – 33.39)	9.56 (7.88 – 11.61)
Youden Index	0.68 (0.57 – 0.77)	0.62 (0.50 – 0.72)	0.50 (0.45 – 0.54)	0.63 (0.52 – 0.73)	0.50 (0.45 – 0.54)
DeLong's P-Value	–	0.23	< 0.001	0.23	< 0.001

Significant p-values (< 0.05) and effect sizes (absolute Wilcoxon $r > 0.30$) are highlighted in red. AUC values > 0.8 and sensitivities/specificities $> 85\%$ are highlighted in green, suggesting strong diagnostic performance. Dataset-specific thresholds are highlighted in yellow. Abbreviations: AUC – area under the curve; DOR – diagnostic odds ratio; FN – false negative; FP – false positive; LR^+ – positive likelihood ratio; LR^- – negative likelihood ratio; NPV – negative predictive value; PPV – positive predictive value; TN – true negative; TP – true positive.

A4.6 k-Nearest Neighbour Model

Table A4.7: Diagnostic accuracy of a k-nearest neighbour model for late-onset sepsis.

	Training Dataset	Testing Dataset	Validation Dataset	Testing Dataset	Validation Dataset
P-Value	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Effect Size	0.57	0.37	0.40	0.37	0.40
AUC	0.836 (0.797 – 0.875)	0.716 (0.669 – 0.762)	0.751 (0.732 – 0.770)	0.716 (0.669 – 0.762)	0.751 (0.732 – 0.770)
Threshold	0.53	0.53	0.53	0.73	0.59
TP	163	174	635	125	594
TN	125	110	1,364	148	1,480
FP	34	62	703	24	587
FN	63	92	252	141	293
Sensitivity (%)	72.1 (65.8 – 77.9)	65.4 (59.4 – 71.1)	71.6 (68.5 – 74.5)	47.0 (40.9 – 53.2)	67.0 (63.8 – 70.1)
Specificity (%)	78.6 (71.4 – 84.7)	64.0 (56.3 – 71.1)	66.0 (63.9 – 68.0)	86.1 (80.0 – 90.9)	71.6 (69.6 – 73.5)
PPV (%)	82.7 (76.7 – 87.7)	73.7 (67.6 – 79.2)	47.5 (44.8 – 50.2)	83.9 (77.0 – 89.4)	50.3 (47.4 – 53.2)
NPV (%)	66.5 (59.3 – 73.2)	54.5 (47.3 – 61.5)	84.4 (82.5 – 86.1)	51.2 (45.3 – 57.1)	83.5 (81.7 – 85.2)
LR^+	3.37 (2.48 – 4.59)	1.81 (1.46 – 2.26)	2.10 (1.96 – 2.26)	3.37 (2.27 – 4.99)	2.36 (2.17 – 2.56)
LR^-	0.35 (0.28 – 0.44)	0.54 (0.44 – 0.66)	0.43 (0.39 – 0.48)	0.62 (0.54 – 0.70)	0.46 (0.42 – 0.51)
DOR	9.51 (5.90 – 15.34)	3.36 (2.25 – 5.01)	4.89 (4.12 – 5.81)	5.47 (3.34 – 8.96)	5.11 (4.31 – 6.06)
Youden Index	0.51 (0.37 – 0.63)	0.29 (0.16 – 0.42)	0.38 (0.32 – 0.43)	0.33 (0.21 – 0.44)	0.39 (0.33 – 0.44)
DeLong's P-Value	–	< 0.001	< 0.001	< 0.001	< 0.001

Significant p-values (< 0.05) and effect sizes (absolute Wilcoxon $r > 0.30$) are highlighted in red. AUC values > 0.8 and sensitivities/specificities $> 85\%$ are highlighted in green, suggesting strong diagnostic performance. Dataset-specific thresholds are highlighted in yellow. Abbreviations: AUC – area under the curve; DOR – diagnostic odds ratio; FN – false negative; FP – false positive; LR^+ – positive likelihood ratio; LR^- – negative likelihood ratio; NPV – negative predictive value; PPV – positive predictive value; TN – true negative; TP – true positive.

A4.7 Alternative Diagnostic Models

Table A4.8: Comparison of five diagnostic models for late-onset sepsis with alternative models incorporating gestational age at birth and postnatal age at sampling across the training and testing datasets.

Model	Parameters	Set	Dataset	AUC	Sens (%)	Spec (%)	PPV (%)	NPV (%)	LR^+	LR^-	DOR	Youden Index	DeLong P-Value
Logistic Regression	Removed: PLR Added: NLR	Original	Train	0.923	82.7	89.3	91.7	78.5	7.74	0.19	40.05	0.72	–
			Test	0.914	83.8	85.5	89.9	77.4	5.77	0.19	30.49	0.69	–
		Alternate	Train	0.922	78.3	93.1	94.1	75.1	11.32	0.23	48.59	0.71	0.71
			Test	0.917	81.2	90.1	92.7	75.6	8.22	0.21	39.39	0.71	0.13
Computer Generated Decision Tree	Removed: MONO#, NEUT-RI, IPF Added: RET#, GA	Original	Train	0.918	90.3	83.7	88.7	85.8	5.52	0.12	47.43	0.74	–
			Test	0.807	85.0	57.0	75.3	71.0	1.97	0.26	7.48	0.42	–
		Alternate	Train	0.897	87.6	87.4	90.8	83.2	6.96	0.14	49.14	0.75	0.01
			Test	0.815	82.7	72.7	82.4	73.1	3.03	0.24	12.72	0.55	0.34
Random Forest	Added: GA, Age	Original	Train	0.988	95.1	95.0	96.4	93.2	18.91	0.05	368.92	0.90	–
			Test	0.911	86.8	71.5	82.5	77.8	3.05	0.18	16.57	0.58	–
		Alternate	Train	0.988	92.9	96.9	97.7	90.6	29.59	0.07	404.85	0.90	0.83
			Test	0.904	85.7	73.8	83.5	77.0	3.28	0.19	16.93	0.60	0.06
Scoring System	Added: GA	Original	Train	0.906	77.0	91.2	92.6	73.6	8.74	0.25	34.66	0.68	–
			Test	0.879	77.8	83.7	88.1	70.9	4.78	0.26	18.04	0.62	–
		Alternate	Train	0.907	85.4	83.7	88.1	80.1	5.22	0.17	29.93	0.69	0.64
			Test	0.887	85.3	76.7	85.0	77.2	3.67	0.19	19.21	0.62	0.006

Continued on next page

Continued from last page

Model	Parameters	Set	Dataset	AUC	Sens (%)	Spec (%)	PPV (%)	NPV (%)	LR^+	LR^-	DOR	Youden Index	DeLong P-Value
K-Nearest Neighbour	Added: GA, Age	Original	Train	0.836	72.1	78.6	82.7	66.5	3.37	0.35	9.51	0.51	–
			Test	0.716	65.4	64.0	73.7	54.5	1.81	0.54	3.36	0.29	–
		Alternate	Train	0.828	64.6	83.0	84.4	62.3	3.80	0.43	8.92	0.48	0.44
			Test	0.712	57.9	71.5	75.9	52.3	2.03	0.59	3.45	0.29	0.73

Significant p-values (< 0.05) are highlighted in red. AUC values > 0.8 and sensitivities/specificities $> 85\%$ are highlighted in green, suggesting strong diagnostic performance. Abbreviations: AUC – area under the curve; DOR – diagnostic odds ratio; GA – gestational age at birth; IPF – immature platelet fraction; LR^+ – positive likelihood ratio; LR^- – negative likelihood ratio; MONO# – monocyte count; NEUT-RI – neutrophil reactive index; NLR – neutrophil-to-lymphocyte ratio; NPV – negative predictive value; PLR – platelet-to-lymphocyte ratio; PPV – positive predictive value; RET# – reticulocyte count; Sens – sensitivity; Spec – specificity.

Appendix A5

Discussion

A5.1 Annual Reagent Consumption

Table A5.1: Annual reagent consumption based on the average number of complete blood count results measured per day (unpublished data)

Annual Reagent Consumption							Annual Reagent Costs						
Average Tests per Day	2	5	10	25	50	100	Price Per Unit	2	5	10	25	50	100
Cellpack DCL 10L	35	40	49	75	119	208	£6.70	£234.47	£267.96	£328.25	£502.43	£797.18	£1,393.39
Sulfolyser 500 mL	6	7	9	15	25	46	£26.05	£158.37	£182.33	£234.43	£390.72	£651.20	£1,198.20
Lysercell WDF 2L	8	10	13	21	36	65	£14.24	£113.94	£142.42	£185.14	£299.08	£512.71	£925.72
Cellpack DFL 1L	10	13	17	28	47	86	£18.46	£184.61	£240.00	£313.84	£516.92	£867.69	£1,587.68
Fluorocell WDF 2x22 mL	3	3	4	7	12	21	£31.76	£95.27	£95.27	£127.02	£222.29	£381.06	£666.86
Fluorocell RET 2x12 mL	5	6	7	12	20	36	£60.28	£301.41	£361.69	£421.98	£723.39	£1,205.65	£2,170.16
Cellclean Auto 20x4 mL	3	3	3	3	3	3	£36.54	£109.62	£109.62	£109.62	£109.62	£109.62	£109.62
XN-L Check L2 6x3 mL	4	4	4	4	4	4	£26.68	£106.73	£106.73	£106.73	£106.73	£106.73	£106.73
							Total	£1,304.41	£1,506.02	£1,827.02	£2,871.17	£4,631.83	£8,158.36
							Cost per Test	£1.79	£0.83	£0.50	£0.31	£0.25	£0.22