



**A MACHINE LEARNING APPROACH TO THE PREDICTION OF POOR
DIARRHEAL OUTCOMES AMONG CHILDREN UNDER FIVE YEARS OF AGE: AN
EVALUATION OF LONGER DURATION DIARRHEA, CHRONIC MALNUTRITION,
AND MORTALITY**

BY
BILLY OGWEL
(Student Number: 16949145)

Submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy (PhD)
in the subject
INFORMATION SYSTEMS
at the University of South Africa (UNISA)

Supervisor: Dr. Vincent H. Mzazi
Co-Supervisor: Prof. Bester Chimbo

March 2026

ABSTRACT

Diarrhea is a public health problem globally, having an incidence of about 1.7 billion childhood diarrheal episodes every year, and an annual mortality of about 1.5 million. The clinical diarrheal outcomes surpass acute dehydration and electrolyte imbalance, and often include: longer duration diarrhea (LDD); chronic malnutrition; and mortality. This diarrheal burden disproportionately affects low- and middle-income countries, and is further exacerbated by delayed care-seeking, inadequate diagnostic capacity, demanding work environments, and provider burnout which can impair clinical judgment and performance. Predictive models can augment clinical decision-making by facilitating the rapid identification of patients at increased risk of poor diarrheal outcomes, facilitating timely and cost-effective interventions to improve prognoses. Existing literature revealed a paucity of research focused on the development of predictive models for diarrheal outcomes. This study aimed to bridge this research gap by: i.) identifying predictors of poor diarrhea outcomes for long duration diarrhea (LDD), chronic malnutrition, and mortality; ii.) deriving and validating patient-level predictive models for poor diarrheal outcomes; iii.) designing an R-shiny product suite for predictive models of poor diarrheal outcomes. A correlational study design that involved a hybrid feature selection strategy to identify predictors, was adopted. Seven machine learning algorithms were utilized for model development and evaluation leveraging data from three pediatric enteric studies conducted in Siaya County, Kenya between 2010 and 2023. Shapley values were estimated to enhance model interpretability and the model with optimal discrimination was selected as the champion model for each outcome. Clinical variables were the primary predictors of poor diarrheal outcomes, although the sets of predictors varied based on the distinct outcome being modeled. The champion models identified were: random forest for LDD

(AUC [95% CI]: 83.0% [78.6–87.5]); gradient boosting for chronic malnutrition (AUC [95% CI]: 83.5% [81.6–85.4]); and random forest for mortality (AUC [95% CI]: 82.6% [77.1–88.1]). The model AUCs declined by 12% for LDD and 18% for chronic malnutrition during temporal validation. An R-Shiny web application was developed, featuring a consolidated interface that dynamically displays risk profiles and outcome-specific Shapley values upon submission of user inputs. Beyond demonstrating the practical utility of machine learning algorithms in rapid identification of high-risk children supporting clinical decision-making, resource prioritization, and improved management, this work contributes to the growing body of literature on the application of machine learning to predict pediatric risks. However, successful implementation and widespread adoption of the developed tool will require further research, collaboration, and ethical oversight. Consequently, future research is recommended to evaluate the clinical acceptability of these models, as well as their impact on clinical practice and patient outcomes. Moreover, it is essential to assess the broader implications of ML integration, including the operational challenges and the cost-effectiveness of model deployment through a before-after study or decision-analytic modeling framework.

Keywords: Diarrhea; pediatric care; machine learning; longer duration diarrhea; chronic malnutrition; mortality; predictive model; explainable ML; deployment.

DECLARATION

I hereby declare that this document: **A MACHINE LEARNING APPROACH TO THE PREDICTION OF POOR DIARRHEAL OUTCOMES AMONG CHILDREN UNDER FIVE YEARS OF AGE: AN EVALUATION OF LONGER DURATION DIARRHEA, CHRONIC MALNUTRITION, AND MORTALITY**, submitted for evaluation towards the requirements of the subject: Information Systems, as part of the PhD qualification at the University of South Africa, is my own original work and has not previously been submitted to any other institution of higher learning or subject for evaluation. *I further declare that I submitted the thesis to originality checking software and that it falls within the accepted requirements for originality.* I further declare that all sources used or quoted in this document are indicated and acknowledged by means of a comprehensive list of references.

Surname, Initials: Ogwel, B

Student Number: 16949145

Signature: 

Date: March-03-2026

DEDICATION

To the countless children whose lives are impacted by diarrheal illness, and to all those dedicated to improving their well-being by the pursuit of innovative solutions. May this research contribute to a healthier future for them.

In loving memory of my mum, Herine Oiko, and dad, Amorce Ogwel, whose lives exemplified diligence and whose memory continue to inspire me each day.

ACKNOWLEDGEMENT

This doctoral thesis would not have been possible without the guidance, support, and encouragement of many individuals and institutions, to whom I am deeply grateful. First and foremost, I acknowledge with gratitude the grace of God and divine inspiration that guided me in completing this endeavor.

I also wish to express my sincere gratitude to my supervisor, Dr. Vincent H. Mzazi, without whom this work would not have made it to the finish line. Thank you for your unwavering support, invaluable guidance, and constructive feedback throughout this journey. This work has been significantly shaped by your expertise and mentorship, which were vital in the successful integration of the public health and informatics perspectives.

I would also like to acknowledge and sincerely thank my co-supervisor, Prof. Bester Chimbo, for her invaluable guidance and contributions, which greatly enhanced the rigor and timely completion of this thesis.

I am particularly grateful to my employer, the Kenya Medical Research Institute as well as collaborating partners including University of Washington and University of Maryland, Baltimore who implemented the primary studies that formed the backbone of this study. Specifically, I am profoundly grateful to Prof. Richard Omore, the principal investigator of the primary studies, for granting me access to these essential datasets. His support extended far beyond data access; he generously dedicated time from his demanding schedule to monitor my progress, ensuring I remained on track through his critical review and mentorship. Furthermore, I deeply appreciate the connections he facilitated with leading scientists in pediatric enteric research, whose insightful contributions not only enhanced the quality of this work but also hold great promise for future collaborations.

Special thanks go to the following: Dr. Karen Kotloff, Dr. Patricia Pavlinac, Dr. Kirkby Tickell, Dr. Bryan Nyawanda, and Dr. Alex Awuor, whose input and technical support have been instrumental. I am particularly grateful to Dr. Tickell who enriched my understanding of machine learning and its application in global health.

On a personal note, I would like to thank my family for their unconditional love and encouragement. To my wife, Esther Tolo, together with our dear children-Amorce, Yolanda and

Yvette, your belief in my abilities and your constant support have been my foundation. You have stood by me, offering words of encouragement and moments of respite during challenging times. A heartfelt thank you to my siblings, Bryan, Oscar and Mary Rose for walking with me through the journey.

Finally, to the countless children and families whose experiences are represented in this research: your resilience is a reminder of the importance of this work. It is my hope that this thesis contributes to improving outcomes for children under five years of age and advancing technological innovation in the public health sector in low- and middle-income countries.

Thank you all for being part of this journey.

LIST OF ACRONYMS

AGE	-	Acute Gastroenteritis
ANN	-	Artificial Neural Networks
AUC	-	Area Under the Curve
DALYs	-	Disability-Adjusted Life Years
DHS	-	Demographic and Health Survey
EFGH	-	Enterics For Global Health Shigella study
EHRs	-	Electronic Health Records
EMA	-	Explanatory Model Analysis
FGR	-	Fetal Growth Restriction
GBM	-	Gradient Boosting Machine
GDPs	-	Gross Domestic Products
GEMS	-	Global Enteric Multicenter Study
GSRH	-	General Self-Rated Health score
HAZ	-	Height-for-Age Z-score
HDSS	-	Health and Demographic Surveillance System
IMCI	-	Integrated Management of Childhood Illness guidelines
KEMRI	-	Kenya Medical Research Institute
KNN	-	K-Nearest Neighbors
LDD	-	Longer Duration Diarrhea
LGF	-	Linear Growth Faltering
LMICs	-	Low-Middle-Income Countries
LR	-	Logistic Regression
MAD	-	Medically-Attended Diarrhea
MALED	-	Malnutrition and the Consequences for Child Health and Development study
MICE	-	Multiple Imputation by Chained Equations
MoH	-	Ministry of Health
MSD	-	Moderate-to-Severe Diarrhea
ML	-	Machine Learning

NB	-	Naive Bayes
NPV	-	Negative Predictive Value
PCS	-	Physical Component Score
PD	-	Persistent Diarrhea
pMSE	-	Propensity score Mean-Squared-Error
PPV	-	Positive Predictive Value
PRAUC	-	Precision-Recall Area Under the Curve
PYO	-	Person Years of Observation
RF	-	Random Forest
ROC	-	Receiver Operating Characteristic curve
SAM	-	Severe Acute Malnutrition
SDGs	-	Sustainable Development Goals
SERU	-	Scientific and Ethics Review Unit
SHAP	-	SHapley Additive exPlanations values
SHC	-	Sentinel Health Center
SIC	-	Seattle Index of Comorbidity
S_pMSE	-	Standardized ration of Propensity score Mean-Squared Error
SSA	-	Sub-Saharan Africa
SVM	-	Support Vector Machine
UNISA	-	University of South Africa
VIDA	-	Vaccine Impact on Diarrhea in Africa study
WASH	-	Water, Sanitation and Hygiene
YLD	-	Years-Lived with-Disability
YLL	-	Years-of-Life Lost

TABLE OF CONTENTS

ABSTRACT.....	ii
DECLARATION.....	iv
DEDICATION.....	v
ACKNOWLEDGEMENT.....	vi
LIST OF ACRONYMS	viii
LIST OF FIGURES	xv
LIST OF TABLES	xviii
LIST OF EQUATIONS.....	xx
LIST OF PUBLICATIONS	xxi
CHAPTER ONE: INTRODUCTION	1
1.1 Background	1
1.1.1 Longer duration diarrhea.....	3
1.1.2 Chronic malnutrition	4
1.1.3 Mortality	5
1.1.4 Machine Learning	7
1.2 Current State of Predictive Modelling in Diarrheal Disease.....	9
1.3 Problem Statement	11
1.4 Research Purpose and Objectives	12
1.5 Research Questions	12
1.6 Scope.....	13
1.7 Ethical Considerations	14
1.8 Thesis Outline	15
CHAPTER TWO: THE PUBLIC HEALTH CONTEXT.....	17
2.1 Introduction.....	17
2.2 Longer Duration Diarrhea.....	17
2.2.1 Burden.....	17
2.2.2 Risk factors	22
2.2.3 Pathogenesis, Diagnosis and Management	29
2.3 Chronic Malnutrition	34
2.3.1 Burden.....	34
2.3.2 Risk factors	43
2.3.3 Pathogenesis, Diagnosis and Management	51

2.4 Mortality	55
2.4.1 Burden.....	55
2.4.2 Risk factors	60
CHAPTER THREE: THE INFORMATICS CONTEXT	68
3.1 Introduction.....	68
3.2 Machine Learning	68
3.2.1 Supervised Learning	71
3.2.2 Unsupervised Learning	72
3.2.3 Semi-supervised learning.....	74
3.2.4 Reinforcement learning.....	74
3.2.5 Machine learning algorithms for predicting poor diarrheal outcomes	75
3.3 Current State of Predictive Modelling	96
3.3.1 Longer Duration Diarrhea (LDD)	97
3.3.2 Chronic malnutrition	103
3.3.3 Mortality	109
3.4 From Development to Deployment.....	117
3.5 Theoretical Frameworks	124
3.5.1 PROGRESS Framework	124
3.5.2 Socio-technical Approach.....	126
3.5.3 Design Science Approach	128
3.6 Study Conceptual Framework.....	129
CHAPTER FOUR: RESEARCH METHODOLOGY	132
4.1 Introduction.....	132
4.2 Study setting.....	132
4.2.1 An Overview of Healthcare Delivery and Referrals in Kenya.....	135
4.2.2 Vaccine Impact on Diarrhea in Africa (VIDA) Study	137
4.2.3 Enterics for Global Health Shigella Surveillance	140
4.2.4 Rotavirus Surveillance Network for Africa	142
4.3 Research Design.....	142
4.3.1 Correlational Research Design.....	143
4.4 Sample Size Calculation	145
4.5 Data Acquisition and Data Preprocessing	147
4.5.1 Handling Missing Data	147

4.5.2 Outcome Variables.....	150
4.6 Modelling Approach	151
4.6.1 Feature Selection.....	153
4.6.2 Model Development.....	154
4.6.3 Model Evaluation.....	159
4.7 Statistical Analysis	167
4.8 Product Development and Deployment	169
4.8.1 Dashboard Design.....	169
4.8.2 Dashboard Deployment.....	170
4.9 Reflections on Study Contribution in Machine Learning and Pediatric Enteric Diseases	171
4.10 Ethical Considerations	171
CHAPTER FIVE: RESULTS AND DISCUSSION	173
5.1 Predictors of Longer Duration Diarrhea	173
5.1.1 Key Findings for RQ1.....	173
5.1.2 Discussion: Predictors of Longer Duration Predictors (RQ1)	181
5.2 Predictors of Chronic Malnutrition	187
5.2.1 Key findings for RQ2.....	187
5.2.2 Discussion: Chronic Malnutrition Predictors (RQ2).....	195
5.3 Predictors of Diarrhea-associated Mortality	200
5.3.1 Key Findings for RQ3.....	200
5.3.2 Discussion: Diarrhea-associated Mortality Predictors (RQ3).....	206
5.4 Model Development and Evaluation for Poor Diarrheal Outcomes	214
5.4.1 Assessing Model Performance for Longer Duration Diarrhea Prediction	215
5.4.2 Assessing Model Performance for Chronic Malnutrition Prediction	232
5.4.3 Evaluating Model Performance for In-hospital Mortality Prediction	249
5.5 Product Development and Deployment	259
5.5.1 Key Findings for RQ7.....	259
5.5.2 Discussion: Development of Decision Support Dashboard	267
CHAPTER SIX: REFLECTIONS ON STUDY CONTRIBUTION IN MACHINE LEARNING AND PEDIATRIC ENTERIC DISEASES	271
6.1 Introduction.....	271
6.2 Contributions to the Methodological Dimension	272
6.2.1 Discussion of Methodological Contributions.....	275

6.3. Contribution in the Health Services Dimension.....	277
6.3.1 Enabling Early Risk Stratification and Proactive Care	277
6.3.2 Blueprint for a Low-Cost Clinical Decision Support Tool	277
6.3.3 Informing Targeted Public Health Interventions	278
6.3.4 Informing National and County-Level Child Health Guidelines	279
6.4. Reflections on Triadic Conceptual Framework	279
6.5 Future Research Directions: Pathways for Implementation, Adaptation, and Integration.....	282
6.5.1 Clinical Implementation and Impact Evaluation.....	282
6.5.2 Addressing Concept Drift and Data Complexity	283
6.5.3 Integrating Multi-Level Determinants	284
CHAPTER SEVEN: CONCLUSION AND RECOMMENDATIONS.....	286
7.1 Research Summary	286
7.2 Accomplishment of the Research Objectives	288
7.2.1 Objective One: <i>Determine the demographic, clinical and socio-economic features that predict longer duration diarrhea, among children aged < 60 months.</i>	288
7.2.2 Objective Two: <i>Determine the demographic, clinical and socio-economic features that predict chronic malnutrition following a diarrheal illness among children aged < 60 months.</i>	289
7.2.3 Objective Three: <i>Determine the demographic, clinical and socio-economic features that predict diarrhea-associated mortality following a diarrheal illness among children aged < 60 months.</i>	289
7.2.4 Objective Four: <i>Develop and evaluate (internal and temporal validation) the performance of predictive models for each poor diarrheal outcome among children aged <5 years using machine learning.</i>	290
7.2.5 Objective Five: <i>Design an R-shiny product suite for predictive models of poor diarrheal outcomes.</i>	291
7.3 Reflection on Study Contributions to the Knowledge Base and Future Research Directions	292
7.4 Limitations	295
7.5 Recommendations.....	296
7.5.1 Features predicting poor diarrheal outcomes	296
7.5.2 Develop and evaluate the performance of predictive models for poor diarrheal outcomes.....	297
7.5.3 Develop an R-shiny product suite for predictive models of poor diarrheal outcomes.....	297
REFERENCES.....	298
APPENDICES	380
Appendix 1: Ethics Approval.....	380
Appendix 2: Glossary of Key Terms	380

Appendix 3: Methodological Appendix..... 380

Appendix 4: Permissions for Use of Images..... 380

LIST OF FIGURES

Figure 1.1: Global Burden of Diarrhea among children aged < 5 years (UNICEF, 2022).....	3
Figure 1.2: The vicious cycle of enteric infections, malnutrition and long-lasting effects (Guerrant et al., 2013)	5
Figure 1.3: Global causes of under-5 deaths in 2015 (L. Liu et al., 2016)	6
Figure 1.4: Schematic diagram of thesis outline.....	6
Figure 2.1: The epidemiologic triad of infectious diseases (Johnson & Kaneene, 2018).....	23
Figure 2.2: A comparative analysis of stunting cases globally and by world regions in 1990 and 2019 (Vaivada et al., 2020; WHO et al., 2020).	36
Figure 2.3: Temporal patterns in in stunting prevalence between 1990 and 2019 (Vaivada et al., 2020; WHO et al., 2020).....	37
Figure 2.4: Trends of childhood malnutrition in Kenya 1993-2022 (Kenya National Bureau of Statistics & Ministry of Health, 2023).	39
Figure 2.5: Mapping childhood stunting in Kenya: County-specific prevalence rates (2022)(Kenya National Bureau of Statistics & Ministry of Health, 2023).	40
Figure 2.6: WHO conceptual framework on childhood stunting (World Health Organization, 2021).....	43
Figure 2.7: Stunting Pathways Across the Developmental Spectrum (Prendergast & Humphrey, 2014).	52
Figure 2.8: Global trends of diarrheal mortality 1990-2017 (Troeger et al., 2019).....	55
Figure 2.9: Diarrheal mortality rate in 2016 (Troeger et al., 2019).	57
Figure 2.10: Age-adjusted diarrheal disability-adjusted life-year rates between 1990 and 2017 (Karambizi et al., 2021).	59
Figure 3.1: Machine Learning Overview (Moubayed et al., 2018)	70
Figure 3.2: ANN Architecture (S. Yang et al., 2021)	93
Figure 3.3: Pathway in the building, evaluation, diffusion and scaling of machine learning models in clinical care (Sendak et al., 2020).....	118
Figure 3.4: Machine learning papers published in PubMed and percentage having deployment/implementation (Drysdale et al., 2020).....	120
Figure 3.5: Prognostic research conceptual framework (Kent et al., 2020).....	125
Figure 3.6: Sociotechnical systems theory (Militello et al., 2013).	127
Figure 3.7: Design Science Approach (Engel et al., 2013).....	129

Figure 3.8: Study Conceptual Framework	131
Figure 4.1: Map of Study Area (Odhiambo et al., 2012)	133
Figure 4.2: The structure of Kenya's healthcare delivery and referral system (Ministry of Health, 2014)	136
Figure 4.3: Flow diagram of key study activities in VIDA (Kotloff et al., 2012).	139
Figure 4.4: Missingness map for VIDA study	148
Figure 4.5: Patterns of missing data in VIDA study	149
Figure 4.6: Missingness map for the rotavirus inpatient surveillance	150
Figure 4.7: Patterns of missing data in rotavirus inpatient surveillance	150
Figure 4.8: Model development and validation schematic diagram for the prediction of poor diarrheal outcomes.	152
Figure 4.9: Visualization of Validation Techniques (Ramspek et al., 2021).	159
Figure 4.10: Map of Explainability Approaches (Belle & Papantonis, 2021).	166
Figure 5.1: Comparative enrolment flow diagram of the development (VIDA) and temporal validation (EFGH) cohorts in longer duration diarrhea prediction.	174
Figure 5.2: Feature selection for longer duration diarrhea in the development cohort (VIDA Study)	181
Figure 5.3: CONSORT diagram of linear growth faltering in development (VIDA) and temporal validation (EFGH) cohorts.....	188
Figure 5.4: Histogram comparison of predictor variable distributions between original (n=1,106) and synthetic (n=8,894) data.....	193
Figure 5.5: Predictor identification for linear growth faltering (Δ HAZ $\geq$ -0.5) in development cohort (VIDA study).	194
Figure 5.6: Predictor identification for linear growth faltering (negative Δ HAZ) in development cohort (VIDA study)	195
Figure 5.7: CONSORT diagram of gastroenteritis-associated deaths among children under 60 months admitted at Siaya county referral hospital, Western Kenya (2010–2020) and data segregation into training and test sets.	201
Figure 5.8: Trends of diarrhea-associated mortality among pediatric patients less than 60 months of age hospitalized for acute gastroenteritis at Siaya County Referral Hospital: 2010-2020.	202
Figure 5.9: Predictor identification for gastroenteritis-associated mortality among children aged < 60 months admitted at SCRH: 2010-2020.	206
Figure 5.10: Nemenyi Post-hoc Test for Performance Differences among LDD Prediction Models	218

Figure 5.11: ROC analysis of the 7 models in the prediction of LDD in the primary use case.....	219
Figure 5.12: Explanatory profile for the top-performing 4 LDD models for the primary use case	220
Figure 5.13: Evaluation of the business value of the champion model in LDD prediction	221
Figure 5.14: Comparative evaluation of model performance in the prediction of LDD between the development (2015-2018) and temporal validation (2022-2023) cohorts.	222
Figure 5.15: Explanatory profile for the top-performing 4 models in the prediction of post-enrolment duration (≥ 7 days).	224
Figure 5.16: Nemenyi Post-hoc Test for Performance Differences among LGF Prediction Models.	234
Figure 5.17: Receiver operating characteristic curve for linear growth faltering prediction models	235
Figure 5.18: Explanatory profile for the two top-performing linear growth faltering models.....	238
Figure 5.19: Evaluation of the business value of the champion model in linear growth faltering prediction	239
Figure 5.20: Comparative evaluation of model performance in the prediction of linear growth faltering between the development (2015-2018) and temporal validation (2022-2023) cohorts.	240
Figure 5.21: Nemenyi Post-hoc Test for Performance Differences among Mortality Prediction Models.	251
Figure 5.22: ROC analysis of the 7 models in the prediction of in-hospital mortality prediction.....	252
Figure 5.23: Explanatory profile for the top-performing 4 models in the prediction of in hospital-mortality.....	254
Figure 5.24: Evaluation of the business value of the champion model in in hospital-mortality prediction	255
Figure 5.25: Online Dashboard Architecture	262
Figure 5.26: Consolidated User Interface for Poor Diarrheal Outcomes Dashboard.	263
Figure 5.27: Processing of Request upon Submission of User Inputs	264
Figure 5.28: Prediction Outputs for Longer Duration Diarrhea.....	265
Figure 5.29: Prediction Outputs for Linear Growth Faltering.	266
Figure 5.30: Prediction Outputs for Diarrhea-associated Mortality.....	266

LIST OF TABLES

Table 2.1: Definitions of diarrheal illnesses (Giannattasio et al., 2016).....	18
Table 2.2: Risk factors for diarrheal duration in children aged < 60 months: A summary of evidence	28
Table 2.3: Factors influencing childhood stunting: A summary of evidence	50
Table 2.4: Determinants of diarrheal mortality among children under five: A summary of evidence	65
Table 3.1: Current State of predictive modelling for LDD: A summary of literature	102
Table 3.2: Current state of predictive modelling for chronic malnutrition: A summary of literature	108
Table 3.3: Current state of predictive modelling for mortality- A summary of literature	115
Table 4.1: VIDA sentinel health centers in Rural Western Kenya, 2015-2018.....	137
Table 4.2: Study procedures at each EFGH study time point.....	141
Table 4.3: Tabulation of Sampling Frame and Sample Size.....	146
Table 4.4: Evaluation metrics for poor diarrheal outcomes	161
Table 5.1: Comparison of baseline characteristics among pediatric patients in development and temporal validation datasets.....	175
Table 5.2: Characteristics of participants in the development cohort (VIDA study) stratified by diarrhea duration status, 2015-2018.....	177
Table 5.3: Profile of children enrolled in the development study (VIDA: 2015-2018), grouped by linear growth faltering.....	189
Table 5.4: Statistical comparison of synthetic data (n = 8,894) and original training data (n=1,106).....	192
Table 5.5: Demographic and clinical profile of pediatric patients aged less than 60 months hospitalized with acute gastroenteritis at Siaya County Referral Hospital: 2010-2020.	203
Table 5.6: Performance of LDD predictive models utilizing over-sampling approach within the resampling procedure.....	216
Table 5.7: Post-hoc analysis of LDD model performance using Nemeyi Test.....	216
Table 5.8: Performance of LDD Predictive Models without correcting for class imbalance....	218

Table 5.9: Calibration results for LDD prediction models.	219
Table 5.10: Predictive capability of models in predicting post-enrolment duration (≥ 7 days) .	223
Table 5.11: Model performance in the prediction of linear growth faltering ^b using composite data (Original/simulated data).....	233
Table 5.12: Post-hoc analysis of LGF model performance using Nemenyi Test.	233
Table 5.13: Model performance in the prediction of linear growth faltering ^b using original data only	236
Table 5.14: Model performance in the prediction of linear growth faltering ^b using composite data.....	236
Table 5.15: Model calibration in the prediction of linear growth faltering.	237
Table 5.16: Predictive capability of existing diarrheal scores in the prediction of in-hospital mortality among children aged < 60 months hospitalized with AGE.....	249
Table 5.17: Model Performance in the prediction of in-hospital mortality among children aged < 60 months hospitalized with AGE.	249
Table 5.18: Post-hoc analysis of mortality model performance using Nemenyi Test.	250
Table 5.19: Calibration results of prediction models applied to test data.....	253
Table 5.20: Summary of Methodology and Model Performance in the Prediction of Poor Diarrheal Outcomes.	260
Table 6.1: Contributions of the Thesis to the Knowledge Base	272

LIST OF EQUATIONS

Equation 1-2 (Sigmoid function)	77
Equation 3 (Logistic function)	77
Equation 4-7 (Naïve Bayes formula)	80
Equation 8 (Gini Index formula).....	83
Equation 9 (Entropy formula).....	83
Equation 10 (Gradient descent formula).....	85
Equation 11 (Bernoulli loss function).....	86
Equation 12 (SVM Linear Kernel function)	88
Equation 13 (SVM Radial Basis Function Kernel).....	88
Equation 14 (Euclidean distance formula).....	90
Equation 15 (Neural network formula).....	93
Equation 16 (Riley sample size estimation formula)	145
Equation 17 (Brier score).....	163
Equation 18 (Spiegelhalter z statistic)	163
Equation 19 (Platt scaling formula)	164

LIST OF PUBLICATIONS

The research produced four journal articles, each addressing the prediction of poor diarrheal outcomes. The publications, which delineate the authors' contributions, are presented herein as key deliverables of the study:

- Ogwel B, Mzazi V.H., Nyawanda BO, Otieno G, Omore R. Predictive modeling for infectious diarrheal disease in pediatric populations: A systematic review. *Learn Health Syst.* 2023 Jul 29;8(1):e10382. <https://doi.org/10.1002/lrh2.10382>

The paper was submitted to *Learning Health Systems* journal on the 7th January, 2023 and accepted on 17th July, 2023. The paper provided a review of current literature on predictive modeling of pediatric infectious diarrheal illness, highlighting existing knowledge gaps and opportunities for future research.

- Ogwel, B., Mzazi, V.H., Awuor, A.O. *et al.* Predictive modelling of linear growth faltering among pediatric patients with Diarrhea in Rural Western Kenya: an explainable machine learning approach. *BMC Med Inform Decis Mak* **24**, 368 (2024). <https://doi.org/10.1186/s12911-024-02779-7>

The paper was submitted to *BMC Medical Informatics and Decision Making* journal on the 8th March, 2024 and accepted on 22nd November, 2024. The paper reported on the predictors and model performance, assessed via internal and temporal validation, for the prediction of chronic malnutrition in children, post-diarrheal episode.

- Ogwel B, Mzazi V.H., Nyawanda BO, Otieno G, Tickell KD, Omore R. A machine learning approach to predicting inpatient mortality among pediatric acute gastroenteritis patients in Kenya. *Learn Health Sys.* 2024; e10478. doi:[10.1002/lrh2.10478](https://doi.org/10.1002/lrh2.10478)

The paper was submitted to *Learning Health Systems* journal on the 24th Feb, 2024 and accepted on 10th December, 2024. The paper reported on the predictors and model performance, assessed via internal validation, for the prediction of mortality among children hospitalized for acute gastroenteritis.

- Ogwel, B., Mzazi, V.H., Awuor, A.O. *et al.* Derivation and Validation of a Clinical Predictive Model for Longer Duration Diarrhea among Pediatric Patients in Kenya using Machine Learning Algorithms. *BMC Med Inform Decis Mak* **25**, 28 (2025). <https://doi.org/10.1186/s12911-025-02855-6>

The paper was submitted to *BMC Medical Informatics and Decision Making* journal on the 8th March, 2024 and accepted on 3rd January, 2025. The paper reported on the predictors and model performance, assessed via internal and temporal validation, for the prediction of longer duration diarrhea in children.

1. CHAPTER ONE: INTRODUCTION

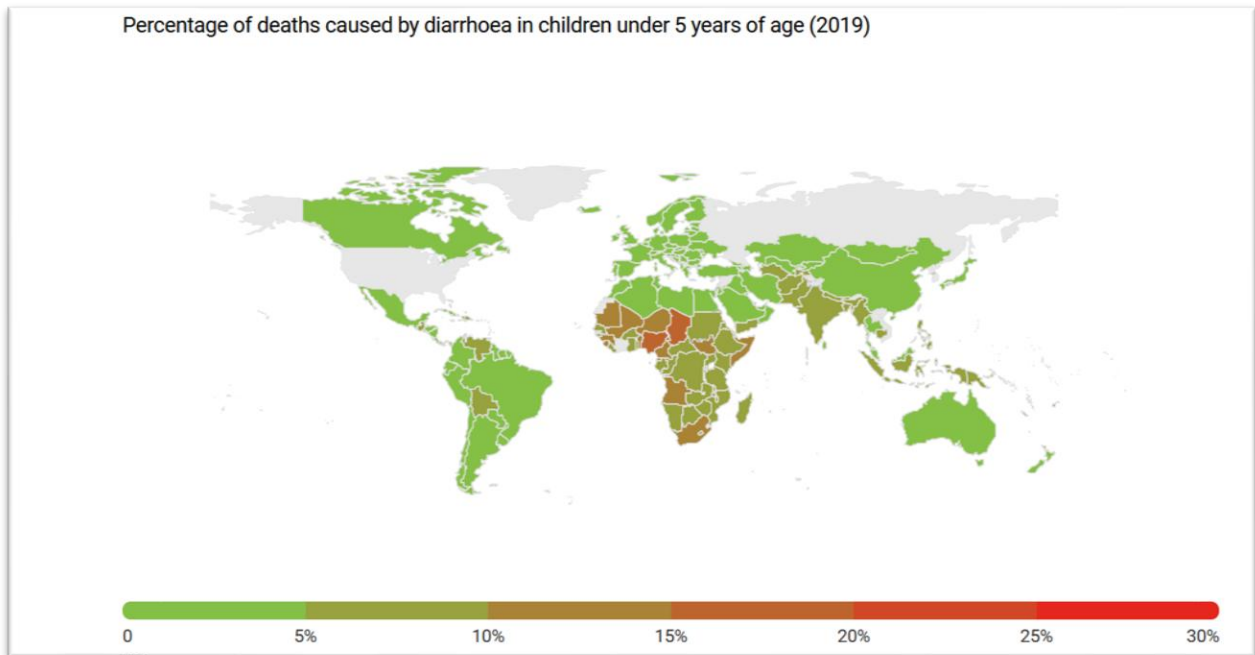
1.1 Background

According to the World Health Organization (WHO), diarrhea is defined as the occurrence of three (3) or more loose or liquid stools within a twenty four (24) hour timeframe, or a more frequent passage than what is typical for an individual (World Health Organization, 2017). Diarrhea is a public health problem globally, having an incidence of circa 1.7 billion childhood diarrheal episodes every year, and a mortality of ~1.5 million making up nearly one in every nine child deaths—more children than AIDS, malaria, and measles put together (Figure 1.1) (CDC, 2018; World Health Organization, 2017). The combination of societal factors (water access, Sanitation, and hygiene variables), and health system factors (healthcare-giver related and drug availability), coupled with limitations on the healthcare system imposed by constraints of resources and infrastructure cause a substantial burden of diarrhea-related morbidity and mortality occurring in low-middle-income countries (LMICs) (Mills, 2014). Diarrhea also poses a great economic burden to households, with the average costs incurred by caretakers for per-episode treatment of diarrhea in LMICS, reported to be \$40 and \$160 for outpatient and inpatients care, respectively (Baral et al., 2020).

The clinical outcomes of diarrhea extend beyond acute dehydration and electrolyte imbalance, and include: longer duration diarrhea; hospitalizations; chronic malnutrition; and mortality. This is further compounded by late care-seeking, care-seeking problems could manifest either as a lack of care-seeking or delayed care-seeking. Delayed care-seeking refers to the time taken from the onset of a symptom (diarrhea) to the first consultation with a healthcare provider (Debsarma et al., 2022). Appropriate and timely care has been reported to reduce preventable morbidity and mortality in

LMICs (Geldsetzer et al., 2014), and improving caretakers care-seeking behavior could contribute significantly to reducing adverse outcomes of childhood diarrhea and the associated mortality in these settings (Sreeramareddy et al., 2006). In LMICs such as Kenya, inadequate management of diarrhea at home (Omore et al., 2013), and delays in care-seeking for moderate-to-severe diarrhea lead to poor outcomes (Kenya National Bureau of Statistics, 2015).

In the backdrop of this diarrheal burden lies health systems that are continually faced with substantial hindrances in offering high quality care, that is reasonably priced, and accessible to everyone (Lewis et al., 2012). eHealth is one of the solutions that can be embraced to mitigate challenges in the health system (Muinga et al., 2020). As per the WHO, eHealth denotes the efficient and secure application of information communication technologies securely and cost-effectively to buttress health and related domains (World Health Organization, 2016). It has been a precedence for WHO from 2005, and it offers the healthcare sector the potential for substantially improving health service efficiency, expansion, and scale up of treatment delivery to many patients in LMICs, resulting in favorable outcomes. The healthcare industry is currently experiencing a significant digital revolution with eHealth at the vanguard of the change (Cihula, 2023). One of the areas driving this transformation is the use of artificial intelligence (AI) which involves computational modeling of human cognitive functions, at the core of AI is machine learning (ML). In this section, poor diarrheal outcomes including longer duration diarrhea, chronic malnutrition and mortality are introduced. Additionally, the paradigm of ML is also discussed.



Source: UN IGME Child and Adolescent Causes of Death Estimation (CA CODE) Project, 2024. © UNICEF.

Figure 1.1: Global Burden of Diarrhea among children aged < 5 years (UNICEF, 2022)

1.1.1 Longer duration diarrhea

Longer duration diarrhea (LDD), a diarrheal illness that extends for seven (7) or more days, is one of the poor outcomes of diarrhea (Schilling et al., 2017). Longer duration diarrhea encompasses prolonged diarrhea which is characterized by a duration of diarrhea between seven (7) and thirteen (13) days, and persistent diarrhea classified as diarrhea episodes lasting for more than fourteen (14) days (Giannattasio et al., 2016). While acute diarrheal illness is often treatable in up to 20% of cases, the illness persists beyond the expected duration of an acute infection eventually progressing to LDD (Alam & Ashraf, 2003). The mean duration of diarrheal episodes in LMICs has been reported to be 4.3 days and 8.4 days among cases in community settings and hospital inpatients, respectively (Lamberti et al., 2012). Infants with LDD face a greater risk of death compared to those with acute diarrheal episodes (Strand et al., 2012; Victora et al., 1992), and it also has dire health outcomes such as: cognitive impairment, growth faltering, and nutrient

deficiency (Bhutta, Nelson, Lee, Tarr, Zablach, Phua, Lindley, Bass, Phillips, & Persistent Diarrhea Working Group, 2008; Niehaus et al., 2002).

1.1.2 Chronic malnutrition

Diarrhea is a prominent etiological factor in childhood malnutrition in LMICs (WHO, 2017), stemming from anorexia, malabsorption, mucosal injury, and nutrient depletion (Ferdous et al., 2013). Evidence suggests that an increased cumulative diarrheal burden in early childhood exerts detrimental effects on nutritional status, and compensatory catch-up growth is generally inadequate to offset the observed deficits (Checkley et al., 2008). Enteric infections result in stunted growth, a condition affecting approximately 20% of children worldwide and 33% of children in LMICs (Guerrant et al., 2013).

Chronic malnutrition is a type of growth faltering caused by an insufficient intake or absorption of essential nutrients over a protracted period of time; its common indicators are stunting and linear growth faltering. Stunting is defined as a height-for-age z-score (HAZ) < -2 (Lenters et al., 2016), whilst linear growth faltering is defined as a decrease of 0.5 HAZ or more ($\Delta \text{HAZ} \geq -0.5$) between enrolment and follow-up. The prevalence of stunting among children aged < 5 years is 29.3% in the East Africa region and 34.5% globally (WHO et al., 2020). In Kenya itself the prevalence is 18% (Kenya National Bureau of Statistics & Ministry of Health, 2023).

Approximately 13.5% of stunting globally is attributable to diarrhea (Danaei et al., 2016). Stunting is associated with a reduced cognitive development and elevated risk of diarrhea, pneumonia, measles and mortality (Benjamin-Chung et al., 2023; Wierzbza & Muhib, 2018). Additionally, the reciprocal relationship between diarrhea and malnutrition is amplified by the immunodeficiency induced by malnutrition, resulting in increased incidence and duration of diarrheal episodes (Figure 1.2). The detrimental effects of this cycle during early childhood may set children on a

path toward increased risk of cardiovascular disease, type 2 diabetes, and obesity in adulthood (Guerrant et al., 2013; ROTA Council, 2019).

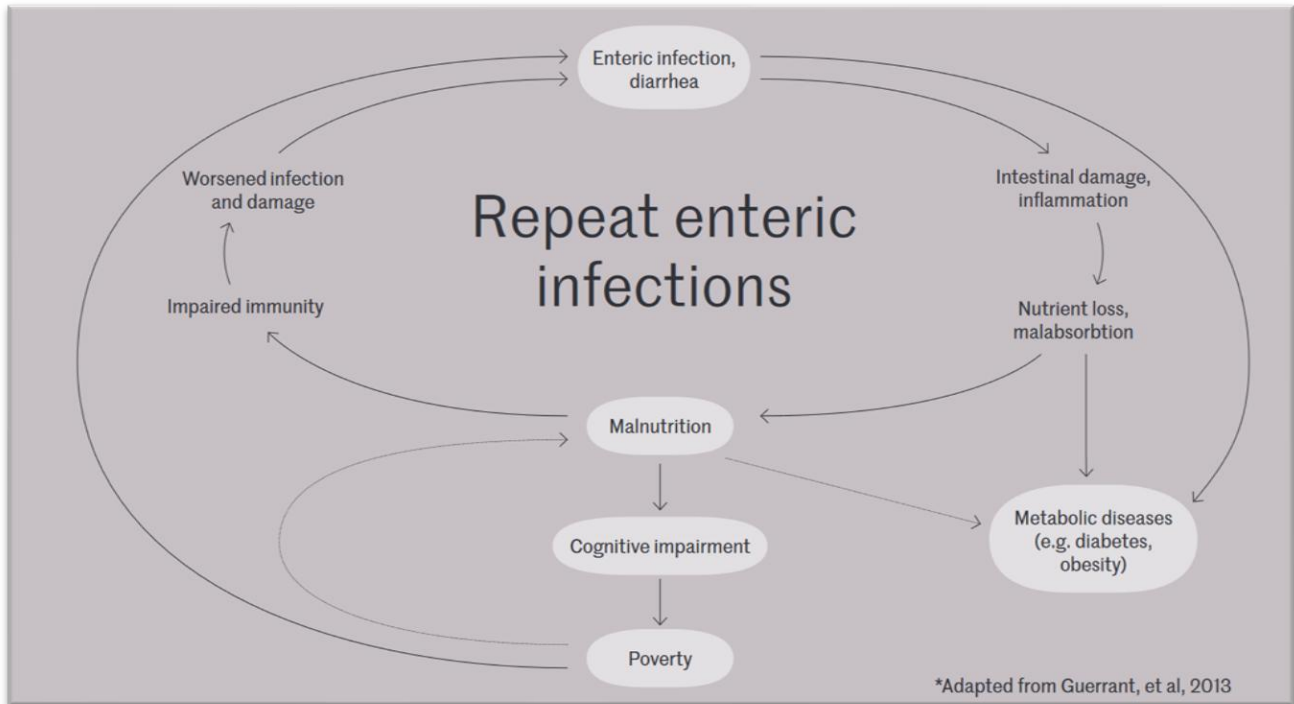
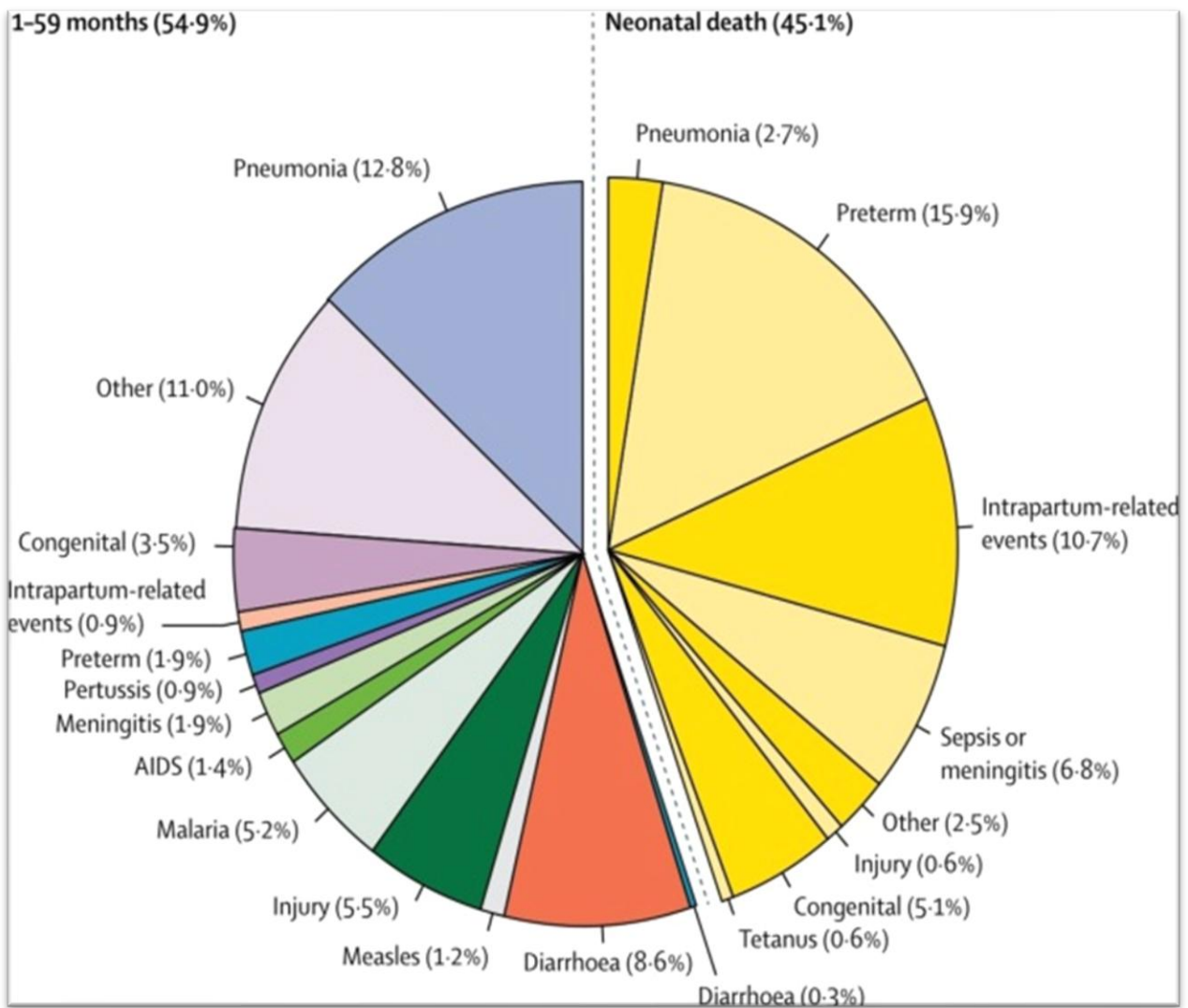


Figure 1.2: The vicious cycle of enteric infections, malnutrition and long-lasting effects (Guerrant et al., 2013)

1.1.3 Mortality

Globally, diarrheal illness is the third leading cause of death among children aged less than five years, causing approximately 9% of deaths within this particular age category (Figure 1.3) (World Health Organization, 2024). This burden of disease is greatest in sub-Saharan Africa (SSA) and South Asia where 90% of all diarrheal deaths occur with rates ranging between 50-150 persons per 100,000 (Troeger et al., 2018; UNICEF, 2022). In Kenya, diarrhea accounts for 16% of all deaths among children aged <5 years (Njuguna & Muruka, 2011). Children presenting with moderate-to-severe diarrhea have an 8.5 fold risk of mortality compared to their counterparts who

are healthy (Pavlinac et al., 2018). Historically, severe dehydration and fluid loss were the predominant drivers of mortality in diarrheal disease; however, recent trends indicate an increasing proportion of deaths attributed to septic bacterial infections (World Health Organization, 2017). Furthermore, malnutrition and impaired immunity increases the risk of life-threatening diarrhea.



Reproduced from Liu et al. (2016) under Creative Commons CC-BY 4.0 license

Figure 1.3: Global causes of under-5 deaths in 2015 (Liu et al., 2016)

1.1.4 Machine Learning

Machine learning (ML) is an arm of artificial intelligence which builds applications that acquire knowledge from data and continuously improve their performance, without requiring additional programming to achieve such improvements (IBM, 2020). Machine learning research is generally structured around several key areas of focus, including:

“ (1) task-driven research—designing and evaluating learning algorithms to enhance performance on specific, predefined tasks, often referred to as the engineering perspective; and (2) cognitive modeling—exploring and simulating human learning processes through computational methods.” (Carbonell et al., 1983).

The prediction tasks for poor diarrheal outcomes were formulated as classification problems with labelled data hence a supervised learning approach was used. The choice of the algorithms used in the study was guided by the knowledge of the dataset, including its structure and complexity, accuracy requirements, processing speed or time and number of features (Sarker, 2021). Since algorithms fall in different families and have different purposes and approaches in the implementation of a classification task, algorithms were selected based on: diversity per family; and prior use in predicting healthcare problems, in order to allow for a comparative evaluation of their performance in the prediction of poor diarrheal outcomes. The algorithms were chosen because of their popularity, potency, wide usage, and variation in performance. Additionally, logistic regression was included as the default model and formed the basis of comparison with other algorithms.

The study employed seven (7) ML algorithms in building predictive models for poor diarrheal outcomes among children aged <5 years. The algorithms used included: white box algorithms [Logistic regression (LR); Naive Bayes (NB)] and black box algorithms [Random Forest (RF), Gradient Boosting (GBM); Support vector machine (SVM); K-nearest neighbors (KNN); Artificial Neural Networks (ANN)]. Among the white box algorithms, LR is a supervised learning

classification algorithm which makes predictions about the probability of an outcome variable which is dichotomous in nature (Shipe et al., 2019). A key characteristic of the NB classifier is its assumption of conditional independence among terms within a category. Based on this assumption, NB estimates category probabilities using the joint probabilities of terms and categories in the dataset. It also exhibits robustness against noise (Langarizadeh & Moghbeli, 2016).

NB shows excellent scalability as the number of features and data points increase, and it offers swift and straightforward implementation and can be utilized in making real-time predictions. Among the black box algorithms, RF is an algorithm that leverages bagging to train multiple decision trees, creating an ensemble that provides enhanced prediction accuracy and stability compared to a single decision tree (Khalilia et al., 2011). RF improves on bagging because it decorrelates the trees with the introduction of splitting on a random subset of features resulting in a model that outperforms any of the individual constituent models. Additionally, it reduces the risk of overfitting and the required training time. GBM is an algorithm that builds strong learners from weak learners in a gradual, additive, and sequential manner offering lots of flexibility allowing for optimization of different loss functions (Zhongheng Zhang, 2019). GBM yields better accuracy, requires less pre-processing and offers higher flexibility due to the availability of many hyperparameter and loss functions.

SVM works through discriminating between two classes through the creation of a hyperplane which perfectly segregates classes following the mathematical transformation of features into a high dimensional space (Yu et al., 2010). SVM can handle both classification and regression on linear and non-linear data. Additionally, it is more effectual in high-dimensional spaces and consumes relatively less memory. SVM is powerful in instances where the dimensions are more than the number of samples. KNN determines the class of a data point by considering the classes

of the k closest data points, using a distance metric (Shee et al., 2014). Its implementation is simple and intuitive to understand. Additionally, it can learn non-linear decision boundaries. ANN are computational networks that simulate the human-brain processes thereby establishing empirical relationships between outcomes and features (Shahid et al., 2019). These algorithms have been previously applied in various classification tasks involving disease prediction in healthcare (Khalilia et al., 2011; Zhongheng Zhang, 2019; Langarizadeh & Moghbeli, 2016; Yu et al., 2010; Shee et al., 2014; Shahid et al., 2019).

1.2 Current State of Predictive Modelling in Diarrheal Disease

Predictive modelling is a technique in which probability, data mining and ML are applied in analyzing patterns in a given set of input data so as to forecast/predict future outcomes or events (Gennatas & Chen, 2021). Predictive modelling in healthcare sets out to determine the likelihood of patient-level or population-level health events. These endeavors assist healthcare workers and policy makers in devising preventative approaches and targeted interventions, leading to a reduction in morbidity and mortality as well as to relieve the societal financial and social burdens (Bandi et al., 2024; Leisman et al., 2020; Wang et al., 2025). Predictive models can support the prompt detection of patients with heightened risk of developing poor outcomes and thereby offer the opportunity for administering enhanced, timely care at a comparable cost, and with outcomes that are more favorable. ML is a paradigm that can be embraced to draw great insights from large datasets using standardized methods consequently providing impartial viewpoints to enhance effectiveness, dependability, and precision (Ayoubi & Foray, 2025; Završnik et al., 2025). ML methodologies are increasingly being employed in public health and clinical settings to expedite the creation of data-driven predictive models, which serve to augment the clinical decision-making process (Chen & Asch, 2017; Jayatilake & Ganegoda, 2021). Furthermore, ML approaches have

demonstrated the ability to offer supplementary insights and comprehension of the data, in such application they can be used to supplement and support traditional statistical analysis. (Ayers, 2016). Notwithstanding the documented successes of ML models in optimizing the implementation of public health interventions on a global scale, currently there are limited research efforts in the development of patient-level predictive models for poor diarrheal outcomes.

Significant strides have been made in predictive modelling for diarrheal disease, with disease incidence forecasts, vaccine-related predictions and disease/pathogen detection leading the endeavors (Ogwel et al., 2023). However, diarrheal outcomes remain under researched with no patient-level predictive models existing in the literature for LDD or diarrheal mortality. While there have been some attempts to develop predictive models for chronic malnutrition, most of these research efforts have looked at all-cause malnutrition and they have been constrained by limited use of clinical data and use of geographical regions as features in these models which creates algorithmic bias (Bitew et al., 2021; Fenta et al., 2021; Talukder & Ahammed, 2020). Two studies have looked at malnutrition as an outcome of diarrhea but it only used logistic regression as the modelling technique (Ahmed et al., 2023; Brander et al., 2019). The current situation stated above presents an opportunity and gap that is filled by the study of the thesis. There exists much room for patient-level predictive modelling of poor diarrheal outcomes to gain ascendancy with the availability of rich, large and diverse diarrheal datasets, from epidemiological studies, that could support these predictions. Additionally, the advancement of digital technologies, and the expected benefits for patient and the healthcare systems by such models are potential stimuli in the development of these predictive models.

1.3 Problem Statement

The burden of disease for diarrhea in Siaya County, Western Kenya is great, especially among children aged <5 years. The prevalence of LDD in this age group has been reported to be 42% (Schilling et al., 2017), whilst chronic malnutrition associated with diarrhea is at 24.2 % (Brander et al., 2019). The diarrheal mortality rate among children aged <5 years was 176 per 100,000; the highest burden was in children aged 6-11 months 558 per 100,000 in the Karemo area of Siaya County (Omore et al., 2019). Finally, 42% of the caretakers of children aged under 5 years did not seek advice or treatment from a healthcare provider for childhood diarrhea (Kenya National Bureau of Statistics, 2015). Reducing the burden of diarrheal diseases is clearly a top priority to both the national and county governments in line with the third goal on health, of the Sustainable Development Goals of the United Nations (SDGs). Currently, close to half of caretakers fail to seek care for their children suffering from diarrheal episodes, and even those who do often delay in doing it. However, after delays in seeking care, treatment tends to have less favorable outcomes, with many complications, and high costs. Additionally, limited diagnostic capacity in many LMICs, such as Kenya (McNerney, 2015), impede the diagnosis and treatment of diarrheal disease. Siaya County also has very low nurse and doctor to population ratios (1:1,697; 1:38,511) (County Government of Siaya, 2017). This is further compounded by a demanding work environment, burnout, and pressure of attending to severely ill patients which might reduce their clinical judgement and performance (Hartzband & Groopman, 2020; Kokonya, 2014; Motluk, 2018). Whilst booklets of the Integrated Management of Childhood Illness (IMCI) guidelines are available at health facilities across the country, no algorithms exist, traditional or digital, to predict poor outcomes of diarrheal illness. The rationale for predictive models is to rapidly identify patients with an elevated likelihood of developing poor diarrheal outcomes and consequently

administer the appropriate care in a timely manner at a comparable cost, for more favorable outcomes.

1.4 Research Purpose and Objectives

The aim of this study was to use explainable ML methods and approaches to develop predictive models for early identification and intervention in children aged < 60 months at increased risk of developing poor outcomes of diarrheal illness in Siaya County. Here are the specific objectives of the study:

1. Determine the demographic, clinical and socio-economic features that predict longer duration diarrhea among children aged < 60 months.
2. Determine the demographic, clinical, and socio-economic features that predict chronic malnutrition following a diarrheal illness among children aged < 60 months.
3. Determine the demographic, clinical, and socio-economic features that predict diarrhea-associated mortality among inpatient children aged < 60 months.
4. Develop and evaluate (internal and temporal validation) the performance of predictive models for each of the three poor diarrheal outcomes among children aged < 60 months using machine learning.
5. Design an R-shiny product suite for predictive models of poor diarrheal outcomes.
6. Make a contribution to the knowledge base on predictive modelling in pediatric enteric diseases from methodological and health services dimensions.
7. Delineate future research directions on predictive modelling in pediatric enteric diseases.

1.5 Research Questions

Here are the research questions:

1. **RQ1:** What features predict longer duration diarrhea among children aged <60 months?
2. **RQ2:** What features predict chronic malnutrition following a diarrheal illness among children aged <60 months?

3. **RQ3:** What features predict diarrhea-associated mortality among inpatient children aged <60 months?
4. **RQ4:** What machine learning model best predicts longer duration diarrhea among children aged <60 months?
5. **RQ5:** What machine learning model best predicts chronic malnutrition following a diarrheal illness among children aged <60 months?
6. **RQ6:** What machine learning model best predicts diarrhea-associated mortality among inpatient children aged <60 months?
7. **RQ7:** Is it possible for R-Shiny to adequately support the design of a product suite for predictive models of poor diarrheal outcomes?
8. **RQ8:** What are the contributions of this study to the knowledge base on predictive modelling in pediatric enteric diseases from methodological and health services dimensions?
9. **RQ9:** What are the future research directions on predictive modelling in pediatric enteric diseases?

1.6 Scope

This study focused on the prediction of poor diarrheal outcomes (LDD, chronic malnutrition, and mortality) using clinical, anthropometric, water, sanitation and hygiene (WASH), and socio-economic data and employing ML algorithms. The study focused on diarrheal disease in children aged < 60 months within the geographical context of Siaya County in rural western Kenya. a region with a great burden of childhood diarrhea. This target population in this study was children <5 years as this age group bears the greatest burden of diarrheal disease. The spectral and temporal scope of the study covered children with moderate-to-severe diarrhea (MSD) enrolled in the time

period 2015-2016 for LDD and chronic malnutrition with temporal validation done, using children aged 6-35 months with medically-attended diarrhea (MAD), enrolled from August 2022- July 2023. Moreover, for the prediction of mortality the focus was on inpatient children admitted for acute gastroenteritis, for the time period 2010-2020.

Methodologically, the study involved feature selection using a blend of statistical measures (chi-square and correlation tests) and the Boruta package (Kursa & Rudnicki, 2020). Formulation of predictive models used ML algorithms (RF, GBM, NB, LR, SVM, KNN and ANN). Subsampling techniques were applied within the resampling procedures to address class imbalance, specifically using oversampling of the minority class. Oversampling was chosen due to the modest sample size, as it increases representation of the minority outcome without discarding information from the majority class, thereby preserving statistical power during model training.

The development of the proof-of-concept clinical decision support dashboard integrated the predictive models and explainer objects for each diarrheal outcome encompassing the design, implementation, and visualization of model outputs within an interactive dashboard capable of facilitating interpretation by clinicians. However, the user testing and real-world integration into clinical workflows were beyond the scope of this study; the emphasis was on demonstrating the feasibility, functionality, and interpretability of the dashboard as a research and development prototype.

1.7 Ethical Considerations

This research study utilized de-identified data from the Vaccine Impact on Diarrhea in Africa (VIDA) study (SERU#2996), Rotavirus surveillance study (SERU# 1801), and Enterics for Global

Health Shigella Surveillance Study (SERU#4362) whose protocols were reviewed and approved by the Scientific and Ethics Review Unit (SERU) of the Kenya Medical Research Institute (KEMRI). No additional consent was required, as the initial consents and approvals for each of these studies covered the use of the collected data for future research. Additionally, ethical approval for undertaking this study was granted by the research ethics committee of the University of South Africa (UNISA), College of Agricultural Sciences (2023/CAES_HREC/2192).

1.8 Thesis Outline

Chapter 2: Presents the literature review of the public health context surrounding each of the poor diarrheal outcomes spanning the burden of disease, the risk factors, and clinical management.

Chapter 3: Presents the literature review of the predictive modeling context and background for each of the poor diarrheal outcomes. This covers key considerations for selecting a deployment framework, as well as the theoretical and conceptual framework guiding the study.

Chapter 4: Discusses the research methodology that was utilized in addressing the study objectives. It outlines the study sites and locations of data collection, the research process and offers justification for the selected methods. It also discusses the data itself, how it was collected and how modelling was done.

Chapter 5: Presents the results of the predictive models that were built and their evaluation for each diarrheal outcome. The discussion of results also takes place here.

Chapter 6: Synthesizes reflections on the study's broader significance, offering a high-level synopsis of its methodological and health services contributions, unexpected findings and a forward-looking agenda.

Chapter 7: Outlines the summary of key findings, conclusions, and recommendations of the study. This chapter also considers future research directions emanating from the study.

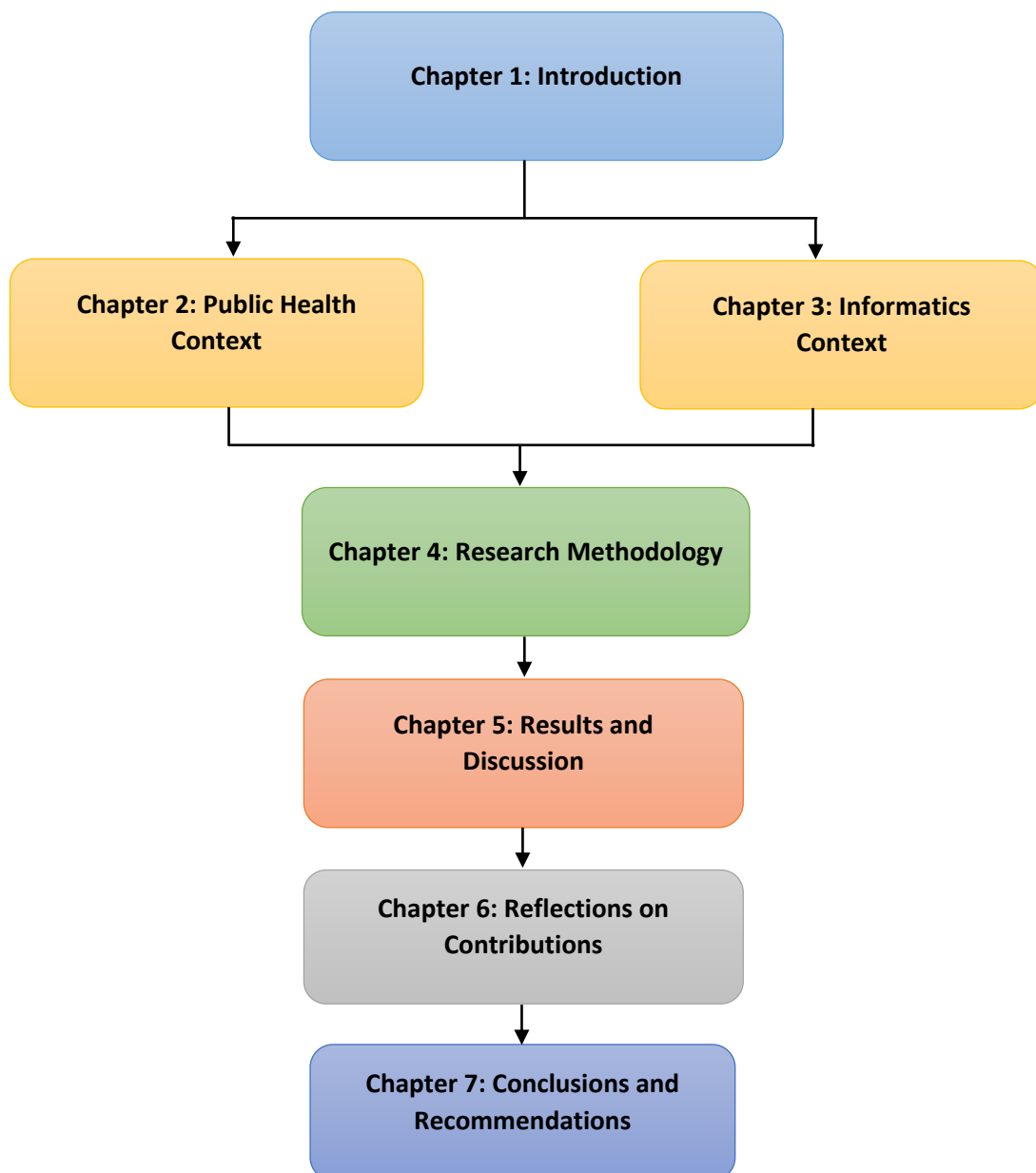


Figure 1.4: Schematic diagram of thesis outline

2. CHAPTER TWO: THE PUBLIC HEALTH CONTEXT

2.1 Introduction

This chapter embarks on a thorough exploration of the public health landscape surrounding poor diarrheal health outcomes. It delves into the health burden these outcomes impose, as well as the multifaceted strategies employed for clinical management and the identification of associated risk factors. Moreover, the chapter systematically identifies pertinent studies, elucidating critical knowledge gaps in the epidemiology and clinical management of poor diarrheal outcomes.

2.2 Longer Duration Diarrhea

2.2.1 Burden

The extent and impact of a health issue is captured in the term burden of disease (Hessel, 2008). This burden is assessed through various indicators, including: morbidity; mortality; and financial costs. When considering diarrhea episodes, longer duration diarrhea (LDD), is described as a diarrheal illness with a duration of seven (7) days or more (Schilling et al., 2017). Typically, acute diarrheal episodes resolve within seven days. However, if the duration extends to between seven (7) and thirteen (13) days, it is referred to as prolonged diarrhea, whereas if it persists for fourteen (14) days or more, it is labeled as persistent diarrhea (Giannattasio et al., 2016; Hossain et al., 2022). It's important to note that LDD encompasses both prolonged diarrhea (mild form of disease) and persistent diarrhea, ≥ 14 days, (severe form) as shown in Table 2.1.

Table 2.1: Definitions of diarrheal illnesses (Giannattasio et al., 2016)

Diarrhea	The most commonly recognized definition of diarrhea is based on World Health Organization parameters and define diarrhea by the passage of 3 or looser than normal stools in the preceding 24-hour period. An episode of diarrhea is defined as lasting 1 day or more and usually ends after at least 2 days without diarrhea.
Acute diarrhea	Episode of self-limiting diarrhea with acute onset, typically lasting 5 to 7 days. In most cases, it is due to an intestinal infection and may be combined to fever and vomiting, meeting the definition of acute gastroenteritis. Acute diarrhea may be also related to extra-intestinal infections (i.e. urinary infection, viral respiratory infections), food-poisoning, iatrogenic intestinal damage (i.e. chemotherapy, radiotherapy) or other intestinal and extra-intestinal diseases such as acute appendicitis.
Prolonged diarrhea	Acute onset diarrhea lasting from 7 to 14 days not covering the definition of persistent diarrhea. It is usually due to persistent infections or to post-infectious intestinal damage (i.e. carbohydrate malabsorption, small intestine bacterial overgrowth) that may prolong the duration of diarrhea behind the expected time. Some experts refer to this as acute-protracted diarrhea.
Persistent diarrhea	Diarrhea lasting 14 days or more, usually associated with weight loss, ultimately leading to severe nutritional impairment and that may require clinical nutrition. The classical definition of persistent diarrhea was intended to exclude some causes of chronic diarrhea such as celiac disease or inflammatory bowel diseases.
Chronic diarrhea	In many contexts chronic diarrhea is a synonymous of persistent diarrhea. The World Health Organization uses this definition rather than persistent diarrhea. However, chronic diarrhea usually does not have an acute onset and is the manifestation of structural and inflammatory bowel disorders. Some experts refer to chronic diarrhea in case of episodes lasting more than 4 weeks.
Post-infectious diarrhea	Acute onset diarrhea lasting 7 to 14 days and following an episode of acute gastroenteritis. This definition is covered by prolonged diarrhea.
Intractable diarrhea	Non-infectious diarrhea lasting more than 14 days, intractable despite extensive hospital therapy. Typical of young infants, usually below 3 months (but not only). Typically needs intravenous fluids or clinical nutrition and is related to high mortality.
Congenital diarrhea	Congenital diarrhea is an inherited enteropathy with a typical onset early in life. For many of these conditions, severe chronic diarrhea represents the main clinical manifestation, while in others, diarrhea is only a component of a more complex multi-organ or systemic disease.

Research indicates that prolonged diarrhea, range 7 – 14 days duration, is twice as prevalent as persistent diarrhea, ≥ 14 days (Moore et al., 2010). The conclusion of an illness of diarrhea is characterized by absence of diarrhea for at least two days in a row, aligning with earlier research (Md et al., 2002; Morris et al., 1994; Platts-Mills et al., 2018; Schilling et al., 2017). Importantly, it has been observed that when an acute diarrhea episode progresses to prolonged diarrhea (7-13 days), the risk of evolving into persistent diarrhea, ≥ 14 days, increases six-fold (Giannattasio et

al., 2016). A systematic review focusing on the episode length and intensity of diarrhea revealed that among children aged less than 60 months, the mean length of a diarrheal episode was 4.3 days in communal setups and 8.4 days when hospitalized (Lamberti et al., 2012). In industrialized nations, the length of an episode was shorter, standing at 2.2 days for the average and 1 day for the 50th percentile.

In developing countries, it has been observed that as many as 20% of diarrhea cases exceed the typical duration expected for acute infectious diarrhea, leading to LDD which has been defined as having duration ≥ 7 days (Alam & Ashraf, 2003; Sarker et al., 2017). Although LDD constitutes a smaller proportion of childhood diarrhea cases, it significantly contributes to more than half of the days marked by diarrhea symptoms (Moore et al., 2010; Strand et al., 2012). An epidemiological investigation done in Brazil highlighted that children contributed significantly to the burden of diarrheal diseases, particularly those who endured diarrhea for more than 14 days in a year (Lima et al., 2000). Furthermore, in an urban slum in Northeastern Brazil, 11% of cases of diarrheal infections were classified as persistent diarrhea (Lima et al., 1992). A prospective longitudinal study aimed at understanding the epidemiological landscape of persistent diarrhea in children from Guatemala found that the attack rate of persistent diarrhea was 0.728 illness episodes per child per year, with the highest occurrence reported in infants aged less than six months (Cruz et al., 1992). Research in Peru estimated an attack rate of 0.25 episodes of persistent diarrhea per child per year (Lanata et al., 1991). A Vietnamese epidemiological investigation with a cohort design estimated the prevalence of persistent diarrhea to be 66.5% among children seeking care for diarrheal illness (Ngan et al., 1992). In Kenya the total percentage of children whose diarrhea was classified LDD, was 42% of all cases of diarrhea recorded in the GEMS study. Of this figure

7% had the more severe persistent diarrhea which is longer in duration, ≥ 14 days duration (Schilling et al., 2017).

Previous studies have documented an elevated risk of death among children with LDD in comparison to those with diarrheal episodes of an acute nature (Strand et al., 2012; Victora et al., 1992). Specifically, persistent diarrhea, ≥ 14 days, is indicative of children with a significant diarrheal morbidity which accounts for between 36% to 54% of all diarrheal deaths (Bhutta et al., 2004). LDD has also been shown to have dire health outcomes: cognitive impairment, growth retardation and deficiencies in essential nutrients (Bhutta, Nelson, Lee, Tarr, Zablach, Phua, Lindley, Bass, Phillips, & Persistent Diarrhea Working Group, 2008; Niehaus et al., 2002). Additionally, sequelae of LDD includes persistent health problems, occasioned possibly by impaired absorption of essential nutrients due to villous atrophy, epithelial damage and submucosal inflammatory response. Persistent diarrhea, ≥ 14 days, accounts for the loss of approximately 3,000,000 disability adjusted life-years every year (Bhutta, Nelson, Lee, Tarr, Zablach, Phua, Lindley, Bass, Phillips, & Group, 2008). Child growth and mucosal immunity are affected by prolonged diarrhea, which leads to dysbiosis, a change in the composition of the gut microbiome, and increases intestinal permeability. Regardless of the etiological agents and correlates, children with prolonged diarrhea are susceptible to nutritional imbalances, lack of essential micronutrients, elevated likelihood of developing persistent diarrhea recurrent infections, and impaired immune system function (Giannattasio et al., 2016).

One of the causes of persistent diarrhea, is recurrent infection, prevalent in environments characterized by inadequate hygiene (Pawlowski et al., 2009). Consequently, recurrent diarrhea is explored because of its connection to LDD. Recurrent diarrhea is defined as subsequent diarrheal episode after the initial episode following two (2) consecutive diarrhea free days (Aluisio et al.,

2015; Morris et al., 1994). Unlike LDD, recurrent diarrhea has limited literature on its burden and epidemiology. A study looking at the correlates of recurrent diarrheal episodes in Afghanistan reported that the rate of new diarrhea episodes was approximately 3.5 episodes per child per year, with between 50% and 48% higher likelihood of experiencing repeated episodes (Aluisio et al., 2015). Another study conducted in Egypt, reported the mean duration of diarrheal episodes as 4.8 days with 33.6% of children having recurrent diarrhea (Yassin, 2000). Despite the limited research, existing literature have reported recurrence of diarrhea as posing a critical public health threat in LMICs (Patel & Rana, 2006; Rodríguez et al., 2011; Saad et al., 2016)

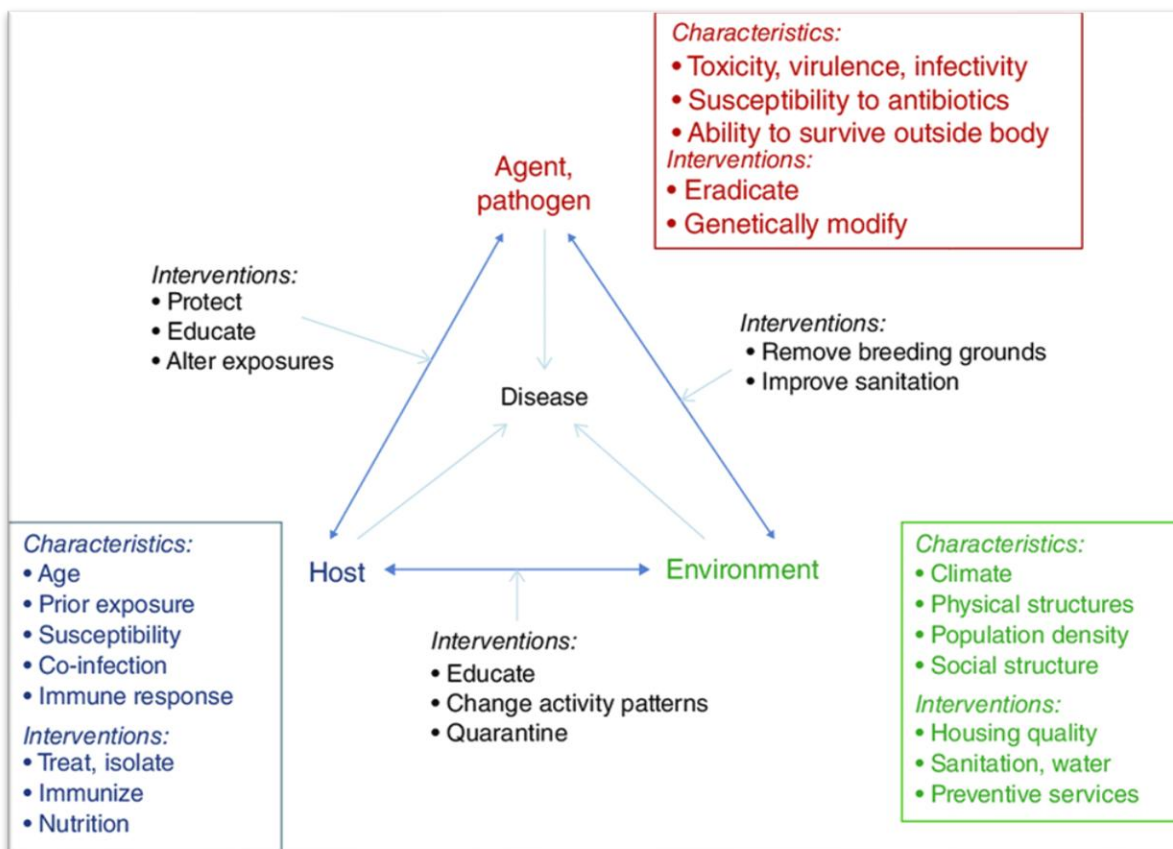
A systematic review of the economic burden of diarrhea in LMICs, between 2006-2018, estimated the mean cost of illness to be \$37 and \$160 for each clinic visit and hospitalization, respectively (Baral et al., 2020). A study in Bangladesh quantified the economic burden of diarrheal disease as \$62 per average cost per episode for caregivers; \$29 and \$34 direct and indirect costs, respectively. While the societal perspective was estimated at averagely \$71 (Hasan et al., 2021). This translated to over 46% of households faced with financial challenges from diarrheal disease. Additionally, in Indonesia the mean direct medical cost for the management of acute diarrheal episode was estimated as \$99.8 (Standard Deviation [SD] $\pm$ \$56.8) and \$7.6 (SD $\pm$ \$4.3) for hospitalizations and clinic visits, respectively (Thobari et al., 2022). Niyibitegeka et al. in their evaluation of the costs associated with diarrheal illness in Burundian children established that outpatient visit and hospitalization costs per episode of diarrhea were \$59.87 and \$292, respectively (Niyibitegeka et al., 2021). Ngabo et al. in their study assessed the costs related to inpatient diarrheal episodes in children in Rwanda (Ngabo et al., 2016). They estimated the overall cost of a diarrheal episode to be \$101, with 65% of costs borne by the household. In this study in Rwanda households, the poorest quintile spent 110% of monthly income according to this costing structure. It is reasonable

to assume that a proportion of these cost estimates for both outpatient and inpatient costs, can be attributed to LDD and recurrent diarrhea.

The review of existing literature reinforces the substantial burden posed by LDD and recurrent diarrhea, particularly in LMICs such as Kenya. This burden encompasses both morbidity and financial costs, emphasizing the urgency of devising innovative interventions aimed at alleviating this impact. Such interventions are crucial in aligning with the third objective of the UN's global framework for sustainable development (SDGs), which focuses on promoting well-being and ensuring healthy lives. This underlines the pressing need for measures to mitigate the adverse effects of LDD and recurrent diarrhea in LMICs.

2.2.2 Risk factors

A risk factor is a characteristic, condition, or behavior that elevates one's susceptibility to a disease or infection (Melissa, 2021). The epidemiological triad (Figure 2.1) of microbial, host, and environmental factors interact to cause diarrheal illness. Disease transmission occurs in this model when a microbial interacts with a vulnerable host, in an environment that is supportive of the spread of the microbial agent from a reservoir to the host (CDC, 2021). Microbial factors refer to microorganisms capable of causing an infection or pathogens, these include: bacteria; viruses; fungi; and protozoa. For an infection to be established in a host, these pathogenic agents colonize the skin or the inner lining of the digestive system and circumvent the body's initial immune response from the epithelial layers and surrounding tissues (Charles A Janeway et al., 2001). Host factors refer to intrinsic characteristics of a human that influence an individual's exposure, susceptibility, or response to a microbial agent.



Reproduced from Johnson & Kaneene (2018), with permission from John Wiley & Sons. © 2018 John Wiley & Sons.

Figure 2.1: The epidemiologic triad of infectious diseases (Johnson & Kaneene, 2018).

These include: demographic characteristics; hygiene; life-style choices; genetic make-up; nutritional and immunological status; and psychological makeup. Lastly, environmental factors are non-inherent drivers that influence microbes and the chance of contact. They include physical, biological, and socioeconomic drivers including but not limited to overcrowding in habitation, sanitation issues, geological and climatic factors, vectors that enable the transmission of the pathogen, and availability of healthcare services (CDC, 2021).

A review of risk factors for LDD and recurrent diarrhea are presented here. Strand et al. in their assessment of risk factors for extended diarrhea among Nepalese children identified age (infants odds ratio [OR] 95%CI 17[3.5-83]; toddlers 9.9 [2.1-47.0]); and Not breastfeeding (OR [95%CI] 9.3 [2.4-35.7]) (Strand et al., 2012). Additionally, increasing stool frequency and monsoon weather season elevated the likelihood of prolonged episodes. A study in India assessing the drivers of diarrhea episode length identified younger children (OR [95% CI]: 1.1 [1.0–1.2]), reported or measured fever (OR [95% CI]: 1.1 [1.0–1.2]), being dehydrated (OR [95% CI]: 1.3 [1.1–1.6]), dysenteric diarrhea (OR [95% CI]: 1.4 [1.1–1.8]), patients who obtained medical treatment (OR [95% CI]: 1.2 [1.1–1.4]), underweight ≤ 2 (OR 1.25, 95% CI 1.07–1.46), females (OR [95% CI]: 2.30 [1.20–4.56]), and length of diarrhea episode pre-enrolment (OR [95%CI]: 1.05 [1.03–1.08]), to increase the risk of prolonged diarrhea (Patel et al., 2012). Schilling et al. in their evaluation of factors associated with duration of diarrheal disease among Kenyan children identified *Cryptosporidium* infection (hazard ratio [HR]: 0.87, $P = 0.04$), lack of drinking water from a source that is protected (HR: 0.88, $P = 0.04$), and stunting at baseline (HR: 0.03, $P < 0.001$), as risk factors (Schilling et al., 2017).

Lins et al. also studied risk factors for persistent diarrhea in Brazilian infants and they identified an increased risk for persistent diarrhea, ≥ 14 days, among infants with: dysentery; fever at the onset of diarrhea, fasting, prior use of antibiotics, perianal hyperemia, and socio-economic status (Lins et al., 2003). Another study in Brazil looked at risk factors for persistent diarrhea, ≥ 14 days, among hospitalized children, they established that maternal education (Point estimate (OR) = 10.70), the number of diarrhea episodes in the three months preceding hospital admission (Point estimate (OR) = 3.70), neonatal weight (Point estimate (OR) = 6.40), being exclusively breastfed (OR = 2.30), underweight (OR = 20.02), wasting (OR = 21.13), and stunting (OR = 6.64) were

predictors of LDD (Pinto et al., 1998). A study in Peru determined being malnourished, suppressed immune response, child's age, and an increment in the preceding diarrheal burden as significant predictors of persistent diarrhea, ≥ 14 days, (Ochoa et al., 2004). The same study identified the following etiologic agents as risk factors: enteropathogenic *E. coli*; enteroaggregative *E. coli*; *Giardia*; *Cryptosporidium*; *Cyclospora*. A study in Cuba identified these variables as drivers of persistent diarrhea, >14 days, in children aged < 60 months: age; previous diarrheal episodes; malnutrition; early introduction of animal milk; prior use of antibiotics ; etiologic agents (enteric pathogens) (Sagaro, 2016).

Shahid et al (1988) also assessed covariates with significant associations with persistent diarrhea, ≥ 14 days, among children in Bangladesh (Shahid et al., 1988). They established that bloody or mucoid diarrhea, pneumonia/infections of the lower respiratory system, being malnourished, hypovitaminosis A, child age and prior antibiotic usage, were strongly linked to the development of persistent diarrhea. Another study in India identified the following risk factors for persistent diarrhea: previous antibiotic usage; hypovitaminosis A; being malnourished; and lower respiratory infection (Umamaheswari et al., 2010). A study, still in India, identified increasing stool frequency ($>10/\text{day}$), weight loss, persistent dehydration, injudicious use of antibiotics, and malnutrition as predictors of PD (Alam et al., 1997).

Another study also assessed risk factors of persistent diarrhea, ≥ 14 days, among children seeking care at a tertiary hospital in Dhaka, Bangladesh. It reported prior recent diarrhea illness, poor nutritional status, children who were not exclusively breastfed as per recommended duration, prior antibiotic use, and access to unimproved drinking water to be significantly associated with persistent diarrhea (Roy et al., 2019). A review of literature also established the etiologic agents associated with PD to be: enteroaggregative *Escherichia (E.) coli*; *Cryptosporidium*; *Shigella*;

Salmonella; enteropathogenic *E. coli*; and *Giardia lamblia* (Lima & Guerrant, 1992a). Durairaj et al. explored the clinical profiles and predictors of persistent diarrhea among children aged < 60 months, in a tertiary care facility in an urban setting in India. They established that dietary protein and energy deficiency (OR [95% CI]: 1.8 [1.42-2.34]), inappropriate antibiotic usage (OR [95% CI]: 2.42 [1.20-4.88]), infections arising from non-gastrointestinal sources (OR [95% CI]: 2.28 [1.17-4.45]), and consuming unimproved drinking water (OR [95% CI]: 2.74 [1.23-6.15]) were significantly associated with persistent diarrhea (Durairaj et al., 2017). They also found *E. Coli* to be the main etiologic agent associated with persistent diarrhea. A study conducted in Guinea-Bissau found weaning; lack of maternal education; and having pigs in the home as risk factors for PD (Mølbaek et al., 1997).

Prolonged diarrhea, lasting 7-13 days, has been shown to be more prevalent in infants and toddlers within the age range of 6 to 23 months peaking in the second semester of life (Giannattasio et al., 2016). Additionally, susceptibility to persistent diarrhea, ≥ 14 days, increased by a factor of two for preschoolers, if a child had prolonged diarrhea in the first year of life. Furthermore, children presenting with severe diarrheal episodes and bloody or mucoid stool had an elevated risk of developing prolonged diarrhea, in comparison with those who had mild and moderate diarrhea. In addition, children not breastfed, early weaning, underlying malnutrition, vitamin deficiency, wasting, WASH variables, caretaker literacy, and socio-economic standing were also, identified as correlates of prolonged diarrhea. A study of prolonged diarrhea, lasting 7-13 days, in Brazilian children showed that age, caretaker education, early weaning, stunting, and wasting were associated with prolonged diarrheal episodes (Moore et al., 2010). They also found *Cryptosporidium* and *Shigella* infections to be predictors of prolonged diarrhea. Another study in Bangladeshi children that assessed the incidence and the determinants of prolonged diarrhea

observed that age, bloody or mucoid stools, warmer season (April through September), household use of antibiotics prior to presentation at health facility, and children with recent diarrheal episodes within the past month, were more likely to experience prolonged diarrhea (Hossain et al., 2022). A study on the risk factors of prolonged diarrheal illness among Kenyan children in an urban slum identified the following behavioral factors as predictors: open water containers (relative risk (RR [95% CI]): 0.8 [0.6–0.9]); non-administration of fluids (RR [95% CI]: 1.5 [1.1–1.8]); and dietary supplementation with diluted cow's milk for the initial three days of the episode (RR [95% CI] 1.3 [1.0–1.6]) (Mirza et al., 1997).

A study investigating the determinants of recurrent diarrhea episodes in Afghani children established that susceptibility to recurrent diarrhea, was reduced among children of caretakers with: secondary education or higher (Adjusted Hazard ratios [aHR] [95% CI]: 0.80 [0.70–0.92]); those in households in which mothers washed their hands with soap (aHR [95% CI]: 0.84 [0.75–0.93]); and with improved waste disposal systems (aHR [95% CI]: = 0.80 [0.64–0.92]). Additionally, malnutrition (aHR [95% CI]: 1.20 [1.0–1.30]) and poor socio-economic status (aHR [95% CI]: 1.30 [1.10–1.40]) increased the probability of recurrence (Aluisio et al., 2015). A study looking at morbidity and predictors of diarrheal illness in Egyptian children identified the following variables to have significant association with recurrent diarrhea: dietary meat intake within the household; maternal age at childbirth; child's age; and paternal literacy status (Yassin, 2000).

The review of literature showed largely similar risk factors for LDD, prolonged diarrhea, PD and recurrent diarrhea. A summary of the microbial, host and environmental factors associated with LDD are presented in Table 2.2. The evidence presented above was undertaken using standard statistical approaches. However, feature selection is conducted prior to the development of the

LDD predictive models in this research, by using a blend of statistical measures (Chi-square and correlation) and Boruta feature selection.

Table 2.2: Risk factors for diarrheal duration in children aged < 60 months: A summary of evidence

Author	Country	Outcome assessed	Risk factors identified
Strand et al., 2012	Nepal	Longer duration diarrhea	Age; not breastfeeding; stool frequency; and monsoon weather season.
A. B. Patel et al., 2012	India	Longer duration diarrhea	Age; fever; dehydration; dysentery; prior antibiotic use; gender; underweight; and diarrheal duration before enrolment.
Schilling et al., 2017	Kenya	Longer duration diarrhea	<i>Cryptosporidium</i> ; unimproved drinking water source; and stunting.
Lins et al., 2003	Brazil	Persistent diarrhea	Dysentery; fever at the onset of diarrhea; fasting; prior use of antibiotics; perianal hyperemia; and socio-economic status.
Pinto et al., 1998	Brazil	Persistent diarrhea	Maternal education; the number of diarrhea episodes in the three months preceding hospital admission; neonatal weight; the exclusive breast feeding; underweight; wasting; and stunting.
Ochoa et al., 2004	Peru	Persistent diarrhea	Malnutrition; immunosuppression; age; an increase in the preceding diarrhea burdens; enteropathogenic <i>E. coli</i> ; enteroaggregative <i>E. coli</i> ; <i>Giardia</i> ; <i>Cryptosporidium</i> ; <i>Cyclospora</i> .
Sagaro, 2016	Cuba	Persistent diarrhea	Age; previous diarrheal episodes; malnutrition; early introduction of animal milk; prior use of antibiotics; and etiologic agents (enteric pathogens).
Shahid et al., 1988	Bangladesh	Persistent diarrhea	Bloody or mucoid diarrhea; pneumonia/infections of the lower respiratory system; being malnourished; hypovitaminosis A; age; and prior antibiotic usage.
Roy et al., 2019	Bangladesh	Persistent diarrhea	Prior recent diarrhea illness; poor nutritional status; children who were not exclusively breastfed as per recommended duration; prior antibiotic; consumption of unimproved drinking water.
Umamaheswari et al., 2010	India	Persistent diarrhea	Previous antibiotic usage; hypovitaminosis A; being malnourished; and lower respiratory infection.
S. Alam et al., 1997	India	Persistent diarrhea	Increasing stool frequency (>10/day); weight loss; being dehydrated over a protracted period; inappropriate usage of antibiotics; and being malnourished.

A. A. M. Lima & Guerrant, 1992	LMICs	Persistent diarrhea	Enteropathogenic <i>Escherichia (E.) coli</i> ; <i>Cryptosporidium</i> ; <i>Shigella</i> , <i>Salmonella</i> , enteropathogenic <i>E. coli</i> ; <i>Giardia lamblia</i> .
Durairaj et al., 2017	India	Persistent diarrhea	Malnutrition; irrational antibiotic use; parenteral infection; use of unsafe drinking water; and <i>E. Coli</i> .
Mølbak et al., 1997	Guinea-Bissau	Persistent diarrhea	Weaning; lack of maternal education; and having pigs in the home.
Giannattasio et al., 2016	-	Prolonged diarrhea	Age; severe diarrheal episodes; bloody or mucoid stool; not breastfed, early weaning; underlying malnutrition; vitamin deficiency; wasting; water sanitation; and hygiene (WASH) variables; caretaker education; and socio-economic status.
Moore et al., 2010	Brazil	Prolonged diarrhea	Age; caretaker education; early weaning; stunting; wasting; <i>Cryptosporidium</i> ; and <i>Shigella</i> .
Hossain et al., 2022	Bangladesh	Prolonged diarrhea	Age; bloody or mucoid stool; warmer season (April through September); household use of antibiotics prior to presentation at health facility; and children with recent diarrhea episodes; within the past month.
Mirza et al., 1997	Kenya	Prolonged diarrhea	Uncovered water containers; non-administration of fluids; and administering diluted cow's milk for the first three days of a diarrheal episode.
Aluisio et al., 2015	Afghanistan	Recurrent diarrhea	Caretaker education; maternal handwashing; improved sanitation facilities; malnutrition; and socio-economic status.
Yassin, 2000	Egypt	Recurrent diarrhea	Dietary meat intake within the household; maternal age at childbirth; child's age; and paternal literacy status.

2.2.3 Pathogenesis, Diagnosis and Management

The pathogenesis of prolonged diarrheal episodes is not only complex but also multifaceted, rooted in several key factors (Bhutta & Syed, 2016; Giannattasio et al., 2016; Pawlowski et al., 2009). Firstly, it involves chronic inflammation of the mucosa caused by distinct pathogens or recurring infections with diverse etiological agents. Secondly, host-related factors play a significant role, encompassing deficiencies in micronutrients or vitamins, undernutrition, and immunodeficiency. Thirdly, previous infectious processes and nutrient deficiencies can lead to heightened mucosal permeability, resulting in malabsorption. Lastly, disruption of the microbiota also contributes to

the pathogenesis. In certain cases, prolonged episodes of diarrhea may serve as a precursor to long-term gastrointestinal conditions such as autoimmune enteropathies, inflammatory bowel disease (IBD), and celiac disease, which commonly exhibit symptoms of persistent diarrhea, >14 days.

Numerous enteropathogens have been linked to prolonged episodes of diarrhea, each presenting its own distinct challenges (Giannattasio et al., 2016; Lanata, Fischer-Walker, Olascoaga, Torres, Aryee, Black, & Child Health Epidemiology Reference Group of the World Health Organization and UNICEF, 2013; Pawlowski et al., 2009; Wyllie et al., 2015). Among bacterial etiologies are: *Shigella*; Enteropathogenic *Escherichia coli* (EPEC); enteroaggregative *Escherichia coli* (EAEC); *Campylobacter*; and *Salmonella*. Protozoal etiologies include: *Giardia*; *Cryptosporidium*; *Cyclospora*; *Entamoeba*; *Isospora*; and microsporidia. Additionally, viral agents such as rotavirus, norovirus, and sapovirus, are implicated in up to half of diarrheal episodes that persist beyond three weeks, particularly in American children (Giannattasio et al., 2016). In certain cases, such as rotavirus infection, prolonged diarrhea may not stem directly from the pathogenic agent itself but rather from heightened susceptibility to other infections or nutrient malabsorption. Astrovirus infections, affecting roughly 10% of cases, can further complicate diarrheal episodes, even in individuals without compromised immune systems (Caballero et al., 2003). Furthermore, the spectrum of pathogens associated with prolonged diarrhea varies between developed and developing nations (Lanata, Fischer-Walker, Olascoaga, Torres, Aryee, Black, & Child Health Epidemiology Reference Group of the World Health Organization and UNICEF, 2013; Wyllie et al., 2015). In developed countries, viruses predominate, whereas specific bacteria and protozoa are more prevalent in developing regions. This divergence highlights the complex interplay between environmental, socioeconomic, and biological factors influencing diarrheal disease dynamics across different global contexts.

The preliminary diagnostic assessment for children exhibiting prolonged episodes of diarrhea should encompass stool testing for bacterial, viral, and protozoal pathogens (Pawlowski et al., 2009). Bacteria in stool specimens are identified by standard culture methods, while advanced enzyme-linked immunosorbent assay (ELISA) and polymerase chain reaction (PCR) techniques are employed for the detection of protozoa in fecal specimen, which despite their value are not yet standard practice (Guarino & Giannattasio, 2011). ELISA and latex agglutination assays are frequently employed to detect enteric viral infections. Molecular genetic techniques, however, offer broader pathogen detection capabilities and characterize their molecular epidemiological profile compared to conventional methods. Unfortunately, these diagnostic approaches remain unavailable in many primary care settings as well as in resource-constrained environments. The breath hydrogen test serves as a valuable tool in identifying abnormal colonization of the small intestine by bacteria, and detecting carbohydrate malabsorption (Giannattasio et al., 2016). However, its utility is tempered by relatively low sensitivity and specificity, making it challenging to administer, particularly in young children. Alternatively, bedside assessments such as stool pH measurement or testing for reducing substances (e.g., Clinitest) offer simpler alternatives for investigating carbohydrate malabsorption, typically lactose intolerance. In cases where diarrhea persists within the realm of pediatric gastroenterology, additional diagnostic modalities become necessary. These may include specialized tests to evaluate: intestinal motility; absorption; and permeability; the presence and severity of inflammation; imaging studies; and endoscopic procedures.

In developing countries where diagnostic capabilities are often limited, conducting tests that are costly and require significant time investment to pinpoint the cause of prolonged episodes of diarrhea poses significant challenges. In light of these constraints, prioritizing optimal

management of acute diarrhea episodes emerges as a crucial strategy to curb the incidence of prolonged diarrhea. Costly laboratory tests often offer minimal guidance in effectively treating prolonged episodes of diarrhea, where an empirical therapeutic approach may prove successful in most cases (Bhutta & Syed, 2016; Guandalini, 2020). When weight loss is absent, a conservative approach is warranted, typically obviating the need for extensive investigations. Instead, attention should be directed towards reviewing the patient's diet, particularly scrutinizing the consumption of sugar-containing beverages as a potential trigger for diarrhea. In cases where weight loss is observed, initiating a lactose-free diet may be beneficial. Interestingly, even in instances where weight loss isn't initially evident, some children with prolonged diarrhea may still find a lactose-free diet beneficial, as disaccharidase deficiency following an acute diarrheal illness is frequently reported.

Antimicrobial agents play a crucial role in the management of specific protozoas and bacterial enteropathogens, including EPEC and EAEC, *Shigella* and *Cryptosporidium* (Bhutta & Syed, 2016; Giannattasio et al., 2016). Nitazoxanide demonstrates potent activity against a diverse array of parasites and bacteria, making it a valuable therapeutic option. It stands as an efficacious treatment option for *cryptosporidium* infection in children. Metronidazole is commonly utilized for *Giardia* infections, while trimethoprim-sulfamethoxazole serves as a suitable therapy for *Cyclospora* and as an alternative when primary antibacterial therapy fails for a variety of pathogens. For *Shigella* infections, ciprofloxacin, azithromycin, and ceftriaxone are recommended. In the realm of viral etiologies of prolonged diarrhea, treatment options remain limited, with scant data available. However, human immunoglobulin, administered via the oral route at a dosage of 0.3g/kg of body weight in a single dose, shows promise as a form of passive immune-based treatment. This treatment employs immunoglobulin preparations enriched with

antibodies that neutralize viruses, has demonstrated efficacy in reducing the duration of severe diarrheal episodes, particularly in immunocompromised patients and those with rotavirus-induced diarrhea. Additionally, probiotics have emerged as a potential therapeutic adjunct in the management of prolonged diarrheal illness. *Lactobacillus* spp. and *Saccharomyces boulardii* have shown efficacy in reducing stool frequency and length of diarrheal episode in children with persistent diarrhea, offering a promising avenue for intervention in prolonged diarrhea cases.

Nutritional intervention stands as a cornerstone in the management of prolonged episodes of diarrhea (Bernaola Aponte et al., 2013; Bhutta & Syed, 2016). However, implementing such interventions can pose significant challenges, particularly in resource-constrained settings. In developing countries, where expensive dietary treatments may not be feasible, the emphasis shifts towards utilizing locally available, cost-effective foods, and vitamin and mineral supplementation. Tailoring nutritional interventions to local contexts proves essential in optimizing outcomes. These interventions possess both economic viability and cultural acceptability, ensuring adequate energy and nutrient intake for malnourished children while facilitating continued dietary therapy at home. Children affected by prolonged episodes of diarrhea and malnutrition often present with deficiencies in key micronutrients, including zinc, vitamin A, copper, selenium and folic acid (Bhutta & Syed, 2016; Giannattasio et al., 2016). Zinc supplementation has been shown to be effective in promoting the healing of small bowel mucosal damage and reducing the length of diarrheal episodes, as demonstrated in numerous clinical trials. Moreover, combining zinc with vitamin A appears to yield even greater benefits compared to either nutrient alone, in reducing the duration of prolonged diarrheal episodes in developing nations.

Furthermore, proactive interventions like rotavirus vaccination, promotion of breastfeeding, advocacy for hand hygiene, increased access to safe and clean water, and community-based

sanitation efforts serve as indirect preventive measures against prolonged episodes of diarrhea (Bhutta & Syed, 2016; Pawlowski et al., 2009). Despite these preventive strategies, the diagnostic and therapeutic management of prolonged diarrhea remains an ongoing challenge, underscoring the need for innovative approaches and resourceful solutions to address this persistent public health issue.

2.3 Chronic Malnutrition

This section outlines the burden, risk factors, pathogenesis, diagnosis and management of chronic malnutrition.

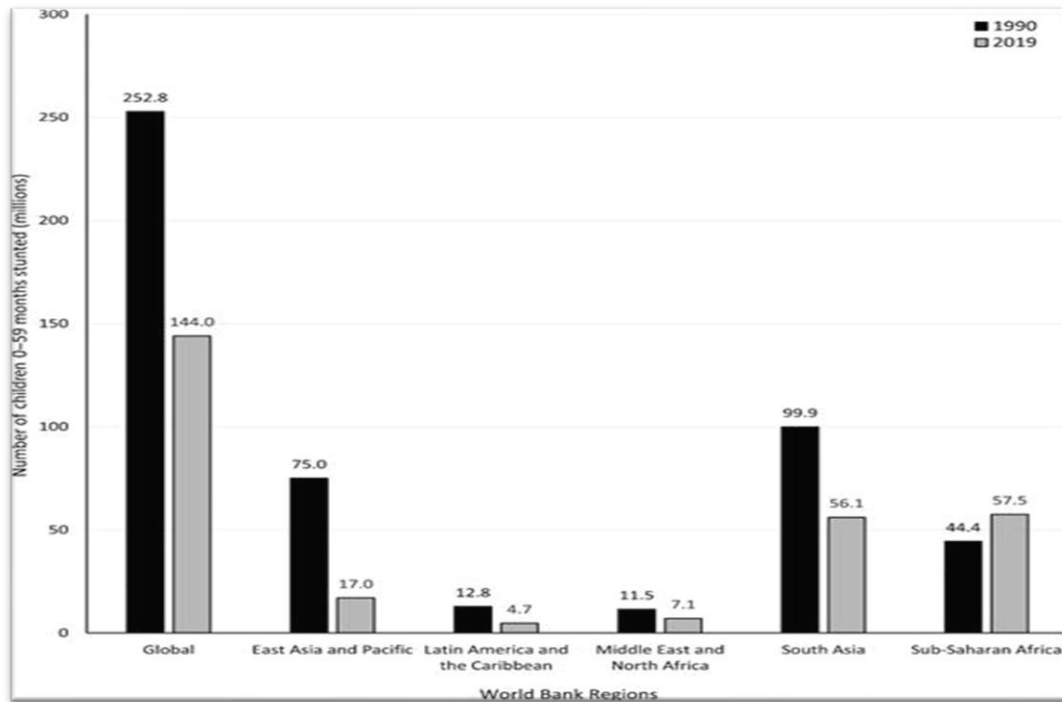
2.3.1 Burden

Chronic malnutrition is a type of growth faltering caused by an insufficient intake or absorption of essential nutrients over a protracted period of time, and its most common indicator is stunting and linear growth faltering. Linear growth faltering (LGF) describes the process whereby a child's height-for-age z-score (HAZ) begins to deviate from recognized WHO child growth standards (Prendergast & Humphrey, 2014). Linear growth faltering is defined as a decrease of 0.5 HAZ or more ($\Delta \text{HAZ} \geq -0.5$) between enrolment and follow-up. Once the HAZ dips below -2 Z-scores, it is identified as stunting (Lenters et al., 2016). Hence, the prevention of LGF stands as the critical primary strategy for effectively reducing stunting prevalence. According to global estimates, approximately 149 million stunting cases were reported among children aged < 60 months (World Health Organization, 2021b), with 13.5% of these cases attributed to diarrheal illness (Danaei et al., 2016). Stunting predisposes children to increased risk of death, morbidity, and developmental delays in both cognitive and motor function (Vaivada et al., 2020). Additionally, the self-perpetuating cycle between diarrheal illness and being malnourished is extended as malnutrition

compromises the body's immune system, increasing susceptibility to future diarrheal episodes, leading to more frequent and prolonged illness. This vicious cycle in early childhood can also potentially predispose children to elevated susceptibility of cardiovascular disease, type 2 diabetes, obesity and other adverse outcomes when they become adults (Guerrant et al., 2013; ROTA Council, 2019). Additionally stunting affects 33.3% of children aged < 5 years in LMICs, accounting for approximately 14% of mortality cases in this age group (Danaei et al., 2016). Research has shown that a greater exposure to diarrheal illnesses is linked to a higher likelihood of stunting. Specifically, the risk of being stunted at 2 years of age increases in a compounding manner with every diarrheal illness and with each day of diarrhea prior to 2 years ($P < 0.001$). The risks of stunting also increased by 13% with each five-fold increase in episodes (aOR [95% CI]: =1.1 [1.1–1.2]), and by 16% for every 5% increment in prevalence over time (aOR [95% CI]: 1.2 [1.1–1.3]). In addition, 25 in every 100 cases of stunting are attributable to experiencing ≥ 5 diarrheal illnesses before 2 years of age, while being sick from diarrhea for a cumulative duration of at least 2% of the time, before 2 years, of age accounted for 18 of every 100 cases of stunting (Checkley et al., 2008).

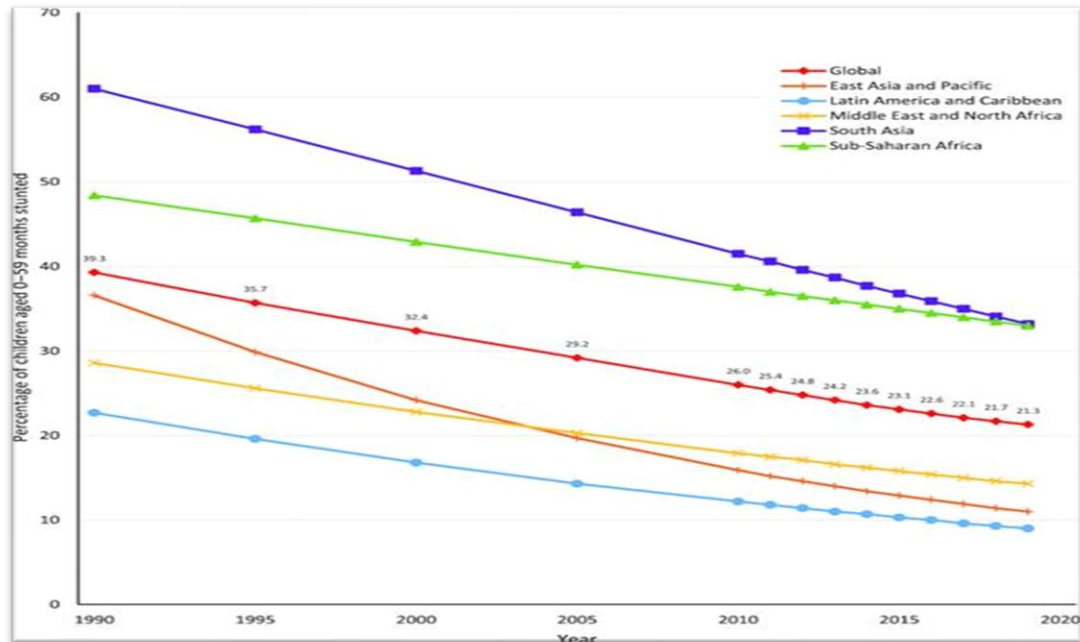
The number of cases reported globally and across World Bank regions are shown in Figure 2.2, and the trends of the prevalence of stunting between 1990-2019 are shown in Figure 2.3. By World Bank regions, East Asia and Pacific regions had a decline in the number of stunting cases from 75 million in 1990 to 17 million, in 2019. Latin America and the Caribbean reported a decline from 12.8 million cases in 1990, to 4.7 million cases in 2019. The Middle East and North Africa also witnessed a decline from 11.5 million cases in 1990 to 7.1 million, in 2019. South Asia also saw a reduction from 99.9 million in 1990 to 56.1 million in 2019. Sub-Saharan Africa has the greatest burden and was the only region that experienced an increase in the prevalence of stunting, from

44.4 million in 1990 to 57.5 million in 2019 (Vaivada et al., 2020). Overall, the global prevalence of stunting has had a downward trend starting at 252 million (39.3%) in 1990 to 144 million (21.3%) in 2019 (Figure 2.3).



Source: Vaivada et al. (2020); Reproduced under the Creative Commons Attribution (CC BY) license.

Figure 2.2: A comparative analysis of stunting cases globally and by world regions in 1990 and 2019 (Vaivada et al., 2020; WHO et al., 2020).



Source: Vaivada et al. (2020); Reproduced under the Creative Commons Attribution (CC BY) license.

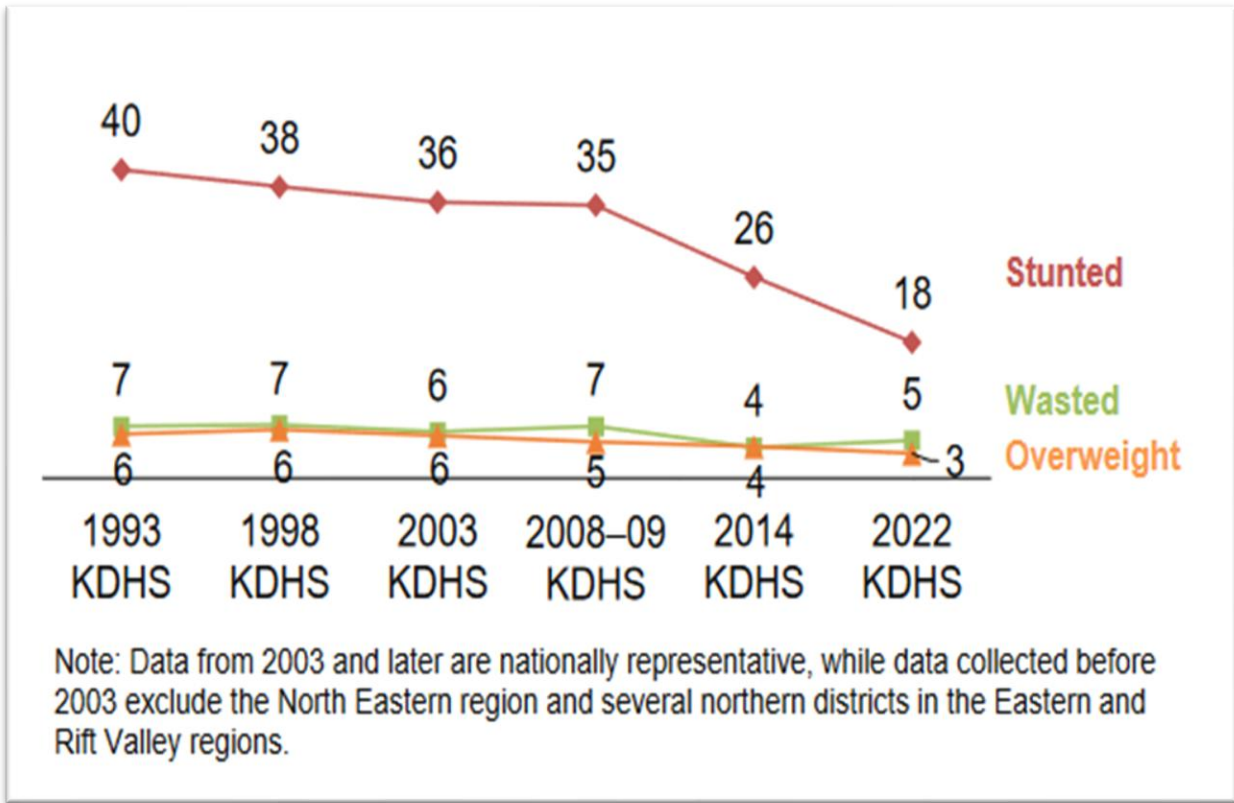
Figure 2.3: Temporal patterns in in stunting prevalence between 1990 and 2019 (Vaivada et al., 2020; WHO et al., 2020).

The global prevalence of stunting among children aged <5 years is 21.3%, in Africa it is 29.3% and 34.5% in Eastern Africa (WHO et al., 2020). Country specific estimates for stunting were also reported. The assessment of the trends of stunting in Brazilian children showed that there was a consistent decline in the national prevalence of stunting over a 33-year period from 37.1% in 1974 to 7.1% in 2007. Specifically the decline was 59.0% to 11.2% and 12.1% to 3.3% in the poorest and wealthiest quintiles, respectively (Monteiro et al., 2010). A study in China assessed the prevalence and drivers of stunting in infants and toddlers aged < 36 months, they reported the estimates of stunted and severely stunted to be 27.0% and 13.2%, respectively (Jiang et al., 2015). Another study in Sri Lanka estimated the prevalence of stunting in infants and toddlers within the age range of 6 to 36 months to be 16.8% (95% CI; 14.1, 18.0) (Senarath & Sujendran, 2015). A

study exploring the prevalence and drivers of stunting among Bangladeshi preschoolers reported that the overall prevalence of stunting was 36.3% and there was a significant difference between children in rural (38.1%) and urban (31.2%) areas (Akram et al., 2018). A study in Indonesia looked at temporal patterns in prevalence and the determinants of stunting among children aged 24-59 month (Rachmi et al., 2016). The cumulative cases of stunting reduced significantly over the study period from a high of 51% to a low of 37%.

A study in South Africa reported that approximately 20.0% of children aged 1-4 years were stunted and 33.0% of those <1 year (Kimani-Murage et al., 2010). A study conducted in Northwest Ethiopia reported the prevalence of stunting to be 49.4% (95% CI: 46.7%–52.3%) (Geberselassie et al., 2018), while in Northeast Ethiopia the prevalence was 39.5% (95% CI: 35.4–43.5%) (Gebreyohanes & Dessie, 2022), and 37.7% in Southern Ethiopia (Mengesha et al., 2021). A different study in Nigeria established the prevalence of stunting among school aged children was 17.1% and 25.3% among those attending private and public schools, respectively (Opara et al., 2010). A Study in the Democratic Republic of Congo (DRC) looked at the determinants of stunting using the demographic and health survey (DHS) 2013-2014. Approximately 43% of the children were stunted in their study with 20% and 23% being moderately and severely stunted, respectively (Kismul et al., 2017). A study in Tanzania reported a prevalence of 28.1% for stunting among school aged children (Gowelee et al., 2021). A trend analysis of the prevalence of stunting in Tanzania showed that the prevalence decreased by 30% over a period of 25 years (1991-2016) (Sunguya et al., 2019).

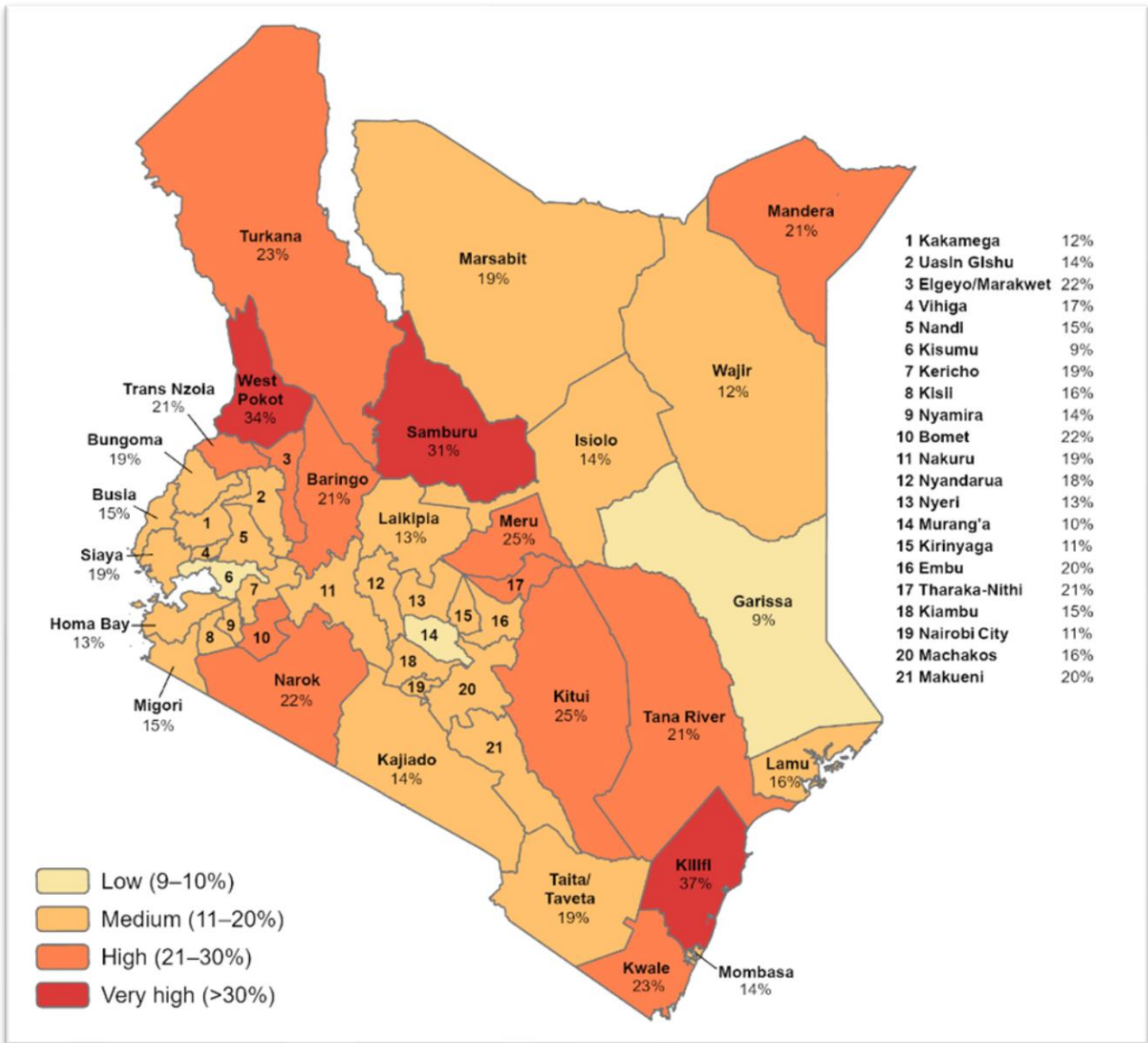
Kenya witnessed a downward trend in the prevalence of stunting among children aged < 60 months from 40.0% in 1993 to 18.0% in 2022 (Figure 2.4) (Kenya National Bureau of Statistics & Ministry of Health, 2023).



Source: Kenya National Bureau of Statistics & Ministry of Health (2023), Kenya Demographic and Health Survey 2022.

Figure 2.4: Trends of childhood malnutrition in Kenya 1993-2022 (Kenya National Bureau of Statistics & Ministry of Health, 2023).

From the KDHS 2022, wide variations were observed in the prevalence of stunting across counties, with counties with highest burden being in Kilifi (37%), West Pokot (34%), and Samburu (31%). (Figure 2.5). Additionally, the lowest burden was seen in Kisumu and Garissa (9% each). Approximately 19% of children in the study area (Siaya county) were stunted (Kenya National Bureau of Statistics & Ministry of Health, 2023).



Source: Kenya National Bureau of Statistics & Ministry of Health (2023), Kenya Demographic and Health Survey 2022.

Figure 2.5: Mapping childhood stunting in Kenya: County-specific prevalence rates (2022)(Kenya National Bureau of Statistics & Ministry of Health, 2023).

A study evaluating the economic burden of stunting in 95 LMICs, showed that childhood stunting costs the private sector $\geq \$135$ billion in revenue from sales yearly (Akseer et al., 2022). The study assessed the impact of childhood stunting on human capital outcomes using longitudinal cohort

data from seven countries and supplemented with regional estimates for other LMICs. Economic losses—including income reduction and private sector sales impacts—were quantified using national and enterprise-level data, and benefit-cost ratios for stunting reduction investments were calculated for the seven countries as regionally representative cases. The heaviest burden fell upon firms located in Latin America, the Caribbean, East Asia, and the Pacific (Akseer et al., 2022). The cumulative impact of sales losses to private enterprises across nations amounted to 1.2% of their respective gross domestic products (GDPs). An evaluation involving the comparison of risks was conducted to quantify the economic consequences throughout an individual's life of the burden of stunting in children in 137 LMICs. They estimated that approximately 7 million cases of stunting in LMICs were attributed to psychological and social factors. The key risk factor being maternal depression which had 3.2 million cases attributable. Furthermore, maternal depression, low maternal education, and IPV, had the greatest economic cost at \$14.5 billion, \$10.0 billion and \$8.5 billion, respectively. The overall cost of these risk factors was \$29.3 billion per birth cohort (Fawzi et al., 2019). McGovern et al. in their study showed that investing in nutrition interventions aimed at reducing stunting offered a remarkable return, with a median benefit-to-cost ratio of 18:1 for every child, as calculated using calibrated estimates. Additionally, the study highlighted the significant wage advantage associated with height, demonstrating that a 1-centimeter increment in stature translated to a 4% and 6% wage boost for males and females, respectively (McGovern et al., 2017). Hoddinott et al. evaluated the return on investment for interventions aimed at reducing stunting. Their results showed that the country-specific benefit-cost ratios for investments that reduce stunting in 17 high-burden countries range from 3.6 (DRC) to 48 (Indonesia) with a median value of 18 (Bangladesh) (Hoddinott et al., 2013).

Montenegro et al. also looked at the economic burden of stunting in Latin America and they reported a decline in the per capita income of the labor force due to stunting, to be at ~5%, highlighting the need for continued extensive approaches to address stunting across the world (Montenegro et al., 2022). In the economic evaluation of the burden of stunting in pregnant women and children < 5 years in Cambodia, it was estimated that the indicators of malnutrition presents a cost \$266 million yearly (1.7% of GDP) to the national economy of Cambodia. Furthermore, stunting is reducing the Cambodian economic output by more than \$ 120 million (Moench-Pfanner et al., 2016). The costs to scaling-up interventions to full coverage in 4 countries (Togo, Nigerai, Mali and DRC) vary substantially across countries, ranging from US\$13m in Togo to US\$837m in Nigeria. The modelling results showed that the investments would save 115 000 Disability-adjusted life years (DALYs) in Togo, 509 000 DALYs in Mali, 2.6 million DALYs in the DRC and 6.3 million DALYs in Nigeria. For each nation, the scale-up would additionally save between 3000 and 180 000 lives and avert between 60 000 and 3 million cases of stunting yearly (Shekar et al., 2016).

The review of literature emphasizes the persistence of substantial morbidity, mortality, and economic burden associated with childhood stunting. This observation holds true even as the global prevalence of stunting has exhibited a gradual decline over time. It is evident that stunting is a substantial and persistent public health concern, demanding the implementation of targeted interventions aimed at alleviating this issue and mitigating the consequent burdens it imposes. In resource-constrained settings such as Kenya, harnessing technological innovations has the potential to tackle this public health challenge and offer support to healthcare professionals in their clinical decision-making processes, ultimately enhancing the quality of care.

2.3.2 Risk factors

The WHO conceptual framework on childhood stunting identifies four levels at which drivers of stunting can be categorized: child, maternal, household, and environmental Figure 2.6.



Source: World Health Organization (2021b). © WHO.

Figure 2.6: WHO conceptual framework on childhood stunting (World Health Organization, 2021b).

Danaei et al. estimated the number of stunting cases among children aged 24–35 months were attributable to 18 risk factors in 137 LMICs (Danaei et al., 2016). The leading risk worldwide was fetal growth restriction (FGR) which accounted for approximately 11 million stunting cases of the global count of 44.1 million. Lack of improved sanitation came in second with approximately 7.2 million cases, while diarrheal illness was third having approximately 6 million cases attributed to it. Across all world regions, the combination of FGR and premature birth emerged as the primary risk factor grouping. Globally and in the South Asia, SSA, and East Asia and Pacific regions, environmental risks were estimated to be the second most influential cluster. However, in the remaining world regions, the combination of poor nutrition and infections was the second leading cause of stunting. Another systematic review looking at environmental exposures linked to stunting in children identified foodborne mycotoxins, poor sanitation, floor type, poor cooking fuel quality, and poor waste management, as elevating the likelihood of stunting in children (Vilcins et al., 2018).

In another systematic review exploring the prevalence and determinants of stunting in SSA, child age (>1 year), low birth weight, shorter interpregnancy interval, feeding practices, and climate data were identified as significant drivers of stunting (Quamme & Iversen, 2022). A different systematic analysis reviewing the determinants of stunting in children in 35 LMICs, highlighted a strong link between stunting and these variables: shorter stature of mother (OR [95% CI]: 4.8 [4.4-5.]), mother having a low level of literacy (OR [95% CI]: 2.0 [1.7-2.1]), lower wealth index (OR [95% CI]: 1.8 [1.5-1.9]), mother with a low maternal body mass index (BMI) (OR [95% CI]: 1.7 [1.5-1.8]), and father's short stature (OR [95% CI]: 2.0 [1.8-2.3]) (Li et al., 2020). Evaluation of LGF among children with MSD in GEMS, showed that younger children, being acutely

malnourished, being hospitalized, watery diarrhea, unimproved sanitation, poor households, fever, comorbidity, or any general danger signs were risk factors for stunting (Brander et al., 2019).

A study investigating the drivers of stunting in Bangladeshi children using the Multiple Indicators Cluster Survey (MICS) 2019 dataset found the following factors to have statistically significant associations with stunting: Age (24-35 months [OR = 2.65 , 95% CI: 2.30, 3.05]; 36-48 months [OR = 2.33 , 95% CI: 2.02, 2.69]), Maternal education ($\geq$ Secondary education [OR = 0.66, 95% CI: 0.56, 0.79]), Paternal education ($\geq$ Secondary education [OR = 0.74, 95% CI: 0.63, 0.87]), and wealth index (Richest [OR = 0.49, 95% CI: 0.41, 0.58]) (Chowdhury et al., 2022). A different study in Vietnam explored the covariates of childhood stunting looking at child, maternal, household, and environmental factors. In infants and toddlers in the age range 6 to 23 months, elevated risks were reported for age (relative risk [RR] 95% CI: 2.5 [2.3-2.7]), mother stature below 1.45 m versus $\geq$ 1.5 m (RR [95% CI]: 2.1 [1.9-2.3]), mother's illiteracy versus university education (RR [95% CI]: 1.8 [1.5-2.2]), and weighing less than the expected weight at birth (< 2.5 kg) (RR [95% CI]: 1.8 [1.6-2.0]). While among children ≥ 24 months, strong associations were observed in maternal illiteracy in comparison to graduate education (RR [95% CI]: 2.1 [1.8-2.4]), and mother stature below 1.45 m in comparison to ≥ 1.5 m (RR [95% CI]: 1.8 [1.7-1.9]) (Beal et al., 2019).

Fenske et al. (2013) investigated the risk factors of stunting in children in India. Key insights from their work showed that a child's age and gender, household wealth index, mother's education, maternal BMI, being a twin or multiple birth, maternal age, sibling rank, and number of ANC visits, were drivers of stunting (Fenske et al., 2013). A study in Sri Lanka identified the following factors to be the predictors of childhood stunting: maternal height, household wealth, duration of breast-feeding, and altitude (Rannan-Eliya et al., 2013). Another study looking at the drivers of

stunting in Indonesian children identified the following factors to have statistically significant associations with stunting: introduction of complimentary feeding before the age of 6 months, lower household socio-economic standing, preterm delivery, shorter length than average at birth, mother's height below standard range, maternal literacy, unimproved waste disposal and water that has not been treated, and limited access to healthcare services and residing in rural regions (Beal et al., 2018).

A study in Myanmar examined the predictors of childhood stunting. Their results showed these predictors: male children, older children 24–35 months (RR [95% CI]: 10.30[6.4-16.7]), mother's height below 1.45 m (RR [95% CI]: 5.1 [3.2--8.2]), perception of diminutive infant size at delivery (RR [95% CI]: 2.1 [1.6- 2.7]), non-health facility delivery (RR[95% CI]: 1.5 [1.2-1.9]), mother's employment status, lack of access to safe drinking water at home, residence (delta, coastal, or upland), and lower household wealth quintiles (Blankenship et al., 2020). Paudel et al. investigated the risk factors of childhood stunting in Nepal (Paudel et al., 2012). They identified the following factors as drivers of stunting in this population: Socio-economic factors associated with stunting included lack of maternal income (OR [95% CI]: 3.1 [1.3-7.7]), food shortages in household (OR [95% CI]: 4.3 [1.7-10.5]), and primary caregiver who is not the mother (OR [95% CI]: 3.0 [1.2-7.7]), and environmental drivers included unventilated kitchens, exposure to pesticides, not being exclusively breastfed (OR [95% CI]: 6.9 [2.8–17.0]), fewer than four complementary feeding per day (OR [95% CI]: 3.6 [1.3-10.0]). and lack of adequately diverse diets (OR [95% CI]: 4.1 [1.7-9.7]). Additionally, children with diarrhea were more likely to be stunted (OR [95% CI]: 7.5, [3.0-18.7]). A study in Bhutan explored the drivers of stunting among preschoolers (Kang et al., 2018). The susceptibility to stunting was elevated with increasing age: 5% at below 6 months (baseline group), 17% at 6–23 months (OR [95% CI]: 3.1 [0.65-14.7]), and 25.0% at ≥ 24 months (OR = 5.1

[1.2-22.3]). Additionally, there was a decline in the likelihood of stunting as a result of improved maternal education. A study in Cambodia identified the following factors to be drivers of childhood stunting: gender and age of child, delivery mode, mother stature, mother BMI, birth spacing, household size, wealth status, using safe sanitation, diarrheal illness, education level of caregivers, and mothers who smoke (Ikeda et al., 2013).

A study in Ethiopia looked at the household, nutritional and healthcare factors that predict childhood stunting. They identified the following environmental factors to be associated with stunting: mother's education (aOR [95% CI]: 0.7[0.5-0.9]), wealth status (aOR [95% CI]: 0.7 [0.5-0.8]), child's gender (aOR [95% CI]: 0.8 [0.7-0.9]), ownership of refrigerator (aOR [95% CI]: 0.6[0.4-0.9]), ownership of television, bigeminal birth, main flooring material of the house, and type of fuel used for cooking. Additionally, the following dietary factors were linked with decreased odds of stunting: introduction of breastfeeding ahead of schedule; being fed milk (aOR [95% CI]: 0.6 [0.5-0.8]); being fed on formula (aOR [95% CI]: 0.4 [0.2-0.8]); organ meat intake (aOR [95% CI]: 0.5 [0.3-0.9]), consumption of foods containing high levels of beta-carotene, ≥ 4 , antenatal care visits, anthelmintic treatment during gestation (aOR [95% CI]: 0.1[0.0-0.7]), health facility births (aOR [95% CI]: 0.6[0.6-0.7]) and weight at delivery (aOR [95% CI]: 5.1[1.6-15.9]) (Ayelign & Zerfu, 2021). Another Ethiopian study utilized data from the 2016 DHS to understand the correlates of stunting. They found child age, gender, birth spacing, maternal BMI, household socio-economic status, maternal literacy, breastfeeding duration, sanitation facility, internet usage and drinking water sources to be key determinants of stunting among Ethiopian children aged < 5 years (Takele et al., 2019). A different study in Rwanda examined the temporal patterns and correlates of stunting in children between 2000 and 2015. The study showed that living in the

lowest wealth quintile, limited maternal education, smoking by mother, male gender, and weighing <2.5 kgs at birth significantly heighten susceptibility to stunting (Binagwaho et al., 2020).

The analysis of the 2013 Namibia DHS to identify the drivers of severe childhood stunting showed that children from male headed households, rural residency, low maternal education, frequent exposure to diarrheal illness, HIV+ status, poor households, child age and duration of breastfeeding were predictors of severe childhood stunting in this setting (Mtambo et al., 2016). A different study looking at childhood stunting in Libya identified the following factors to be significantly associated with stunting: younger children (12–23 months: OR [95% CI]: 2.3 [1.7–3.2]; 24–35 months: OR [95% CI]: 1.6 [1.2–2.2]), paternal education (Non-literate: OR [95% CI]: 2.1 [1.2–3.8]; preparatory education: OR [95% CI]: 1.7 [1.1–2.7]), insufficient psychosocial support (Absence of familial travel or outings: OR [95% CI]: 1.5 [1.1–2.2]; limited paternal play interaction with the child: OR [95% CI]: 2.2 [1.2–4.2]), water that has undergone filtration (OR [95% CI]: 8.5 [2.3–31.0]), poor waste disposal (OR [95% CI]: 13.8 [2.3–81.7]), diarrhea illness (OR [95% CI]: 1.6 [1.1–2.3]) and neonatal weight < 2.5 kgs (OR [95% CI]: 1.8 [1.2–2.4]). Furthermore, advanced paternal age (OR [95% CI]: 0.5 [0.3–0.9]) and storing water (OR [95% CI]: 0.7 [0.5–0.9]) decreased the odds of stunting (Taguri et al., 2009).

The analysis of the 2008 Nigeria DHS to establish the individual and contextual correlates of childhood stunting established the following factors to be drivers of stunting: male sex, children aged ≥ 1 year, multifetal pregnancy, weight at birth < 2.5 kgs, limited mother's education, underweight mothers', inappropriate health care-seeking behavior of mothers', lower wealth index of household, and closely spaced births. Additionally, at communal level a child living in an area characterized by a significant level of illiteracy and the northernmost regions, encompassing the Northwest and Northeast, were found to be predictors (Adekanmbi et al., 2013). Kismul et al.

studied the drivers of stunting among children in the DRC, they identified the following factors to be drivers of stunting: rural residency, gender, birth interval, lower socio-economic status, introduction of breastfeeding ahead of schedule, mother's age, maternal height, maternal BMI, improved drinking water, improved sanitation, maternal education, and how many children live in a home (Kismul et al., 2017). A study looking at the temporal patterns and predictors of stunting among Ugandan children showed that lower wealth quintiles, caretaker's ill health, residence situated at a considerable distance from a healthcare facility, and utilizing an unsafe water source, were significantly associated with stunting (Biondi et al., 2011). A study conducted in Kenya to evaluate the risk factors of stunting in children identified the following to be drivers of stunting: age (OR =7.02, 95%CI (5.37-9.18)), female (OR =0.67, 95%CI (0.61-0.73)), birth weight (OR =0.43, 95%CI (0.34-0.55)), birth interval (OR=1.54, 95%CI (1.29-1.84)), mothers BMI, sanitation and wealth quintiles (Oluoko, 2021).

The examination of the scientific literature concerning risk factors affecting childhood stunting revealed that relevant articles were published within the timeframe spanning 2009 to 2022, with a predominant focus on the geographical regions of Asia and SSA. The predictors of childhood stunting exhibited consistency across various settings and over time. A comprehensive overview of the predisposing, enabling, and need-based factors influencing childhood stunting has been encapsulated in Table 2.3. The evidentiary foundation presented herein has been constructed through the application of established statistical methodologies. However, an innovative approach that integrates a combination of statistical metrics and ML algorithms for the purpose of conducting feature selection is utilized in the current study. This approach preceded the development of models aimed at understanding chronic malnutrition. Moreover, an in-depth

exploration of explanatory models is undertaken to discern the relationships between features and childhood stunting, shedding light on the nature and extent of these associations.

Table 2.3: Factors influencing childhood stunting: A summary of evidence

Author	Country	Risk factors identified
Danaei et al., 2016	137 LMICs	Fetal growth restriction; unimproved sanitation; diarrheal illness; and child nutrition and infection.
Z. Li et al., 2020	35 LMICs	Shorter stature of mother; mother having a low level of literacy; lower wealth index; mother with a low maternal body mass index (BMI); and father's short stature.
Quamme & Iversen, 2022	SSA	Child age (>1 year); low birth weight; shorter inter-pregnancy interval; feeding practices; and climate data.
Brander et al., 2019	Multi-Site (GEMS)	Younger children; acutely malnourished; hospitalized; watery diarrhea; unimproved sanitation; low socio-economic status; fever; comorbidity; or any general danger sign.
T. R. Chowdhury et al., 2022	Bangladesh	Age; maternal education; paternal education; and wealth index.
Beal et al., 2019	Vietnam	Child age; maternal height; maternal education; and low birthweight.
Rannan-Eliya et al., 2013	Sri Lanka	Maternal height; household wealth; duration of breast-feeding; and altitude.
Beal et al., 2018	Indonesia	Introduction of complimentary feeding before the age of 6 months; lower household socio-economic standing; preterm delivery, shorter length than average at birth; mother's height below standard range; maternal literacy; unimproved waste disposal and water that has not been treated; and limited access to healthcare services and residing in rural regions.
Blankenship et al., 2020	Myanmar	Male children; child age; maternal height; perception of diminutive infant size at delivery; lack from health facility delivery; mother's employment status; lack of access to safe drinking water at home; residence (delta, coastal, or upland); and lower household wealth quintiles.
Fenske et al., 2013	India	Child age; gender; household wealth index; maternal education; maternal BMI; being a twin or multiple birth; maternal age; birth order; and number of ANC visits.
Paudel et al., 2012	Nepal	Lack of maternal income; food shortages in households; and primary caregiver who is not the mother; unventilated kitchens; exposure to pesticides; not being exclusively breastfed; fewer than four complementary feeding per day; and lack of adequately diverse diets; and diarrheal illness.
Ikeda et al., 2013	Cambodia	Gender; age of child; delivery mode; mother stature; mother's BMI; birth spacing; household size; wealth status; using safe sanitation; diarrheal illness; education level of caregivers; and mothers who smoke.
Taguri et al., 2009	Libya	Young age; having a less educated father; poor psychosocial stimulation; filtered water; poor garbage disposal; diarrhea illness; low birth weight; age of father; and water storage.

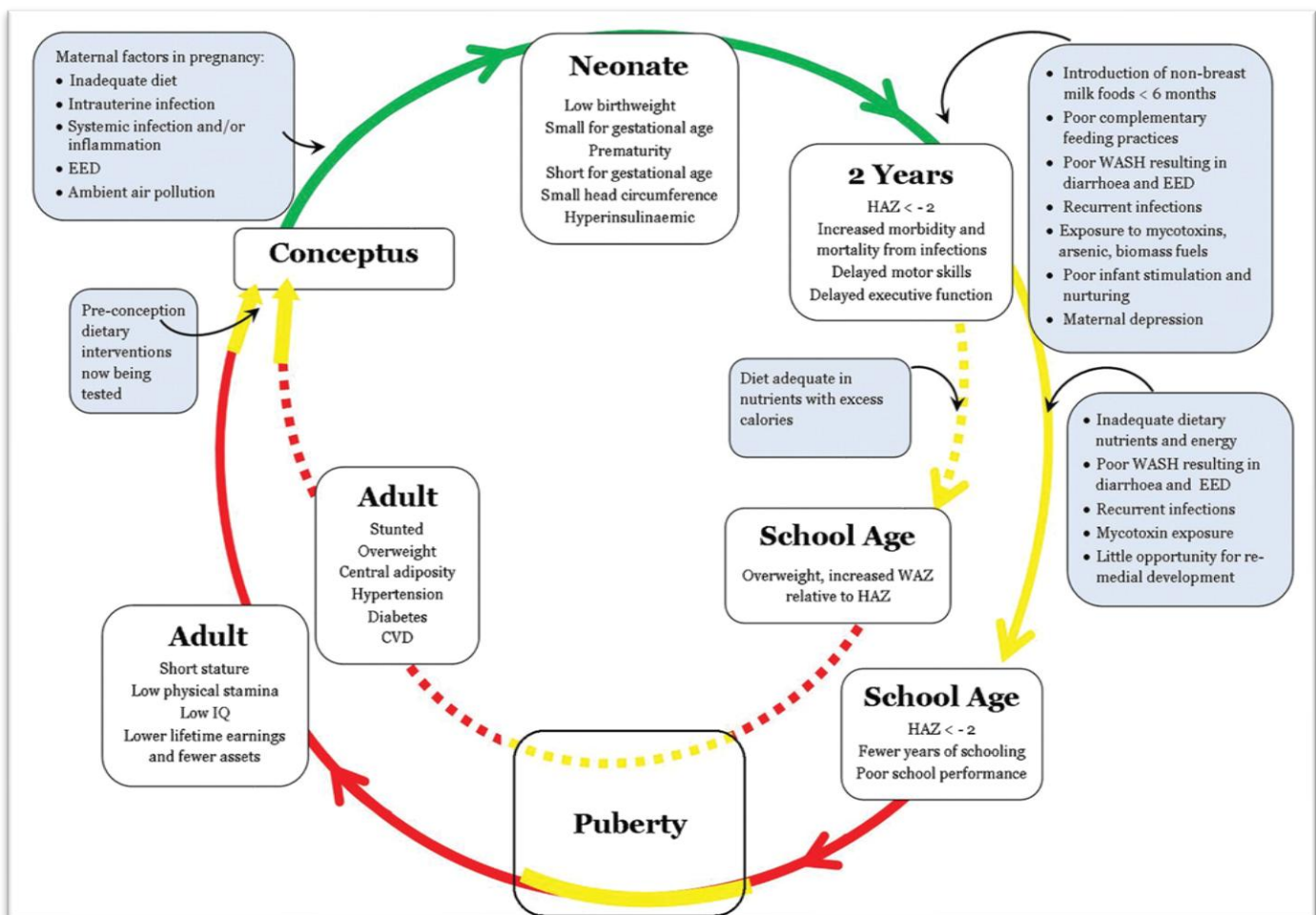
Ayelign & Zerfu, 2021	Ethiopia	Maternal education; wealth index; gender of child; possession of refrigerator; possession of television; bigeminal birth; main flooring material of the house; and type of fuel used for cooking; introduction of breastfeeding ahead of schedule; being fed milk; being fed on formula; organ meat intake; consumption of foods containing high levels of beta-carotene; ≥ 4 antenatal care visits; anthelmintic treatment during gestation; health facility births; and weight at delivery.
Takele et al., 2019	Ethiopia	Child age; gender; birth spacing; maternal BMI; household socio-economic status; maternal literacy; breastfeeding duration; sanitation facility; internet usage; and drinking water sources.
Binagwaho et al., 2020	Rwanda	Lowest wealth quintile; limited maternal education; smoking by mother; male gender; and low-birth weight.
Mtambo et al., 2016	Namibia	Male headed households; rural residency; low maternal education; frequent exposure to diarrheal illness; HIV+ status; poor households; child age; and duration of breastfeeding.
Kismul et al., 2017	DRC	Rural residency; gender; birth interval; lower wealth index; early initiation of breastfeeding; maternal age; maternal height; maternal BMI; improved drinking water; improved sanitation; maternal education; and how many children live in a home
Adekanmbi et al., 2013	Nigeria	Male gender; children aged ≥ 1 year; multifetal pregnancy; weight at birth < 2.5 kgs; limited mother's education; underweight mothers'; inappropriate health care-seeking behavior of mothers'; lower wealth index of household; closely spaced births; and high level of communal illiteracy.
Vilcins et al., 2018	-	Foodborne mycotoxins; poor sanitation; floor type; quality of cooking fuels; and inadequate waste disposal
Oluoko, 2021	Kenya	Age; gender; birth weight; birth interval; mothers' BMI; sanitation; and wealth index

2.3.3 Pathogenesis, Diagnosis and Management

Despite the alarming prevalence of stunting globally, the precise mechanisms driving linear growth failure remain elusive. While epidemiological investigations point towards being sub-optimally breastfed and less than ideal complementary feeding practices, recurring infections, and inadequate intake of micronutrients as key proximal determinants of stunting, the phenomenon is embedded within a multifaceted interplay of broader societal and community factors (de Onis & Branca, 2016; Prendergast & Humphrey, 2014; Tang et al., 2021). These encompass the availability of healthcare and educational services, a stable political environment, rate of urban growth, dynamics of population per unit area, and the resilience of social support systems. Such intricate dynamics

have been captured within the WHO Conceptual Framework on Childhood Stunting. The initial two years of life represent a phase of heightened susceptibility, carrying the greatest potential for enduring physical and developmental consequences. It is within this critical timeframe that interventions hold the greatest promise for addressing stunting and likely all associated pathology.

A comprehensive overview of the pathogenesis of stunting, including both biological and environmental factors is shown in Figure 2.7.



Source: Prendergast & Humphrey (2014). Reproduced under the Creative Commons Attribution-NonCommercial (CC BY-NC) license.

Figure 2.7: Stunting Pathways Across the Developmental Spectrum (Prendergast & Humphrey, 2014).

In summary, the importance of early intervention during the first 1000 days is highlighted, and the potential for catch-up growth among older children and adolescents, as well as the impact of environment-based exposures on growth trajectories (Prendergast & Humphrey, 2014). Additionally, the role of dietary interventions and the varying effectiveness of interventions at different stages of development is shown.

Linear growth faltering, indicative of a deceleration in physical development, may signal compromised health status including infections or undiagnosed birth defects, alongside challenges stemming from inadequate diet or feeding practices (Cooke et al., 2023; Smith et al., 2023). It is imperative to identify and tackle the fundamental drivers of this phenomenon to promote the child's health and developmental trajectory in a sustainable way. While nutritional deficiencies are often implicated as the primary cause, particularly in affluent settings, in low-income regions, it frequently coexists with gastrointestinal disorders and other underlying pathologies. Additional medical investigations to unveil potential contributors to LGF are warranted primarily when children exhibit discernible symptoms or when concerns regarding underlying medical issues arise.

The initial crucial step in addressing LGF is to optimize the patient's nutrition and overall feeding regimen while simultaneously pinpointing the most probable underlying causes (Lezo et al., 2020; Selbuz et al., 2019). Clinicians should commence by conducting a comprehensive evaluation, incorporating a detailed patient history and thorough physical examination, along with selective employment of laboratory and radiology tests, and seeking subspecialty consultation when necessary. Upon identification of the underlying etiologies, treatment strategies should be customized, aiming for therapies tailored to address these specific causes wherever feasible (Smith et al., 2023).

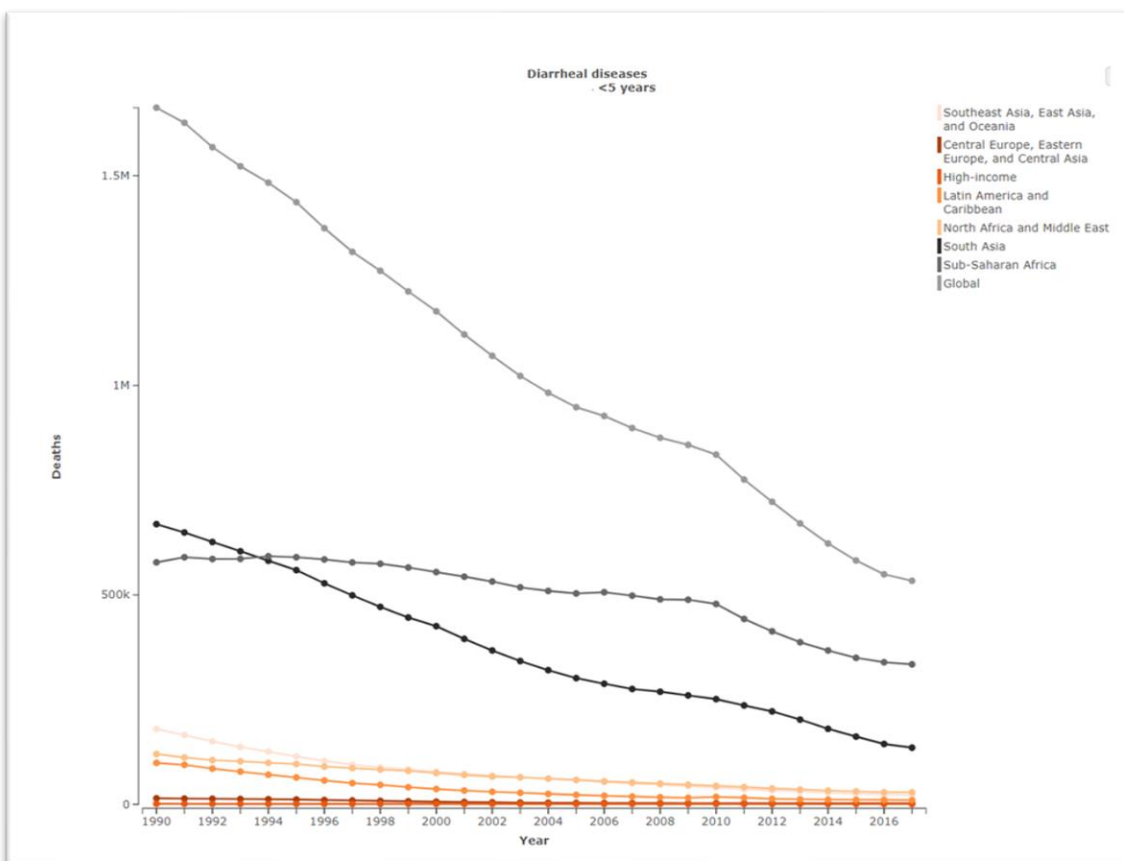
The cornerstone of early interventions lies in promoting optimal breastfeeding practices (Cooke et al., 2023). At the community level, it's evident that single-mode interventions demonstrate limited efficacy; thus, a combination of approaches utilizing both direct nutrition interventions and those that address the root causes of malnutrition is recommended. In LMICs, all nutritional interventions should strive to rectify growth faltering by striving to normalize growth patterns as much as possible. These efforts should be guided by WHO recommendations on infants and toddlers feeding, emphasizing the provision of adequate amounts of age-appropriate, nutrient-dense foods. When managing faltering growth, it's essential to tailor interventions to address the underlying causes (Mehta et al., 2013). Regardless of etiology, nutritional management must focus on achieving an appropriate energy and protein balance, supplemented with micronutrients to optimize growth and development in children who have experienced growth delays. For disease-related growth faltering, nutritional needs tailored to the specific health conditions diagnosed. Similarly, for non-disease-related growth faltering, addressing underlying factors such as feeding challenges, inadequate dietary practices, and food insecurity is paramount (Mank et al., 2020). Support for breastfeeding is crucial in growth faltering cases, irrespective of etiology (Rana et al., 2020). This entails assessing breastfeeding techniques and ensuring an adequate milk supply. Where necessary, consider supplementing with fortified milk, transitioning to cup feeding, or using supplemental formula as appropriate. Subsequently, close monitoring is indispensable to track the patient's response to interventions and observe trends in catch-up growth over time.

On a broader scale, addressing faltering growth requires addressing macroeconomic factors and reducing socio-economic disparities. Economic development and efforts to mitigate inequalities serve as potent mechanisms to combat growth faltering on a global scale.

2.4 Mortality

2.4.1 Burden

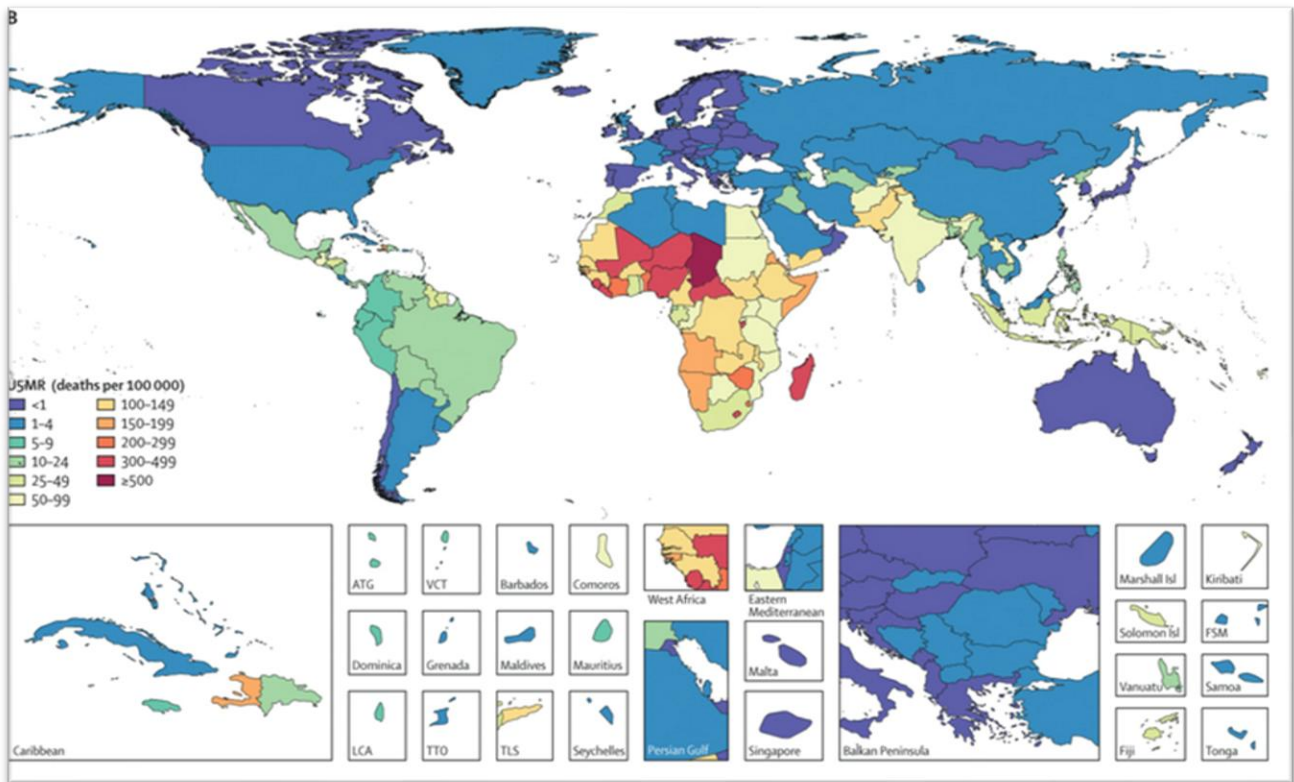
Diarrheal disease is the third primary cause of mortality in children aged below 60 months across the world, causing approximately 9% of all deaths among these children (Liu et al., 2016; World Health Organization, 2024). This burden is greatest in SSA and Southern Asia where 90% of all diarrheal mortality occur with rates ranging between 50-150 persons per 100,000 (Troeger et al., 2018; UNICEF, 2022). Despite the significant decline in diarrheal mortality globally in the last 3 decades (Figure 2.8), the burden remains unacceptably high in SSA (Wolde et al., 2022).



Source: Troeger et al. (2019). Reproduced under the Creative Commons Attribution (CC BY) license.

Figure 2.8: Global trends of diarrheal mortality 1990-2017 (Troeger et al., 2019)

Projections of all-cause mortality shows that more than 4 million child deaths will occur among children aged < 60 months in 2030, with SSA contributing 60% (Liu et al., 2015). Drawing from the 2016 comprehensive analysis of the global impact of diarrheal disease, diarrheal deaths in children aged <5 years was estimated to be 446 000 deaths globally, with a 95% uncertainty interval (UI) (390 894–504 613), and the incidence per 100,000 was 70·6 (61·9–79·8) (Troeger et al., 2018). Furthermore, they identified rotavirus to be the foremost etiology for diarrheal deaths among children aged < 60 months accounting for approximately 129,000 deaths. The number of deaths and the incidence per 100,000 by region was: High income countries had 761 95%UI (693–839) and 1·3 (1·2–1·5); Central Europe, eastern Europe, and central Asia had 1 943 (1 411–2 716) and 6·9 (5·0–9·6); Latin America and Caribbean had 8 828 (7 589–10 522) and 17·8 (15·3–21·2); Southeast Asia, East Asia, and Oceania had 15 443 (13 267–18 208) and 12·5 (10·8–14·8); North Africa and Middle East had 26 373 (19 539–34 818) and 41·7 (30·9–55·1); South Asia had 101 927 (85 817–122 100) and 66·4 (55·9–79·5); and Sub-Saharan Africa had 290 724 (243 545–342 557) and 185·7 (155·5–218·8). The distribution of <5 mortality rate across world regions are shown in Figure 2.9.



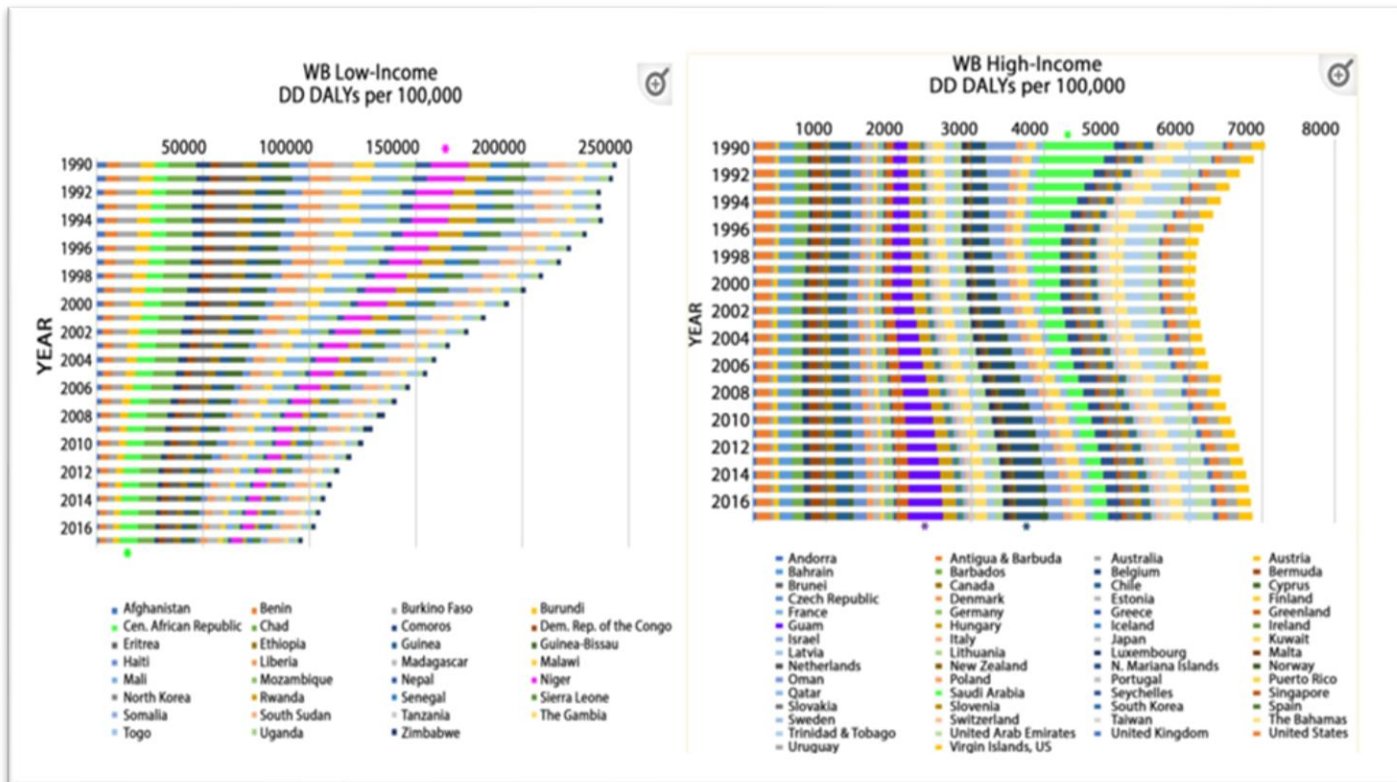
Source: Troeger et al. (2019). Reproduced under the Creative Commons Attribution (CC BY) license.

Figure 2.9: Diarrheal mortality rate in 2016 (Troeger et al., 2019).

Children with diarrheal illness had an increased risk of death, 8.5 times higher compared to their counterparts who were healthy (Pavlinac et al., 2018). The diarrheal mortality rates among children aged < 60 months was 176 per 100,000; the highest burden was in children aged 6-11 months 558 per 100,000 in Karemo area of Siaya County, Kenya (Omore et al., 2019). A disability-adjusted life-year (DALY) incorporates data on years-lived with-disability (YLD) and years-of-life lost (YLL) is a robust indicator to monitor the impact of disease in terms of lost healthy years. A DALY quantifies the loss of the equivalent of one year of full health. Diarrheal disease was responsible for the third highest number of DALYS contributing to a global burden of around 74 million DALYs, with a substantial proportion, roughly 63%, affecting children under the age of five (Troeger et al., 2018). In another study, the global DALYs for diarrheal disease were

approximated to be 81 million while the rate of age-adjusted DALY was approximately 12.6 per 1000 population (Figure 2.10). This represented an 85.43% decline in DALY rates from 1990 to 2017 (Karambizi et al., 2021). From their study, World Bank (WB) high per capita income economies, the USA territories (Guam and Northern Mariana Islands) reported the greatest burden of disease with the estimated DALYs rates of 4.8 and 3.3 per 1000 in 2017, respectively. While Malta had the least burden reporting DALYs rate per 1000 of 0.3. WB low-income countries had an estimated DALY rate of 1009 per 1000 population, a decline of 41.0% from 1990. Central African Republic reported the highest DALY rate, in 2017 (85.9 per 1000), while North Korea had the least DALYs rate for all years of the study (1990-2017). Despite the decline in DALY rates in low per capita income economies, the ratio of DALY rates between the later and high per capita income economies in 2017 remains high (14.8:1).

The annual number of DALYs in low-income countries, as classified by the World Bank, showed a reduction from 1990 to 1998. However, after 1998, these rates experienced a steady surge each year until 2017 (Figure 2.10). Additionally, YLL decreased steadily over the period under evaluation, while YLD showed a gradual rise. Contrastingly, in WB high-income countries, the YLL as well as the YLD increased over the study period (Karambizi et al., 2021). From the global estimation of health estimates by WHO using DALYs, Kenya reported a rate of 3.43 DALYS and YLL of 3,21 per 1000 people for diarrheal illness among children under the age of five (World Health Organization, 2020). Despite, the decline of diarrheal disease burden over the years, DALY rates still remain high.



Source: Karambizi et al. (2021). Reproduced under the Creative Commons Attribution (CC BY) license.

Figure 2.10: Age-adjusted diarrheal disability-adjusted life-year rates between 1990 and 2017 (Karambizi et al., 2021).

The extensive review of literature highlights the substantial burden of diarrhea mortality and its economic implications, particularly in LMICs like Kenya. This highlights the pressing demand for effective interventions to mitigate this burden, aligning with the third objective within the UN's global framework for sustainable development. Leveraging ML predictive models for mortality can significantly enhance patient outcomes by enabling early identification of high-risk individuals, personalized treatment strategies, and real-time monitoring of disease trends. These models empower healthcare professionals with valuable insights to make timely, data-driven decisions, ultimately leading to improved patient care and a reduction in the burden of diarrhea mortality.

2.4.2 Risk factors

Mehal et al. explored predictors of diarrheal mortality among infants in USA for the period 2005-2007 (Mehal et al., 2012). They found the following factors to be associated with death: Low birth weight (risk ratio [95% CI]: 92.0 [77.5–109.2]); younger age (median: 7 vs. 15 weeks, $P < 0.001$); comorbidities (septicemia=26% and hypovolemia=20%); Male sex; black race; unmarried status; An Apgar score below the standard range at 5 minutes post-delivery; younger mothers (below 25 years); and a higher birth order rank (≥ 3). A study looking at the effect of birth weight on risk factors of diarrheal deaths in children in the USA found the following factors to associated with mortality: very low birth weight, age, prematurity, black race, low mother education, and an Apgar score below the standard range at 1 minute post-delivery (Parashar et al., 1998). Additionally, approximately 88% of deaths from diarrhea are accounted for by water that is not safe for consumption, poor waste management, and poor hygienic conditions (CDC, 2022).

Hartman et al. studied the drivers of death among children aged < 60 months having severe diarrheal episodes in LMICs (Hartman et al., 2022). They identified, from multivariable analysis, the following to be predictors of mortality: infants (below 6 months in comparison to children aged ≥ 6 months, OR [95% CI]: 3.6 [2.7–4.6]); females (OR [95% CI]: 1.2 [1.1–1.8]); experiencing protracted diarrheal illness (OR [95% CI]: 1.9 [1.1–3.3]); absence of vomiting (OR [95% CI]: 1.2, [0.9–1.4]); severely dehydrated (OR [95% CI]: 3.8 [3.1–4.9]); rotavirus negative test result (OR [95% CI]: 2.3 [1.8–2.7]). An exploration from the global burden of disease study 2017 to quantify the impact of interventions in reducing the risk of diarrheal deaths in children aged < 5 years found that reduced access to unimproved hygiene facilities (13.3% [11.2%-15.5%]), being wasted (9.9% [9.6%-10.2%]), and inadequate ORS uptake (6.9% [4.8%-8.4%]), accounted for the greatest declines in diarrheal death rates (Global Burden of Disease 2017 collaborators, 2017). In their

evaluation of the determinants of diarrheal mortality reduction among children in LMICs for the time period 1980-2015, Black et al. established that correct diarrhea management (ORS, zinc therapy, antibiotics for dysenteric diarrhea and clinical management of PD) and rotavirus vaccination were responsible for approximately 50% of the decline in diarrheal deaths between 1980 to 2015. Furthermore, nutritional improvements, encompassing reductions in stunted and wasted children, increased rates of optimal breastfeeding, and vitamin A supplementation, contributed to 38.8% of the observed decline while improvements in access to safe water, sanitation facilities, and handwashing practices accounted for 11.5% of the decline in mortality (Black et al., 2019).

In their assessment of the factors influencing the likelihood and course of diarrhea episodes complicated with dehydration in children aged < 12 months in Brazil, Fuchs et al. identified neonatal weight < 2.5 kgs, being stunted, as well as inadequate breastfeeding to be predictors of mortality in the study setting (Fuchs & Victora, 2002). Page et al. conducted a study to determine the predictors of death during the Haitian cholera public health emergency (2010-2011) (Page et al., 2015). They identified residing in geographically remote locations, advanced age, severity of disease and failure to seek care as predictors of death.

Mitra et al. studied the determinants and gender-based disparities of mortality in Bangladeshi children admitted in hospitals for diarrheal illness (Mitra et al., 2000). They observed a higher likelihood of death among children having altered consciousness, hypoglycemia, septicemia, loss of bowel motility, toxic inflammation of the colon, inflammatory bowel necrosis, thrombotic microangiopathy with hemolysis and renal failure, invasive or protracted diarrhea as well as those who were dehydrated, malnourished or had alterations in their electrolyte levels. Another study in Bangladesh explored the incidence and drivers of shigellosis deaths among hospitalized patients

(Bennish et al., 1990). The following variables had significant associations with deaths in their study: Age <1 year; suboptimal breastfeeding in toddlers aged 12–23 months; hypothermic state; severely malnourished; severely dehydrated; impaired consciousness; distended abdomen; platelet deficiency; decreased serum protein levels; decreased serum sodium concentration; hypoglycemic state; kidney failure; and bacterial sepsis. Uysal et al. evaluated the clinical drivers for diarrheal mortality in Turkish children (Uysal et al., 2000). They identified young age (< 6 months) (OR [95% CI]:10.5, [1.8-62.8]), being dehydrated (OR [95% CI]: 8.2 [1.5-43.7]), being severely malnourished (OR [95% CI]: wasted (0.04 [0-0.22]) and stunted (0.03 [0.00,0.36])), septicemia (OR [95%CI]: 37.3 [6.9-200.1]), shigellosis (OR [95%CI]: 23.1 [3.1-172.0]), reduced albumin levels (OR [95%CI]: 0.1 [0.02- 0.5]), and metabolic acid-base disturbance (OR [95%CI]: 0.03 [0.00- 0.34] of $\text{HCO}_3^- < 20 \text{ mMol/L}$), to be predictors of mortality.

In a study looking at predictors of death among Pakistani children admitted for diarrhea, the following factors were identified as predictors of mortality: anorexia (OR [95% CI]: 3.8 [1.3-11.0]); somnolence (OR [95% CI]: 4.5 [1.4–15.4]); impaired respiratory function (OR [95% CI]: 7.1, [1.5- 36.7]); hemoglobin deficiency (OR [95% CI]: 5.9 [2.1–16.7]); and confirmation of a bloodborne infection (OR [95% CI]: 8.8 [2.6–30.8]) (Bhutta et al., 1996). Conducting a review of the prevalence and correlates of diarrheal morbidity amongst children aged less than 60 months in India, Ganguly et al. identified malnutrition (OR [95% CI]: 1.72 [1.54–1.97]), hemoglobin deficiency (OR [95% CI]: 1.72 [1.30–2.29]), and lower wealth index (OR [95% CI]: 7.15 [2.20–23.33]), as risk factors for severe diarrhea and its resulting mortality. Additionally, gender, religious belief, advanced maternal education, and seasonal trends were shown to be intermittently associated in some studies. Conversely, children aged < 2 years, suboptimal breastfeeding practices, maternal illiteracy, and water without treatment, showed no statistical significance

(Ganguly et al., 2015). Sachdev et al. also assessed the correlates of diarrheal mortality in hospitalized Indian children (Sachdev et al., 1991). They found comorbidities (pneumonia, septicemia, or meningitis), severe wasting, severe stunting, and protracted diarrheal illness to have a significant association with death. The estimates (aOR [95%CI]) for these predictors were 4.7 (3.9, 5.6), 3.3 (2.7, 4.0), 1.9 (1.6, 2.3), and 1.5 (1.3, 1.8), respectively.

As part of GEMS, predictors of death in Mozambican children with MSD were evaluated (Acácio et al., 2019). The following factors were determined to be drivers of mortality in this setting: Typical enteropathogenic *Escherichia coli* (typical EPEC) (HR = 4.2, $p = 0.046$); *Cryptosporidium* (HR = 2.8, $p < 0.001$); malnutrition (HR = 4.1, $p < 0.001$), lower respiratory tract infection (HR = 3.5, $p < 0.001$); systemic bacterial infection (HR = 6.8, $p = 0.001$); lethargic or unconscious at presentation (HR = 1.7, $p = 0.03$); lineated skin (HR = 1.7, $p = 0.039$); and HIV (HR = 5.1, $p = 0.001$). Contrarily, the study found a protective effect of rotavirus infection against mortality (HR = 0.5, $p = 0.02$). Data from the DHS waves done between 1990 to 2013 in Burkina Faso, Mali, Nigeria, and Niger was utilized in evaluating the trends and drivers of diarrhea associated mortality (Bado et al., 2016). From their review, they identified the following as risk factors for mortality: child age, sex, breastfeeding modality, low weight-for-age, being acutely malnourished, sanitary deficiencies, access to safe drinking water, and lack of safe and hygienic waste disposal methods, lower wealth index, food shortage, poor handwashing practice, water storage modality, household socio-economic standing; a large household size or a large number of children under five residing in the home, mother's age, literacy level, and occupational status.

In the evaluation of temporal patterns, etiology, and determinants of death among Ethiopian children aged less than five years, in the time period 1990–2013, the following factors were identified to be predictors of diarrheal mortality: being wasted, low weight-for-age, being stunted,

not exclusively breastfeeding, and hypovitaminosis A (Deribew et al., 2016). In the assessment of factors associated with the decline of diarrheal-associated mortality among children <5 years in Tanzania, Masanja et al. identified the following factors: ORS therapy, variations in the prevalence of stunting, enhancement of vitamin A intake, rotavirus vaccination, shifts in the prevalence of wasting and suboptimal breastfeeding practices (Masanja et al., 2019). Griffin et al. investigated the drivers of diarrheal mortality among children in Lesotho. They showed that the variables enumerated below, significantly heightened the risk of death: co-infection, age < 6 months, ≥ 7 days of illness prior to hospitalization, oral candidiasis or stomatitis diagnosed at admission, severely dehydrated, previous episodes of emesis, persistent dehydration despite 12 hours of inpatient management, and pyrexia or hypothermia (Griffin et al., 1988).

O'Reilly et al. also assessed predictors of mortality among Kenyan children aged < 60 months admitted for diarrheal illness (O'Reilly et al., 2012). They identified the following factors to be associated with mortality: non-typhoidal *Salmonella* infection (aOR [95% CI: 6.7 [3.2–14.8]]); *Shigella* infection (aOR [95% CI]: 5.5 [2.2–14.0]); negative for rotavirus (aOR [95% CI]: 0.4 [0.2–0.8]); suffering from malnutrition (aOR [95% CI]: 4.3 [2.2–8.6]); presenting with oral thrush (aOR [95% CI]: 2.4 [1.3–3.9]); prior sought care previously at a health facility (aOR [95% CI]: 2.1 [1.3–3.7]); and dehydration diagnosis (aOR [95% CI]: 2.6 [1.6–4.2]). However, a clinical determination of malarial illness, and laboratory-confirmed malaria, did not heighten the risk of mortality in this population. Another study in Kenya, examined mortality during hospitalization as well as after hospital discharge among children admitted for diarrhea in Kilifi County Hospital (Talbert et al., 2019). Co-infections of both diarrhea and severe pneumonia, HIV infection, evidence of malnutrition using mid-upper arm circumference (MUAC), and bacterial bloodstream infection, had increased the risk of the two outcomes. Circulatory problems, septicemia, and

electrolyte imbalances were significant predictors of in-hospital deaths only. Previous hospitalization and chest indrawing had associations with death in the period following discharge from hospital. While child’s age, stunting, and protracted diarrheal illness/dysentery were not predictors of in-hospital or post-discharge deaths.

In a systematic review of global causes of diarrheal mortality, Lanata et al. identified the following etiologies to cause more than 50.0% of diarrheal mortalities in children aged less than five years: approximately 200, 000 deaths were attributable to rotavirus; enteropathogenic *E. coli* accounted for approximately 80,000; calicivirus caused circa 70,000; and enterotoxigenic *E. coli* 40,000 deaths (Lanata, Fischer-Walker, Olascoaga, Torres, Aryee, Black, & Unicef, 2013).

The literature review addressing the risk factors related to diarrheal mortality exhibited a broad temporal range, spanning from 1988 to 2022, and encompassed diverse countries across expansive geographical regions. Despite this temporal and geographical diversity, the observed risk factors remained largely consistent across countries and over time. A comprehensive summary detailing the microbial, host, and environmental factors associated with diarrheal mortality is provided in Table 2.3. The evidence presented in this context was derived through conventional statistical methodologies. Nevertheless, the thesis advocates for an amalgamation of statistical measures and ML algorithms to facilitate feature selection as a preparatory step in the development of mortality prediction models.

Table 2.4: *Determinants of diarrheal mortality among children under five: A summary of evidence*

Author	Country	Risk factors identified
Mehal et al., 2012	USA	Age; low birth weight; sepsis; volume depletion; gender; race; unmarried status; maternal age; high birth order; and Apgar score below the standard range at 5 minutes post-delivery.

Parashar et al., 1998	USA	Very low birth weight; age; prematurity; black race; low mother education; and an Apgar score below the standard range at 1 minute post-delivery.
Lanata et al., 2013	Global	Rotavirus; enteropathogenic <i>E. coli</i> ; calicivirus; and enterotoxigenic <i>E. coli</i> .
Global Burden of Disease 2017 collaborators, 2017	Global	Unsafe sanitation; wasting; and low ORS coverage.
Hartman et al., 2022	LMICs	Infants; females; experiencing protracted diarrheal illness; absence of emesis; severely dehydrated; and rotavirus negative test result.
Black et al., 2019*	LMICs	Correct diarrhea management; rotavirus vaccine; enhancements in food quality and availability; coupled with advancements in sanitation infrastructure and hygiene practices.
Fuchs & Victora, 2002	Brazil	Low birth weight; stunting; and lack of breastfeeding.
Page et al., 2015	Haiti	Residing in geographically remote locations advanced age; severity of disease; and failure to seek care.
Uysal et al., 2000	Turkey	Age; dehydration; wasting; stunting; <i>Shigella</i> infection; hypoalbuminemia; and metabolic acidosis.
Ganguly et al., 2015	India	Malnutrition; anemia; low socioeconomic status; gender; religion; higher education of mothers; and seasonality
Sachdev et al., 1991	India	Comorbidities (pneumonia, septicemia, or meningitis); severe wasting; severe stunting; and protracted diarrheal illness.
Mitra et al., 2000	Bangladesh	Altered consciousness; hypoglycemia; septicemia; loss of bowel motility; toxic inflammation of the colon; inflammatory bowel necrosis; thrombotic microangiopathy with hemolysis and renal failure; invasive or protracted diarrhea as well as those who were dehydrated, malnourished or had alterations in their electrolyte levels.
Bennish et al., 1990	Bangladesh	Age; breastfeeding; hypothermic state; severely malnourished; severely dehydrated; impaired consciousness; distended abdomen; platelet deficiency; decreased serum protein levels; decreased serum sodium concentration; hypoglycemic state; renal failure; and bacterial sepsis.
Z. Bhutta et al., 1996	Pakistan	Anorexia; drowsiness; respiratory distress; anemia; and a positive blood culture.

Acácio et al., 2019	Mozambique	Typical EPEC infection; Cryptosporidium infection; malnutrition; pneumonia/lower respiratory infection; invasive bacterial disease; lethargy or overt unconsciousness; wrinkled skin; HIV; and negative for rotavirus.
Bado et al., 2016	Burkina Faso, Mali, Nigeria, and Niger	Child age; sex; breastfeeding modality; low weight-for-age; being acutely malnourished; sanitary deficiencies; access to safe drinking water; and lack of safe and hygienic waste disposal methods; lower wealth index; food shortage; poor handwashing practice; water storage modality; household socioeconomic standing; a large household size or a large number of children under five residing in the home; mother's age; literacy level; and occupational status.
Acácio et al., 2019	Mozambique	Typical EPEC infection; Cryptosporidium infection; malnutrition; pneumonia/lower respiratory infection; invasive bacterial disease; lethargy or overt unconsciousness; wrinkled skin; HIV; and negative for rotavirus.
Griffin et al., 1988	Lesotho	Co-infection; age; ≥ 7 days of illness prior to hospitalization; oral candidiasis or stomatitis diagnosed at admission; severely dehydrated; previous episodes of emesis; persistent dehydration despite 12 hours of inpatient management; and pyrexia or hypothermia.
Deribew et al., 2016	Ethiopia	Being wasted; low weight-for-age; being stunted; not exclusively breastfeeding; and hypovitaminosis A.
Masanja et al., 2019*	Tanzania	Oral rehydration solution (ORS) utilization; variations in the prevalence of stunting; enhancement of vitamin A intake; rotavirus vaccination; shifts in the prevalence of wasting; and suboptimal breastfeeding practices
O'Reilly et al., 2012	Kenya	Non-typhoidal <i>Salmonella</i> infection; <i>Shigella</i> infection; negative for rotavirus; malnourished; oral thrush; prior care-seeking; and dehydration
Talbert et al., 2019	Kenya	Co-infections of both diarrhea and severe pneumonia; HIV; lower MUAC; bacteremia; circulatory problems; septicemia; and electrolyte imbalances; prior admission; and chest indrawing.

* Factors that contributed to the decline of diarrheal mortality

3. CHAPTER THREE: THE INFORMATICS CONTEXT

3.1 Introduction

This chapter explores the informatics context of the study. Through an extensive review of current literature, its primary objective is to offer an in-depth understanding of the prevailing landscape of predictive modeling for each of the poor diarrheal outcomes. This chapter commences with a detailed exploration of the ML paradigm. It systematically identifies relevant literature, analyzes past studies and trends in predictive modeling, and discusses critical knowledge gaps in predicting poor diarrheal outcomes. Furthermore, this chapter critically evaluates key considerations for selecting a deployment framework, offering insightful perspectives into the diverse panorama of frameworks available in the deployment of ML models, particularly in the context of R. Additionally, it explores the theoretical and conceptual framework guiding this study, outlining the evidence that served as the foundation for the findings presented within.

3.2 Machine Learning

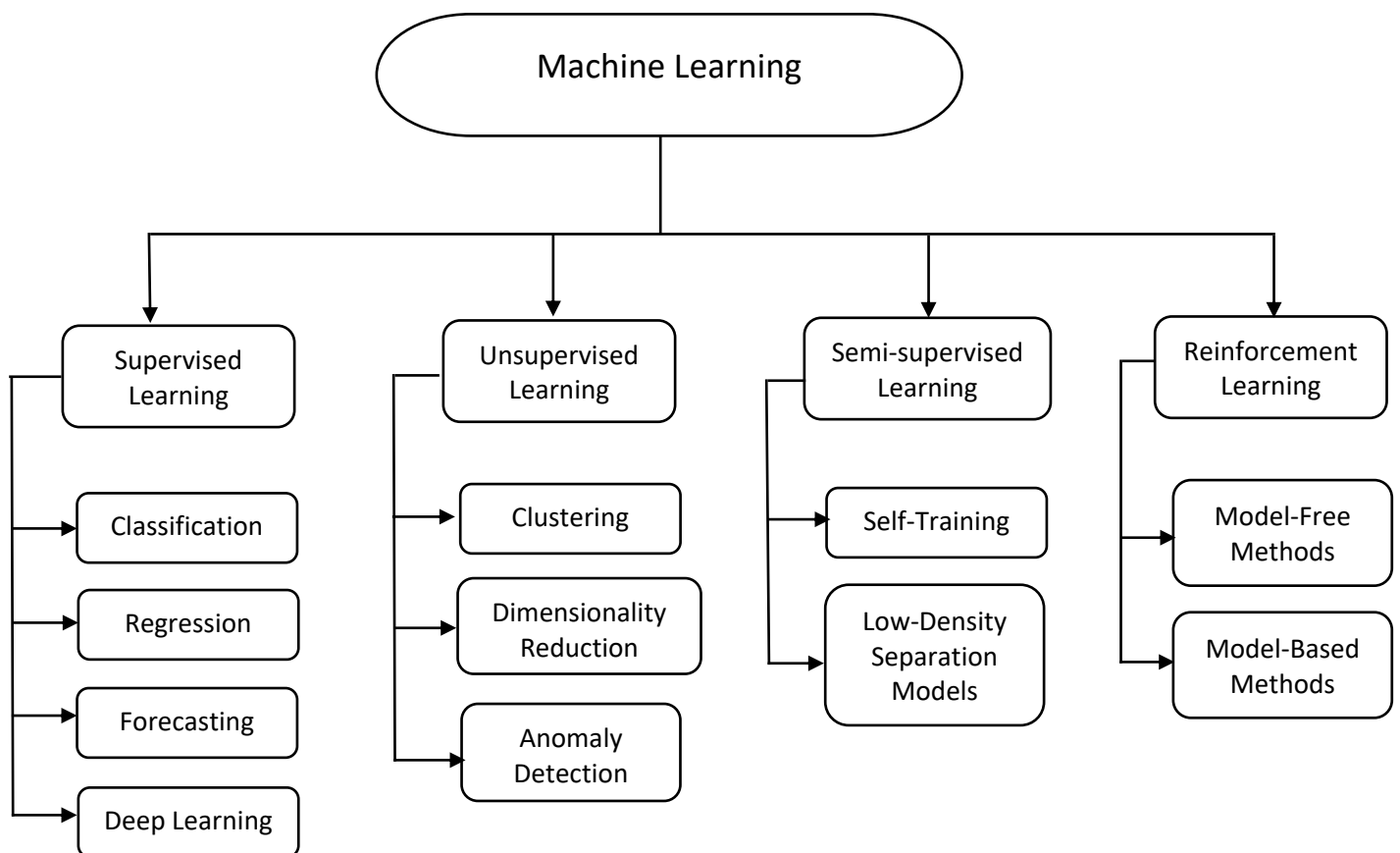
As presented in Chapter One, ML is an arm of artificial intelligence responsible for constructing applications that glean knowledge from data and continuously improve their predictive power autonomously, without needing additional programming in order to achieve these improvements (IBM, 2020). The history and adoption of ML in healthcare represents a transformative journey through time, and is indeed one of the sagas from the tech space. In the early years, from the 1950s to the 1970s, ML techniques were cautiously explored for rudimentary tasks like diagnostic support. As the field progressed into the 1980s, the emergence of expert systems like MYCIN and Dendral demonstrated the promising role of AI in the health domain by incorporating medical

knowledge into reasoning systems. The 1990s saw the development of knowledge-based systems, leveraging methods such as Bayesian networks and decision trees for medical diagnosis. With the dawn of the 21st century, the adoption of electronic health records (EHRs) provided a treasure trove of healthcare data, offering new possibilities for ML applications. The decade of the 2010s was characterized by a significant transition in AI, importantly the emergence and rapid advancement of deep learning, specifically convolutional neural networks and recurrent neural networks. These breakthroughs enabled the remarkable progress of ML in healthcare, particularly in medical imaging and natural language processing (Habeheh & Gohel, 2021; Jones et al., 2018).

In terms of adoption, ML has permeated diverse areas in healthcare. It significantly enhances medical imaging, aiding radiologists to detect and diagnose diseases. Disease diagnosis and risk assessment have also benefited from ML's predictive capabilities, contributing to early disease detection. ML accelerates drug discovery, and it utilizes patient-specific data to develop individualized treatment regimens, advancing the field of precision medicine. It is a vital tool in health monitoring through wearable devices and remote patient management systems. Natural language processing (NLP) enables the extraction of meaningful information from the vast corpus of unstructured medical data like clinical notes and research papers. Predictive analytics assist decision makers in matters of measuring patient outcomes and deciding about resource allocation, optimizing hospital operations. Telemedicine platforms, chronic disease management, patient engagement, and other types of e-health systems have integrated ML for improved healthcare delivery.

The adoption of ML in healthcare is expanding, with the potential to improve diagnostic precision, the effectiveness of treatments, and the quality of patient care. However, it is important to navigate

the ethical and regulatory considerations in handling sensitive patient data and to ensure the highest standards of privacy and security in healthcare applications. ML algorithms offer numerous categorization possibilities, including learning style, task type, model complexity, data availability, use cases, algorithm family, scalability, and interpretability. In this context, the study will focus on the primary classification based on learning style (Sarker, 2021). ML can be broadly delineated into four fundamental learning styles: supervised learning; unsupervised learning, semi-supervised learning; reinforcement learning (Figure 3.1).



Adapted from Moubayed et al. (2018).

Figure 3.1: Machine Learning Overview (Moubayed et al., 2018)

3.2.1 Supervised Learning

Supervised learning is a ML paradigm, which utilizes labeled data to train algorithms. These algorithms learn to establish a rule that maps each input to its corresponding output. Subsequently, they apply this learned functioning to make predictions on unseen data (Cunningham et al., 2008; Sarker, 2021). Supervised learning can be broadly classified into four distinctive subgroups, each catering to different task types: Classification, Regression, Forecasting, and Deep Learning, which are discussed below:

- **Classification-** It involves the prediction of predefined classes or categories using a set of known values. This technique is applicable to both organized and free-flowing data. Classification problems can be categorized into three distinct types: binary, multi-class, and multi-label classification. In summary, binary classification is used when the target variable is binary or has two classes and involves the categorization of data into one of two possible classes or categories (Kumari & Kr., 2017). On the other hand, multi-class classification is used when the target variable has more than two class labels and it involves the classification of data into more than two distinct classes (Grandini et al., 2020; Sahare & Gupta, 2012). Lastly, multi-label classification allows for instances to have multiple labels simultaneously (Herrera et al., 2016; Tarekegn et al., 2021). Multi-label classification differs from regular classification by allowing items to be assigned to multiple non-exclusive classes, rather than just one.
- **Regression-** Refers to ML methods that focuses on predicting a target variable that is continuous as a function of given input features (Saleh & Layous, 2022; Sarker, 2021). Whereas classification entails the prediction of distinct class labels, regression predicts a continuous quantity. Some common regression algorithms include: Linear Regression;

Polynomial Regression; Ridge Regression; Lasso Regression; Support Vector Regression (SVR); Elastic Net; Bayesian Regression; Quantile Regression; and Robust Regression.

- **Forecasting-** Refers to the process of predicting future events or trends using algorithms and historical data (Bojer, 2022; Bontempi et al., 2013). Some common forecasting algorithms include: Autoregressive Integrated Moving Average (ARIMA); Seasonal ARIMA (SARIMA); Exponential Smoothing Methods; Prophet; and Vector Autoregression (VAR).
- **Deep Learning-** Refers to the process of learning representations of data with high-level abstractions using computational models composed of multi-processing layers made up of non-linear transformations or complex structures (Hao et al., 2016; LeCun et al., 2015). Prominent deep learning algorithms and architectures include: Multi-layer Perceptron; Convolutional Neural Networks; Recurrent Neural Networks; Feedforward Neural Network; Generative Adversarial Networks; and Long Short-Term Memory.

3.2.2 Unsupervised Learning

Unsupervised learning refers to a set of techniques whose primary goal is uncovering patterns, structures, and relationships in data without relying on annotated or predefined outcome variables (James et al., 2023). In contrast to supervised learning that entails training a model to predict specific outcomes, unsupervised learning involves use of algorithms to uncover underlying patterns, detect significant trends and structures, categorize data into clusters, and support exploratory analysis. Typical applications of unsupervised learning include: clustering; dimensionality reduction; and anomaly detection. Specifically:

- Clustering- Refers to learning approach that entails organizing analogous or closely related data points into categories called clusters without worrying about a specific outcome (Chander & Vijaya, 2021; Serra & Tagliaferri, 2019). Objects within a cluster exhibit greater similarity among themselves compared to those assigned to a different cluster. It is usually utilized to discover patterns and trends in data. Common clustering algorithms include: K-Means; Hierarchical Clustering; and DBSCAN.
- Dimensionality reduction- Refers to techniques that set out to eliminate redundant or irrelevant features in a dataset while retaining essential information and underlying patterns (Espadoto et al., 2021; Nanga et al., 2021). It aims to simplify data representation by mapping it into a space with fewer dimensions, making it more manageable and computationally efficient. Dimensionality reduction is especially beneficial when handling datasets with a large number of features, since it has the potential to improve model performance, reduce overfitting, and enhance visualization. Some common methods for dimensionality reduction include: Principal Component Analysis (PCA); t-Distributed Stochastic Neighbor Embedding (t-SNE); and autoencoders.
- Anomaly detection- Refers to a ML technique employed in the identification and flagging of data points or observations that deviate significantly from the majority of the dataset (Ahmed et al., 2016; Omar et al., 2013). These deviating data points are often referred to as anomalies, outliers, or novelties. The primary goal of anomaly detection is to uncover unusual or unexpected patterns or behaviors within a dataset. Common anomaly detection methods include: Isolation Forest; One-Class SVM (Support Vector Machine); and Autoencoders.

3.2.3 Semi-supervised learning

Semi-supervised learning is a machine learning approach that bridges the gap between supervised and unsupervised learning, leveraging both labeled and unlabeled data for improved model performance (van Engelen & Hoos, 2020). In semi-supervised learning, the training dataset comprises both labeled and unlabeled instances allowing ML models to leverage the benefits of both labeled data, which provides explicit supervision, and unlabeled data, which can help discover underlying patterns and relationships. Some common semi-supervised approaches include: self-training; and low-density separation models. Self-training is an iterative process where the model starts with a small amount of labeled instances and utilizes them in making predictions on the unlabeled instances (Oymak & Gulcu, 2020). The predictions with high confidence are added to the labeled data, and the process repeats. Semi-supervised low-density separation models are ML models that combine the principles of semi-supervised learning with low-density separation techniques (Burkhart & Shan, 2020; van Engelen & Hoos, 2020). These models are used for the task of identifying rare or anomalous instances in a dataset when both labeled and unlabeled data are available and the boundaries between the classes or clusters are characterized by regions of low data density. By incorporating the benefits of semi-supervised learning, they can enhance the predictive power and robustness of low-density separation in scenarios where labeled data is limited.

3.2.4 Reinforcement learning

Reinforcement learning is a paradigm for teaching agents to learn optimal decision-making strategies in an environment by maximizing cumulative rewards (Ding et al., 2020; Woergoetter & Porr, 2008). The agent's learning process is characterized by experiential acquisition of

knowledge, derived from observing the consequences of its actions, instead of being explicitly instructed. It makes decisions based on an evaluation of both its historical experiences (exploitation) and novel choices (exploration), in a procedure that is analogous to trial-and-error learning. This approach draws inspiration from the psychology of behavior, where learning is driven by interactions with the environment and the consequences of actions. Reinforcement learning is commonly used in scenarios wherein an agent is tasked with learning sequential decision-making to achieve a specific goal or maximize a cumulative reward. Reinforcement learning can be categorized into two main approaches (Kaiser et al., 2020; Moerland et al., 2023), specifically:

- **Model-Free Methods:** It involves methodologies that are agnostic to environmental dynamics or transition probabilities. Instances of this approach include: Q-Learning; SARSA; and deep reinforcement learning algorithms.
- **Model-Based Methods:** Model-based approaches involve learning or estimating the dynamics of the environment and then using this model to plan and make decisions. These methods can be more sample-efficient but require knowledge of the environment's structure. These methods include: Monte Carlo Tree Search (MCTS); Dynamic Programming; Probabilistic Models; Simulated Annealing; Model Predictive Control; and Cross-Entropy Method.

3.2.5 Machine learning algorithms for predicting poor diarrheal outcomes

The predictive tasks related to poor outcomes in diarrheal disease were structured as binary classification problems utilizing labeled data, thus adopting a supervised learning methodology. The selection of algorithms for this study was meticulously guided by factors such as data

complexity, accuracy requirements, processing efficiency, time constraints, and feature dimensions (Sarker, 2021). Given that algorithms encompass diverse families, each with distinct objectives and methodologies for classification tasks, a spectrum of algorithms representing various families that have previously demonstrated their efficacy in healthcare-related predictions were chosen. This approach enabled us to perform a comprehensive comparative assessment of their predictive performance in the context of poor diarrheal outcomes.

The choice of algorithms was informed by their popularity, robustness, widespread application, and variance in performance characteristics. Logistic regression serves as the default model and serves as the benchmark against which the other selected algorithms are evaluated. To construct predictive models for identifying poor diarrheal outcomes in children under the age of five, the potential of seven distinct ML algorithms was harnessed, each contributing its unique strengths to the analysis. These algorithms encompass Logistic Regression (LR), Naive Bayes (NB), Random Forest (RF), Gradient Boosting (GBM), Support Vector Machine (SVM), K-nearest neighbors (KNN), and Artificial Neural Networks (ANN). In this section, a detailed description of each of these algorithms is given as well as outlining their application in healthcare.

3.2.5.1 Logistic Regression (LR)

LR is a fundamental and extensively utilized statistical approach in ML, particularly for binary classification tasks. This model aims to predict a binary outcome based on ≥ 1 predictors by approximating the likelihood of an event occurring (Schein & Ungar, 2007; Shipe et al., 2019; Thabtah et al., 2019). In LR, the sigmoid function is used to generate a linear transformation of the predictors into a probability value between 0 and 1. LR, as described by Hosmer and Lemeshow, and can be understood as originating from a Bernoulli model. When provided with a

set of predictors represented as X_n , the model is designed to estimate the probability of a dichotomous dependent variable, denoted as Y_n . The probability model is defined in Equation 1 (Schein & Ungar, 2007):

$$P(Y_n = 1 | X_n = \sigma(W \cdot X_n)) \dots\dots\dots (1)$$

and the corresponding likelihood function is given by Equation 2:

$$\begin{aligned} P(\mathbf{y} | \mathbf{x}_n, n = 1 \dots N) &= \prod_n \sigma(\mathbf{w} \cdot \mathbf{x}_n)^{y_n} (1 - \sigma(\mathbf{w} \cdot \mathbf{x}_n))^{(1-y_n)} \\ &= \prod_n \sigma(\mathbf{w} \cdot \mathbf{x}_n)^{y_n} \sigma(-\mathbf{w} \cdot \mathbf{x}_n)^{(1-y_n)} \end{aligned} \dots\dots\dots (2)$$

And the logistic function represented by $\sigma(\theta)$, is a sigmoidal curve that transforms a linear combination of predictors to a probability score ensuring that the predictions are bounded between 0 and 1 consistent with the probabilistic nature of the model's output as shown in Equation 3:

$$\sigma(\theta) = \frac{1}{1 + \exp[-\theta]} \dots\dots\dots (3)$$

LR is a parametric method because the contribution of parameters in logistic regression (coefficients and intercept) can be interpreted. It is valuable in the healthcare sector due to its interpretability and ability to model the probability of binary outcomes, making it a key tool for making informed decisions in medical diagnosis, treatment planning, and patient management. Specifically, LR has been successfully implemented in a number of classification tasks in the healthcare domain: the detection of Autistic Spectrum Disorder in adults and adolescents (Thabtah et al., 2019); the prediction of faltered growth in children under five years seeking care for moderate or severe diarrhea (Brander et al., 2019); the prediction of comorbidities in children with cerebral palsy (Bertoncelli et al., 2020); predicting lung cancer in indeterminate pulmonary

nodules (Deppen et al., 2014); prediction of mortality following surgery (Bilimoria et al., 2013); the prediction of admission (Patel & Saini, 2021).

LR offers several advantages in the field of classification (Boateng & Abaye, 2019; Hu et al., 2025; Rahaman, 2024). First, its simplicity is a key asset. The model is easy to comprehend and implement. Additionally, LR is highly interpretable with the coefficients associated with independent variables directly indicating the impact of each variable on the logarithmic transformation of the odds of the outcome, providing clear insights about the relationship between variables. Moreover, efficiency is another important advantage. LR has a low computational cost, making it suitable for handling large amounts of data and high-dimensional feature spaces, which makes it well-suited in situations that require rapid analysis. Furthermore, it scales well with the number of observations, proving its versatility for both small and large datasets. One of the distinctive strengths of LR is its ability to provide probabilistic outputs. It estimates the likelihood of an event's occurrence, allowing for an express interpretation of the likelihood of a particular outcome. This probabilistic nature of the model is valuable for decision-making and risk assessment. In addition, LR can be extended to handle multiclass classification tasks. This flexibility broadens its applicability in a range of scenarios.

While LR boasts several advantages, it is not without its limitations (Boateng & Abaye, 2019; Hu et al., 2025; Rahaman, 2024). The model posits a linear association between predictors and the logarithmic transformation of the odds of the outcome that does not capture complex, non-linear dependencies between variables within the dataset. The aforementioned is a fundamental limitation in scenarios where the true relationship is not linear. LR also works under the assumption that predictors are independent and they do not have a direct impact on each other, an assumption

that is rarely met in practical applications with real data. Correlations between features can challenge the model's accuracy and make its assumptions invalid. Furthermore, LR may not perform as well as more complex models when it comes to capturing intricate patterns in the data. In situations where relationships are nonlinear or intricate, other models may be more suitable. Moreover, LR is sensitive to outliers, and extreme values can substantially alter the coefficients and predictions, potentially reducing its performance. Additionally, logistic regression does not handle missing data gracefully and necessitates additional preprocessing steps to address missing values. Its performance may also be affected by the bias-variance tradeoff, and choosing an inappropriate model complexity can lead to either underfitting or overfitting. Finally, the model is less effective at modelling decision boundaries that are not linear, which can be a drawback in tasks characterized by complex, non-linear relationships between variables (Hu et al., 2025).

3.2.5.2 Naïve Bayes (NB)

NB is a probability-based classifier which utilizes the likelihood of terms and categories co-occurring in the estimation of the probabilities of distinct classes represented in the data assuming conditional independence between terms within each category and is insensitive to noise (Berrar, 2019; Langarizadeh & Moghbeli, 2016; Webb, 2016). This algorithm is grounded in Bayesian principles and is termed "naive" because it assumes conditional independence between the predictors used for classification. Despite this simplification, it is a powerful and widely used classification algorithm in various applications, including healthcare.

NB belongs to the family of Bayesian Network Classifiers, utilizing Bayes' theorem for probability estimation as shown in Equation 4.

$$P(y | \mathbf{x}) = P(y)P(\mathbf{x} | y)/P(\mathbf{x}) \dots\dots\dots (4)$$

Alongside the presumption that the attributes exhibit conditional independence when considering the distinct classes as depicted in Equation 5. For datasets with attributes and their corresponding values, the assumption implies:

$$P(\mathbf{x} | y) = \prod_{i=1}^n P(x_i | y) \dots\dots\dots (5)$$

Where x_i represents the value of the i th feature in the input vector $\mathbf{X}$, and n is the dimensionality of the feature space as shown in Equation 6.

$$P(\mathbf{x}) = \prod_{i=1}^k P(c_i) P(\mathbf{x} | c_i) \dots\dots\dots (6)$$

Where k denotes the cardinality of the class set and c_i represents the i -th element of the class set.

However, in many cases, a binary classification isn't the preferred outcome. For instance, in ranking tasks that involve both positive and negative classes, the primary interest lies in how effectively a model orders the instances of one class in relation to those of the other class. The estimated class posterior probabilities naturally serve as ranking scores as depicted in Equation 7 (Berrar, 2019).

$$P(y_j | \mathbf{x}) = \frac{\prod_{k=1}^p P(x_k | y_j) P(y_j)}{\prod_{k=1}^p P(x_k | y_j) P(y_j) + \prod_{k=1}^p P(x_k | y_j^c) P(y_j^c)} \dots\dots (7)$$

NB classifiers are valuable in healthcare for tasks where simplicity and efficiency are crucial. They provide interpretable results and possess broad applicability in improving patient care and operational processes. Specifically in healthcare, NB has been deployed in various tasks including: heart disease detection (Gupta et al., 2020; Jabbar & Samreen, 2016); skin cancer diagnosis

(Balaha & Hassan, 2023); breast cancer detection (Kharya et al., 2014); predicting infection among close contacts with COVID-19 (Yoshikawa, 2022).

NB offers several advantages (An et al., 2023; Pajila et al., 2023; Phatcharathada & Srisuradetchai, 2025). Its foremost strength is attributable to its straightforward design and simple operationalization, making it accessible to both novice and experienced practitioners. This simplicity enables quick deployment and straightforward understanding. Efficiency is another key advantage. NB is computationally efficient, performing admirably even with vast datasets and high-dimensional feature spaces. Its ability to handle large-scale, real-time applications is valuable. The model's interpretability is a significant asset. It produces probabilistic outputs, allowing for a clear understanding of the likelihood of specific outcomes, making it a valuable tool for decision-makers. Furthermore, NB can readily handle multiclass classification tasks, offering versatility across a range of applications. Handling missing data is an additional advantage. NB copes well with missing values without necessitating complex imputation methods. It simplifies data preprocessing and reduces the burden of data cleaning.

While NB offers multiple advantages, it is not without drawbacks (An et al., 2023; Pajila et al., 2023; Phatcharathada & Srisuradetchai, 2025). The "naive" assumption of feature independence is one of its primary drawbacks. In real-world data, features are often correlated, and this assumption may not hold, decreasing the model's ability to make correct predictions. The zero-frequency issue is another concern. If a category or value in the test data has not been observed in the training data, NB assigns a probability of zero, potentially leading to inaccurate predictions. Robust smoothing techniques are often required to address this limitation. Sensitivity to outliers is a disadvantage. Extreme values can significantly influence the model's performance, introducing

bias and potential inaccuracies in the results. Limited expressiveness is also a concern. NB may not capture complex relationships between variables as effectively as more advanced ML models. This can be a limitation when dealing with intricate, nonlinear patterns in the data. Lastly, NB can exhibit high bias, potentially oversimplifying relationships in the data and leading to underfitting, particularly in scenarios where the true relationships between features are more complex.

3.2.5.3 Random Forest (RF)

RF is an ensemble learning method that employs bootstrap aggregation technique to train multiple decision trees to get a more accurate and stable prediction (Abdulkareem & Abdulazeez, 2021; Khalilia et al., 2011; Reis et al., 2018). RF is a comprehensive learning model that incorporates the Decision Tree as its fundamental classifier, and it was introduced by Leo Breiman in 2001. RF operates by using a bagging approach to train multiple decision trees, each on a random sample of the data and predictors, and the ultimate prediction is an aggregation of the predictions from each tree, usually by a majority vote for classification tasks. It improves bootstrap aggregation, minimizing the covariance between the trees with the introduction of randomly splitting a sample of predictors resulting in a model that outperforms any of the individual constituent models.

The Gini Index is the primary formula employed to determine the splitting of decision tree branches when utilizing RF for classification as shown in Equation 8 (Amamra, 2025; Reis et al., 2018; Salman et al., 2024). This formula utilizes both the category and its associated likelihood to calculate the Gini impurity metric of each branch at a node, enabling the identification of the branch with a higher chance of happening. Within this framework, p_i denotes the empirical probability of observing a given class within the dataset, and 'c' represents the cardinality of the class set.

$$Gini = 1 - \sum_{i=1}^c (p_i)^2 \dots\dots\dots (8)$$

Entropy is an alternative method for guiding node splits in a decision tree, relying on the probability of specific outcomes to inform branching decisions. Differing from the Gini index, it involves advanced mathematical complexity because of the use of logarithmic functions in its calculation as depicted in Equation 9.

$$Entropy = \sum_{i=1}^c -p_i * \log_2(p_i) \dots\dots\dots (9)$$

RF's ability in handling data with high complexity and dimensionality and delivering precise predictions has made it a valuable tool in healthcare, contributing to improved diagnosis, treatment, and patient care. Precisely, RF has been utilized in a variety of predictive applications within the healthcare domain, such as: prediction of COVID-19 severity and patient outcomes (Iwendi et al., 2020); feature ranking based approach for medical data classification (Alam et al., 2019); predicting pressure ulcer (Song et al., 2021); the prediction of disease risk of individuals using their diagnosis history (Khalilia et al., 2011); identifying features and predicting susceptibility to type 2 diabetes (Ooka et al., 2021); exploring drivers of patients' overall experience with their care (Simsekler et al., 2021); multi disease prediction system (Shanthakumari et al., 2022).

The RF algorithm offers a combination of strengths and limitations that enhance its utility as a highly effective tool across a broad spectrum of ML tasks (Abdulkareem & Abdulazeez, 2021; Amamra, 2025; Fawagreh et al., 2014; Pensa et al., 2025; Rodriguez-Galiano et al., 2012). On the positive side, RF is prized for its high predictive accuracy. By aggregating multiple decision trees, it excels at capturing intricate relationships within data, resulting in reliable and accurate

predictions. Furthermore, it effectively mitigates the risk of overfitting, thanks to its ensemble approach, this makes it a robust and versatile solution for a variety of tasks. The algorithm's capacity to evaluate variable importance contributes to feature selection and facilitates a deeper understanding of the relationships between variables. Additionally, it exhibits resilience to noisy data and can handle missing values gracefully, simplifying the data preprocessing process. Its compatibility with parallel processing leverages modern hardware for faster model building.

However, RF is not without its challenges. Its inherent complexity, stemming from the ensemble of decision trees, can exhibit reduced interpretability compared to simpler models such as LR. This complexity might make it challenging to understand the inner workings of the algorithm. Furthermore, building a large number of decision trees, as is often the case with RF, can be resource-intensive, consuming both time and memory, particularly for extensive datasets. The algorithm may also exhibit bias toward dominant classes in imbalanced datasets, necessitating additional techniques to address class imbalance effectively. While it is robust to some extent, extremely noisy datasets with a high degree of randomness can still pose challenges for RF. Finally, the size of the model, especially when employing a large number of trees, can be substantial, which should be considered when deploying the algorithm in resource-constrained environments.

3.2.5.4 Gradient Boosting Machines (GBM)

GBM is an algorithm that builds strong learners from weak learners in a gradual, additive and sequential manner offering lots of flexibility allowing for optimization of different loss functions (Bentéjac et al., 2021; Natekin & Knoll, 2013; Zhongheng Zhang, 2019). It was introduced by Jerome Friedman in the late 1990s and has since become a powerful and widely used algorithm in

ML. This algorithm is based on the principle of constructing new base learners that exhibit a strong correlation with the negative gradient of the loss function to compensate for the inaccuracies of the preceding models with the choice of loss functions being flexible. The formula uses gradient descent to find the optimal model parameters by repeatedly adjusting them in a direction that reduces the loss function.

$$F(x) = F_0(x) + \lambda \cdot h_1(x) + \lambda \cdot h_2(x) + \dots + \lambda \cdot h_m(x) \dots\dots\dots (10)$$

Where: $F(x)$ denotes the model's final prediction for the input data instance x ; $F_0(x)$ represents the first prediction (typically a constant value) before any trees are added; λ denotes the learning rate, which controls the step size when iteratively optimizing the model; $h_1(x)$, $h_2(x)$... $h_m(x)$ are the contributions of individual trees added to the ensemble.

The prediction $F(x)$ is the result of combining the initial prediction with the contributions from each tree, scaled by the learning rate. The goal of the GBM algorithm is to minimize a specified loss function by iteratively adding new trees ($h_1(x)$, $h_2(x)$, ...) that correct the errors made by the previous trees. The specific form of $h_i(x)$ for each tree depends on the choice of the loss function and the tree-building algorithm used within the GBM implementation. Each tree is constructed to minimize the loss associated with the ensemble's previous predictions.

Depending on the response variable family, the choice of loss function for categorical responses, where $y \in \{0, 1\}$, can be either the Binomial loss function or the Adaboost loss function. Specifically, for binary outcome variables y , where 1 represents success and 0 represents failure, the Binomial loss function is employed, assuming a Bernoulli distribution. For simplification, let's transform the labels by $y=2y-1$, which places $y \in \{-1, 1\}$. In this context, the class-conditional response probabilities are estimated via minimization of the negative log-likelihood function with

respect to the newly assigned class labels as shown in Equation 11. (Natekin & Knoll, 2013). The Bernoulli loss is the standard name for this loss function:

$$\Psi(y, f)_{Bern} = \log(1 + \exp(-2\bar{y}f)) \dots \dots \dots (11)$$

Where y : the true class label and f : the model's prediction function output. The term yf is the margin (how correct the prediction is)

Furthermore, GBM has demonstrated significant success in its application in the healthcare sector. Specifically, it has been used in the prediction of: risk of sepsis (Delahanty et al., 2019); readmissions following lumbar laminectomy (Kalagara et al., 2018); delirium risk (Wong et al., 2018); risk of cardiovascular disease (Theerthagiri, 2022; Zhao et al., 2019); length of stay among inpatients (Zelege et al., 2023); risk stratification on day 5 of COVID-19 patients with ICU admission (Montomoli et al., 2021); mortality prediction in an emergency department (Klug et al., 2020); COVID-19 associated deaths (Wichmann et al., 2023); Type-2 diabetes risk (L. Wang et al., 2020).

GBM yields better accuracy, requires less pre-processing and offers higher flexibility due to the availability of many hyper-parameter and loss functions. GBM is a powerful and versatile algorithm with several strengths and some inherent weaknesses (Alwanin et al., 2025; Kyrychek, 2025; Natekin & Knoll, 2013). One of its primary strengths lies in its high predictive accuracy, making it well-suited for complex tasks where capturing intricate patterns is essential. GBM is also robust against outliers and exhibits applicability to both regression and classification tasks effectively. GBM provides flexibility in terms of loss functions, allowing it to adapt to diverse data types and structures. Moreover, GBM provides insights into feature importance, aiding in variable selection. However, GBM has a number of shortcomings, its computational demands can be substantial and may require careful hyper parameter tuning which can be time-consuming. When

not configured properly, GBM models can be prone to overfitting, leading to poor generalization on unseen data. The ensemble nature of GBM can make it less interpretable compared to simpler models, which might be a concern in some applications. Furthermore, the need for substantial computational resources can be a drawback when working with large datasets or in resource-constrained environments. Despite these limitations, GBM remains a popular choice in the ML community due to its outstanding predictive capabilities.

3.2.5.5 Support Vector Machine (SVM)

SVM works through discriminating between two classes through the creation of a hyperplane which perfectly segregates classes following the mathematical transformation of features into a high dimensional space (Yu et al., 2010). The SVM algorithm is founded on the theoretical framework pioneered by Vapnik and Chervonenkis and subsequently refined by other researchers (Bui, 2021; Yousefi et al., 2022). Linear SVM learning seeks to discover separating hyperplanes that effectively divide the dataset into distinct data classes. In an optimal scenario, where the data is entirely linearly separable, the hyperplanes will be positioned as far away as possible from the nearest data points of the classes. In the broader context, SVM strives to approximate this condition as closely as possible. In the case of nonlinear SVM, hyperplanes are replaced with different classes of manifolds, while the fundamental principle remains consistent. SVM can handle both classification and regression on linear and non-linear data. Additionally, this method demonstrates superior performance in high-dimensional settings, exhibiting particular strength when the dimensionality surpasses the sample size. SVMs offer the flexibility to employ a variety of kernels, with the Linear Kernel and the Gaussian Radial Basis Function (RBF) kernel being the most commonly used options (Maggioni & Spinelli, 2025; Wijayanti et al., 2022). The linear Kernel is

defined by the function as shown in Equation 12 (Maggioni & Spinelli, 2025; Wijayanti et al., 2022):

$$f(x) = w^T x + b \dots\dots\dots (12)$$

Where: w represents the weight vector defining the hyperplane's orientation; x denotes the input data vector; b represents the bias term or intercept.

The RBF Kernel stands out as one of the most potent and frequently adopted kernels in SVMs, typically the preferred option for handling non-linear data. The equation for the RBF kernel is depicted in Equation 13:

$$f(X_1, X_2) = \exp(-\gamma \|X_1 - X_2\|^2) \dots\dots\dots (13)$$

Where X_1, X_2 is the feature vectors of two observations, $\|X_1 - X_2\|^2$ is the squared Euclidean distance; $\gamma > 0$ is the kernel width parameter given by $1/2\sigma^2$.

SVMs serve as a potent tool for numerous healthcare applications, where precise classification and prediction are integral in enhancing patient management and advancing medical research. They have found successful applications in various healthcare scenarios, encompassing: forecasting medication adherence in heart failure patients (Son et al., 2010); predicting diabetes and pre-diabetes (Yu et al., 2010); addressing Sickle Cell Disease (Khalaf et al., 2019); prognosticating severe acute myocardial infarction outcomes (Zhou et al., 2022); detecting and diagnosing cancer (Zhang et al., 2019); assessing heart disease, cancer, and diabetes (Janardhanan et al., 2015).

SVMs offer a robust and versatile approach to ML, with distinct advantages and limitations (Akkaya & Çolakoğlu, 2019; Auria & Moro, 2008; Jun, 2021; Karamizadeh et al., 2014). On the

positive side, SVMs excel in high-dimensional spaces, making them ideal for problems with many features, including applications like computer vision or NLP. SVMs exhibit robust generalization, preventing overfitting and producing reliable predictions on unseen data. SVMs can tackle both classification and regression problems, having a flexibility that is extended by using various kernel functions to handle non-linear data. Furthermore, they aim to find the global optimum, enhancing the model's stability and performance. However, SVMs are not without their drawbacks. Training SVMs requires significant computing resources, especially for large datasets, which limit their use in time-critical applications. The selection of the right hyper-parameters, including the choice of the kernel and regularization strength, can be complex, requiring thorough parameter tuning techniques. The decision boundaries created by SVMs may not be easily interpretable, especially with complex kernels. This characteristic can compromise the model's transparency and render its decision-making process less readily interpretable. Additionally, SVMs are susceptible to distortion from outlier data points, and their presence can significantly compromise the model's predictive accuracy, often necessitating preprocessing of the data. Therefore, while SVMs offer powerful tools for many ML tasks, it is essential to carefully weigh their pros and cons when applying them to specific applications.

3.2.5.6 K-Nearest Neighbors (KNN)

KNN is a classifier that approximates how likely an observation is assigned to a group based on the group membership of observations within close proximity (Shee et al., 2014). The nearest neighbor rule, originally proposed by Fix and Hodges in 1951, laid the foundation for subsequent advancements in the field (Li et al., 2019). In 1967, Cover and Hart made significant contributions by generalizing key characteristics of KNN, including the determination of the upper bound for the asymptotic error rate. Since its inception, KNN has seen rapid development and continuous

progress, driven by the contributions of numerous researchers. KNN is a non-parametric classification algorithm operating under the assumption that comparable points tend to be in close proximity, and its implementation is simple and intuitive to understand. Additionally, it can learn non-linear decision boundaries.

The KNN algorithm classifies or predicts an observation using the following procedure: 1.) Calculate the distance, typically using the Euclidean method, between a specific observation to every other point in the dataset; 2.) Identify K observations closest to the target observation; 3.) For classification, assign the category which appears most often amongst the closest observations identified in step 2. The Euclidean distance stands as the most commonly employed distance metric, representing the actual shortest distance between two points within the Euclidean space. The distance between points x and y calculated as the square root of the sum of the squared differences between their coordinates gives the Euclidean distance as shown in Equation 14.

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \dots\dots\dots (14)$$

The parameter "k" in the k-NN algorithm determines the number of neighbors considered when classifying a particular data point (IBM, 2023). For instance, if only the nearest neighbor is considered (k=1), the point receives the same classification. Using a small k can lead to greater variability in the results but less systematic error, whereas larger k values can introduce more systemic errors (bias) and less variability. The choice of k should be driven by the characteristics of the data points; noisy or outlier-heavy data might be better classified using a larger k. It is generally advisable to use an odd value for k to avoid ties in classification, with the ideal k value determined by utilizing cross-validation approaches.

KNN has found successful applications in the healthcare domain, addressing a variety of tasks, including predicting rehabilitation potential (Zhu et al., 2007), developing monitoring and advisory systems for diabetes (Lee et al., 2010), classifying medical health big data (Shee et al., 2014; Xing & Bei, 2019), predicting disease risk (Uddin et al., 2022), diagnosing heart disease (Shouman et al., 2012), and categorizing lymph node metastasis in gastric cancer (Li et al., 2012).

KNN is a versatile ML algorithm with distinct advantages and limitations (IBM, 2023; Shouman et al., 2012; Uddin et al., 2022). One of its primary advantages is its simplicity and ease of implementation. KNN doesn't require complex model training, making it an accessible choice for beginners and quick experimentation. Additionally, it is effective in handling multi-class classification problems and can adapt to different data distributions since it doesn't assume any specific data distribution. KNN is especially effective at handling outlier-heavy or data with noise, as it can be robust to these irregularities. Furthermore, KNN's ability to consider local patterns by focusing on nearby data points makes it suitable for applications where data may not have a globally consistent structure.

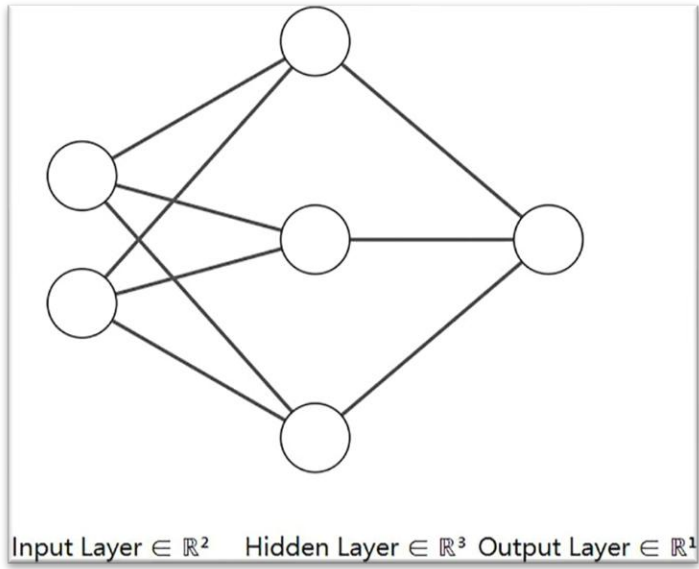
However, KNN also has its limitations. One of the main drawbacks is its computational inefficiency, especially with large datasets. KNN's computational cost scales with the size of the dataset because it needs to measure distances to every data point, a process that can be time-consuming. Choosing the optimal number of nearest neighbors (k) is crucial for model performance, as this parameter significantly affects results. KNN is sensitive to the scaling of features, meaning features with larger scales can dominate the distance calculations. Moreover, the performance of KNN degrades in high-dimensional spaces due to the sparsity of data points making it challenging to find meaningful neighbors. Despite these limitations, KNN continues to

be a useful method for a range of classification and regression applications, especially in scenarios where understanding the model and ease of use are more important than computational efficiency.

3.2.5.7 Artificial Neural Networks (ANN)

ANN are computational networks that simulate the human-brain processes thereby establishing empirical relationships between outcome and features (Shahid et al., 2019; Supriya & Deepa, 2020). ANNs draw inspiration from the densely interconnected, parallel structure of the human brain in processing information. This information-processing paradigm encompasses a diverse range of mathematical models that mimic certain observed characteristics of biological information-processing systems, with a focus on the concept of 'emergent learning'. In emergent learning, the system dynamically adapts to perform specific stimulus-response mapping tasks. The quantitative input, representing the stimulus, is provided to an ANN, which, in turn, produces a quantitative response as its output. A common feature across all ANN models is the presence of a network of interconnected processing nodes, often called neurons, each performing elementary mathematical computations (Figure 3.2).

Within an ANN, a neuron receives inputs from other neurons connected to it and generates an output, which is subsequently transmitted to the neurons it feeds into. These interconnections carry weighted values that regulate the interactions between neurons, and the ANN's capacity for emergent learning arises from the ability to adjust these weights during the learning process until it reaches completion. ANNs learn intricate data patterns without explicit programming, making them versatile tools for various applications like image and speech recognition, and NLP.



Source: Yang et al. (2021). Reproduced under the Creative Commons Attribution (CC BY) license.

Figure 3.2: ANN Architecture (Yang et al., 2021)

An ANN consists of two fundamental components: the summation function and the activation function (Morales-Morales et al., 2025; Petersen & Zech, 2026). When inputs are supplied to a perceptron, the summation function processes these inputs, and subsequently the activation function plays a pivotal role in transforming a combination of input signals where each signal's contribution is weighted within a neuron into the neuron's output signal. The fundamental formula in a neural network is the weighted sum of inputs, which is then transformed using an activation function as shown in Equation 15:

$$y = f(\sum_{i=1}^n w_i x_i + b) \dots \dots \dots (15)$$

Where: y is the output of the neuron; $f(\cdot)$ is the activation function; w_i are the weights assigned to each input x_i ; b represents the bias term.

The activation function adds non-linear terms to the model, enabling neural networks to learn intricate patterns in the data. Several activation functions are available, including ReLu, the sigmoid, and tanh functions.

Training and utilizing ANNs rely on the essential processes of forward and backward propagation (Morales-Morales et al., 2025; Petersen & Zech, 2026). Forward propagation is a fundamental process within ANNs that drives the flow of information from the input layer through to the output layer. It is the mechanism by which input data is processed and predictions are generated. The process commences with the input layer, where data is initially introduced to the network. As the data traverses through the network's hidden layers, every neuron calculates a weighted sum of its inputs and then passes the result through an activation function and subsequently transmits the output to the next layer. The final hidden layer's output is then forwarded to the output layer, which ultimately produces the network's prediction or output for the given input. Essentially, forward propagation allows ANNs to conduct inference and generate predictions based on the learned relationships between the inputs and the desired outputs.

Backward propagation, also known as backpropagation, plays a pivotal role in training ANNs. It is the iterative process used to adjust the network's internal parameters, primarily its weights, to reduce the difference between the network's predictions and the true values. The process initiates with the calculation of the error or loss, which quantifies the disparity between the model's output and the actual target values. Subsequently, gradients of the error concerning the network's weights are determined, signifying how changes in each weight influence the error. These gradients inform the modification of the weights against the direction of the gradients. The size of these weight updates is modulated by a learning rate, dictating the step size during the optimization process.

The entire process, encompassing error calculation, gradient computation and weight updates, is repeated for each training sample, facilitating the network's learning process. The network's performance improves iteratively over many training cycles, and it converges toward a state where prediction errors are minimized, making it capable of making more accurate predictions. Backpropagation is a cornerstone in the training of ANNs, underpinning various optimization algorithms that drive the learning process and improve the performance of the network.

ANNs are proving highly effective in many healthcare applications, demonstrating their efficacy in addressing critical tasks. These applications span a broad range of medical domains, including the prediction of cardiologic auscultation in pediatrics (Bhatikar et al., 2005), breast cancer prediction (Supriya & Deepa, 2020), forecasting hospital length of stay (Ippoliti et al., 2021), anticipating intensive care unit (ICU) outcomes (Gholipour et al., 2015), predicting the outbreak of diseases like COVID-19 (Majhi, 2023; Niazkar & Niazkar, 2020), forecasting pediatric mortality (Pienaar et al., 2022), foreseeing outcomes in trauma patients (Hassanipour et al., 2019), diagnosing pulmonary tuberculosis among hospitalized patients (Aguiar et al., 2016), and classifying conditions such as Crohn's Disease using medical images (S. Ahmed et al., 2017). These applications underscore the versatility and potential of ANNs in improving healthcare outcomes through predictive modeling and diagnosis.

ANNs present a spectrum of advantages and disadvantages in their applications (Foukalas, 2024; Kariri et al., 2023; Mijwil, 2021). On the positive side, ANNs offer remarkable versatility, capable of addressing a diverse array of tasks across different domains, such as healthcare, image recognition, and NLP. They excel in capturing intricate patterns and associations in the data, which makes them powerful tools for applications such as diagnosing illnesses, predicting prognosis of diseases, and analyzing medical images. ANNs also exhibit scalability, effectively handling large

datasets and high-dimensional feature spaces. Their adaptability to evolving data and continuous learning capabilities is particularly valuable in healthcare, where patient information is dynamic. Additionally, advancements in hardware and software have facilitated efficient parallel processing, enabling ANNs to serve in real-time and large-scale applications.

Conversely, ANNs come with their share of challenges. The complexity involved in training and optimizing neural networks can be daunting, demanding specialized knowledge and significant computational resources. Effective training of ANNs requires large datasets, a factor that is a limitation in healthcare settings where obtaining labeled data can be arduous. The interpretability of ANNs remains a concern, as understanding the reasoning behind their predictions is difficult. Overfitting, particularly in complex models with scant data, is a persistent issue that necessitates proper regularization. Furthermore, training large neural networks can be computationally intensive, requiring substantial time and computational power. As healthcare applications increasingly rely on ANNs, these pros and cons should be carefully weighed to harness their full potential while mitigating their limitations.

3.3 Current State of Predictive Modelling

Predictive modelling is an approach in which probability, data mining and ML are applied in identifying trends and relationships within a dataset so as to forecast/predict future outcomes or events (Gennatas & Chen, 2021). The purpose of predictive modeling in healthcare is to estimate the probability of health events occurring in individuals or populations. These insights enable healthcare workers and policymakers to develop preventive measures and implement targeted local interventions, ultimately lowering rates of illness and deaths while alleviating the associated economical and societal burdens (Bandi et al., 2024; Leisman et al., 2020; Wang et al., 2025). This

section provides a review of literature on the current research efforts on predictive models for LDD, chronic malnutrition and Mortality.

3.3.1 Longer Duration Diarrhea (LDD)

A systematic review on predictive modelling for diarrheal disease established that diarrheal outcomes were under researched with no predictive models for LDD reported in literature (Ogwel et al., 2023). This section reviews current research effort in similar prediction tasks for other diseases to establish the methodology, data used and the key findings. Prediction models addressing the length of an illness episode or duration of hospitalization (length of stay) are reviewed in this section. Length of stay refers to the cumulative days spent in a health facility during a patient's single hospitalization (Stone et al., 2022).

A study in the Netherlands explored the prediction of the length of mental health-related work absence (Nieuwenhuijsen et al., 2006). They applied Cox's regression models to this prediction task utilizing clinical, demographic, and socio-economic data to predict duration of sickness absence for 3, 6, and 12 months. Their predictive model yielded areas under the curves (AUC) ranging from 68.0% to 73.0%. A study (2019) conducted in the United States of America (USA) evaluated the prediction of duration of hospitalizations among Corona Virus Disease (COVID-19) patients (Ebinger et al., 2021). They used ML algorithms in the prediction of the probability of a hospitalization stay exceeding 8 days, using electronic health record data. They investigated a wide range of ML algorithms (n=13), including linear models like Elastic-net and logistic regression, tree-based methods like gradient boosted trees and random forest, support vector machines, neural networks, and ensemble-based models. Their models achieved an accuracy of 76.5%, with an AUC of 81.9%. Another study in Sweden examined the length of time patients with knee osteoarthritis were off work sick (Holm et al., 2021). They developed Piecewise constant hazards regression

models utilizing detailed microdata connected from five national registries in Sweden. With a 95% confidence interval of 0.72 to 0.78, their model's AUC was 0.75. Another study in the U.S. developed predictive models to flag intensive care unit (ICU) patients at high risk for prolonged stays (Kramer & Zimmerman, 2010). They predicted ICU length of stay using a multivariable linear regression model built on electronic health data. Their model achieved an r^2 of 20.2% across individuals and 44.3% across units.

Lisk et al. developed a predictive model for the duration of hospitalization among older patients within the United Kingdom (UK) (Lisk et al., 2019). They applied logistic and linear regression to demographic and clinical data to predict duration. They estimated the duration of stay as 11.6, 15, and 9-14 days for discharge to other places, stroke, and acute and recurrent fractures, respectively. Combined, these factors accounted for 32% variability in the duration of hospitalization. Another study in Belgium looked at the prediction of hospitalization duration among patients with severe illness and organ dysfunction (Houthoofd et al., 2015). They explored various ML algorithms utilizing electronic health records of patients with SVM emerging as the champion model for the prediction of prolonged stay with an AUC of 82.0%. The regression of duration of stay achieved a mean absolute error and a median absolute error of 1.8 and 1.2 days among patients with a non-prolonged hospitalization, respectively. Another study in Taiwan predicted the duration of hospitalization among first-ever ischemic stroke patients (Chang et al., 2002). They applied multiple regression to demographic and clinical data to predict duration of hospitalization. Mild-to-moderate neurological impairment increased the duration of hospitalization by approximately one day, while severe impairment shortened it by about one day. A different study in Taiwan investigated the duration of hospitalizations for burn patients using linear regression, model-tree-based regression and support vector machine regression with electronic health records data (Yang

et al., 2010). Their findings showed that predicting the duration of hospitalization at the acute phase showed no improvement in predictive power, compared to those made at admission. However, a significant improvement in the accuracy of length-of-stay prediction was observed in the post-treatment phase. Gabriel et al. in their study built and evaluated predictive models for prolonged stay in post anesthesia care for outpatient surgeries in USA (Gabriel et al., 2017). They built a logistic regression model using electronic health record data and achieved an AUC value of 72.2% (95% CI 69.8%–74.7%). In another study conducted in the USA, Gabriel et al. investigated the prediction of patients not needing a protracted stay after elective primary total hip arthroplasty (Gabriel et al., 2019). They built and evaluated various ML algorithms (Logistic regression, Ridge regression, Lasso and random forest) using electronic health records with the champion model achieving an AUC of 73.5% (95% CI, 67.5%–78.7%). A different study, predicted the duration of hospitalization among individuals who experienced stroke seeking care at a tertiary Saudi Arabian hospital (Al Taleb et al., 2017). They utilized various ML algorithms on patient data in this prediction task. Their Bayesian-network based model correctly predicted the outcome in 81.3% of cases. Another study, still in Saudi Arabia, predicted the duration of stay among patients with cardiac arrest (Daghistani et al., 2019). They employed different machine algorithms to electronic medical records in this prediction task, with a random forest model achieving the best results: accuracy (80.0%), and AUC (94.0%).

Verburg et al. researched various regression techniques in modelling the duration of ICU stay in the Netherlands (Verburg et al., 2014). They explored the following methods using electronic health records data: ordinary least squares regression; a generalized linear model; Poisson regression; negative binomial regression; Gamma regression; and the APACHE IV model, both before and after recalibration. The models' predictive power was as follows: R^2 (0.09 and 0.20);

root mean squared prediction error (RMSPE) (7.3 and 8.7 days); mean absolute prediction error (MAPE) (3.0 and 4.4 days). Alghatani et al. built and evaluated predictive models for the duration of ICU stay, by leveraging patient vital signs from the MIMIC dataset in the USA (Alghatani et al., 2021). They employed six machine algorithms to this prediction task with random forest model, yielding the best accuracy of 65.0%. Alsinglawi et al. studied the prediction of duration of stay among lung cancer patients using an explainable ML framework leveraging the MIMIC dataset (Alsinglawi et al., 2022). Their random forest model outperformed the other algorithms achieving an AUC of 98.0%, with CI 95%: 95.3–100.0%. A study conducted in Iran explored the prediction of duration of hospitalization among individuals with trauma (Gholipour et al., 2015). They utilized ANN to predict duration of hospitalization using clinical and laboratory data achieving 75% and 96.26% for sensitivity and specificity, respectively. Cai et al. built and evaluated predictive models for readmission and length of stay across various diseases (Cai et al., 2016). They built a Bayesian Network model using electronic health data from a hospital in Sydney Australia achieving an accuracy of 80% and AUC of 82.0%.

Levin et al. explored the prediction of all-cause duration of hospital stay among American patients (Levin et al., 2021). They applied ML algorithms to electronic health records data, achieving an AUC of 70.0% to 80.0% for same-day and next-day predictions, sensitivity was between 63.0%-83.0% and specificity between 48.0%-80.0%. Ling et al. also studied length of hospital stay among septic patients using the MIMIC dataset (Ling et al., 2018). They built a random forest model that achieved the following performance: AUC= 86.0%; sensitivity=88.0%, and specificity =71.0%. Ettema et al. conducted a systematic review and validation study within the predictive models for extended ICU stays post-cardiac surgery (Ettema et al., 2010). They used patient data from surgical

cardiac patients in the Netherlands where the Parsonnet score had the best discriminatory ability (AUC=75.0% [95% CI, 73.0% to 76.0%]).

The synthesis of the aforementioned evidence is presented comprehensively in Table 3.1. To provide a concise overview, the prediction of the duration of stay was approached either as a classification task with a binary outcome variable or as a regression task with a continuous variable, which subsequently dictated the type of metrics reported. ML algorithms emerged as the predominant modeling technique, although certain studies also incorporated regression methods. Only one study adopted an explainable ML approach. The majority of these investigations drew upon patient data extracted from electronic health records, encompassing demographic and clinical information. The geographical scope of data collection was predominantly focused on the USA and Europe. The publication period spanned from 2002 to 2022, with a noticeable concentration of publications in the past decade. The area under the curve (AUC) constituted the most frequently reported metric, with estimates ranging from 65.0% to 98.0%. These observations align closely with the findings of Lequertier et al., who reported that 98.6% of studies in their review employed patient data, with regression methods and ML being the prevalent modeling techniques (Lequertier et al., 2021). Particularly in the last decade, ML has emerged as the dominant approach. The review of the burden, determinants, and the present state of predictive modeling provides a robust rationale for leveraging clinical, demographic, anthropometric, and socio-economic data in the prediction of LDD using diverse ML algorithms. Moreover, the incorporation of explanatory model analysis to enhance the understanding of the predictive mechanisms is merited, as only a solitary study has ventured into this territory.

Table 3.1: Current State of predictive modelling for LDD: A summary of literature

Author; Country	Data used	Outcome assessed	Modelling technique	Model performance
Ebinger et al., 2021; USA	Electronic health records (EHR) data	Duration of hospitalization for COVID-19	A wide range of machine learning algorithms (n=13) including: linear models like Elastic-net and logistic regression; tree-based methods like gradient boosted trees and random forest; support vector machines; neural networks; and ensemble-based models.	Accuracy of 76.5%, with an AUC of 81.9%.
Kramer & Zimmerman, 2010; USA	EHR data	Duration of ICU stay	Multivariable linear regression	R ² of 20.2% across individuals and 44.3% across units.
Gabriel et al., 2017; USA	EHR data	Duration of stay among outpatient surgery patients	Logistic regression (LR)	AUC of 72.2% (95% CI 69.8%–74.7%).
Gabriel et al., 2017; USA	EHR data	Duration of stay among outpatient surgery patients	Machine learning (ML): LR, Lasso, ridge regression and random forest (RF)	AUC of 73.5% (95% CI, 67.5%–78.7%
Alghatani et al., 2021; USA	EHR data	Duration of stay in ICU	Machine learning	Accuracy of 65.0%.
Alsinglawi et al., 2022; USA	EHR data	Duration of hospitalization for lung cancer patients	ML: LR, RF; GBM	AUC of 98.0%, with CI 95%: 95.3–100.0%.
Levin et al., 2021; USA	EHR data	All-cause duration of hospitalization	ML: Bayesian Network Model	AUC of 70.0% to 80.0%
Ling et al., 2018; USA	EHR data	Duration of hospitalization for septic patients	ML: RF	AUC= 86.0%; sensitivity=88.0%, and specificity =71.0%.
Nieuwenhuijsen et al., 2006; the Netherlands	Clinical, demographic and socio-economic data	Length of mental health-related work absence	Cox's regression	AUC ranging from 68.0% to 73.0%.
Holm et al., 2021; Sweden	Comprehensive microdata from national registry	Length of time patients with knee osteoarthritis were off work sick	Piecewise constant hazards regression	AUC of 0.75 (95 % CI 0.72–0.78).
Lisk et al., 2019; UK	Demographic and clinical data	Duration of hospitalization for older patients	Logistic and linear regression	The estimated duration of stay was 11.6, 15, and 9-14 days for discharge to other places, stroke, and acute and recurrent fractures, respectively Combined, these factors accounted for 32% variability in length of hospital stay.
Houthoofd et al., 2015; Belgium	EHR data	Duration of stay among patients with severe illness	ML; SVM	AUC of 82.0%; Mean absolute error of 1.79 days

		and organ dysfunction		
Verburg et al., 2014; Germany	EHR data	Duration of stay in ICU	Ordinary least squares regression; a generalized linear model; Poisson regression; negative binomial regression; Gamma regression;	R^2 (0.09 and 0.20); root mean squared prediction error (RMSPE) (7.3 and 8.7 days); mean absolute prediction error (MAPE) (3.0 and 4.4 days).
Ettema et al., 2010; the Netherlands	EHR data	Duration of hospitalization for cardiac surgery patients	Parsonnet; EuroSCORE	AUC=75.0% [95% CI, 73.0% to 76.0%].
Cai et al., 2016; Australia	EHR data	All-cause duration of hospitalization	Machine learning: Bayesian Network Model	Accuracy of 80% and AUC of 82.0%
Chang et al., 2002; Taiwan	Demographic and clinical data	Duration of hospitalization for first-ever ischemic stroke	Multiple regression	Mild-to-moderate neurological impairment increased duration of hospitalization by approximately one day, while severe impairment shortened it by about one day
Yang et al., 2010; Taiwan	EHR data	Duration of hospitalization for burn	Linear regression; model-tree-based regression; and support vector machine regression	Length of stay predictions at the post-treatment stage were significantly more accurate
Al Taleb et al., 2017; Saudia Arabia	EHR data	Duration of hospitalization for stroke	ML	Accuracy of 81.3%
Daghistani et al., 2019; Saudia Arabia	EHR data	Duration of hospitalization for cardiac arrest	ML	Accuracy (80.0%), and AUC (94.0%).
Gholipour et al., 2015; Iran	Clinical and laboratory data	Duration of hospitalization for traumatic patients	ML: ANN	75% and 96.26% for sensitivity and specificity, respectively

3.3.2 Chronic malnutrition

A comprehensive literature review on the ongoing research endeavors related to predictive models for chronic malnutrition is presented in this section. The review encompasses an examination of existing research initiatives that focus on the prediction of stunting and LGF. This analysis serves to elucidate the methodologies employed, the sources of data utilized, and the significant findings generated within this field of study.

Rahutomo et al. conducted a systematic literature review to provide a preliminary understanding on the landscape of ML applications within stunting domain, serving as a foundation for shaping an effective research roadmap (Rahutomo et al., 2023). They reviewed 22 meticulously selected manuscripts. The primary focus of their review was to delineate the ML models employed in research pertaining to stunting, as well as the datasets utilized in studies conducted between 2017 and 2022. Their results revealed the integration of ML techniques in investigations into stunting as early as 2017, with a proliferation and diversification of ML implementations observed during 2021-2022. Among the ML models employed, XGBoost and RF emerged as the two predominant choices, with stunting prediction being the most prevalent ML application. Additionally, it was found that a significant proportion of ML-infused stunting studies originated from Indonesia. While data from nationwide surveys represents the most frequently employed data in stunting studies in the rest of the world, Indonesian researchers primarily use data from regional or independent surveys.

Looking at individual studies, Brander et al. harnessed multisite data spanning Africa and Asia, drawing from the GEMS study (Brander et al., 2019). Their objective was to predict the occurrence of LGF in children aged less than 24 months, subsequent to experiencing episodes of moderate-to-severe diarrhea. They adopted a logistic regression approach as their modeling technique. Their predictive model incorporated clinical and sociodemographic factors, ultimately achieving an AUC of 67.0% (95% CI: 64.0%- 69.0%). In a recent study, Ahmed and his team made use of data from both the GEMS study and the Etiology, Risk Factors, and Interactions of Enteric Infections and Malnutrition and the Consequences for Child Health and Development (MALED) study. Their objective was to construct and externally validate a clinical prediction tool designed for the swift identification of children under 60 months of age who are at an elevated risk of experiencing LGF,

subsequent to episodes of moderate-to-severe diarrhea (Ahmed et al., 2023). To achieve this, they harnessed ML techniques, employing both random forest and logistic regression algorithms to analyze demographic, clinical, and socio-economic factors. Their model yielded an AUC of 75.0% (95% CI: 75.0, 75.0) in the development study (GEMS), while the external validation conducted in the MALED study resulted in an AUC of 68.0% (95% CI: 63.0%, 74.0%).

Furthermore, in a study by Hanieh et al., the prediction of all-cause stunting in preschool children was conducted. The study leveraged data from two Vietnamese longitudinal cohort studies conducted, which followed children from embryonic period through to 36 months of age (Hanieh et al., 2019). The analysis encompassed a wide array of potential predictors, including mother-related, father-related, child-specific, socio-economical, environment-based, laboratory-assessed, and nutrition-based factors, spanning a data collection period from pregnancy to six months post-partum. The modeling approach employed in this study was logistic regression. Their predictive model achieved an AUC of 85.0% (95% CI, 80.0–90.0) in the validation set. Furthermore, in a parallel study, Rahman et al. undertook predictive modeling for all-cause stunting among children under 60 months of age in Bangladesh. They based their analysis on data derived from the Demographic Health Survey conducted in the country in 2014 (Rahman et al., 2021). The study incorporated the utilization of three ML algorithms, namely SVM, RF, and LR, for the prediction of stunting within this context. These algorithms were applied to evaluate a range of socio-economic and demographic factors as potential predictors. The RF model exhibited the most exceptional performance, boasting an accuracy rate of 88.3%. Another research investigation conducted in Bangladesh focused on the prediction of the risk of stunting in children aged 12-23 months (Haque et al., 2023). This study utilized data from two cross-sectional studies, both at baseline and end line, and undertook an assessment of various child, parental, and household

characteristics as potential predictors. To achieve this, logistic and probit regression models, alongside a decision tree model, were employed. Among these models, the best-performing one attained an AUC of 61.0%.

Additionally, Ndagijimana et al. conducted a predictive analysis focusing on all-cause stunting in children aged less than 60 months, utilizing data extracted from the Rwanda DHS conducted between 2019 and 2020. To achieve this, a range of ML algorithms, including SVM, NB, RF, LR, and XGBoost were applied in assessing child, parental, and household characteristics as potential predictors of all-cause stunting (Ndagijimana et al., 2023). Among these algorithms, the GBM model emerged as the most effective, yielding an AUC of 89.0%. Similar to the study by Ndagijimana et al., Shen and their team conducted predictive modeling for all-cause stunting in children under 60 months of age (Shen et al., 2023). They drew upon data from the Papua New Guinea DHS conducted between 2016 and 2018. In their analysis, they harnessed ML algorithms, specifically LR, a conditional decision tree, a SVM, and XGBoost. These algorithms were applied to assess child, parental, and household characteristics obtained from the survey. The XGBoost model demonstrated superior performance in predicting stunting, achieving an AUC of 76.5% (95% CI: 71.4–81.9).

An alternative investigation conducted in Zambia harnessed data from the 2018 DHS in the country to perform predictive modeling for all-cause stunting in children under 60 months of age (Chilyabanyama et al., 2022). In this study, five distinct ML algorithms, specifically RF, LR, SVM, NB and XGBOOST were employed to predict the risk of stunting based on a comprehensive evaluation of child, parental, and household characteristics. The RF model outperformed other models, correctly predicting 79% of the cases. Moreover, a separate research endeavor conducted in Tanzania drew upon data from the Tanzania 2015 DHS to undertake predictive modeling for

all-cause stunting. This study employed five distinct ML algorithms, namely RF, Decision tree, KNN, SVM, and LR (Lucy Lawrence, 2021). These algorithms were employed to assess an array of socio-economic and demographic factors as potential predictors. With a superior AUC of 92.0%, the RF model showed a substantially better performance than other models

The synthesis of the aforementioned evidence is presented in Table 3.2. In summary, the predominant focus of predictive modeling pertained to the prediction of all-cause stunting, with the majority of studies relying on country-specific Demographic Health Surveys as data sources. These investigations extensively employed a variety of ML algorithms, with only two exceptions utilizing logistic regression. Stunting was consistently framed as a classification task, with a binary outcome variable, and the AUC was the most frequently reported metric, with estimates ranging from 61.0% to 92.0%. The predictors considered for stunting were typically child, parental, and household factors. The publication period for these predictive studies of all-cause stunting extended from 2019 to 2023, and the geographical focus predominantly encompassed Africa and Asia.

However, it is noteworthy that only two studies approached stunting as an outcome of diarrhea. These studies aimed to predict linear growth faltering following episodes of diarrheal illness in children aged less than 60 months. The first of these studies employed logistic regression as the modeling technique, while the latter utilized logistic regression and Random Forest. Both studies leveraged multi-site data derived from the GEMS and MALED studies. A marginal increase in model performance was observed, with the first study reporting an AUC of 67.0% and the second achieving 75.0%, the second study also conducted external validation of the developed model.

Although current models offer a useful starting point, their predictive power may be compromised by shifts in the study population's characteristics over time (Rahmani et al., 2023; Sahiner et al., 2023). Moreover, utilizing updated, relevant data to build new models offers a chance to improve predictive accuracy and gain valuable new insights into this public health challenge. Additionally, the integration of explanatory model analysis to improve model interpretability is warranted, as no study predicting LGF has explored this aspect. The comprehensive review of the burden, drivers, and the current state of predictive modeling provides a compelling rationale for harnessing clinical, demographic, anthropometric, and socio-economic data in the prediction of chronic malnutrition, following a diarrheal episode through the application of diverse ML algorithms.

Table 3.2: Current state of predictive modelling for chronic malnutrition: A summary of literature

Author; Country	Data used	Outcome assessed	Modelling technique	Model performance
Brander et al., 2019; Multi- Site (GEMS)	Clinical and sociodemographic factors	Linear growth faltering following diarrheal episode	LR	AUC of 67.0% (95% confidence interval [CI] 64.0%, 69.0%).
S. M. Ahmed et al., 2023; Multi-site (GEMS & MALED studies)	Demographic, clinical and socio-economic data	Linear growth faltering following diarrheal episode	LR and RF	AUC of 75.0% (95% CI: 75.0, 75.0).
Hanieh et al., 2019; Vietnam	Mother-related, father-related, child-specific, socio-economical, environment-based, laboratory-assessed, and nutrition-based factors	All cause-stunting	LR	AUC of 85.0% (95% CI, 80.0–90.0).
Rahman et al., 2021; Bangladesh	Socio-economic and demographic factors	All cause-stunting	ML: SVM; RF; and LR.	Accuracy of 88.3%
Haque et al., 2023; Bangladesh	Child, parental, and household data	All cause-stunting	Logistic Regression; Probit Regression; and Decision Tree	AUC of 61.0%.
Shen et al., 2023; Papua New Guinea	Child, parental, and household data	All cause-stunting	ML: SVM; NB; RF; LR; and XGBoost.	AUC of 76.5% (95% CI: 71.4–81.9)

Chilyabanyama et al., 2022; Zambia	Child, parental, and household data	All cause-stunting	ML: RF; LR; SVM; NB; and XGBOOST	Accuracy score of 79%.
Ndagijimana et al., 2023; Rwanda	Child, parental, and household data	All cause-stunting	ML: SVM; NB; RF; LR; and XGBoost	AUC of 89.0%.
Lucy Lawrence, 2021; Tanzania	Socio-economic and demographic factors	All cause-stunting	ML: RF; Decision tree; KNN; SVM; and LR	AUC of 92.0%.

3.3.3 Mortality

The current section presents a comprehensive review of the literature on recent research advancements in predictive models for mortality. A systematic review on predictive modelling for diarrheal disease established that diarrheal outcomes, were under researched with no predictive models for diarrheal mortality reported in literature (Ogwel et al., 2023). This section reviews current research efforts in the prediction of mortality for other diseases to establish the methodology, data sources, and the key findings.

A study predicting all-cause mortality among army veteran patients in the USA used an age-adjusted, LR algorithm in the prediction of mortality. The study investigated the predictive validity of single- and multi-item general self-rated health (GSRH) measures for mortality, using electronic health record data. The GSRH, physical component score (PCS), and Seattle index of comorbidity (SIC) had similar AUC for predicting mortality (AUC 74.0%, 73.0%, and 73.0%, respectively) (DeSalvo et al., 2005). A different study looked at the prediction of 21-day all-cause mortality among Dutch individuals seeking care for COVID-19 (Otenhoff et al., 2021). They applied LR and GBM algorithms to patient history, clinical presentation, and diagnostic imaging and lab data identifying 10 predictors from a data-centric variable selection using ANOVA and the two models yielded an AUC of 81.0% (95% CI 77.0% to 85.0%) and 82.0% (79.0% to 85.0%), respectively.

Another study investigated the prediction of the risk of death among patients with SARs-CoV-2 infection globally. Pourhomayoun and Shakibi used ML algorithms (SVM, ANN, RF, Decision Tree, LR, and KNN) in the prediction of death using demographical and physiological data. They got an overall accuracy of 89.8% (Pourhomayoun & Shakibi, 2021). Bello-Chavolla et al. studied the prediction of mortality among Mexican patients infected with SARS-CoV-2 (Bello-Chavolla et al., 2020). Cox proportional hazards models, incorporating demographic and clinical data, were used to predict the likelihood of death. Their model had a good discrimination (C-statistic = 0.823).

Tsui et al. explored the prediction of mortality in acutely hospitalized older patients in London, UK. They built tree-based prognostic models leveraging a gradient boosting approach to predict death in a 600 day window following a hospitalization, using data from medical records, hematological parameters, cognitive assessments, and head scans (Tsui et al., 2023). Their models yielded AUCs ranging between 67.0%-87.4% for the most comprehensive model. A study in Ireland, investigated the prediction of 4-year mortality among community-dwelling aged patients. They built multivariable regression models using 6 baseline predictor categories: population characteristics; heart-related conditions and other illnesses; health behaviors and lifestyle factors; functional attributes; self-reported measures (including well-being and social connections); and access to and use of healthcare services. Their model yielded an AUC of 78.0% (Matthews et al., 2022).

Chiu et al. studied the prediction of mortality among ICU patients in the USA using 5 ML algorithms (Latent Dirichlet Allocation (LDA) model, classification and regression trees (CART), LR, multivariate adaptive regression splines (MARS), RF, and GBM) fitted to electronic health record data. The GBM model had the highest predictive power achieving 92.8%, 93.2% and 83.3% for AUC, specificity and sensitivity, respectively (Chiu et al., 2022). In the evaluation of predictive

models for hospital-based deaths among individuals hospitalized with sepsis in the USA, Alnsour et al. fitted a tree-based model using chi-square automatic interaction detection on electronic health records data. Their model correctly predicted 85.3% of the outcomes and had a good discriminatory ability (AUC= 84.0%) (Alnsour et al., 2019). Heber et al. studied the prediction of 4-day mortality among hospitalized Covid-19 patients in Austria. They used logistic regression to build a predictive model, using demographic and clinical data (age, fever at admission and 5 hematological parameters), which achieved an AUC of 92.0% (Heber et al., 2022). A different study built and evaluated the predictive model and mortality risk score among patients hospitalized with Ischemic Stroke in Spain (Lea-Pereira et al., 2022). They observed moderate predictability (AUC 74.2%, 95% CI: 73.7%–74.7%) from fitting a logistic regression model to electronic health record data.

Nopour et al. studied the prediction of mortality among COVID-19 Patients in Iran. They fitted ANN and LR models using electronic health record data to predict mortality. Their models achieved AUC of 96.0% and 65.0% for ANN and LR, respectively (Nopour et al., 2022). A study conducted in Kenya (2009–2012) evaluated an adapted mRISC (modified respiratory index of severity in children) risk score as a standardized method in the identification of admitted pediatric patients with elevated heightened susceptibility for respiratory-associated mortality in Kenya. Additionally, a multivariable LR model was utilized in the identification of predictors of hospital-based deaths. They achieved the best performance at mRISC cut-off of ≥ 1 : Sensitivity (80.7%) and specificity (72.9%) (Emukule et al., 2014).

Di et al. conducted a systematic review of the prediction models of Mortality in acute pancreatitis among adult patients in which most scores demonstrated low prognostic accuracy (Di et al., 2016). Gutiérrez-Gutiérrez et al. developed prognostic models to predict 2-week overall mortality among

individuals with bloodstream infections. They fitted a LR model using demographic and clinical data achieving an AUC of 80.0% (95% CI, 73.0%-88.0%) (Gutiérrez-Gutiérrez et al., 2016). Another systematic review that investigated prognostic models for mortality following a cardiac surgery in patients with infective endocarditis (Fernandez-Felix et al., 2021). They identified 11 prognostic models that were aggregated into a meta-model using stacked regression achieving an AUC of 79.0% (95% CI 76.0%–82.0%).

Naemi et al. also conducted a systematic review of ML approaches for predicting death in emergency departments (Naemi et al., 2021). The review identified a diverse set of popular algorithms in this domain, including linear models, tree-based methods, support vector machines, and neural networks. Additionally, they observed that most studies did not report essential steps in model development and validation such as how the data was preprocessed and missing data handled. From their findings, they recommended that scholars in this field should examine the predictive power of the models by externally validating them to understand more comprehensively how well they generalize to new cases. Furthermore, they proposed the need to utilize appropriate approaches in handling noise in the data and missing values rather than excluding such observations since how noise and missing data are handled are fundamental to model development. They also advocated for the use of right sub-sampling techniques (under sampling, over sampling or hybrid techniques (SMOTE or ROSE) to handle class imbalance in the target variable because if it is not addressed it could lead to development of models that are biased and overfitted. Furthermore, they highlighted the need for researchers to develop methods that enhance the transparency and interpretability of their predictions and workflows. This could be achieved by conducting an explanatory model analysis. Lastly, they observed the absence of standardized reporting and publication guidelines for ML in healthcare as a significant obstacle and to this end

they recommended the use of available checklists such as Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) (Collins et al., 2015).

To bridge the gap between complex mathematical modeling and real-world clinical utility, the TRIPOD and PROBAST frameworks, along with their AI-specific extensions, provide a dual-layered guarantee of scientific excellence. TRIPOD statement is a globally recognized set of reporting guidelines, comprising a 22-item checklist, designed to ensure that studies developing or validating clinical prediction models are reported with absolute clarity (Collins et al., 2015). Its extension, TRIPOD+AI, specifically addresses the unique complexities of machine learning and artificial intelligence in healthcare (Collins et al., 2024). It mandates the disclosure of details regarding data preprocessing, algorithmic architecture, hyperparameter tuning, and potential biases, ensuring that the "black box" of AI is opened for scientific scrutiny. Adhering to these guidelines enhances methodological rigor and transparency of studies by ensuring that every phase of model development from participant selection to performance metrics is documented. This transparency allows peer reviewers and the broader scientific community to evaluate the robustness of the findings. They also facilitate reproducibility by standardizing the reporting of variables and model coefficients, the guidelines allow other researchers to replicate the study or validate the models in diverse geographic and clinical settings, which is essential for global health applications. Additionally, the TRIPOD+AI guidelines provides the structured evidence required to prove that a model is not just mathematically accurate, but also clinically sound and free from significant training bias thereby fostering clinical trust and safety. The guidelines also support regulatory alignment as many journals and regulatory bodies now require TRIPOD-compliant reporting as a prerequisite for publication and subsequent approval of AI-driven medical devices.

PROBAST (Prediction model Risk Of Bias Assessment Tool) is the gold-standard framework used to critically appraise the quality of clinical prediction model studies (Wolff et al., 2019). It evaluates research across four key domains: Participant Selection, Predictors, Outcome, and Analysis. Each domain is assessed for Risk of Bias (to determine if the findings are likely to be correct) and Applicability (to determine if the model is relevant to the intended clinical setting). The PROBAST-AI extension adapts this framework for AI, introducing specialized signaling questions to detect AI-specific risks, such as data leakage, overfitting, and the lack of appropriate validation techniques. Complementing TRIPOD+AI guidelines, PROBAST+AI serves as a critical safeguard against methodological flaws, systematically screening for risks of bias such as overfitting or data leakage that could otherwise compromise patient safety (Moons et al., 2025). Studies adhering to these global benchmarks, elevate the research on their predictive models from technical exercises to robust, evidence-based medical tools, providing the transparency and validity required for policymakers to confidently integrate the insights into pediatric care within resource-limited settings.

Another systematic review of the models for predicting death among extremely preterm infants was conducted (Medlock et al., 2011). Their review found 41 studies focused on building prediction models and 18 studies focused on validating developed models. The reporting of key methodological elements, including performance metrics, how the models were internally and externally validated, and how missing data was handled, was suboptimal. The included studies were from heterogeneous populations (21 countries) and the time period was from 1982–2010. The most used modelling technique was LR algorithm (85.4%), but the remaining studies reported classification trees, a log linear regression model and discriminant analysis. Most studies formulated the prediction task as classification task (mortality) or as regression (survival). Majority

of the studies reviewed involved predictive models focused on predicting death (33 [80.5%]) while the rest predicted survival chance (8 [19.5%]). The Hosmer-Lemeshow p-value and AUC were the primary performance metrics reported in the reviewed studies at 46.3% and 43.9%, respectively.

A review of COVID-19 prediction models for severe outcomes and death evaluated 79 studies and the reported AUC's ranged between 52.0%-98.0%. However, risk of bias was unclear or high among these studies, due to limited transparency on the features and target assessment and limited number of samples and outcomes (Miller et al., 2022). Most of the studies did not seem to follow TRIPOD guidelines in reporting their models.

The comprehensive synthesis of the aforementioned evidence is systematically presented in Table 3.3. In essence, the prediction of mortality was framed as a classification task, with the outcome variable being binary (death), or as a regression task, where the variable took on continuous values (survival). Consequently, the choice of modeling technique in each study influenced the performance metrics reported. While ML algorithms were the predominant modeling approach, some studies also incorporated regression methods, specifically logistic regression and Cox regression. None of the reviewed studies adopted an explainable ML approach to enhance interpretability.

Table 3.3: Current state of predictive modelling for mortality- A summary of literature

Author; Country	Data used	Outcome assessed	Modelling technique	Model performance
Pourhomayoun & Shakibi, 2021; Global	Demographic and physiological data	Mortality among COVID-19 patients	Machine learning (ML): Support Vector Machine (SVM), Artificial Neural Networks (ANN), Random Forest (RF), Decision Tree, Logistic Regression (LR), and K-Nearest Neighbor (KNN)	Accuracy of 89.8%

Gutiérrez-Gutiérrez et al., 2016; Multi-site	Electronic health record (EHR)	Mortality among patients with bloodstream infections	LR	AUC of 80.0% (95% CI, 73.0%–88.0%)
Fernandez-Felix et al., 2021; Systematic review	EHR	Mortality among patients with infective endocarditis	Stacked regressions	AUC: 79.0% (95% CI 76.0%–82.0%).
DeSalvo et al., 2005; USA	EHR	All-cause mortality among veterans	LR	AUC: The GSRH (74.0%), PCS (73.0%), and SIC (73.0%)
Chiu et al., 2022; USA	EHR	Mortality in ICU patients	ML algorithms: Latent Dirichlet Allocation model, classification and regression trees (CART), LR, multivariate adaptive regression splines, RF, and grGBM.	AUC (92.8%), specificity (93.16%), and sensitivity (83.25%) (Chiu et al., 2022).
Alnsour et al., 2019; USA	EHR	Mortality among sepsis patients	Chi-square automatic interaction detection model	Accuracy of 85.3% and an AUC of 84.0%.
Ottenhoff et al., 2021; The Netherlands	Premorbid, clinical presentation and laboratory and radiology data	Mortality among COVID-19 patients	LR and GBM algorithms	AUC of 81.0% (95% CI 77.0% to 85.0%) and 82.0% (79.0% to 85.0%)
Tsui et al., 2023; UK	Clinical data	Mortality among hospitalized older patients	Gradient boosted trees-based models	AUC ranging between 67.0%–87.4%
Matthews et al., 2022; Ireland	Demographic and clinical data	Mortality among older patients	Multivariable regression	AUC=78.0%
Lea-Pereira et al., 2022; Spain	Electronic health record	Mortality among patients with ischemic stroke	LR	AUC 74.2%, (95% CI: 73.7%–74.7%)
Bello-Chavolla et al., 2020; Mexico	Demographic and clinical data	Mortality among COVID-19 patients	Cox proportional risk regression	C-statistic = 0.823
Emukule et al., 2014; Kenya	Clinical data	Mortality among children hospitalized with respiratory illness	LR	Sensitivity (80.7%) and specificity (72.9%)

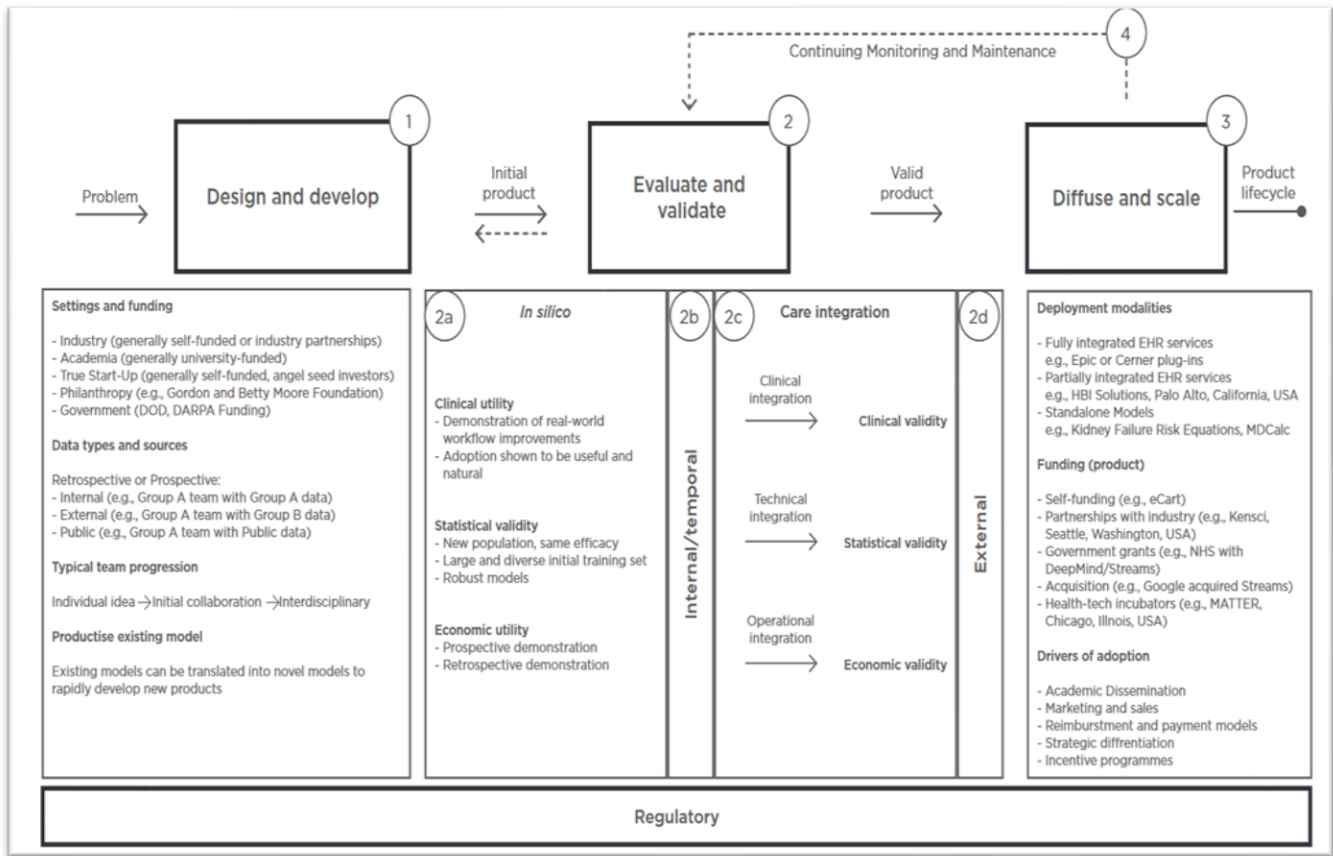
The majority of these studies made use of patient data extracted from electronic health records, encompassing demographic and clinical information. Furthermore, most of the data sets used in these studies were collected in the US and Europe. The publication time-period spanned from 2005 to 2022, with a surge in publications over the past decade. The AUC was the most commonly reported performance indicator, with estimates ranging from 67.0% to 92.8%. These observations

align with findings from Medlock et al., where ML emerged as the most prevalent technique in the past decade.

The comprehensive review of the burden, determinants, and the current state of predictive modeling strongly justifies the utilization of clinical, demographic, socio-economic and anthropometric data to predict hospital-based deaths among children hospitalized for acute gastroenteritis (AGE), leveraging a variety of ML algorithms. Additionally, there is a compelling need for the incorporation of explanatory model analysis to better understand the underlying decision-making dynamics, as none of the reviewed studies have ventured into this critical aspect. Additionally, the gaps in previous studies support the use of TRIPOD+AI and PROBAST+AI guidelines to ensure rigorous reporting and critical appraisal of research quality, including risk of bias and applicability. Applying these frameworks across all three outcomes, not only the mortality models, strengthens methodological transparency and supports the operationalization of clinical prediction models for effective integration into real-world clinical decision-making.

3.4 From Development to Deployment

The pathways in the building, evaluation, diffusion, and scaling of ML models in clinical care are a multifaceted and meticulously planned process (Figure 3.3). It begins with the conceptualization and development of ML models, often rooted in robust research and data analysis, ensuring that the models are not only accurate but also ethically sound.



Source: Sendak et al. (2020). Reproduced under the Creative Commons Attribution-NonCommercial 4.0 (CC BY-NC 4.0) license.

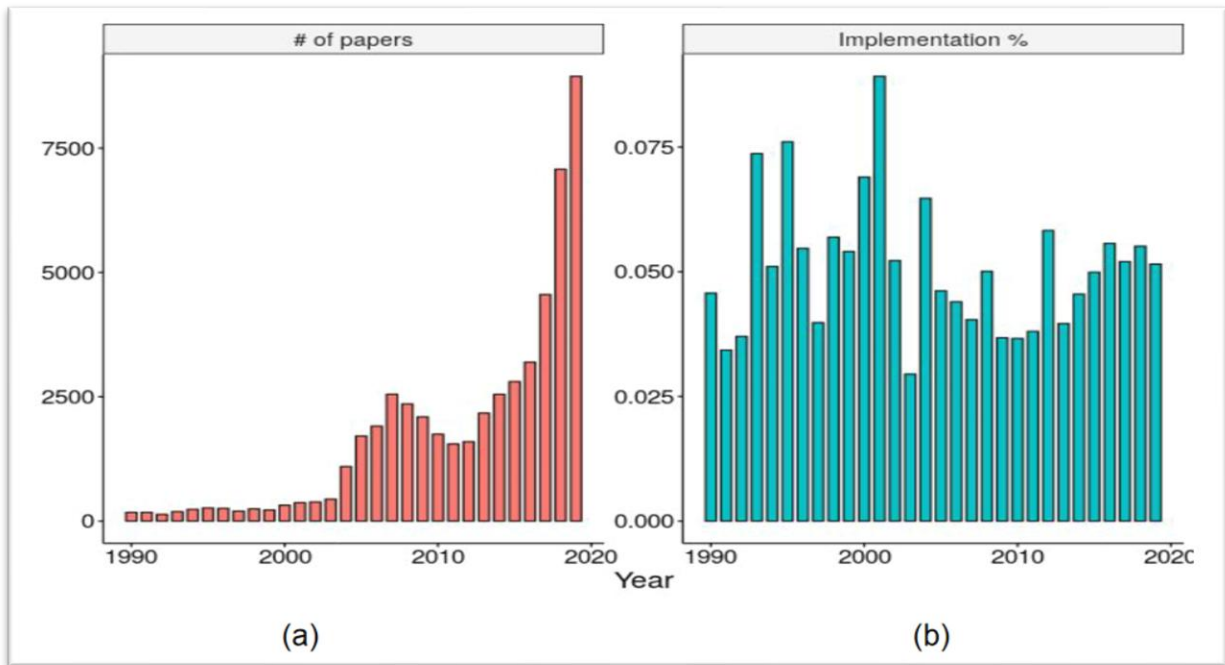
Figure 3.3: Pathway in the building, evaluation, diffusion and scaling of machine learning models in clinical care (Sendak et al., 2020)

Following development, extensive validation is crucial in demonstrating the efficacy and safety of the models in actual clinical scenarios. This validation process involves rigorous testing, refinement, and adaptation to the unique needs of different healthcare settings. Once validated, the models are diffused into clinical practice, integrating seamlessly into healthcare workflows. However, the scaling of ML models across a broader healthcare landscape is a challenge that requires addressing issues of data interoperability, privacy, and standardization. Successful scaling hinges on a collaborative effort between healthcare providers, technology developers, regulators,

and policy makers to harness fully the capabilities of ML in clinical care while safeguarding patient well-being and data security (Sendak et al., 2020).

During the recent past, the healthcare sector has undergone a significant transformation, thanks to the incorporation of ML techniques. ML has the capability to revolutionize healthcare by leveraging evidence-based approaches and data to optimize patient management, improve diagnostic accuracy, and improve treatment outcomes through better-informed medical decision-making, and streamlining healthcare operations. Moreover, these algorithms have the potential to bridge existing gaps in healthcare delivery, unlock new opportunities for medical research, and address key challenges facing the sector. While the development of powerful ML models has been a cornerstone of this evolution, the current focus has shifted from model development to model deployment, marking a pivotal phase in the use of this cutting-edge technology (Jackson et al., 2019; Mbunge & Batani, 2023; Zhang et al., 2022). The transition from development to deployment in healthcare ML is emblematic of the industry's commitment to practical, real-world applications. Here, the reasons behind this shift are explored, the challenges and opportunities it presents, and the broader implications for healthcare. For years, much of the research efforts in healthcare ML has been on the development of sophisticated algorithms and models (Habeheh & Gohel, 2021; Javaid et al., 2022) with only ~ 7% of ML published papers addressing deployment (Figure 3.4) (Drysdale et al., 2020).

Researchers and data scientists have invested their time and expertise in creating a suite of ML systems, these models, trained on extensive datasets, have shown remarkable potential. There is now a growing need to investigate deployment issues of these systems. Deploying ML models in clinical settings allows healthcare professionals to leverage these tools for real-time decision-making, diagnosis, and patient care.



Source: Drysdale et al. (2020). Whitepaper posted on ResearchGate. © 2020 Drysdale et al. All rights reserved.

Figure 3.4: Machine learning papers published in PubMed and percentage having deployment/implementation (Drysdale et al., 2020)

Model deployment looks into issues of seamless integration into existing healthcare workflows, issues of data quality, interoperability, regulatory compliance, ethical considerations and the need for clinician training (Cohen et al., 2021; Mbunge & Batani, 2023). The methodologies and strategies employed to bridge the gap between innovation and implementation are crucial, as they dictate the success of these cutting-edge technologies within the healthcare ecosystem.

Moreover, the selection between proprietary and open-source frameworks for deploying ML models is a critical decision with multifaceted considerations (Berberi et al., 2025; Harish et al., 2022; Vake et al., 2025). Proprietary frameworks, characterized by vendor-specific ownership, often offer dedicated support, seamless integration within proprietary ecosystems, and user-friendly interfaces. These frameworks may provide advanced features and optimizations, yet they

are accompanied by licensing costs, potential vendor lock-in, and limited transparency. In contrast, open-source frameworks foster collaboration through a community-driven development model, granting users flexibility, customization, and cost-effectiveness. However, open-source systems may require greater technical expertise from users and adopters, lack comprehensive support, and challenges related to security and privacy auditing. The dichotomy extends to considerations of vendor support, integration within ecosystems, ease of use, advanced features, licensing costs, community collaboration, flexibility, cost, vendor neutrality, and transparency. Proprietary frameworks excel in streamlined support, seamless integration, user-friendly interfaces, and access to advanced features but are encumbered by licensing costs and potential dependency on specific vendors. Open-source frameworks, characterized by community collaboration, flexibility, and cost-effectiveness, empower users with customization capabilities, yet they may lack comprehensive support, require technical proficiency, and necessitate independent security evaluations.

The decision-making process involves meticulous evaluation of these factors based on project-specific requirements, organizational goals, budget constraints, and the imperative of balancing ease of use with customization options. Hybrid approaches, combining proprietary and open-source components, are also possible. The overarching consideration is aligning the framework choice with the project's intrinsic needs, overarching budgetary considerations, and the strategic vision of the organization.

Several frameworks support deploying ML models developed in R, including Plumber, Shiny, RServe, and OpenCPU. In summary, Plumber is a streamlined R framework for deploying machine learning (ML) models as RESTful APIs, enabling seamless integration of R-based analytics into broader software ecosystems (Karamitsos et al., 2020; Schloerke et al., 2022). By

exposing R functions as API endpoints, Plumber allows developers and data scientists to create interactive and interoperable ML services accessible via standard HTTP methods. Its automated documentation feature enhances transparency and collaboration, especially important in healthcare settings where clear communication of model behavior is vital. Tools like RStudio Connect further bolster its utility by providing version control, authentication, and deployment support. In practice, Plumber has been effectively used to develop web-accessible predictive models for clinical applications, including cardiovascular risk and hospital readmission, highlighting its adaptability and usefulness for operationalizing ML insights in healthcare (Murphree et al., 2018; Schreier et al., 2022).

Shiny empowers users to convert R scripts into interactive web applications, providing an intuitive platform for deploying and interacting with ML models (Eddington et al., 2022; Piccolo et al., 2020). Designed for ease of use, Shiny removes the need for web development expertise, enabling dynamic visualization and manipulation of model outputs through user-friendly interfaces. Its ability to handle real-time input adjustments makes it especially suitable for clinical environments where decision-making depends on flexible data exploration. Shiny's interoperability with other R packages facilitates complex data workflows, and its deployability via web servers enhances accessibility for non-technical stakeholders. Widely adopted in healthcare, Shiny has supported tools for pneumonia risk prediction (Mohammed et al., 2022), COVID-19 mortality estimation (Das et al., 2013), and surgical risk calculators (Corey et al., 2018), demonstrating its value in making ML outputs understandable and actionable across clinical settings.

RServe is a server-based framework that connects R computations with external systems using a client-server architecture, enabling scalable deployment of ML models across diverse applications (Shah, 2016; Urbanek, 2022). Functioning as a background R process, RServe supports multiple

protocols—including HTTP—and bridges R with other languages, making it ideal for integration in multi-language environments or cloud-based solutions. Its capacity to manage complex statistical workflows remotely allows for efficient deployment in settings with high computational demands or limited local resources. Security features such as authentication and encryption make it well-suited for sensitive healthcare applications, where patient data protection is essential. RServe's versatility and robustness make it a preferred choice for deploying sophisticated R models that require seamless communication with other platforms.

OpenCPU offers a stateless, HTTP-based API layer for R, making it a flexible and standardized solution for deploying ML models as web services (Ooms, 2014). Its RESTful architecture allows remote execution of R functions, facilitating interoperability and scalability across applications. By supporting Docker containers and complex package dependencies, OpenCPU ensures reproducibility and simplifies model deployment in varied environments. The platform also enables interactive visualization and real-time exploration of ML predictions via browser-based interfaces, which is especially beneficial in healthcare for enabling clinicians to interpret outputs dynamically. OpenCPU has been used in applications ranging from personalized medicine tools to drug discovery platforms, illustrating its capability to integrate R-based ML models into both research and operational healthcare systems effectively (Arslan et al., 2020; Chandra et al., 2017; Södergård et al., 2021).

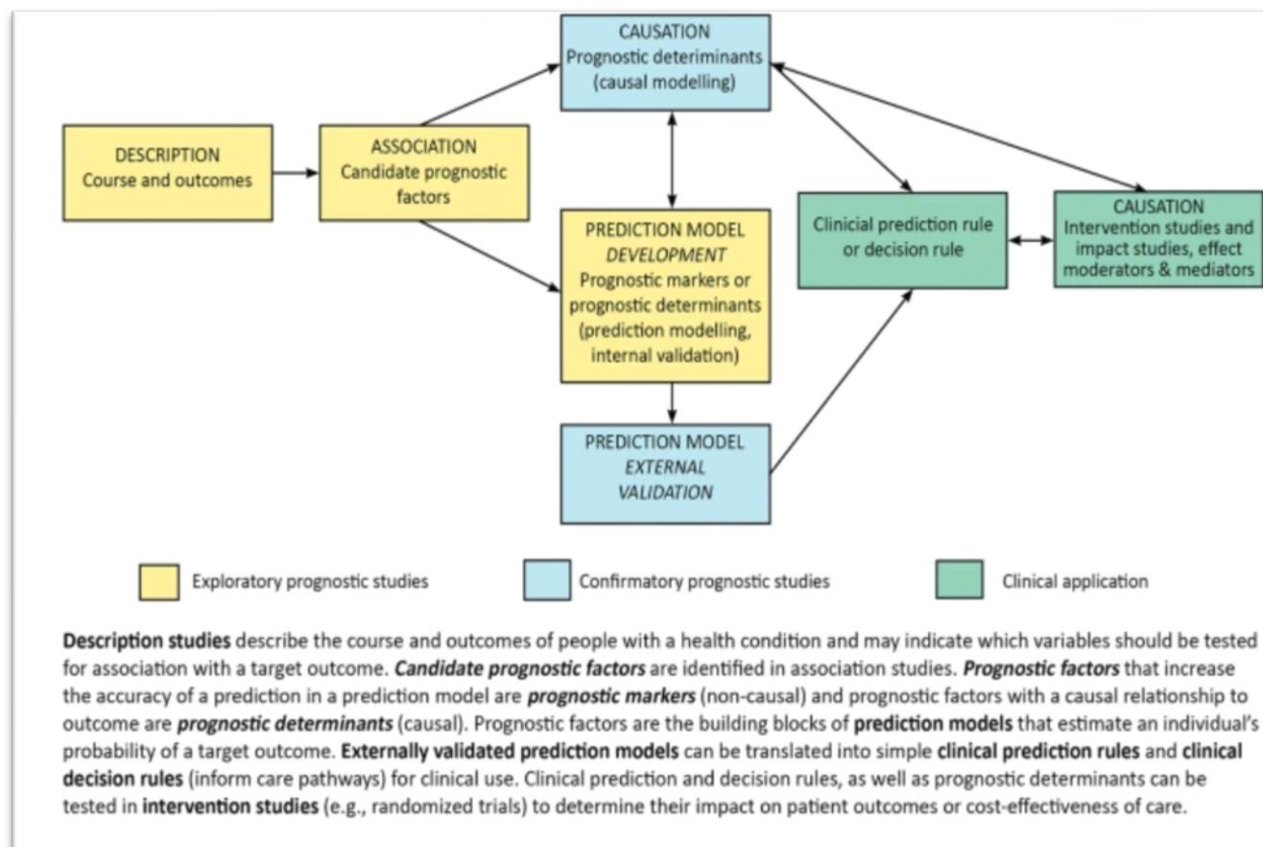
The Shiny framework was the preferred platform for deploying predictive models related to poor diarrheal outcomes, due to its user-friendliness, seamless R integration, robust deployment capabilities and broad adoption in healthcare.

3.5 Theoretical Frameworks

In essence, the theoretical framework provides the organizational structure within which the theoretical foundations are articulated and applied, offering a comprehensive framework for understanding the research problem (Varpio et al., 2020). In the pursuit of unraveling the complexities surrounding predictive models for poor diarrheal outcomes, the theoretical frameworks stand as cornerstones guiding the intellectual trajectory of this study. Three foundational approaches are considered in this research: the Prognosis Research Strategy (PROGRESS) framework, the socio-technical approach, and the design science approach.

3.5.1 PROGRESS Framework

Prognostic research has many salient objectives which include: a.) The characterization of disease progression and clinical outcomes; b.) The investigation of factors associated with health outcomes; c.) The estimation of a person's likelihood of developing an outcome; d.) The investigation of the use of predictive models in clinical practice (Kent et al., 2020). The development of the PROGRESS framework was necessitated by the absence of a coherent structure for prognosis research, evident in inconsistent terminology, ambiguous concepts, and deficient standards in design, analysis, and reporting. These issues had rendered the prognosis field susceptible to challenges like small-scale studies, publication bias, and the failure to replicate previous findings. Consequently, the PROGRESS framework (Figure 3.5) was conceived to enhance the design, execution, and reporting of prognostic research by elucidating distinctions between association, prediction, and causation.



Source: Kent et al. (2020). Reproduced under the Creative Commons Attribution 4.0 (CC BY 4.0) license.

Figure 3.5: Prognostic research conceptual framework (Kent et al., 2020)

It addresses how these factors influence the parameters and interpretation of specific research inquiries. Examining overall prognosis enables a comprehensive understanding of disease burden, facilitating comparisons across diverse settings and temporal assessments to gauge improvements (Hemingway et al., 2013).

This study adopted a prognostic approach to explore factors linked to poor diarrheal outcomes. It did not only investigate the features associated with each poor diarrheal outcome but also employed predictive models to assess the probability of a child developing such an outcome.

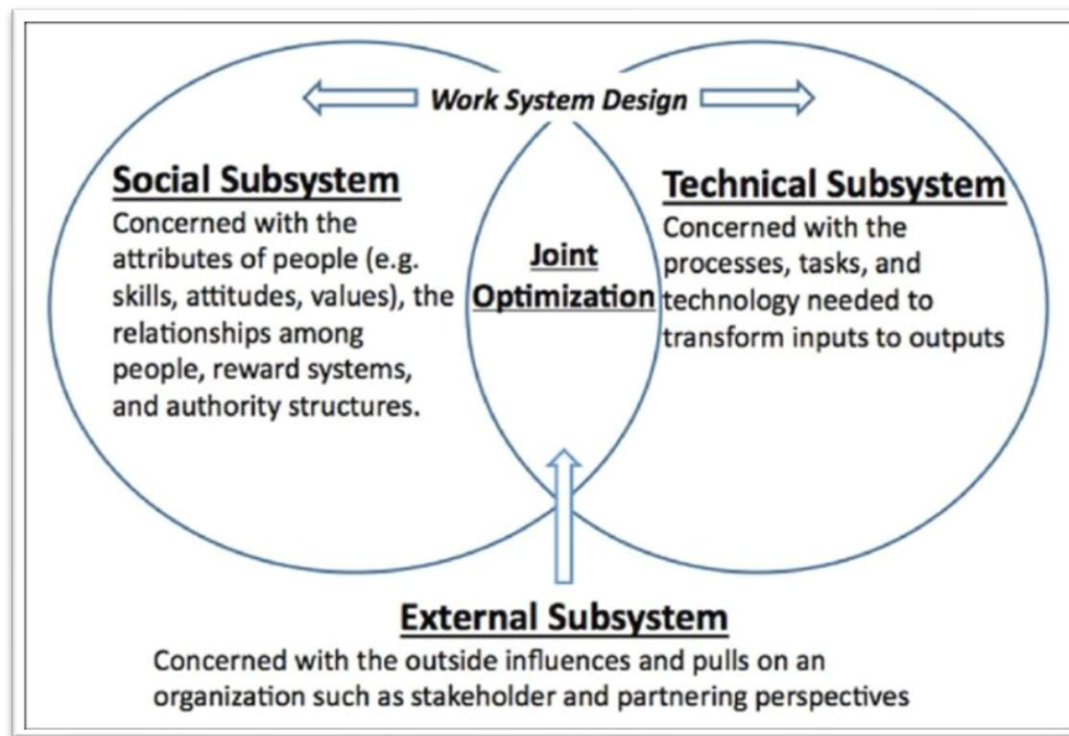
3.5.2 Socio-technical Approach

The socio-technical approach, originating in the 1950s, evolved as a response to the realization that an exclusive concentration on technical aspects in the design of organizational systems produced suboptimal results (Mumford, 2000; Patnayakuni & Ruppel, 2010; Scacchi, 2003). Pioneered by Eric Trist and researchers affiliated with the Tavistock Institute, this paradigm represents a departure from conventional perspectives, acknowledging the inseparable interconnection between technological and social systems. Gaining prominence as a holistic framework, the socio-technical approach aims to optimize the performance and adaptability of organizations and systems, by recognizing the intricate relationship between technology and social factors. The foundational premise asserts that organizations achieving joint optimization of these subsystems have a higher likelihood of generating positive outcomes (Militello et al., 2013). Consequently, the integrated design of organizational work systems necessitates complementary adjustments to both the social and technical subsystems (Figure 3.6).

The socio-technical approach encompasses several fundamental principles and concepts (Bauer & Herder, 2009; Patnayakuni & Ruppel, 2010). The first principle is the interdependence of social and technical systems, positing that the social and technical subsystems are intricately linked, and alterations in one system can have consequential impacts on the other. Additionally, the approach advocates for mutual adjustment, diverging from independent optimization of each system.

Healthcare is inherently centered on individuals, encompassing interactions that extend beyond patients and healthcare providers. Moreover, there is a discernible shift in the healthcare domain towards a patient-centered approach to care, accompanied by a growing demand for enhanced collaboration (IBM Institute for Business Value, 2018). The socio-technical approach to healthcare is particularly pertinent as it aligns with three critical elements of healthcare quality: the safety of

patients, the efficiency of healthcare delivery, and the emphasis on client-driven care. This approach ensures the development of a software solution that is not only technologically robust but also socially acceptable within the healthcare framework.



Reproduced from Militello et al. (2013), with permission from Sage Publications. © Sage Publications.

Figure 3.6: Sociotechnical systems theory (Militello et al., 2013).

The socio-technical approach stands out as an exceptionally fitting theoretical framework for the study of predictive models for poor diarrheal outcomes. This approach excels in providing a comprehensive understanding of the intricate interplay between technological systems and the social context in which they operate. In the realm of healthcare, particularly in predicting and preventing poor diarrheal outcomes, human factors, socio-demographic elements, and environmental conditions are pivotal contributors. By adopting a socio-technical perspective, the

study can comprehensively account for these human elements in the development and implementation of predictive models.

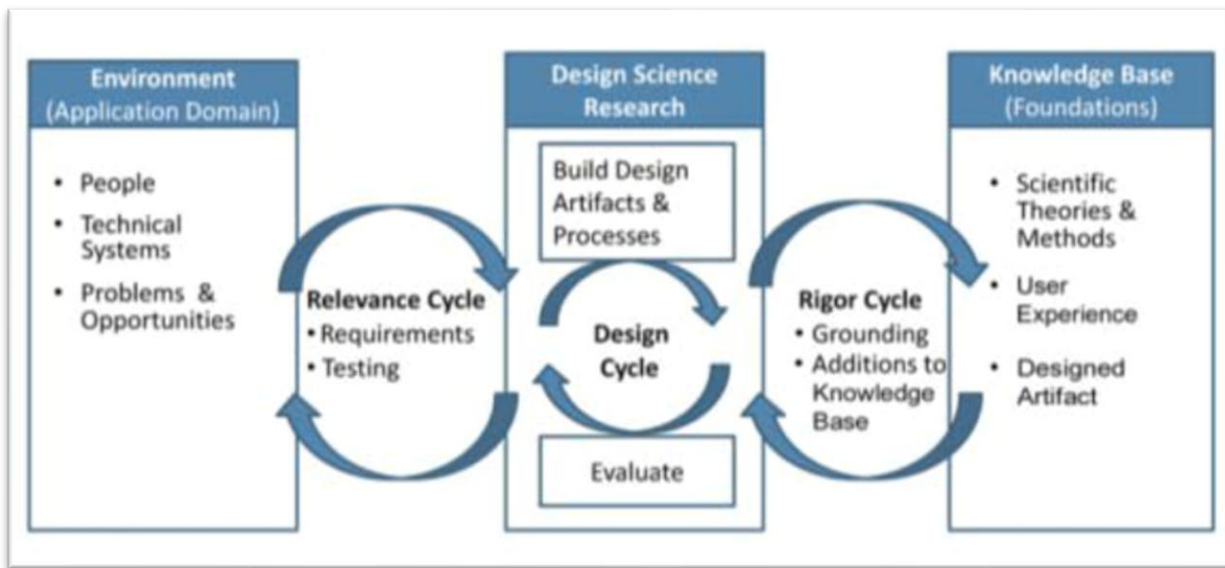
3.5.3 Design Science Approach

Design science approach, a framework designed to solve problems, was developed with the goal of advancing the creation of innovative artifacts (constructs, models, methods and instantiations) that tackle issues and enhance the settings in which they are deployed (Adomavicius et al., 2008; Brocke et al., 2020; Hevner & Wickramasinghe, 2018; Holmström et al., 2009; Raptis et al., 2012). This approach seeks to produce knowledge about how human agents can design and construct things to achieve specific goals (design knowledge) often resulting in economic and societal impact.

The design science approach in healthcare emerged as a response to the growing complexity of healthcare systems and the need for creative approaches to tackle evolving problems. Rooted in the broader field of design science, which emphasizes the development of elements to address real-world challenges, the design science approach in healthcare integrates principles of design, technology, and interdisciplinary collaboration. The approach gained prominence as healthcare increasingly recognized the importance of not just understanding problems but actively designing and implementing solutions to improve patient outcomes and healthcare delivery. Specifically in healthcare, the goal is to improve the viability and efficiencies of products, services and delivery yielding maximum value for care (Hevner & Wickramasinghe, 2018). Figure 3.7 is an illustration and representation of the design science paradigm.

This emphasis on tangible outcomes aligns seamlessly with the pragmatic requirements of healthcare research. Given that poor diarrheal outcomes result from a complex interplay of socio-

demographic, environmental, and healthcare-related variables, the iterative nature of the approach allows continual refinement of predictive models based on feedback and real-world application. This iterative process ensures the adaptability and efficacy of the models across diverse healthcare contexts.



Source: Engel et al. (2013). Article posted on ResearchGate. © 2013 Engel et al. All rights reserved.

Figure 3.7: Design Science Approach (Engel et al., 2013)

3.6 Study Conceptual Framework

The aim of this study was the prediction of poor diarrheal outcomes, utilizing a workflow implemented at each healthcare facility and harnessing the capabilities of ML technology. Figure 3.8 is the illustration and representation of the conceptual framework of the study. Each facility's healthcare system constituted the physical infrastructure under examination. Within the scope of this research, the dependent variable focused on poor diarrheal outcomes, assessed in terms of LDD, chronic malnutrition, and mortality. Acting as an intermediate variable, care-seeking played a pivotal role. The independent variables encompassed demographic factors (child age and

gender), clinical indicators (such as blood in stool, poor drinking habits, inability to drink, fever, restlessness, lethargy, unconsciousness, rectal prolapse, difficulty breathing, convulsions, sunken eyes, home zinc use, capillary refill, chest indrawing, bipedal edema, abnormal hair, dehydration, rotavirus vaccination, ORS use, zinc use, Intravenous (IV) rehydration, antibiotic use, malaria, dysentery, stunting, wasting, belly pain, cough, vesikari score, vomiting, respiratory rate, oxygen saturation, fast breathing, dry mouth, mental status, skin turgor, undernutrition, and stool frequency). Additionally, socio-economic factors included caretaker education, religion, occupation, number of people in the household, number of children under 5 years in households, total assets, animal ownership, improved water and sanitation, shared facilities, and floor type.

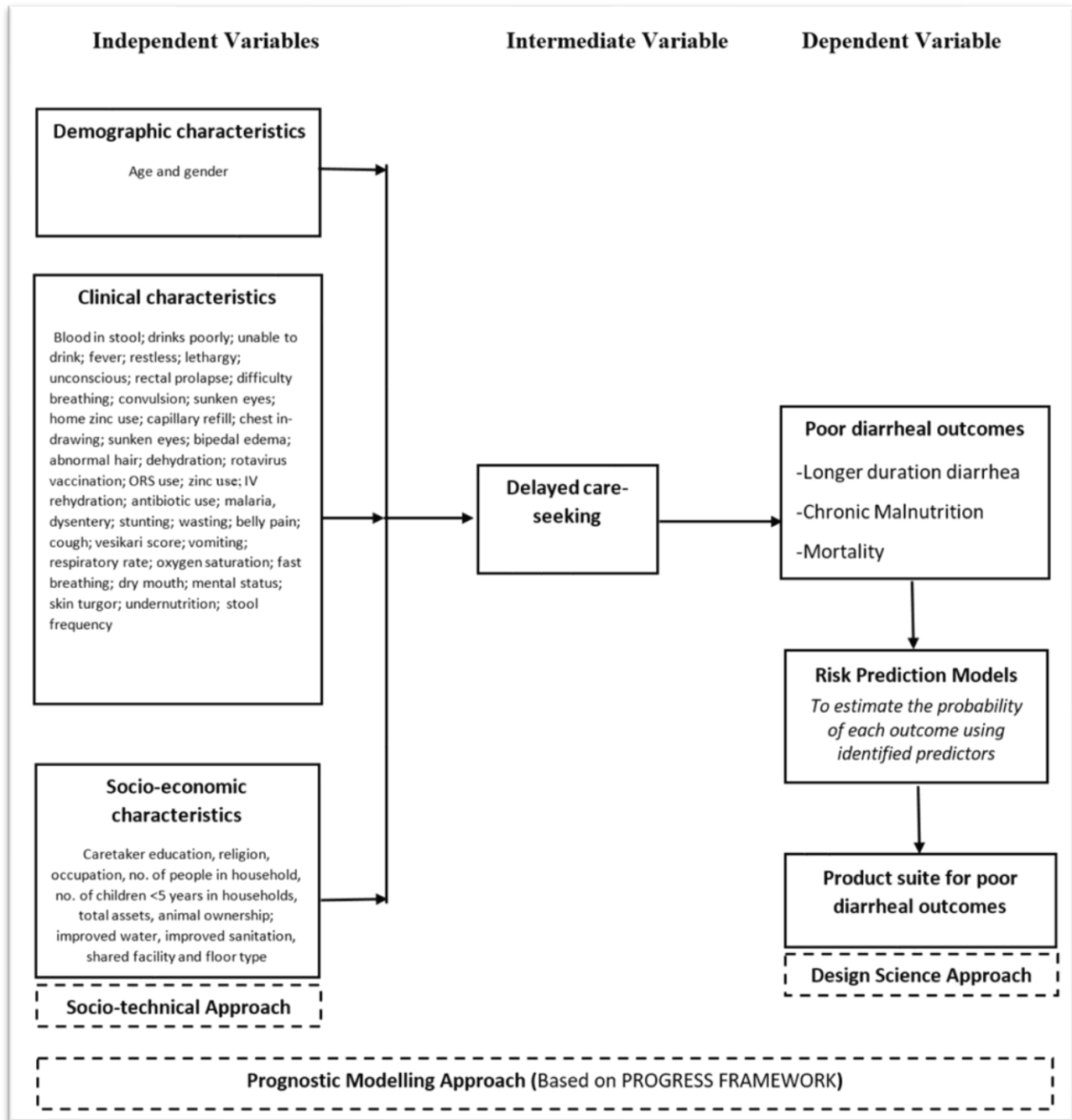


Figure 3.8: Study Conceptual Framework

4. CHAPTER FOUR: RESEARCH METHODOLOGY

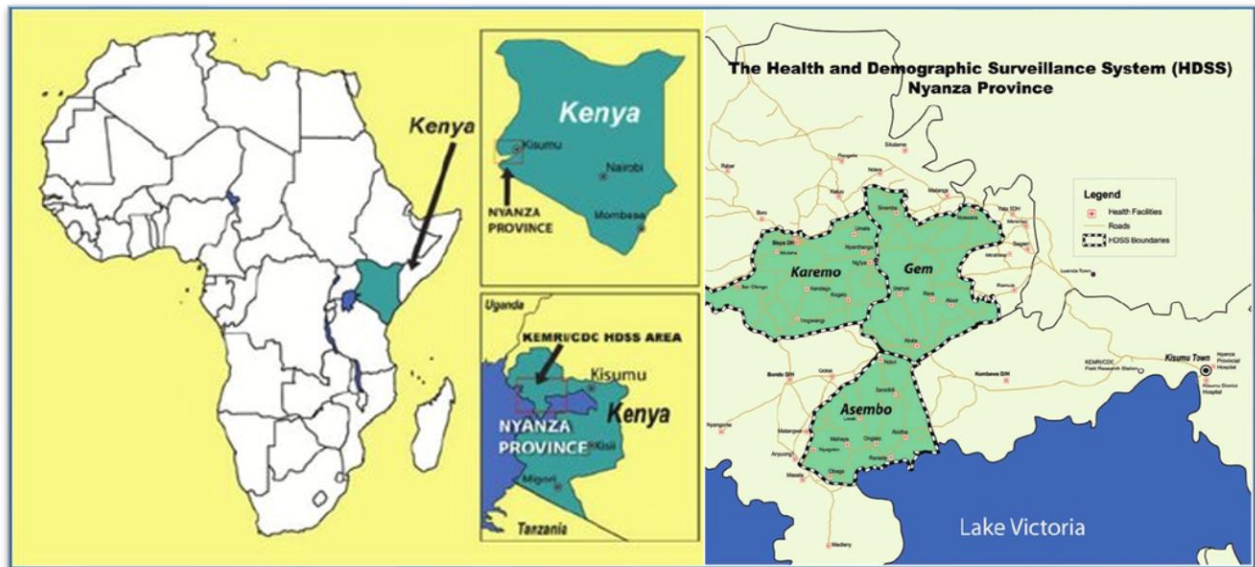
4.1 Introduction

This chapter discusses the research methodology utilized in the study in addressing the research objectives and research questions. It outlines the study sites and locations of data collection, the research process and offers justification for the selected methods. Furthermore, it describes the data used in the study, how it was collected and how the modelling was done.

4.2 Study setting

The study setting is shown in Figure 4.1. The data utilized in this study were from studies conducted within a Kenya Medical Research Institute (KEMRI) operated Health and Demographic Surveillance System (HDSS) area in rural Western Kenya, Siaya County. The KEMRI-HDSS was set up to provide longitudinal population data on demographics, leading causes of illness, and death as well as the burden of diseases, and healthcare utilization and access (Odhiambo et al., 2012). Furthermore, the HDSS provides a strong sampling framework for conducting epidemiological research and evaluating public health programs. The HDSS study is situated in Asembo, Gem and Karemo areas of Siaya County a distance of ~40 km from the central office at Kisian Campus in Kisumu County (Odhiambo et al., 2012; Omore et al., 2019). These are adjacent administrative areas located between -0.22°S and -0.23°N latitude, and 34.53°W and 34.28°E longitude. The rainfall follows a seasonal pattern, characterized by two peak periods, one from April to May and the other from November to December. The annual precipitation measures 2,155mm, with the lowest occurring in July at 114mm and the highest in April, averaging 290 mm. The average temperature is 21.4°C at a mean altitude of ~ 1070 meters above sea level. February

records the highest temperature, averaging 22.7 °C, while July is the coldest month, with a mean temperature of 20.7 °C (Climate-Data, 2023).



Reproduced from Odhiambo et al. (2012), with permission from Oxford University Press. © Oxford University

Figure 4.1: Map of Study Area (Odhiambo et al., 2012)

The catchment area of the study is made up of a culturally homogenous, majority Luo ethnic group, a predominantly rural population. The HDSS follows a population of ~225,825 individuals, reflecting a demographic profile characteristic of typical rural African regions, where 40.2% of the population is aged below 15 years and approximately 6.0% is older than 64 years (KEMRI-CGHR, 2020). Residents within the HDSS are defined as individuals who have resided in the study area for a continuous period of $\geq$ four months or infants born to such residents. The HDSS surveillance began with the enumeration of a baseline census. Subsequently, field workers conduct two-three visits to the enumerated households each year. During these visits, they perform population enumeration for updating the number of people in the study area, which are then utilized for estimating and extrapolating incidence. These routine household visits are conducted to update

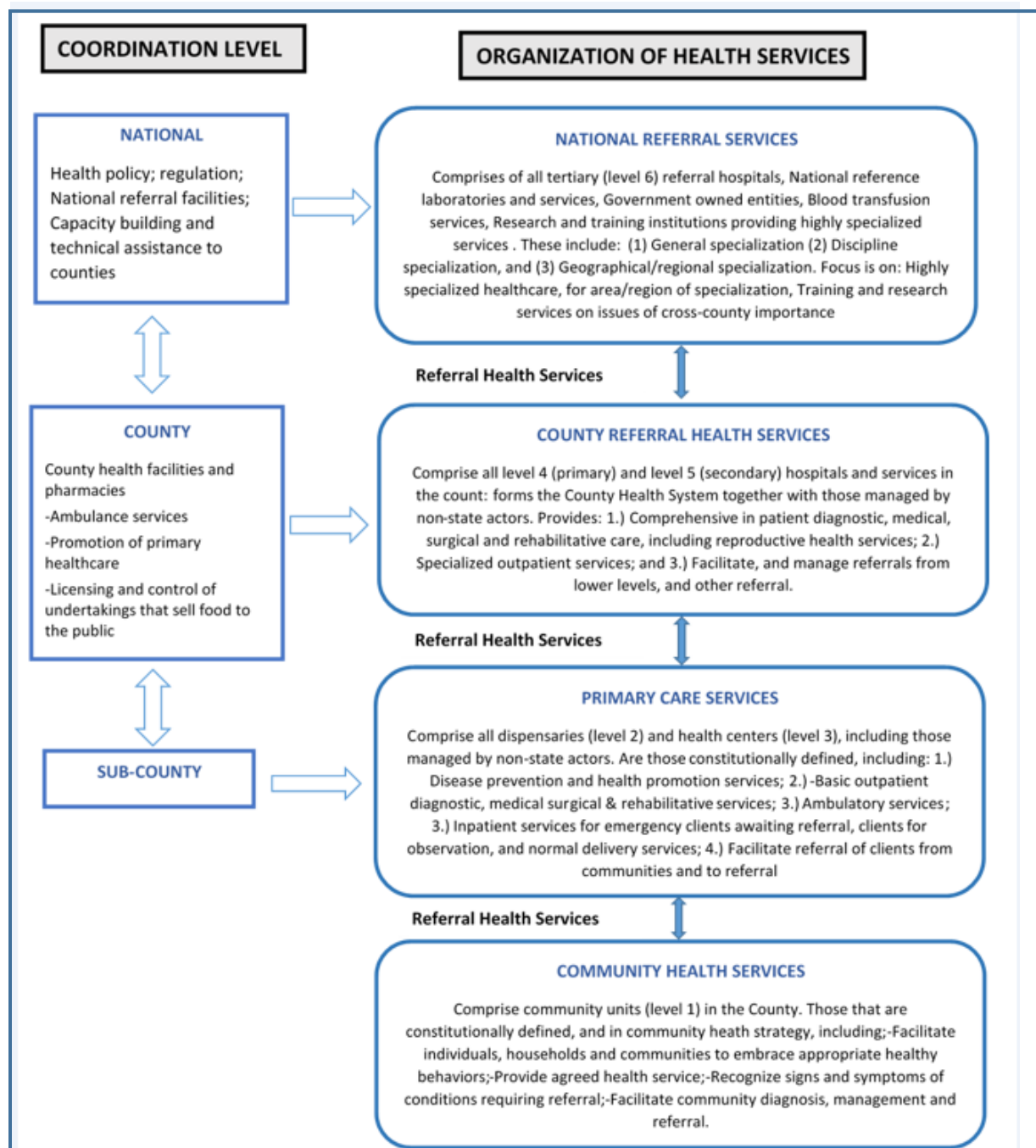
resident demographic information and record key events, encompassing, but not limited to, births, deaths, education levels, household possessions, status of marriage, childhood vaccination, mortality causes determined by verbal autopsy, migrations (in/out), and other health incidents, including diarrheal illness in both adults and children under 60 months. This population resides in 385 villages covering an area of 700 km², distributed across 66,660 households and 45,719 compounds. The majority of the houses are constructed using brick, cement, or mud, with grass-thatched or iron sheet roofs. Most of the households use unsafe water sources (~52%), although the drinking water is treated in ~65% of the households. Moreover, 77.4% of households use traditional pits as the main waste disposal means. The level of education among residents was 58.5% and 30.3%, for primary and secondary education, respectively. The main economic activity is subsistence farming, 60% of households.

The area is a malaria lake-endemic region with more than 70% of the population at risk (Were et al., 2019). In 2020, the prevalence of malaria by microscopy among children <15 years of age was 19% in the lake-endemic region of Western Kenya (Ministry of Health et al., 2021). The area also reported a high HIV prevalence of 15.3% (National AIDS and STI Control Programme (NASCOP), 2020) compared to the national prevalence of 4.6%. The nurse and doctor to population ratios are 1:1,697 and 1:38,511, respectively (County Government of Siaya, 2017). The under-five mortality ratio is 61/1,000 live births) (Institute for Health Metrics and Evaluation, 2020; H. Wang et al., 2020). The life expectancy is 69.4 years. From the 2022 Kenya DHS, the estimates of stunting, wasting and underweight among children < 5 years in Siaya County were, 19.2%, 1.7% and 7.0%, respectively (Kenya National Bureau of Statistics & Ministry of Health, 2023). Additionally from the same survey, vaccination coverage was reported as 88.1 % for fully-vaccinated to basic antigens (BCG, DPT-HepB-Hib [n=3], polio [n=3] besides birth dose, and MR

[n=1]) and the full vaccination rate was 69.9% based on the national schedule (BCG, DPT-HepB-Hib[n=3], OPV [n=4], IPV [n=1], pneumococcal vaccine [n=3], rotavirus vaccine [n=3], and MR [n=1]).

4.2.1 An Overview of Healthcare Delivery and Referrals in Kenya

Healthcare in Kenya operates under a decentralized governance system primarily overseen at the county level yet closely intertwined with the national administration. The provision of essential health services falls under the purview of the County governments in each of the 47 devolved units. Meanwhile, the central government assumes responsibility for coordinating and executing the national health policy, regulatory functions, providing technical assistance, enhancing capacity, and managing national healthcare facilities and programs. This structure ensures a collaborative approach to healthcare management, with both county and national entities playing integral roles in the overall healthcare system. The healthcare system comprises six tiers, encompassing various healthcare facilities: communal health services (tier 1), dispensaries, and clinics (tier 2), health centers, maternity, and nursing homes (tier 3), sub-county hospitals, and medium-sized private hospitals (tier 4), county referral hospitals, and large private hospitals (tier 5), as well as national referral hospitals and large private teaching hospitals (tier 6) (Muga et al., 2005; Toroitich et al., 2022). Primary Health Care (PHC) services are predominantly offered in tier 1, 2 and 3, underscoring their pivotal role in delivering essential healthcare services to communities. Primary healthcare in Kenya is provided through a multi-faceted system: the Ministry of Health (MoH), parastatal entities, and the private sector, which includes a spectrum of providers, from for-profit businesses to non-profit NGOs and faith-based groups [16]. The hierarchical arrangement of healthcare delivery and the referral system in Kenya is illustrated in Figure 4.2.



Source: Kenya Health Policy (Ministry of Health, 2014).

Figure 4.2: The structure of Kenya's healthcare delivery and referral system (Ministry of Health, 2014)

4.2.2 Vaccine Impact on Diarrhea in Africa (VIDA) Study

The Vaccine Impact on Diarrhea in Africa (VIDA) study was a prospective, age-segmented, matched case/control study of moderate-to-severe diarrhea (MSD) amongst children aged less than 60 months residing in censured populations (Powell et al., 2023). The VIDA study was established to evaluate the causes of diarrhea, the effectiveness of the rotavirus vaccine, and the population-level impact of introducing the rotavirus vaccine among children aged less than 60 months living in surveyed populations across three African countries. Moderate-to-severe diarrhea cases were determined as children under 60 months old who visited a sentinel health center with diarrhea (at least three loose stools in 24 hours) that started within the previous week after at least seven diarrhea-free days and displayed at least one of these symptoms: sunken eyes, reduced skin elasticity, bloody diarrhea, need for IV fluids, or being hospitalized. VIDA used a consistent approach by enrolling approximately 8-9 eligible children in each of three age groups (0-11, 12-23, and 24-59 months) every two weeks over 36 months, ensuring a year-round sample with the Kenyan site recruiting between May 2015-April 2018. Moreover, the Kenyan site conducted recruitment at 15 sentinel health centers (SHCs) (Table 3.1) where children aged less than 60 months from the HDSS sought care when they experienced MSD.

Table 4.1: VIDA sentinel health centers in Rural Western Kenya, 2015-2018

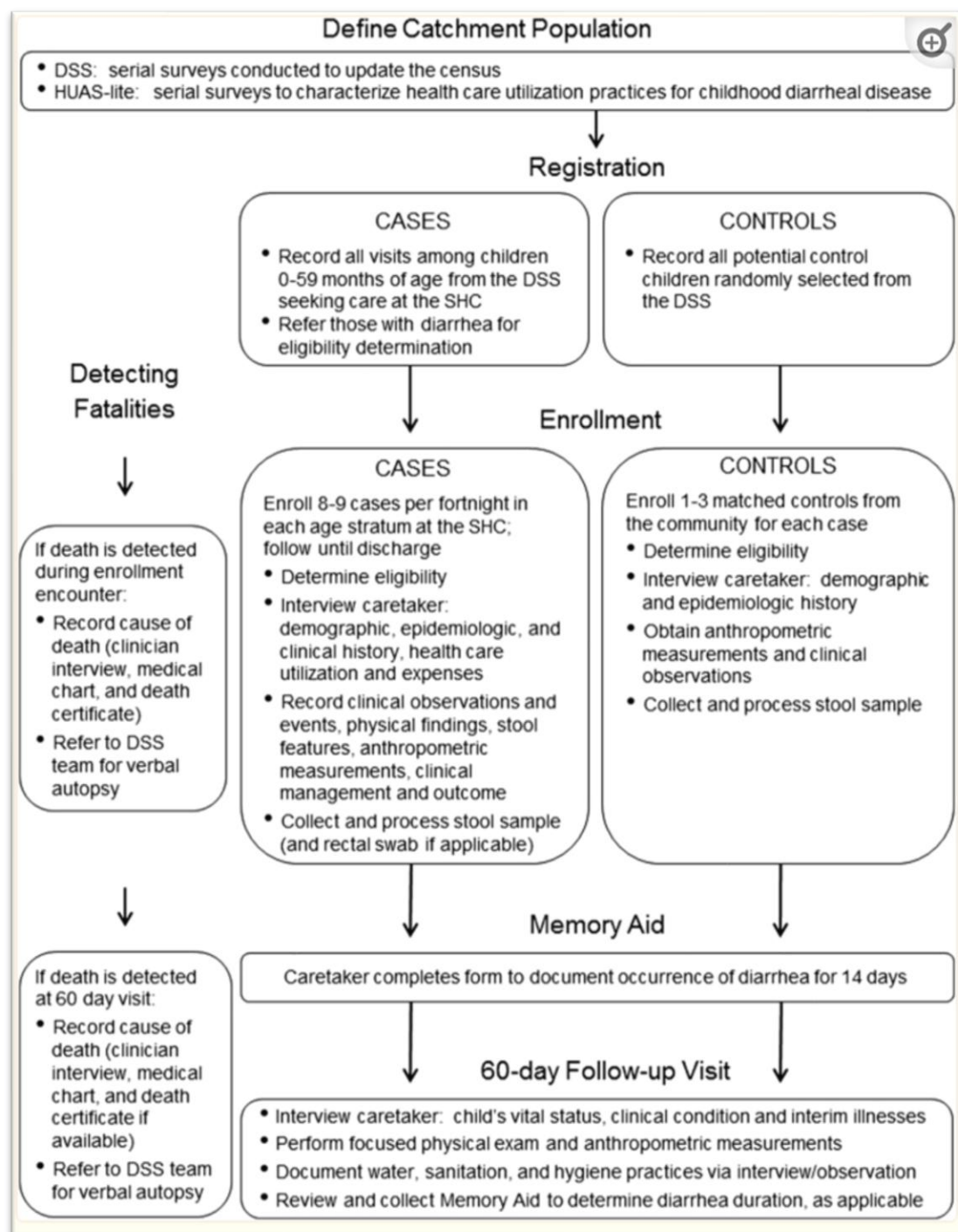
Sentinel Health Centre	TYPE	KEPH Level	AREA	LONGITUDE	LATITUDE	ELEVATION
Abidha Health Centre	G.O.K	Level 3	Asembo	34.411124000	-0.162158000	1098.00
Akala Health Centre	G.O.K	Level 3	Gem	34.425384000	-0.059145000	1286.24
Dienya Health Centre	G.O.K	Level 3	Gem	34.407073248	-0.001744321	1280.64
Lwak Mission Hospital	Mission	Level 3	Asembo	34.357334000	-0.140667000	1317.10
Mahaya Dispensary	G.O.K	Level 3	Asembo	34.341053000	-0.165152000	1209.30
Mulaha Clinic	Private		Karemo	34.260422963	0.049405771	1281.43

Ndori disp	Mission	Level 2	Asembo	34.378309000	-0.072347000	1147.00
Nyangoma Kogelo Disp	G.O.K		Karemo	34.344288635	-0.006929401	1283.97
Ongielo Dispensary	G.O.K	Level 3	Asembo	34.378348000	-0.171392000	1208.10
Siaya County Referral Hospital	G.O.K	Level 4	Karemo	34.287147000	0.062987000	1314.00
Ting wang'I H/C	G.O.K	Level 3	Karemo	34.282897698	-0.024535994	1246.26
Wagai Private	Private		Gen	34.437671272	0.052241118	1415.28

*KEPH- Kenya Essential Package for Health; G.O.K- Government of Kenya

For each case enrolled, 1-3 diarrhea-free controls were recruited during a home visit, using a strategy that adjusted the required number of controls based on the cases enrolled in that fortnight. Controls were matched to each case based on age (a margin of ± 2 and ± 4 months for cases aged 0-11 months and those aged ≥ 12 months, respectively), gender, and residential proximity to the case (same or a neighboring village) were randomly selected from the KEMRI-operated HDSS database and subsequently enrolled within two weeks of the case. During enrollment, VIDA meticulously collected a wealth of data, including clinical and epidemiological information, along with precise anthropometric measurements. Additionally, a stool specimen was obtained to identify enteropathogens. Approximately 60 days after enrolment, a home-based follow-up evaluation was performed to gather crucial information on vital status, clinical outcomes, and growth progression during the interim period.

The flow diagram of key study activities in VIDA are shown in Figure 4.3. As part of the collaborative effort between VIDA and HDSS, healthcare utilization and coverage surveys (HUCS) were conducted in the 3 HDSS areas. At least twice a year, whole population HUCS for children aged less than 60 months were conducted, which entailed asking the primary caretakers of participants' additional questions during the routine HDSS interviews.



* HUAS,-Health Care Utilization and Attitudes Survey.

Reproduced from Kotloff et al. (2012), with permission from Oxford University Press. © Oxford University

Figure 4.3: Flow diagram of key study activities in VIDA (Kotloff et al., 2012).

The HUCS served several purposes: i) to determine the proportion of children that failed to seek care at the SHCs when they had MSD, which was utilized to estimate the incidence of diarrhea in the population taking into account children who sought care at the SHCs and adjusting for those with MSD who did not; ii) to allow determination of where children from the HDSS population sought care when they had MSD, to guide the choice of SHCs for case enrollment during the case-control study; iii) to assess rotavirus vaccine coverage; and iv) to determine whether vaccinated and unvaccinated children differed with regards to their propensity in seeking care at a SHC for MSD. The clinical, epidemiological and laboratory methods of VIDA follow the same approach as the Global Enteric Multicenter Study (GEMS) and have been described in detail elsewhere (Kotloff et al., 2012; Panchalingam et al., 2012). Data from the VIDA study was used to build and internally evaluate predictive models for LDD and LGF.

4.2.3 Enterics for Global Health *Shigella* Surveillance

The EFGH study employed a hybrid surveillance approach that entailed cross-sectional and longitudinal study designs to determine incidence and sequelae of *Shigella* medically attended diarrhea (MAD) in seven sites spanning Africa (Kenya, Malawi, Mali and The Gambia), Asia (Bangladesh, and Pakistan), and Latin America (Peru). The primary objective of the EFGH study was to ascertain the incidence of *Shigella*-associated MAD among children aged 6 to 35 months across the various EFGH sites. This investigation aimed to delineate the incidence by factors such as specific strain of *Shigella* (serotype), the seriousness of the illness (severity), and the method used to identify the bacteria (either traditional culture or a more sensitive quantitative polymerase chain reaction [qPCR]), age group, and seasonal variations (Atlas et al., 2024). Secondary objectives encompassed describing the prevalence of antimicrobial resistance in *Shigella* spp. isolates within each EFGH site, evaluating the likelihood of adverse outcomes (including

mortality, hospital admission, PD, diarrhea recurrence, and LGF) within three months of a *Shigella* attributed MAD, and assessing the effectiveness of different severity criteria in distinguishing *Shigella* from non-*Shigella* attributed diarrhea and predicting subsequent risks of death or hospitalization. Additional objectives related to quantifying costs and addressing specific laboratory inquiries are detailed elsewhere (Horne et al., 2024; Liu et al., 2024; Morozoff et al., 2024).

Children aged 6 to 35 months who presented with MAD at designated SHCs were enrolled over two years and then followed up over a three-month period. Eligible MAD cases included infants and toddlers in the age range of 6-35 months who visited a SHC with a diarrheal illness (at least three loose stools within 24 hours), starting within the previous week and following at least two diarrhea-free days. Further requirements to be eligible included: be resident in the designated study area; do not intend to out-migrate from the study area for at least four months; have a legal guardian consent to allow the child participate in the study including follow-ups for a duration of three months following enrolment; no referral to a non-EFGH facility during screening, and the site enrollment limit must not have been reached (Atlas et al., 2024). During enrollment, a standardized questionnaire was administered to the caregiver at the SHCs to collect data, which included the child's recent clinical history, anthropometric measurements, and basic demographic and social characteristics of the household (Table 4.2) (Feutz et al., 2024).

Table 4.2: Study procedures at each EFGH study time point

Activity	Study Time point			
	Enrollment	4-week follow-up	3-month follow-up	Unscheduled visit (if applicable)
Caregiver interview	X	X	X	X
Anthropometry	X	X	X	
Child Physical Exam	X	X	X	X

Medical record abstraction	X			X
Rectal swabs	X			
Whole stool	X**			
Dried Blood Spots	X	X*		

* If 4-week follow up is being conducted at the home, DBS will not be collected.

**If whole stool was not produced at the enrollment visit, at select sites, home collection within the 24-hour period following leaving the enrollment visit, took place.

4.2.4 Rotavirus Surveillance Network for Africa

As a component of the African-based surveillance network coordinated by the WHO for Rotavirus disease (Mwenda et al., 2014), a prospective hospital-based surveillance was carried out for acute gastroenteritis (AGE) caused by Rotavirus, within the KEMRI-HDSS area in Karembo. This surveillance initiative was implemented at SCRH, which has a pediatric bed capacity of 60 and serves a culturally homogenous, predominantly rural population. The study enrolled pediatric patients (0-59m) hospitalized with AGE, defined as having at least three looser than normal stools and/or at least one instance of unexplained vomiting followed by loose stool within a 24-hour timeframe, where these symptoms began within the 7 days prior to care-seeking at the facility. The study site, procedures and population have been previously described (Khagayi et al., 2014; Nyawanda et al., 2016; Omore et al., 2019). Data collected from January 2010 to December 2020 was used to build and internally validate the AGE-associated mortality prediction models.

4.3 Research Design

According to McCombes and Bhandari, research design is considered as a strategy or framework employed to address research question using empirical data (McCombes & Bhandari, 2021). It delineates the comprehensive structure and roadmap for gathering, analyzing, and interpreting data aimed at addressing particular research questions or hypotheses. In this study, the aim was to assess

the relationship between demographic, clinical, and socio-economic factors and poor diarrheal outcomes— LDD, LGF, and mortality. Predictor variables were utilized to develop a series of predictive models for each diarrheal outcome. The design of tools used to examine the relationships between these variables was informed by existing theories. Subsequently, models were constructed and assessed to determine their efficacy in predicting these poor diarrheal outcomes.

4.3.1 Correlational Research Design

Correlational research design is a non-experimental approach where a researcher assesses two variables, examines and evaluates the statistical association between them, without the influence of any external variables (Devi et al., 2023; Lau, 2017; Mekonnen, 2020). It explores associations between variables without the researcher exerting control or manipulation over any of them. Seeram further defines correlational research design as a non-experimental approach that enables the prediction and explanation of relationships among variables (Seeram, 2019). This approach reveals interacting variables and the nature of their interaction, thereby empowering the researcher to make predictions based on the established relationships. At its essence, this study aimed to identify variables associated with poor diarrheal outcomes, without accounting for external factors that could influence these associations. Causation was not a focus, and the relationship between the study variables was not manipulated. Although correlations between variables are established without implying causality, correlational research design remains within the realm of quantitative research methods.

Explanatory and predictive designs are two distinct approaches within correlational research design, each serving different purposes and addressing different types of research questions (Devi et al., 2023; Jackson, 2008). Explanatory design seeks to understand underlying mechanisms

between variables and explain why certain relationships exist, while predictive design focuses on forecasting or predicting outcomes based on those relationships. Both designs are valuable in advancing the understanding of the complex phenomena of childhood diarrhea and informing decision-making in its therapeutic management.

The research design for this study is thus categorized as correlational. This classification stems from the exploration of the relationships between variables using explanatory research design. Additionally, models were developed aimed at predicting poor diarrheal outcomes—namely, LDD, LGF, and mortality—thus aligning with the principles of prediction research design.

In general, science pursues three main goals: description, explanation, and prediction. Different research designs are tailored to achieve these overarching objectives. The correlational design exemplifies this by simultaneously describing the relationships between variables through explanatory research methods and facilitating the prediction of outcomes through predictive design strategies. The rationale for correlational design for this study is justified by the following reasons. Firstly, correlational research analyzes relationships between variables as they exist, without establishing cause-and-effect through manipulation, which is appropriate for exploring complex phenomena such as childhood diarrheal outcomes. The study evaluated the relationship between predictor variables (demographic, clinical, and socio-economic) and predicted variables (LDD, LGF and mortality). Examining correlations between various predictors and poor diarrhea outcomes, allows for the identification of potential predictors without imposing interventions on the vulnerable population. These variables were then used to build predictive models for poor diarrheal outcomes. Additionally, the correlational design offers models that align with both explanatory and predictive research designs, which fit the desired analytical approach for this study.

While correlational research design offers valuable insights into relationships between variables, it also has inherent limitations regarding causality, directionality problem, assumption of linearity and generalizability (Crawford, 2014; Lau, 2017). In certain instances, correlational research may allow for limited inferences regarding potential causal relationships between variables, provided that reverse causation and the influence of common-causal variables are effectively mitigated. This could further be reinforced by considering multiple lines of evidence when drawing conclusions from correlational data.

4.4 Sample Size Calculation

Constructing a prediction model demands a development dataset of sufficient size, enabling the formulation of a robust prediction equation that maintains reliability when extrapolated to new individuals within the target population. The widely accepted guideline recommending a sample size of 10 events per predictor parameter (Peduzzi et al., 1996), has been demonstrated to be overly simplistic (Riley et al., 2020). Consequently, the determination of sample size was conducted utilizing a formula developed by Riley et al., which requires the following inputs: the overall likelihood of outcomes or average outcome value within the target population, along with the number of candidate predictor parameters and the expected model performance regarding overall fit as shown in Equation 16 (Riley et al., 2020).

For each of the diarrheal outcomes studied sample size was estimated using the formula below:

$$n = \frac{P}{(S-1)\ln\left(1-\frac{R^2_{cs}}{S}\right)} \text{ (Riley et al., 2020) 16}$$

Where P= Candidate predictor parameters; S= 1- shrinkage; R^2_{cs} - Cox-Snell R squared statistic;

P was **17** for the LDD and LGF prediction and **25** for mortality prediction; desired shrinkage level was $\leq 10\%$ $S=0.9$ and R^2_{cs} is at least 0.1

For the prediction of LDD and LGF

$$n = P / ((S-1)\ln(1-(R^2_{cs}/S)))$$

$$= 17 / ((0.9-1)\ln(1-(0.1/0.9)))$$

$$= \mathbf{1,443}$$

For mortality prediction:

$$n = P / ((S-1)\ln(1-(R^2_{cs}/S)))$$

$$= 25 / ((0.9-1)\ln(1-(0.1/0.9)))$$

$$= \mathbf{2,122}$$

The determined sample sizes were at least 2,122 and 1,443 observations for mortality outcome, and LDD and LGF outcomes respectively. The sample size calculation closely aligns with the total population count (Table 4.3). Consequently, the entire population in the primary studies were included as participants for each diarrheal outcome in this particular study.

Table 4.3: Tabulation of Sampling Frame and Sample Size

No	Participants in the study	Study	Total population	Sample Size
1	Children presenting with Moderate-Severe-Diarrhea and have Diarrhea duration data available in Rural Western Kenya	VIDA Study (2015-2018)	1,484	1,443
2	Children presenting with Moderate-Severe-Diarrhea and have Diarrhea duration data available in Rural Western Kenya	VIDA Study (2015-2018)	1,484	1,443

3	Children hospitalized with acute gastroenteritis at Siaya County Referral Hospital and discharge outcome available	Rotavirus Inpatient Surveillance (2010-2020)	2,271	2,122

4.5 Data Acquisition and Data Preprocessing

The current study leveraged data from three enteric studies conducted among pediatric populations within the study setting, to develop and validate prediction models for three diarrheal outcomes. Specifically, data from the VIDA study (2015-2018) was used as the development cohort for the LDD and LGF outcomes with data from the EFGH study (2022-2023) being used to temporally validate the champion models for these 2 outcomes. Moreover, data from the rotavirus inpatient surveillance (2010-2020) was used to develop and internally validate prediction models for AGE-associated mortality. These studies collected demographic, clinical, and socio-economic data of the participants at enrolment, and these were considered potential predictors of the diarrheal outcomes.

4.5.1 Handling Missing Data

Patterns of missing data were analyzed and gaps in the predictor variables addressed by performing imputations using the Multiple Imputation by Chained Equations (MICE) package (Azur et al., 2011; van Buuren et al., 2021). MICE is a versatile and comprehensive framework for performing data imputation and includes several functions for identifying the missing data patterns present in a dataset. In summary, it provides a robust solution for handling missing data through multiple imputation. It operates by creating several complete datasets with plausible imputations for missing values, employing a chained equations approach where variables are imputed iteratively based on observed values of other variables. Users can specify imputation models for different

variable types, and MICE supports various statistical models for imputation, including linear and logistic regression. After multiple iterations, MICE pools result from imputed datasets to obtain valid estimates and standard errors, accounting for uncertainty due to missing data. Diagnostic tools are available for assessing imputation quality, ensuring reliable analysis. Generating multiple imputations, rather than just one, addresses the statistical uncertainty inherent in imputed values. Moreover, the chained equations method offers remarkable flexibility, capable of accommodating variables with diverse characteristics (e.g., continuous or binary) and handling complexities like boundaries or survey skip patterns.

The VIM (Visualization and Imputation of Missing Values) package in R was used to visualize missing data patterns (Kleinke et al., 2020; Templ et al., 2022). The *aggr()* function, was used to produce summary plots showing the proportion of missing values in each variable in the datasets used for model development. The missingness map and the patterns of missing data in the VIDA study, used in LDD and LGF predictions, are shown in Figure 4.4 and 4.5 respectively.

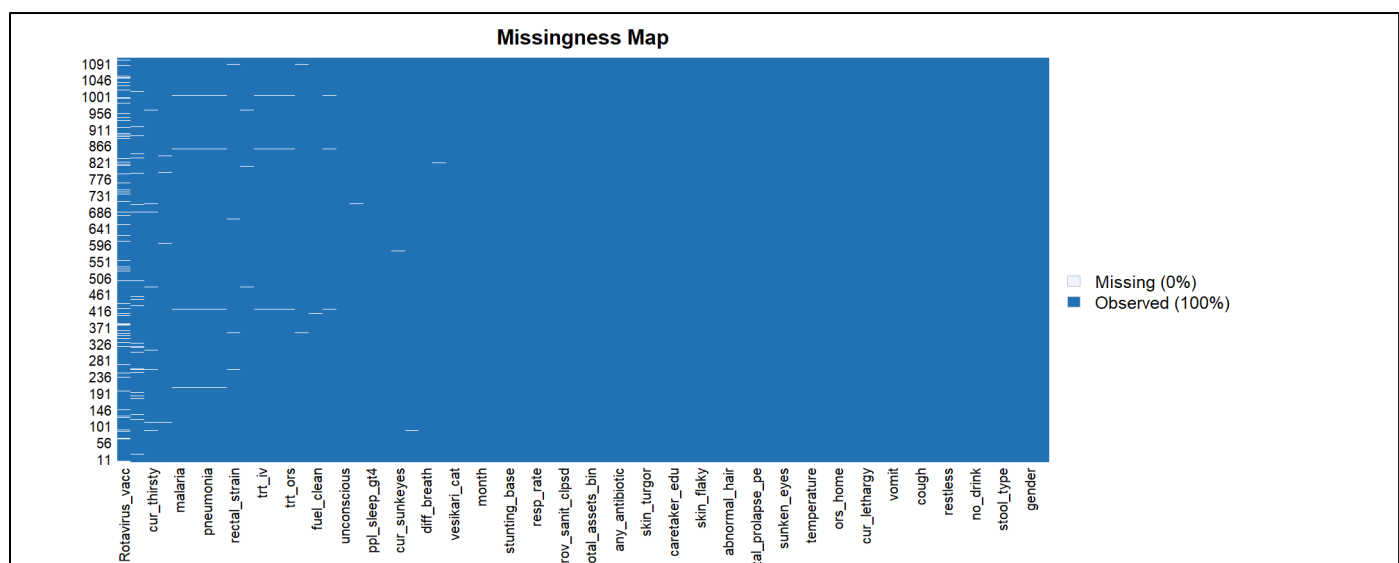


Figure 4.4: Missingness map for VIDA study

Overall, VIDA had <1% of missing data points with rotavirus vaccination (11.2%), belly pain (4.5%) and currently thirsty (0.9%) having the highest missing values.

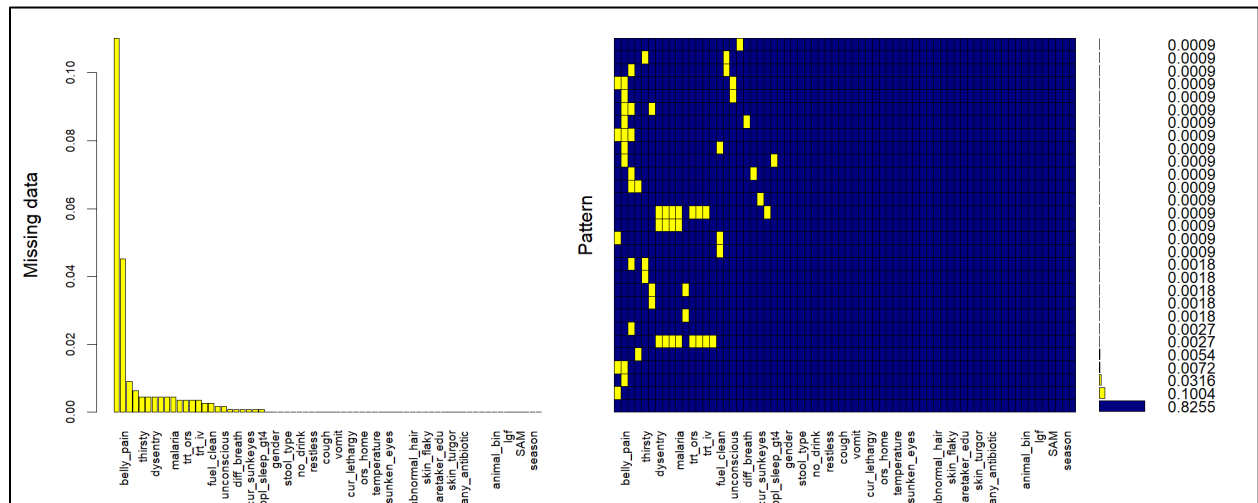


Figure 4.5: Patterns of missing data in VIDA study

The missingness map and the patterns of missing data in the rotavirus inpatient surveillance, used in the mortality prediction, are shown in Figure 4.6 and 4.7 respectively. Overall, the rotavirus inpatient surveillance had 8% of missing data points with prior admission (34.0%), pulse (30.8%), bulging fontanelle (28.8%), red eyes (26.6%), and frequency of vomiting (16.1%), having the highest missing values.

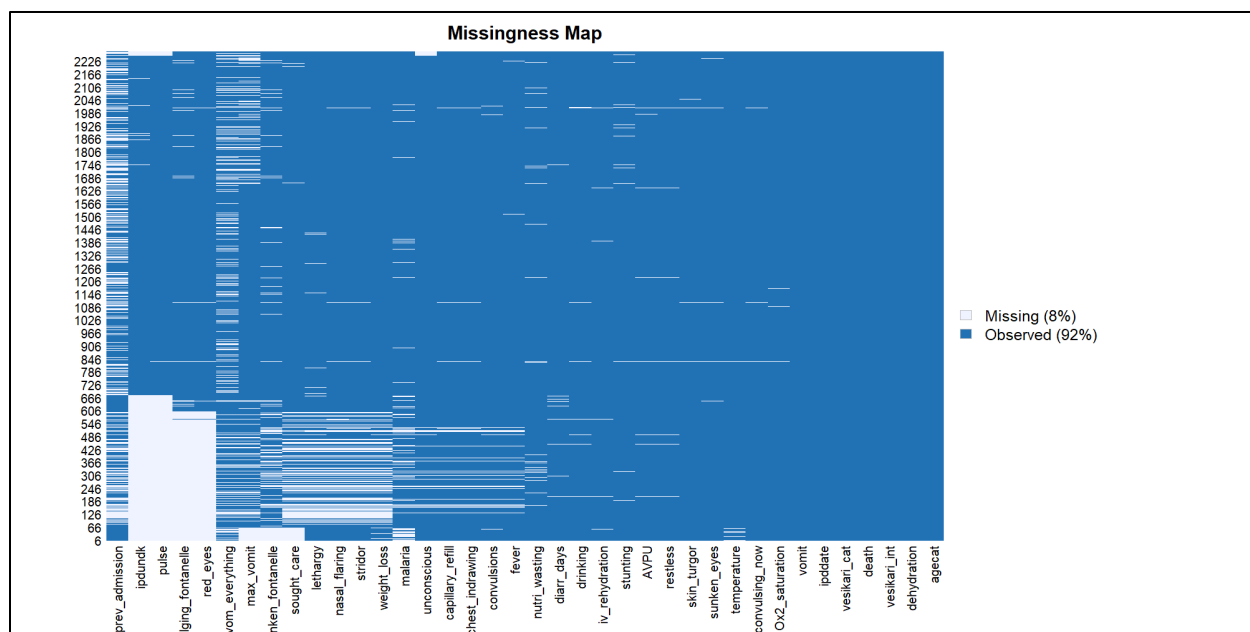


Figure 4.6: Missingness map for the rotavirus inpatient surveillance

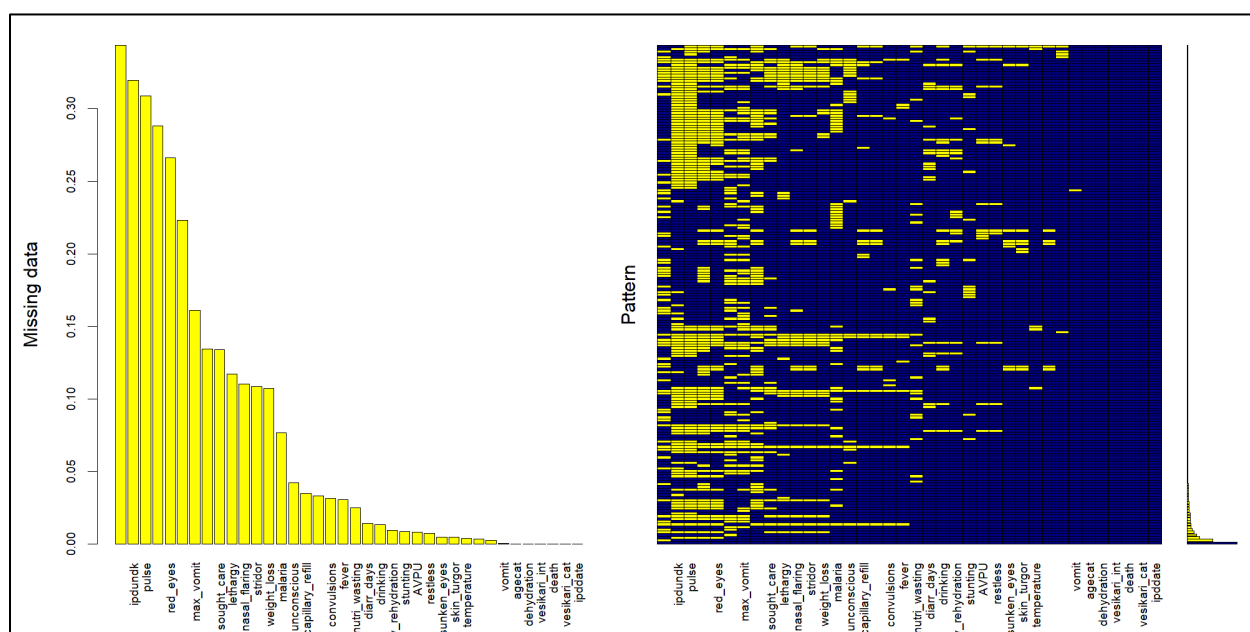


Figure 4.7: Patterns of missing data in rotavirus inpatient surveillance

4.5.2 Outcome Variables

For longer duration diarrhea (LDD), the duration of diarrhea was assessed using two distinct data sources. Prior to enrollment, caregivers provided information on the duration, with the reported

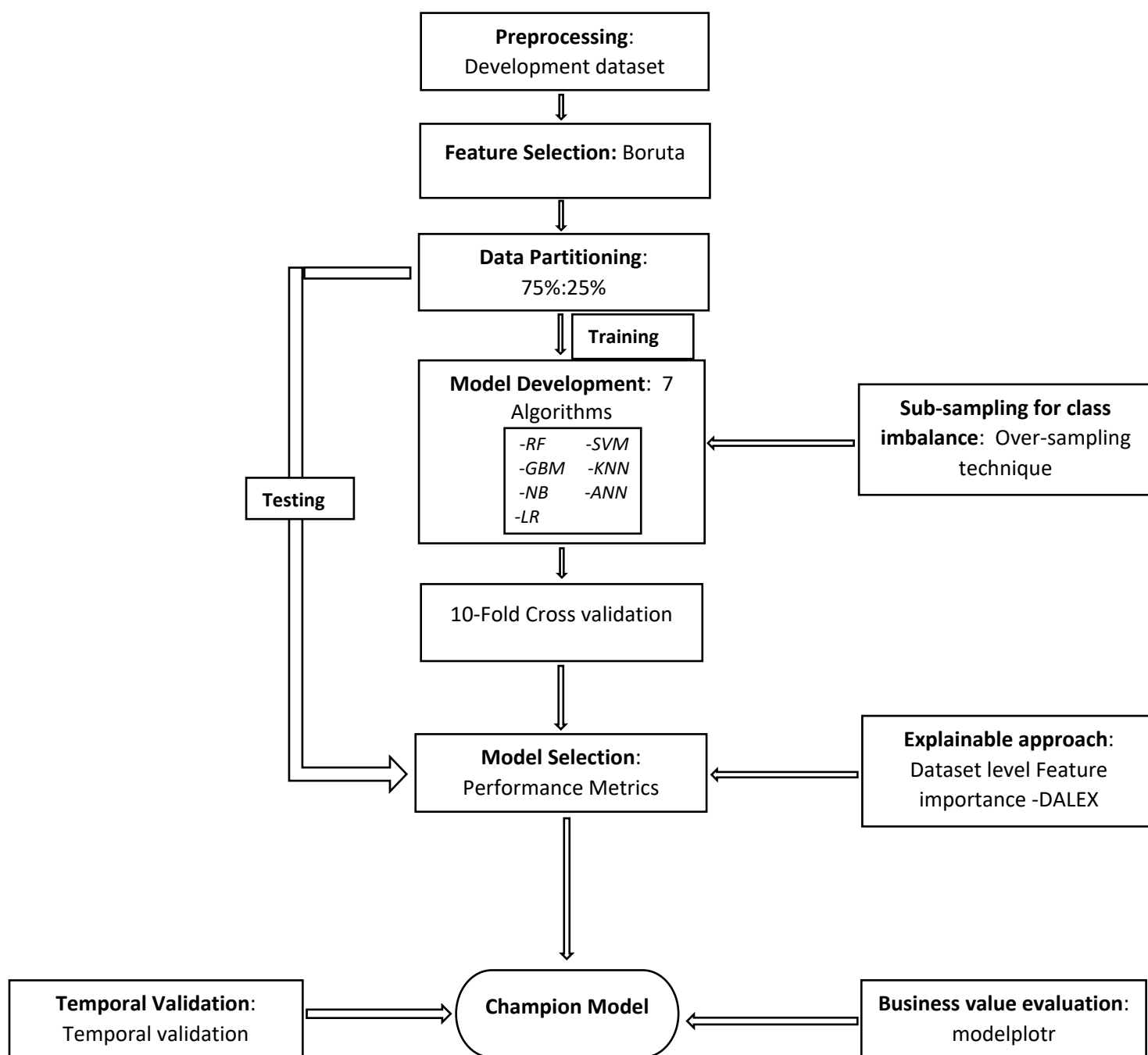
days of diarrhea considered continuous. Post-enrollment, data was collected for 14 days using memory aid/ diarrhea diary. Data from the VIDA study as well as the EFGH study was used. The details of the protocol used on deciding on LDD have been elaborately detailed elsewhere (Ogwel et al., 2025).

For linear growth faltering (LGF), trained personnel conducted measurements of length or height both upon recruitment and during follow-up visits. Age information was gathered from various sources such as birth certificates and vaccination records. For children aged 2 or more years, standing height was captured, while children under 2 years had their length measured while lying down, utilizing a Shorr board (Weigh and Measure, LLC). The details of the protocol used on deciding on LGF have been elaborately detailed elsewhere (Ogwel, Mzazi, Awuor, et al., 2024).

Prior to modelling the mortality outcome, an assessment was carried out of the predictive performance of existing diarrheal scores in predicting mortality to avoid duplication of research efforts. Specifically, the WHO dehydration scores, the modified Vesikari score for severity of diarrheal illness and the IMCI general danger signs were evaluated. The details of the protocol used on deciding on mortality have been elaborately detailed elsewhere (Ogwel, Mzazi, Nyawanda, et al., 2024)

4.6 Modelling Approach

In the study, the modeling approach adopted involved several key steps: feature selection, model development, model validation, explanatory model analysis, and business value evaluation. This approach is delineated in Figure 4.8.



*RF-Random Forest; GBM-Gradient Boosting; NB- Naïve Bayes; LR-Logistic Regression; SVM- Support vector machine; KNN-K-nearest neighbors; ANN-Artificial Neural Networks

Figure 4.8: Model development and validation schematic diagram for the prediction of poor diarrheal outcomes.

4.6.1 Feature Selection

Feature selection is a commonly used method to reduce dimensionality concerns whose primary objective is to identify and select a compact subset of pertinent features from the original set based on specific criteria for assessing relevance (Liu, 2010; Tang et al., 2014). Beyond the built-in selection capabilities of specific models, the feature selection landscape can be broadly be classified into two primary approaches: wrapper and filter methods (Chandrashekar & Sahin, 2014; Dhal & Azad, 2022; Khalid et al., 2014; Pudjihartono et al., 2022). Wrapper methods iteratively build and evaluate models with different combinations of features with the goal being to identify the subset of features that yields the highest performance on a given metric. While filter methods on the other hand assess the importance of features using statistical measures, without considering the specific model being used and ranking schemes to assess the individual contribution of each feature, filtering out irrelevant or redundant predictors. Common filter techniques include: correlation-based methods; variance-based methods; information gain; and Chi-squared test.

Both wrapper and filter methods for feature selection offer distinct strengths and drawbacks (Nnamoko et al., 2014; Suto et al., 2016). Filter methods, while computationally efficient, often lack a direct link between their selection criteria and the model's effectiveness. Conversely, wrapper methods, although computationally more expensive and prone to overfitting, excel at finding highly accurate feature combinations tailored to the specific model and dataset. Their model-specific approach considers the intricate interactions between features, often leading to a higher likelihood of selecting features that truly contribute to optimal predictive performance. However, the disadvantage of this approach is the need to evaluate many models, often requiring parameter tuning, which significantly increases computation time and the likelihood of overfitting to the training dataset.

Based on the aforementioned literature, a hybrid feature selection strategy was chosen, blending a filtering approach with a wrapper algorithm. This combined method leveraged the strengths of both approaches to effectively handle the moderate sample size and potentially complex feature interactions within the study datasets. Specifically, a filtering step was applied first, retaining features with a $p\text{-value} < 0.2$, which helped to reduce the dimensionality of the datasets, removing features with weak associations to the target variables. Subsequently, the Boruta algorithm, a wrapper method built upon random forests, was employed to further refine the feature set because of its ability to handle complex interactions and mitigate overfitting, while maintaining computational efficiency (Kursa & Rudnicki, 2020). This approach aimed to achieve a balance between computational efficiency, robustness, and accuracy in the feature selection process.

4.6.2 Model Development

This section outlines the model development process, including data partitioning and model training using seven ML algorithms.

4.6.2.1 Data Partitioning

Data partitioning, also known as split sampling, involves breaking down a dataset into distinct, non-overlapping subsets, typically utilized for training, validation, and testing ML models ultimately leading to the creation of robust and generalizable models. This stratification allows for a systematic approach to model development and evaluation as it facilitates a more precise assessment of model efficacy and serves to mitigate the risk of overfitting. While there is no established guidelines on data split ratios, extensive analysis across diverse datasets has consistently demonstrated that adopting an equidistributed data partitioning scheme falling within the intervals of (20%, 80%) and (30%, 70%), yields superior outcomes in building classification

models (Li, 2022). This optimal sampling strategy holds true across a spectrum of domains, dataset sizes, and class distribution imbalances, establishing it as a robust approach for model construction. In this study, a 75%:25% split was selected for the development study to ensure sufficient data were available for model training while reserving a substantial and independent test set to allow reliable assessment of out-of-sample performance and generalizability. This 75:25 split was not arbitrary. It was a specific, optimal instantiation of the general guideline (H. L. and M. Li, 2022), and supported by direct empirical precedent (Nguyen et al., 2021).

4.6.2.2 Model Training

Model training in ML involves the methodology of developing a model's ability to generate predictions or make decisions through the utilization of input data (Google for Developers, 2022; Hanchuk & Semerikov, 2025). During training, the model learns the mapping of input features to target outputs by adapting its internal parameters through optimization algorithms. A loss function is defined to measure the model's performance, and an optimization algorithm is chosen to minimize this loss function during training. Hyper-parameters are tuned to optimize performance, while the model's predictive power on the validation set is tracked to ensure it generalizes well and avoids overfitting. Early stopping may be implemented to halt training when performance deteriorates. The prediction tasks for poor diarrheal outcomes were formulated as classification problems with labelled data hence a supervised learning approach was used. The choice of the algorithms to be used in the study was guided by the knowledge of data structure and complexity, the accuracy requirements, processing speed or time and number of features (Sarker, 2021). Since algorithms fall in different families and have different purposes and approaches in the implementation of a classification task, algorithms were selected based on diversity by family and

prior use in predicting healthcare problems to allow for a comparative evaluation of their performance in the prediction of poor diarrheal outcomes. The algorithms were chosen because of their popularity, potency, wide usage, and variation in performance.

In the process of model development, a comprehensive range of seven ML algorithms were applied to the training datasets. These encompassed Random Forest (RF), Gradient Boosting (GBM), Naive Bayes (NB), Logistic Regression (LR), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Artificial Neural Networks (ANN). Each algorithm was leveraged to explore various aspects of the data and address the unique complexities inherent in the predictive task at hand. This diverse selection allowed for thorough analysis and comparison, ultimately leading to the identification of the most effective model for each diarrheal outcome.

The training dataset was subjected to K-fold cross-validation ($k=10$) to mitigate the risks of underfitting or overfitting the model (Irizarry, 2019; Kassambara, 2018). K-fold cross-validation serves as a pivotal methodology for rigorously assessing predictive models (James et al., 2022; Lumumba et al., 2025; Wilimitis & Walsh, 2023). Here, the dataset is segmented into k partitions, with the model being developed and evaluated k times. During each iteration, one specific fold is assigned as the validation set, while the other folds are utilized to train the model. By sequentially cycling through the folds, each data point is used once for validation and multiple times for training, ensuring comprehensive assessment. The performance metrics from each fold are combined and averaged to provide an overall assessment of the model's ability to generalize. This robust methodology significantly contributes to model assessment, selection, and hyper-parameter tuning, providing a more reliable measure of model's performance across diverse datasets. Typically, the parameter k in k -fold cross-validation is commonly set to 10, a value empirically determined to yield the estimation of the model's predictive capability characterized by minimal

bias and a reasonable degree of variability. However, this choice is not governed by strict rules. As k grows larger, the disparity in size between the training dataset and the resampled subsets is reduced, and accordingly, the bias inherent in the technique is also diminished. This phenomenon stems from the decreasing discrepancy in data distribution between the training and validation sets, resulting in a more realistic evaluation of the model's predictive power across diverse datasets.

To address the class imbalance in the outcome variable, sub-sampling was used during resampling to mitigate potential adverse effects of unequal class frequencies on model fitting (Kuhn, 2019). Class imbalance arises when one class, referred to as the minority class, has substantially fewer instances in comparison to the majority class within the dataset. In numerous applications, the class with fewer instances often holds greater significance and interest (Elrahman & Abraham, 2013). This imbalance issue becomes particularly pronounced when the class of interest is rare and possesses only limited data points compared to the predominant class. Furthermore, the repercussions of misclassifying the minority class are considerably more severe than those of misclassifying the majority class. For instance, consider scenarios such as mortality and survival, where accurately identifying the minority class—such as instances of mortality—holds paramount importance. In such cases, addressing class imbalance becomes imperative to ensure the model effectively captures the nuances of the less prevalent but crucial class, consequently improving the overall predictive capability and dependability of the predictive system. Broadly, three categories of solutions exist to tackle the challenge of class imbalance: algorithmic, data-centric, and cost-sensitive methods (Huang et al., 2021). Data-centric approaches, typically prioritized, are independent of the classifier and primarily revolve around data sampling techniques. These methods center on recalibrating the provided training set, aiming to restore balance between classes. By strategically manipulating the dataset through techniques such as under-sampling,

over-sampling, or hybrid approaches, these methods strive to equip the model with a more balanced and representative training data, thereby fostering improved performance and generalization across diverse class distributions.

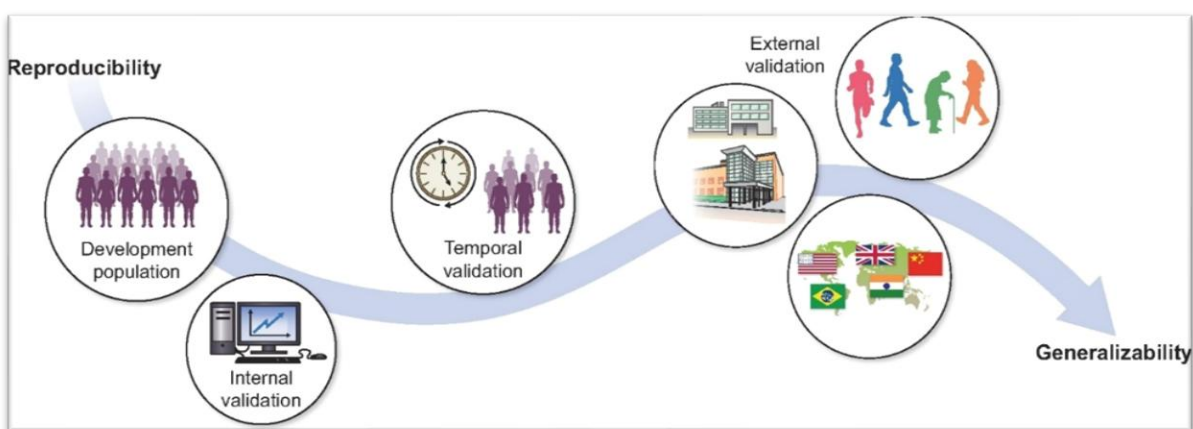
Sub-sampling, a core strategy within resampling procedures in ML, plays a pivotal role in tackling the issue of unequal class distribution—a common hurdle in classification tasks where one class significantly outweighs the other (Nejatian et al., 2018; Qu et al., 2020). These methods strategically rebalance class distribution: either by the random removal of instances from the majority class to approximate the representation of minority class (under-sampling) or by generation of new instances for the minority class to achieve parity with the majority class (over-sampling). Hybrid approaches amalgamate both techniques, leveraging their strengths and minimizing their weaknesses for enhanced model performance. In this study, an over-sampling technique was specifically employed due to the limited sample size of the primary studies. Implementation involved specifying "up" in the sampling option within the trainControl function. Over-sampling entailed replicating or generating new samples for the minority class, thereby providing the model with more representative examples and improving its ability to discern underlying patterns and make accurate predictions.

Model training was executed using the caret package in R, presenting a seamless and effective methodology for building ML models (Kuhn et al., 2022). This versatile package furnished a cohesive platform for training and fine-tuning the seven algorithms named above, facilitating straightforward comparison and selection of the most optimal model for each diarrheal outcome. Its intuitive interface enabled the specification of the desired modeling pipeline through a straightforward syntax, thereby expediting the process of experimentation and iteration. Moreover, caret proved instrumental in automating numerous laborious tasks, including cross-validation, sub-

sampling and parameter tuning, thereby significantly reducing the time and effort required for model development.

4.6.3 Model Evaluation

Receiver Operating Characteristic (ROC) curve and Precision-Recall curve were plotted to evaluate the model's discriminatory ability across different thresholds (Li & Guo, 2021; Saito & Rehmsmeier, 2023). These curves help to visually illustrate compromise between sensitivity measured by the true positive rate (sensitivity) and the false positive rate measured by (1-specificity), or precision and recall, respectively. A model's true predictive power cannot be determined solely by showing that it performs well on the training dataset, since the prediction model is optimized for the specific dataset used for its creation, it might perform well on that dataset but might not generalize well to unseen data. Various validation approaches—internal, temporal, and external—can be used, each with differing levels of stringency as shown in Figure 4.9 (Ramspek et al., 2021).



Reproduced from Ramspek et al. (2021), with permission from Oxford University Press. © Oxford University

Figure 4.9: Visualization of Validation Techniques (Ramspek et al., 2021).

4.6.3.1 Internal Validation

Internal validation stands as a pivotal step in scrutinizing the efficacy of a predictive model, employing the very dataset utilized for its development (Sperrin et al., 2022). Its primary aim lies in mitigating the in-sample optimism phenomenon, wherein models tend to overfit the development data compared to other data from the same population. Given its reliance on the same dataset, internal validation is consistently executable. However, it's crucial to acknowledge that this method lacks data from new patients, thereby predominantly offering insights into the prediction model's replicability (Ramspek et al., 2021). Typically, a patient cohort is subjected to random partitioning, resulting in the creation of development and internal validation sets. The predictive model is then built on the development cohort (training set) and subsequently subjected to internal validation on the testing dataset. This iterative process is indispensable for meticulously gauging the model's performance and discerning any overfitting issues during its developmental phase.

The assessment of model performance for each diarrheal outcome was conducted on the test datasets, employing a comprehensive suite of evaluation metrics. These included sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), F1-Score, as well as Area under the curve (AUC) and Precision-recall AUC (PRAUC) scores. Table 4.4 provides a summary of these metrics giving the formula and evaluation focus. This holistic approach ensured a thorough comprehension of the model's predictive power and provided insights into its effectiveness across diverse performance aspects.

Table 4.4: Evaluation metrics for poor diarrheal outcomes

Metrics	Formula	Evaluation Focus
Sensitivity (sen)	$TP/(TP+FN)$	It represents the percentage of correctly predicted positives out of all actual positives.
Specificity (spec)	$TN/(TN+FP)$	It represents the percentage of correctly predicted negatives out of all actual negatives.
Positive Predictive Value (PPV)	$TP/(TP+FP)$	It represents the percentage of correctly predicted positives out of all positive predictions.
Negative Predictive Value (NPV)	$TN/(FN+TN)$	It represents the percentage of correctly predicted negatives out of all negative predictions.
F1-Score (F1)	$TP/(TP+1/2(FP+FN))$	It's a balanced metric that summarizes the trade-off between precision and recall, the harmonic mean, to give an overall sense of model effectiveness.
Area Under the Curve (AUC)	$\int_0^1 (\frac{TP}{TP+FN})d(\frac{FP}{FP+TN})$	It measures the model's discriminative power using the relationship between true and false positive rates across varying decision thresholds.
Precision-Recall - Area Under the Curve (PRAUC)	$\int_0^1 (\frac{TP}{TP+FP})d(\frac{TP}{TP+FN})$	It measures a model's performance by evaluating the precision-recall trade-off, offering a more accurate assessment, especially for imbalanced classes.

4.6.3.2 Temporal Validation

Temporal validation serves as a critical intermediary step, bridging the gap between internal and external validation processes (Austin et al., 2017; Ramspek et al., 2021). This validation method entails assessing the model's performance on new patient data collected from the same study but at different time points. By doing so, temporal validation offers important perspectives into both the model's reproducibility and generalizability, especially in effectively capturing temporal patterns and trends. In this study, champion models developed in the derivation cohort underwent rigorous temporal validation using a separate cohort. Specifically, for the LDD and LGF prediction tasks, the VIDA study data (2015-2018) served as the validation cohort, while data from the EFGH

Shigella surveillance (2022-2023) acted as the temporal validation cohort. The evaluation metrics used in temporal validation mirrored those employed during internal validation. However, due to the unavailability of a suitable temporal validation dataset, the mortality prediction champion model was not subjected to temporal validation.

4.6.3.3 Model calibration

Model calibration has been defined as the alignment between predicted class probabilities and observed outcomes or consistency between the calculated risk and the actual likelihood of an event (Huang et al., 2020; Manokhin & Grønhaug, 2026; Paquette-Rufiange et al., 2026). Essentially, a properly calibrated model is one with the ability to minimize residuals accurately assessing the likelihood of an event occurring, reflected in the predicted probabilities it assigns to different outcomes. The importance of model calibration lies in its impact on decision-making processes. In many applications, especially those related to risk assessment or medical diagnosis, decision thresholds are set based on predicted probabilities. If a model is poorly calibrated, these thresholds may not accurately reflect the true likelihood of an event, leading to suboptimal decisions and potentially harmful consequences.

While model calibration is typically assessed by plotting predicted probabilities against observed outcomes using a calibration curve or reliability diagram, the level of calibration can be quantified by metrics such as Brier score. In this study, calibration was assessed using Brier scores, which measure the average squared difference between predicted likelihoods and observed outcomes as shown in Equation 17, Spiegelhalter's z-test, a formal metric derived from the decomposition of the Brier score as depicted in Equation 18, and serving as an alternative measure of calibration along with its corresponding p-value (Huang et al., 2020).

$$\text{Brier Score} = \frac{\sum_{i=1}^N (E_i - O_i)^2}{N} \dots\dots\dots (17)$$

In this context, N represents the total count of patients; E_i signifies the model's prediction for a particular patient ' i '; and O_i denotes the actual result observed for that same patient ' i '.

$$\text{The Spiegelhalter z statistic} = Z(E, O) = \frac{\sum_{i=1}^N (O_i - E_i)(1 - 2E_i)}{\sqrt{\sum_{i=1}^N (1 - 2E_i)^2 E_i (1 - E_i)}} \dots\dots\dots (18)$$

If $|Z(E, O)| > q_{1 - \alpha/2}$, If the result exceeds the threshold defined by $q\alpha$, which is the 0.05 quantile of the standard normal distribution, it has statistical significance, indicating the model is likely poorly calibrated.

The Brier score was computed utilizing the specialized R package "rms" with the function "val.prob" (Harrell Jr, 2023). Moreover, Spiegelhalter's z statistic was derived using the same "val.prob" function within the "rms" package.

Predicted probabilities were generated on an independent test dataset to evaluate out-of-sample calibration performance. For probabilistic classifiers predicted class probabilities were obtained using the model-specific probability prediction functions. Only the predicted probability corresponding to the positive outcome class was retained for calibration analyses. Observed outcomes were encoded as binary indicators, with events coded as 1 and non-events as 0. The Brier score ranges from 0 to 1, with lower values indicating better calibration and discrimination. A score of 0 corresponds to perfect probabilistic predictions, whereas a score of 0.25 corresponds to non-informative predictions in a balanced binary outcome setting. Graphical calibration was assessed using calibration plots generated with the `val.prob()` function from the `rms` package in R. Predicted probabilities were plotted against observed outcome frequencies, with optional

nonparametric smoothing to visualize departures from the 45-degree line of perfect calibration. These plots provide an intuitive assessment of calibration across the full spectrum of predicted risk.

Several techniques are available to convert model predictions into posterior probabilities including regression-based calibration, isotonic calibration, general updating approaches and similarity-binning averaging (Ojeda et al., 2023). Platt scaling was used to correct poorly calibrated models, which is a popular regression-based calibration technique that was first proposed by John Platt in the context of SVMs in 1991 but can be applied to various classifiers (Manokhin & Grønhaug, 2026; Nikolić et al., 2025). Platt scaling approach involves transforming model estimates to posterior probabilities by processing the estimates through a calibrated sigmoid function (Huang et al., 2020). It was chosen because of its wide applicability, effectiveness in significantly improving the calibration of models that produce skewed or poorly calibrated probabilities, and the simplicity in its implementation. To get calibrated probabilities, the prediction generated by the learning algorithm $f(x)$ was passed through a sigmoid as shown in Equation 19:

$$P(y=1|x) = 1 / (1 + \exp(-(a * f(x) + b))) \dots \dots \dots (19)$$

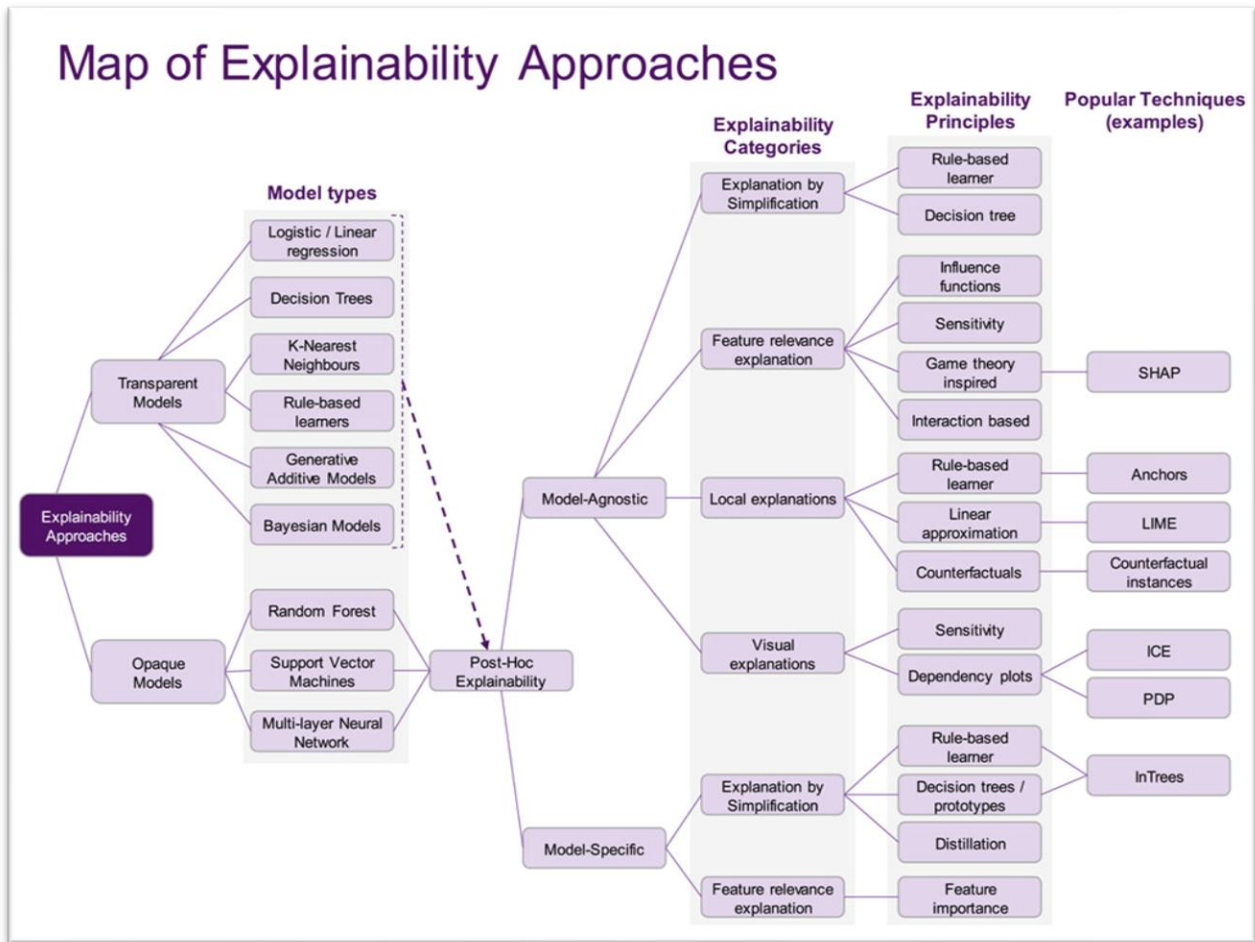
Where: $p(y=1|x)$ represented the model's posterior probability estimate of class 1, given input x , after calibration; $f(x)$ was the raw score or decision value produced by the classifier for input x ; a and b were the parameters learned by the logistic regression model through a fitting that used maximum likelihood estimation from a fitting training set (f_i, y_i) .

The procedure for executing Platt scaling involved setting aside a calibration dataset that is separate from the data used for training the original model to ensure unbiased estimation. Thereafter, a logistic regression model was fitted to the calibration dataset with the original model's

predicted probabilities as the independent variable and the true class labels (0 or 1) as the dependent variable. The calibration parameters a and b representing the intercept and slope of the calibration function were then learned by logistic regression. Finally, calibrated probabilities were generated by applying the calibration function to new data once the parameters had been learned.

4.6.3.4 Explanatory Model Analysis

Explanatory model analysis (EMA) in ML refers to the process of understanding and interpreting the decisions made by a ML model (Belle & Papantonis, 2021; Biecek & Burzykowski, 2020; Linardatos et al., 2021). While many ML algorithms are highly accurate in making predictions, they often operate in an opaque manner making it hard to discern the reasoning behind the model's output. Explanatory model analysis aims to shed light on this opacity by giving us visibility into the model's decision-making process. This involves examining various aspects of the model, such as feature importance, model structure, and decision-making processes. Explainability methods are increasingly championed as pragmatic tools aimed at fostering transparency in ML models (Vale et al., 2022). By shedding light on the inner workings of these models, explainability techniques serve as a vital means to uncover potential biases and other social injustices embedded within predictive algorithms. Through transparency, stakeholders gain insight into how these models arrive at their decisions, enabling them to identify and address instances of discrimination and other harmful outcomes. Embracing explainability in ML not only enhances accountability but also facilitates the development of fairer and more equitable systems. A taxonomy-based explainability approaches are shown in Figure 4.10.



Source: Belle & Papantonis (2021); Reproduced under the Creative Commons Attribution (CC BY) license.

Figure 4.10: Map of Explainability Approaches (Belle & Papantonis, 2021).

In this study, the DALEX package was used to generate feature importance explanations for the best-performing models using a model-agnostic approach to compute SHapley Additive exPlanations (SHAP) values (Biecek et al., 2023). SHAP values were visualized as bar plots, arranged from most to least important. Red and green bars represented negative and positive associations, respectively. An explainer object was created using the explain() function from the DALEX package by passing the trained models and the corresponding datasets to this function. With the explainer object created, SHAP attributions were computed using the predict_parts()

function and specifying the type argument as "shap" to generate SHAP values. The SHAP attributions were then visualized to determine how each feature contributes to the model's predicted values.

4.6.3.5 Business value evaluation

To offer a non-technical explanation of how the model performs and its impact on business value, the features of the modelplotr package were employed (Nagelkerke, 2020). Through this analysis, a suite of informative evaluation plots were generated, including cumulative gains, cumulative lift, response, and cumulative response plots (Nagelkerke, 2018). The cumulative gains plot offers a visual representation of the proportion of target class members captured by selecting up to a specified percentile. Within the modelplotr package, three categories of essential functions streamline the evaluation process: data preparation, plot parameterization, and plotting. The prepare_scores_and_ntiles function constructs a dataframe containing actuals and predictions for each dataset and model combination. Meanwhile, the plot_plotting_scope function formats a dataframe tailored to the selected evaluation scope, ensuring relevance to the analysis. A key feature is the plot_multiplot function, which generates a comprehensive canvas showcasing four essential evaluation plots - cumulative gains, cumulative lift, response, and cumulative response.

4.7 Statistical Analysis

For each diarrheal outcome, patient characteristics were evaluated, comparing children who experienced the outcome against those who did not. This analysis involved reporting proportions for categorical variables and employing either chi-square or Fisher's exact tests accordingly to gauge significance. Continuous variables underwent comparison using Wilcoxon rank sum tests where applicable, ensuring a comprehensive understanding of differences.

The trend of AGE-associated mortality over the study period (2010-2020) was assessed using the Cochran-Armitage trend test. The same statistical test was employed to compare the distribution of both LDD and LGF in the development cohort (VIDA cohort) and the temporal validation cohort (EFGH study). The Cochran-Armitage test is a statistical method utilized to investigate the presence of a linear trend within binary response data (Park et al., 2019; Zhuo et al., 2017).

In the comparative evaluation of the performance of ML models for each diarrheal outcome, initially the logistic regression model was considered as the default choice. However, to rigorously assess the performance differences among multiple algorithms, the Friedman test was employed. This non-parametric test serves as the equivalent of a repeated-measures ANOVA in scenarios where normality assumptions may not hold (Demšar, 2006; Gardner & Brooks, 2017; Rajić, 2026). Following a significant result from the Friedman test ($\alpha < 0.05$), indicating differences in performance among multiple models, the Nemenyi test serves as a post-hoc analysis to pinpoint specific pairwise differences (Gardner & Brooks, 2017; Rajić, 2026). This non-parametric test operates under the null hypothesis of no difference between any two models, contrasted with the alternative hypothesis that at least one pair exhibits a significant difference.

The Friedman test and Nemenyi post-hoc test were implemented using the *PMCMRplus* package. The *friedman.test* function computed the Friedman test while the Nemenyi test was conducted using the *frdAllPairsNemenyiTest* function. The results were then visualized using the *plot* function.

A significance level of 0.05 was used for all statistical tests, with two-sided tests being employed. All analyses, including descriptive statistics, predictive modeling, and visualizations, were conducted using R version 4.2.2 (R Core Team, 2021).

4.8 Product Development and Deployment

4.8.1 Dashboard Design

The R Shiny package (Chang et al., 2023), complemented by *Shinydashboard*, facilitated the creation of an web-accessible dashboard that is interactive without requiring extensive web development. The dashboard comprises two essential components: the user interface file (ui.R) and the Shiny server file. The dashboard development process involved three key steps (Eddington et al., 2022; Ji & Kattan, 2018) :

4.8.1.1 User Interface Design

The user interface file (ui.R) was responsible for creating an interactive layout, allowing users to input data features and view model outputs. Within this file, the Shiny dashboard was organized into a single main tab that consolidated data entry for the three diarrheal outcomes and displayed model results for each outcome. This tab hosts an integrated input fields tailored for all three outcomes under consideration. Input widgets and interactive elements were specified, enhancing user engagement. The *Shinydashboard* package provided a user-friendly framework for developing apps with multiple tabs, streamlining the presentation of inputs, results, and additional information. Within this framework, the Shiny user interface incorporated three key functions: `dashboardSidebar` that serves as the container for `menuItem`s, effectively organizing them into tabs that structure the application's content. The `dashBoardBody` function, which functions as a container for `tabItems`, housing predictors and other elements that will be displayed on a specific tab. Lastly, the `dashBoardPage` function acts as the overarching host for both tabs and `tabItems`, ensuring a cohesive and structured layout for the entire application. The ui.R also displays the model results for each diarrheal outcome, which includes an individual risk gauge reflecting the

risk profile of the patient and an EMA plot elucidating the mechanisms by which the model arrives at its decisions.

4.8.1.2 Server Functionality

The Shiny server file established a server function that utilizes pre-trained models to make predictions based on user inputs. This file loaded pre-trained champion models and the associated explainer objects for each outcome. Importantly, the pre-trained model formulas were not hardcoded into the server.R file; instead, pre-trained models were directly used for predictions. These pre-trained champion models and explainer objects were stored in the same directory as the server.R and ui.R files, simplifying their referencing. Observer events were set up to respond to user interactions, specifically when the "Submit" button is clicked. Upon submission, the relevant model predicted the probability, categorizing the result into low, moderate, or high risk. Visualizations, including risk gauges and SHAP plots, enhanced the degree to which the model's decisions are transparent.

4.8.1.3 Shiny App Integration

The Shiny app framework unified the UI and server logic through the shinyApp function. This step involved connecting the UI and server functions, ultimately executed using runApp. The app allowed users to input data into the designated fields and dynamically displays results upon clicking the "Submit" button. The reactive{} function ensured the application remains dynamic, refreshing its display whenever user inputs change.

4.8.2 Dashboard Deployment

The shinyapps.io provided the most straightforward path for rapid deployment, debugging, and user-testing of the application and was used to deploy the champion models. The *rsconnect*

package (Atkins et al., 2023) was used to link *rsconnect* and the shinyapps.io account allowing for the deployment of the application by setting the working directory using `setwd()` to the directory the server.R and ui.R files are located and running the `deployApp()` function from *rsconnect* in the RStudio console to publish the dashboard. During deployment, RStudio displayed progress messages indicating the status of the deployment process. Once completed, RStudio provided the URL where the dashboard was hosted on shinyapps.io. The dashboard was then tested on shinyapps.io, to ensure it functioned correctly in the cloud environment with various inputs and scenarios tested, to verify that the dashboard behaved as expected and the predictive models generated accurate predictions.

4.9 Reflections on Study Contribution in Machine Learning and Pediatric Enteric Diseases

Objective six and seven are reflections on the empirical findings of the study. These objectives move beyond empirical findings to offer a broader reflection on the study's contributions and future research potential. They pivot from what was found to why it matters.

4.10 Ethical Considerations

Ethical approval for this study was obtained from the University of South Africa's College of Agricultural Sciences research ethics committee (2023/CAES_HREC/2192), ensuring compliance with ethical standards. Moreover, the research utilized de-identified data from previous enteric studies, including the GEMS (scientific and ethics review unit (SERU)#1155), VIDA study (SERU#2996), Rotavirus inpatient surveillance (SERU#1801), and EFGH Shigella surveillance (SERU#4362). The protocols of these studies underwent thorough review and approval by the SERU of KEMRI. Importantly, the initial consents and approvals obtained for each of these studies encompassed the future use of collected data for research purposes, obviating the need for

additional consenting procedures. This approach ensured that ethical considerations were upheld while utilizing existing data resources.

5. CHAPTER FIVE: RESULTS AND DISCUSSION

This section delves into the key findings from the study, which hinged upon the correlation research design delineated in the methodology section. The research traversed through a structured journey comprising descriptive analysis, feature selection, predictive modeling, and an exhaustive explanatory model analysis for top models. The key findings of the study are presented in tables and charts using a blend of descriptive and inferential statistics.

5.1 Predictors of Longer Duration Diarrhea

To fulfill the first study objective to “*Determine the demographic, clinical and socio-economic features that predict Longer Duration Diarrhea among children aged < 60 months.*”, the first research question was formulated as follows:

- **RQ1-** What features predict longer duration diarrhea among children aged less than 60 months?

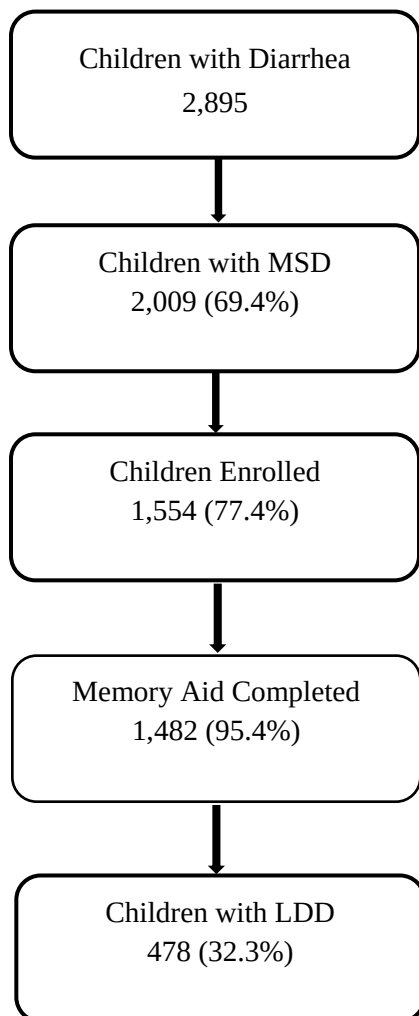
5.1.1 Key Findings for RQ1

To address the first research question (RQ1), the enrolment flowchart is presented in both the development study (VIDA) and the temporal validation study culminating in the estimation of the prevalence of LDD in both studies. Thereafter, the characteristics of children with LDD were compared to those who did not have them, and p-values were computed. This was followed by boruta feature selection, where confirmed and tentative features were used for model development.

The flow of study enrolment for this research question is shown in Figure 5.1. In the VIDA study, a total of 2,895 children under the age of five presented for care due to diarrhea at the study SHCs. Of this number, 1,554 met study criteria and were enrolled, but 72 were lost to follow up, leaving a count of 1482. Of this number 478 (32.3%) met the criteria for LDD. Whereas the EFGH study documented 1,879 instances of children seeking care for diarrhea, of this number only 706 met

enrolment criteria, and 21 were lost to follow up, leaving a count of 685. Of this number 69 (10.1%) had LDD. The prevalence of LDD differed significantly between the development cohort (VIDA) and the temporal validation cohort (EFGH) (32.3% [478/1,482] vs 10.1% [69/685], $p < 0.001$). The substantial variation in LDD prevalence between the VIDA and EFGH studies necessitated further exploration of their baseline characteristics.

Development Data (VIDA)-2015-2018



Temporal Validation Data (EFGH) 2022-2023

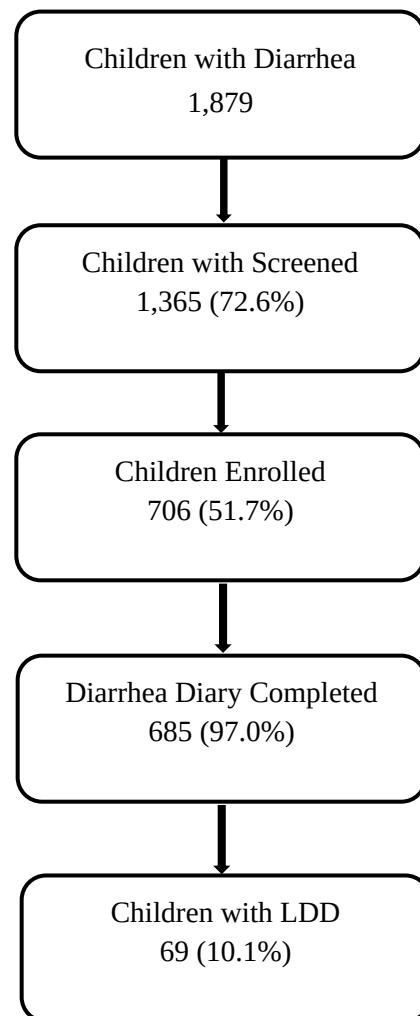


Figure 5.1: Comparative enrolment flow diagram of the development (VIDA) and temporal validation (EFGH) cohorts in longer duration diarrhea prediction.

The comparison of baseline characteristics between the two studies is shown in Table 5.1. Disparities in the participant profiles at enrolment were evident between the two cohorts. Particularly, participants in the VIDA study had a significantly older median age (15.0 months [IQR: 9.0-25.0]) than participants in the EFGH study (13.6 months [IQR: 8.9-20.4], $p = 0.0361$).

Table 5.1: Comparison of baseline characteristics among pediatric patients in development and temporal validation datasets.

Characteristics	VIDA n=1,482 n (%)	EFGH n=685 n (%)	P-value
Median Age in Months [IQR] Age Category	15.0 [9.0-25.0]	13.6 [8.9-20.4]	0.036
0-11 months	557 (37.6)	295 (43.1)	<0.001
12-23 months	508 (34.3)	274 (40.0)	
24-59 months	417 (28.1)	116 (16.9)	
Stool Count			
3	267 (18.0)	162 (23.7)	0.001
4-5	818 (55.2)	324 (47.3)	
≥ 6	397 (26.8)	199 (29.0)	
Median diarrhea days [IQR]	3 [2-4]	3 [2-3]	0.025
Vomiting	843 (56.9)	346 (50.5)	0.006
No. of vomit			
0	639 (43.1)	339 (49.5)	0.001
1	167 (11.3)	95 (13.9)	
2-4	541 (36.5)	208 (30.4)	
≥ 5	135 (9.1)	43 (6.2)	
Median vomit days [IQR]	1 [0-2]	1 [0-2]	0.001
Rotavirus vaccination	1,175 (84.6)	452 (76.5)	<0.001
Skin turgor (slow/very slow)	614 (41.4)	137 (20.0)	<0.001
Median Respiratory rate [IQR]	36.5 [31.5-41.0]	33.0 [28.0-39.0]	<0.001
Vesikari Score			
Mild	145 (9.8)	402 (58.7)	<0.001
Moderate	630 (42.5)	114 (16.6)	
Severe	707 (47.7)	169 (24.7)	
Median Vesikari score [IQR]	10 [8-13]	8 [6-10]	<0.001
Dehydration			
None	82 (5.5)	180 (26.3)	<0.001
Some	1,012 (68.3)	483 (70.5)	

Severe	388 (26.2)	22 (3.2)	

Furthermore, the VIDA participants endured a higher severity of diarrheal illness as indicated by a higher median Vesikari score (10 [IQR: 8-13] compared to 8 [IQR: 6-10], $p < 0.001$) among EFGH participants. They also presented with elevated respiratory rates, with medians of 36.5 [IQR: 31.5-41.0] compared to 33.0 [IQR: 28.0-39.0], $p < 0.001$. Additionally, VIDA participants exhibited a higher prevalence of several symptoms compared to EFGH participants, including vomiting (56.9% versus 50.5%, $p = 0.006$), diminished skin turgor (41.4% versus 20.0%, $p < 0.001$), and severe dehydration (26.2% versus 3.2%, $p < 0.001$).

Table 5.2 presents the comparison of baseline characteristics of participants in the development cohort, according to whether they had LDD or not. Children who experienced LDD were younger compared to those without LDD, with a median age of 13 months [IQR: 7-20] versus 16 months [IQR: 10-27], exhibiting a statistically significant difference ($p < 0.001$). Additionally, participants with LDD demonstrated an elevated respiratory rate (median [IQR]: 37.5 [33-42.5] versus 36 [31-40], $p < 0.001$), a longer duration of diarrhea pre-enrolment (median of 4 days [IQR: 3-5] versus 2 days [IQR: 2-3], $p < 0.001$) and more severe diarrheal illness, as indicated by a higher median Vesikari score (11 [IQR: 9-13] versus 10 [IQR: 8-12], $p < 0.001$) compared to those without LDD.

Additionally, the following factors were positively associated with LDD: partial breastfeeding (289 [60.5%] versus 538 [53.6%], $p < 0.001$), stool frequency in 24 hours (≥ 6 stools) (147 [30.7%] versus 250 [24.9%], $p < 0.001$), belly pain (299 [65.0%] versus 568 [58.7%], $p = 0.024$), rectal straining (155 [32.6%] versus 235 [23.5%], $p < 0.001$), cough (280 [58.6%] versus 526 [52.4%], $p = 0.025$), prior home oral rehydration salts use (58 [12.1%] versus 77 [7.7%], $p = 0.005$), rotavirus vaccination (414 [90.6%] versus 761 [81.7%], $p < 0.001$), fast breathing (63 [13.2%] versus 98 [9.8%], $p = 0.048$) and reduced skin turgor (222 [46.4%] versus 392 [39.0%], $p = 0.007$). On the

contrary, the following variables were negatively associated with LDD: caretaker education ($\geq$ Secondary education) (53 [11.1%] versus 153 [15.3%], $p=0.03$) and frequency of vomiting episodes (≥ 5) (30 [6.3%] versus 105 [10.5%], $p=0.026$).

Table 5.2: Characteristics of participants in the development cohort (VIDA study) stratified by diarrhea duration status, 2015-2018.

Longer Duration Diarrhea (LDD)			
Characteristics	Yes (n=478)	No (n=1,004)	p-value
	n (%)	n (%)	
Demographic			
Median age [IQR]	13 [7-20]	16 [10-27]	<0.001
Age Category			
0-11 months	225 (47.1)	332 (33.1)	<0.001
12-23 months	152 (31.8)	356 (35.5)	
24-59 months	101 (21.1)	316 (31.5)	
Gender: Female	213 (44.6)	463 (46.1)	0.574
Household Details			
Caretaker education ($\geq$ Secondary)	53 (11.1)	153 (15.3)	0.03
≤ 2 children under 60 months	429 (89.8)	912 (90.8)	0.505
≤ 3 Total Assets	429 (89.8)	880 (87.7)	0.239
Natural Floor	323 (67.7)	644 (64.1)	0.177
Refined/Electric Primary Fuel Source	14 (2.9)	44 (4.4)	0.183
Animal ownership	443 (92.7)	924 (92.0)	0.664
Improved water			
Safely managed	212 (44.4)	473 (47.1)	0.393
Basic	64 (13.4)	105 (10.5)	
Limited	65 (13.6)	135 (13.4)	
unimproved/Surface water	137 (28.7)	291 (29.0)	
Improved Sanitation			
Safely Managed and Basic	53 (11.1)	118 (11.8)	0.851
Limited	163 (34.1)	329 (32.8)	
Unimproved/Open Defecation	262 (54.8)	557 (55.5)	
Shared facility	339 (72.8)	693 (72.0)	0.779
Clinical characteristics			
Breastfeeding before diarrhea onset			<0.001
None	147 (30.8)	411 (40.9)	
Partial	289 (60.5)	538 (53.6)	

Exclusive	42 (8.8)	55 (5.5)	
Median pre-enrolment diarrhea days [IQR]	4 [3-5]	2 [2-3]	<0.001
Stool Type			
Simple watery	265 (55.4)	569 (56.7)	0.294
Rice watery	6 (1.3)	23 (2.3)	
Sticky/Mucoid	195 (40.8)	377 (37.6)	
Bloody	12 (2.5)	35 (3.5)	
Stool Count			
3	75 (15.7)	192 (19.1)	0.037
4-5	256 (53.6)	562 (56.0)	
≥ 6	147 (30.7)	250 (24.9)	
Blood in stool	63 (13.2)	115 (11.5)	0.332
Very Thirsty	384 (80.8)	789 (79.5)	0.535
Drinks poorly	120 (25.1)	245 (24.5)	0.785
Unable to drink	13 (2.7)	22 (2.2)	0.531
Belly Pain	299 (65.0)	568 (58.7)	0.024
Fever	372 (77.8)	752 (74.9)	0.219
Restless	372 (77.8)	765 (76.2)	0.488
Lethargy	301 (63.0)	665 (66.2)	0.218
unconscious	12 (2.5)	38 (3.8)	0.202
Rectal straining	155 (32.6)	235 (23.5)	<0.001
Rectal prolapse	6 (1.3)	14 (1.4)	0.827
Cough	280 (58.6)	526 (52.4)	0.025
Difficulty breathing	74 (15.5)	146 (14.6)	0.64
Convulsion	7 (1.5)	15 (1.5)	0.965
Vomiting	255 (53.4)	588 (58.6)	0.058
No. of vomit			
0	223 (46.7)	416 (41.4)	0.026
1	48 (10.0)	119 (11.9)	
2-4	177 (37.0)	364 (36.2)	
≥ 5	30 (6.3)	105 (10.5)	
Median vomit days [IQR]	2 [1-3]	2 [1-2]	<0.001
Currently			
Very Thirsty	350 (74.2)	695 (70.0)	0.1
Drinks poorly	97 (20.5)	192 (19.1)	0.582
Sunken Eyes	412 (86.4)	861 (85.8)	0.75
Wrinkled skin	124 (26.1)	234 (23.6)	0.3
Restless	288 (60.3)	604 (60.2)	0.991
Dry mouth	344 (72.0)	706 (70.4)	0.532
Fast breathing	63 (13.2)	98 (9.8)	0.048
Home ORS use	58 (12.1)	77 (7.7)	0.005
Home Zinc use	21 (4.4)	34 (3.4)	0.338

Rotavirus vaccination	414 (90.6)	761 (81.7)	<0.001
4 B			
Fever (Temp>37.0)	179 (37.5)	396 (39.4)	0.461
Median Respiratory rate [IQR]	37.5 [33-42.5]	36 [31-40]	<0.001
Capillary refill (slow/very slow)	118 (23.0)	231 (24.7)	0.477
Chest indrawing	6 (1.3)	10 (1.0)	0.652
Sunken eyes	440 (92.1)	931 (92.7)	0.643
Dry mouth			
Normal	4 (0.8)	23 (2.3)	0.053
somewhat Dry	441 (92.3)	891 (88.7)	
Very dry	33 (6.9)	90 (9.0)	
Skin turgor (slow/very slow)	222 (46.4)	392 (39.0)	0.007
Mental Status			0.093
Normal	199 (41.6)	437 (43.5)	
Restless/Irritable	266 (55.7)	555 (55.3)	
Lethargic/Unconscious	13 (2.7)	12 (1.2)	
Rectal prolapse	2 (0.4)	1 (0.1)	0.202
Bipedal edema	5 (1.1)	7 (0.7)	0.484
Abnormal hair	20 (4.2)	46 (4.6)	0.729
Under Nutrition	65 (13.6)	109 (13.6)	0.125
Dehydration			
None	25 (5.2)	57 (5.7)	0.669
Some	321 (67.2)	691 (68.8)	
Severe	132 (27.6)	256 (25.5)	
Vesikari Score			<0.001
Mild	15 (3.1)	130 (13.0)	
Moderate	204 (42.7)	426 (42.4)	
Severe	259 (54.2)	448 (44.6)	
Median Vesikari score [IQR]	11 [9-13]	10 [8-12]	<0.001
ORS at facility	475 (99.8)	999 (99.9)	0.591
Zinc at facility	463 (97.3)	963 (96.3)	0.336
IV rehydration	49 (10.3)	115 (11.5)	0.491
Any antibiotic	222 (46.4)	438 (43.6)	0.308
Ciprofloxacin/ceftriaxone	20 (4.2)	62 (6.2)	0.117
Diagnosis			
Malaria	180 (37.8)	399 (39.9)	0.435
Dysentery	34 (7.1)	74 (7.4)	0.855
Stunting			0.394
None	197 (41.5)	445 (44.6)	
Mild	168 (35.4)	315 (31.6)	
Moderate	71 (14.9)	165 (16.5)	
Severe	39 (8.2)	73 (7.3)	
Wasting			

None	369 (77.5)	802 (80.2)	0.233
Mild	85 (17.9)	145 (14.5)	
Moderate/Severe	22 (4.6)	53 (5.3)	

Moreover, the following variables did not have a significant association with LDD ($P \geq 0.05$): gender; number of children aged < 60 months in household; aggregate assets owned by household; refined primary fuel source; owning animals; water quality; sanitation status; joint facility; fecal consistency; dysentery; very thirsty; dry mouth; drinking poorly; not able to drink; reported or clinical fever; restlessness; lethargic; unconsciousness; proclivita; difficulty breathing; convulsions; sunken eyes; mental status; under nutrition, zinc usage at household; evaluation of capillary blood flow; thoracic retractions; bilateral lower extremity edema; atypical hair; being dehydrated; ORS administration at SHCs; zinc administration at SHCs; iv rehydration; prescription of antibiotics; malaria diagnosis; dysentery diagnosis; stunted ; wasted.

The results of the feature selection are shown in Figure 5.2. The Z scores of confirmed features are shown in green boxplots; tentative features are shown in yellow boxplots; rejected features are shown in red boxplots; and shadow features are shown in blue boxplots. Based on the feature selection process conducted for LDD, the confirmed variables, listed in ascending order of importance, included the following: rotavirus vaccination status (6.1%), frequency of vomits in the last 24 hours (6.2%), occurrence of vomiting (6.4%), respiratory rate (6.5%), breastfeeding status (8.4%), duration of vomiting (8.8%), age category (10.7%), modified Vesikari score (18.2%), and diarrhea duration prior to enrolment (55.1%). Additionally, rectal straining (3.4%), poor drinking (2.4%), and stool frequency (2.4%) with an importance of 3.4%, 2.4% and 2.4%, respectively, were tentative features. Features that were identified either as confirmed or tentative were used to build the models. However, because breastfeeding and rectal straining were not

collected in EFGH study, they were excluded from model development to allow for comparison of model performance across the two studies.

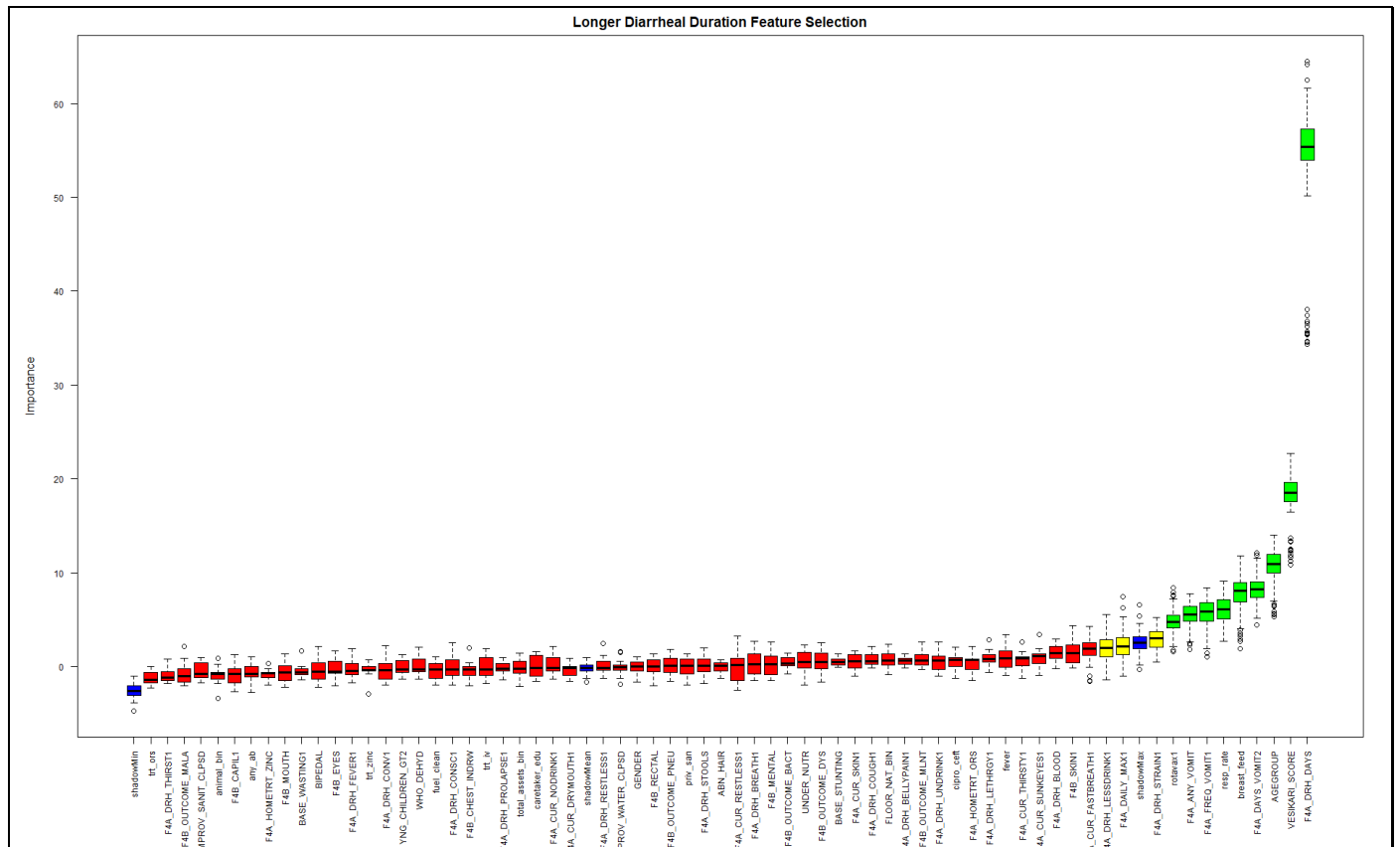


Figure 5.2: Feature selection for longer duration diarrhea in the development cohort (VIDA Study)

5.1.2 Discussion: Predictors of Longer Duration Predictors (RQ1)

The results were presented as counts and proportions and the accompanying chi-square p-values. While the feature selection was presented as color-coded boxplots. Additionally, the key findings for this research question are discussed in the context of existing literature.

The GEMS (Global Enteric Multi-Centre Study) study estimated the prevalence of LDD at 42% (Schilling et al., 2017), the VIDA (Vaccine Impact on Diarrhea in Africa) study recorded a prevalence of 32% of LDD in the development cohort, while the temporal validation cohort extracted from the Enterics for Global Health Shigella Surveillance (EFGH) data recorded a

moderately low prevalence of 10%. This observed difference in prevalence between the two studies was statistically significant. A possible explanation for this difference could be attributed to the baseline differences in severity of diarrhea with the VIDA study enrolling more severe diarrhea cases based on its MSD criteria compared to EFGH study, which enrolled all diarrhea cases. This was confirmed using 11 LDD predictors and is a possible explanation for this difference in LDD prevalence between the two samples. This observed sub-optimal resemblance between the temporal validation and development cohorts carry significant implications for the reproducibility and transportability of models, potentially leading to diminished discrimination and calibration during validation. These implications can manifest in three primary ways: first, heterogeneity in predictor effects, where the same predictor may exhibit varying prognostic value across different settings and populations, thereby influencing model performance (Riley et al., 2016). Second, differences in outcome incidence predominantly result in poor calibration (Cowley et al., 2019; Ramspek et al., 2021). Third, changes in predictor measurement criteria also contribute to poor calibration, thus impacting model performance (Luijken et al., 2020). Recognizing and addressing these disparities is vital for maintaining the robustness and applicability of predictive models across diverse contexts and populations.

From the descriptive analysis, LDD was shown to have significant association with child age, respiratory rate, severity of disease based on Vesikari score, caregiver literacy, being breastfed, frequency of stool, abdominal pain, rectal straining, coughing, frequency of vomiting, prior usage of ORS at household, rotavirus vaccination, fast breathing and reduced skin turgor. But upon further feature selection analysis using boruta, only the following features, that were either confirmed or tentative, were used in model development: diarrhea duration prior to enrolment, modified Vesikari score, age category, duration of vomiting, respiratory rate, emesis, frequency of

vomiting, rotavirus vaccination status, rectal straining, poor drinking, and stool frequency. Despite breastfeeding status and rectal straining emerging as predictors of LDD, they were not used in model development as they were not collected in the temporal validation cohort. The feature selection methodology did not show sensitivity to socio-economic features, it rather showed sensitivity to clinical symptoms.

This finding from the feature selection, that uses a random forest wrapper algorithm to compute feature importance, resonates closely with results obtained from epidemiological studies that have been conducted using standard statistical approaches. This alignment accentuates the efficacy and validity of employing ML approaches in uncovering risk factors for diseases. It emphasizes the complementary nature of ML techniques alongside conventional statistical methodologies in identifying and understanding disease determinants, thereby enhancing the ability to address health challenges more comprehensively and effectively.

In the evaluation of potential predictors for LDD, spanning demographic, clinical, and socio-economic variables, clinical predictors emerged as particularly influential. Out of the 11 predictors identified, clinical factors constituted a substantial majority, encompassing 10 of them. This dominance accentuates the pivotal role of clinical variables in forecasting LDD occurrences, highlighting their importance in predictive modeling efforts. The type of modeling and feature selection methodology that was followed in the study displayed a dominance of clinical variables rather than socio-economic variables. Perhaps because the study was handling acute episodes for an individual child, rather than population level modeling, the clinical characteristics have thus dominated. However, a conclusion can't be made about the role of the socio-economic variables from the results of this study. The matter of the role and influence of socio-economic variables on

the model has been designated as a future research direction. From the results of the study the model has shown good performance despite not including many socio-economic variables.

Expositing the identified predictors, the severity of diarrheal illness emerged as a significant driver of LDD based on the findings. Both the aggregated severity, as measured by the modified Vesikari score, and individual components of the severity score—including episode duration prior to enrolment, emesis, vomit frequency within 24 hours, stool frequency, and decreased skin turgor—emerged as predictors of LDD. The results are congruent with past studies that have shown episodes of longer duration to exhibit greater severity upon presentation (Alam et al., 1997; Giannattasio et al., 2016; Lima & Guerrant, 1992b; Patel et al., 2012; Strand et al., 2012). This severity may induce intestinal inflammation, prolonging the duration of illness and recovery time as has been shown by existing literature (Lo Vecchio et al., 2021). Moreover, the pronounced diarrhea severity, characterized by frequent vomiting and increased stool frequency, can result in substantial fluid and electrolyte loss, exacerbating the illness and extending its duration as the body requires additional time to achieve fluid repletion and normalization of electrolyte levels. Additionally, an assertive immune response to diarrheal etiological agents may contribute to an escalation of symptom severity and a protracted period of illness.

Moreover, this study showed that younger children had a heightened susceptibility to LDD, a result that corroborates findings from earlier research (Giannattasio et al., 2016; Hossain et al., 2022; Moore et al., 2010; Ochoa et al., 2004; Patel et al., 2012; Strand et al., 2012). This observation could be attributed to their immature immune system, which are still developing, making them more susceptible to microbial infections and experiencing more severe symptoms than their older counterparts (PrabhuDas et al., 2011; Simon et al., 2015; Zhang et al., 2017). Their bodies consequently take longer to clear the infection and recover. Existing research in epidemiology and

clinical studies has revealed that protective immunity specific to certain strains can develop following enteric infections (Lockhart et al., 2022). However, there remains conflicting evidence regarding the degree and duration of this protective effect. Moreover, the gut microbiota, which undergoes age-related variations, is essential for the gut's early formation and maturation as well as in the differentiation of the immune system (Li et al., 2021; Saeed et al., 2022; Zheng et al., 2020). This symbiotic relationship between the microbiome and the host's immune system is pivotal. The first 1000 days of life are particularly influential, during which microbiota changes have profound effects. A healthy gut microbiota serves as a barrier, restraining the proliferation of pathogenic and potentially harmful microbes while impeding the colonization of enteropathogens. However, an imbalance in gut microbiota or dysbiosis, especially in younger children, can lead to diarrhea, and therefore it takes longer to restore a healthy gut flora thereby prolonging diarrheal duration. In addition, younger children are more susceptible to fluid loss through insensible means, due to their higher metabolic requirements and larger body surface area relative to their size that makes them more vulnerable to dehydration (D'Anci et al., 2006; Vega & Avva, 2024), which can worsen and prolong diarrheal episodes.

This study also showed that not being vaccinated against rotavirus increased the likelihood of LDD. Research has demonstrated that receiving the rotavirus vaccination significantly lowers the likelihood of experiencing rotavirus diarrhea, and the extent of this risk reduction becomes more pronounced as the severity of the illness increases (Khagayi et al., 2020; Patel et al., 2009). In populations where vaccination coverage is low or nonexistent, children are more susceptible to rotavirus infection, which can lead to severe diarrhea, vomiting, fever, and dehydration, particularly in young children. Without proper management, the symptoms can persist and worsen, leading to prolonged diarrhea and potentially life-threatening complications. While some studies

have highlighted the concept of herd or indirect immunity stemming from the rotavirus vaccine (Bennett et al., 2018; Otieno et al., 2020), various caretaker characteristics have been linked to the absence of vaccination. These include: insufficient knowledge; misconceptions; financial constraints; lack of support from partners; and lack of trust in healthcare systems (Bangura et al., 2020). Additionally, factors such as the number of children, lifestyle choices, parental forgetfulness, single motherhood, language barriers, and feelings of shame related to poverty have been associated with the decision not to vaccinate. These factors not only elevate the likelihood of a child contracting diarrhea but also signal suboptimal healthcare utilization and inadequate home management by caretakers. Such circumstances may lead to prolonged duration of diarrheal episodes.

As indicated above, future studies could further explore the role of socio-economic factors by expanding the range of variables under investigation and examining how health policies—such as universal access to care—may modify the association between these factors and diarrheal outcomes. Additionally, incorporating qualitative approaches could offer deeper insight into household-level coping strategies and healthcare-seeking behaviors, helping to uncover the nuances surrounding socio-economic influences on child health. Moreover, future research should aim to incorporate breastfeeding and rectal straining data. While identified as predictors, their absence in the temporal validation dataset necessitated their exclusion from the final model, and incorporating them holds potential for improved model performance in future studies.

5.2 Predictors of Chronic Malnutrition

To fulfill the second study objective of “*Determine the demographic, clinical and socio-economic features that predict chronic malnutrition among children aged < 60 months.*”, the second research question was formulated as follows:

- **RQ2-** What features predict chronic malnutrition following a diarrheal illness among children aged less than 60 months?

5.2.1 Key findings for RQ2

The second research question (RQ2) sought to establish predictors of chronic malnutrition following a diarrheal illness among children aged < 5 years. Specifically, linear growth faltering a precursor of stunting was used as the outcome variable to assess chronic malnutrition. The same methodology applied to RQ1 was also used for RQ2. The enrolment flowchart in both the development study (VIDA) and the temporal validation study culminating in the estimation of the prevalence of linear growth faltering (LGF) in both studies. Thereafter, the characteristics of children with LGF were compared to those who did not have them, and p-values were computed. This was followed by boruta feature selection, where confirmed and tentative features were used for model development.

The CONSORT diagram for this research question is shown in Figure 5.3. The development cohort had 1,554 recruited children, of whom 1,106 (71.2%) were aged 6-35 months and had plausible HAZ data, while the temporal validation cohort enrolled 706 children, of whom 655 (92.7%) had HAZ data deemed plausible. Among those with plausible HAZ data, 16.9% (n=187) and 22.4% (n=147) exhibited LGF in the development cohort and temporal validation cohorts, respectively. The variance in LGF prevalence between the two studies exhibited statistical significance

($p=0.0042$). Both the development and temporal validation cohorts experienced a median decrease in height-for-age Z-score (Δ HAZ) from enrollment to follow-up. The development cohort's median change was -0.21 (IQR: -0.42 to -0.01), while the temporal validation cohort's was -0.24 (IQR: -0.48 to -0.02). In the development cohort, the prevalence of LGF was 1,051 (28.7%) when using the threshold for negative Δ HAZ in the sensitivity analysis.

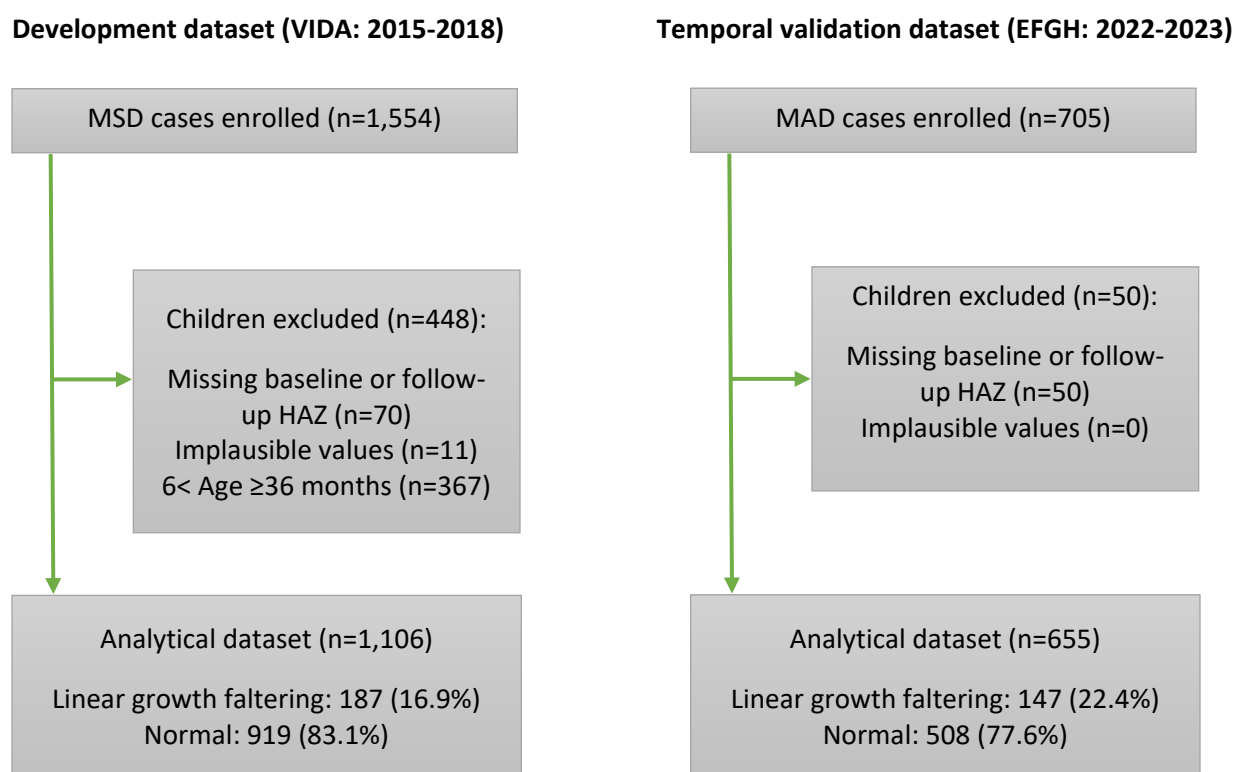


Figure 5.3: CONSORT diagram of linear growth faltering in development (VIDA) and temporal validation (EFGH) cohorts

Table 5.3 illustrates the demographic, clinical and socio-economic profile of participants enrolled in VIDA, categorized by whether they developed LGF or not. Children exhibiting LGF were younger (Median: 11 months [IQR: 8-14]) compared to their counterparts who did not (Median: 17 months [IQR: 11-24], $p < 0.001$). Moreover, those with LGF displayed elevated respiratory rates

(median: 38.5 breaths/min [IQR: 34.0-42.5] versus 36.0 breaths/min [IQR: 31.5-39.5], $p < 0.001$), higher temperatures (median: 37.1°C [IQR: 36.6-37.8] vs. 36.8°C [IQR: 36.4-37.5], $p < 0.001$), and heightened severity as indicated by Vesikari scores (Median [IQR]: 11 [9-12] vs 10 [8-12], $p < 0.001$), compared to their counterparts who did not. Furthermore, these factors were associated with and increased likelihood of LGF: caregiver education level (78 [41.7%] vs 305 [33.2%], $p=0.026$), partial breastfeeding (159 [85.0%] vs 563 [61.2%], $p<0.001$), vomiting (127 [67.9%] vs 531 [57.8%], $p=0.01$), skin turgor (56 [30.6%] vs 211 [23.0%], $p=0.029$), restlessness (134 [71.7%] vs 557 [60.6%], $p=0.004$), hospital admission (27 [14.4%] vs 87 [9.5%], $p=0.043$), and intravenous rehydration (31 [16.6%] vs 92 [10.1%], $p=0.01$).

Table 5.3: Profile of children enrolled in the development study (VIDA: 2015-2018), grouped by linear growth faltering.

Linear Growth Faltering			
Characteristics	Yes (n=187)	No (n=919)	p-value*
	n (%)	n (%)	
Demographic			
Median age [IQR]	11 [8-14]	17 [11-24]	<0.001
Age Category			
0-11 months	104 (55.6)	259 (28.2)	<0.001
12-23 months	74 (39.6)	428 (46.6)	
24-59 months	9 (4.8)	232 (25.2)	
Gender: Female	83 (44.4)	428 (46.6)	0.584
Household Details			
Caretaker education ($\geq$ Secondary)	78 (41.7)	305 (33.2)	0.026
≤ 2 children under 5 yrs	167 (89.3)	839 (91.2)	0.387
≤ 4 people sleeping	77 (41.2)	400 (43.6)	0.546
≤ 3 Total Assets	158 (84.5)	812 (88.4)	0.142
Refined/Electric Primary Fuel Source	5 (2.7)	39 (4.3)	0.313
Animal ownership	176 (94.1)	836 (91.0)	0.159
Improved water			
Safely managed	83 (44.3)	431 (46.9)	0.15
Basic	14 (7.5)	112 (12.2)	
Limited	28 (15.0)	125 (13.6)	
unimproved/Surface water	62 (33.2)	251 (27.3)	

Improved Sanitation			
Safely Managed and Basic	20 (10.7)	106 (11.5)	0.392
Limited	72 (38.5)	306 (33.3)	
Unimproved/Open Defecation	95 (50.8)	507 (55.2)	
Clinical characteristics			
Reported by caretaker			
Breastfeeding before diarrhea onset			
None	25 (13.4)	248 (37.9)	<0.001
Exclusive	3 (1.6)	8 (0.9)	
Partial	159 (85.0)	563 (61.2)	
Median diarrhea days [IQR]	3 [2-3]	3 [2-4]	0.7196
Stool Type			
Simple watery	113 (60.4)	532 (57.9)	0.41
Rice watery	5 (2.7)	12 (1.3)	
Sticky/Mucoid	65 (34.8)	347 (37.8)	
Bloody	4 (2.1)	28 (3.1)	
Stool Count			
3	27 (14.4)	165 (18.0)	0.489
4-5	101 (54.0)	506 (55.0)	
6-10	55 (29.4)	228 (24.8)	
> 10	4 (2.1)	20 (2.2)	
Blood in stool	15 (8.0)	108 (11.8)	0.138
Vomiting	127(67.9)	531 (57.8)	0.01
Very Thirsty	156 (83.9)	752 (82.2)	0.582
Drinks poorly	47 (25.1)	232 (25.3)	0.962
Unable to drink	2 (1.1)	27 (2.9)	0.145
Belly Pain	109 (61.2)	508 (57.9)	0.404
Fever	142 (75.9)	709 (77.2)	0.720
Restless	151 (80.8)	710 (77.3)	0.295
Lethargy	123 (65.8)	600 (65.3)	0.898
unconscious	7 (3.7)	32 (3.5)	0.864
Rectal straining	55 (29.4)	211 (23.1)	0.066
Rectal prolapse	2 (1.1)	15 (1.6)	0.565
Cough	103 (55.1)	482 (52.5)	0.511
Difficulty breathing	32 (17.1)	124 (13.5)	0.197
Convulsion	3 (1.6)	17 (1.9)	0.818
Currently			
Very Thirsty	145 (78.4)	653 (71.7)	0.062
Drinks poorly	40 (21.5)	188 (20.5)	0.747
Sunken Eyes	171 (91.4)	792 (86.3)	0.054
Wrinkled skin	56 (30.6)	211 (23.0)	0.029
Restless	134 (71.7)	557 (60.6)	0.004
Lethargy/unconscious	23 (12.3)	151 (16.4)	0.157

Dry mouth	142 (75.9)	658 (71.7)	0.235
Fast breathing	24 (12.8)	100 (10.9)	0.44
Home ORS use	21 (11.2)	86 (9.4)	0.43
Home Zinc use	8 (4.3)	34 (3.7)	0.706
Assessed by Clinician			
Temperature [IQR]	37.1 [36.6-37.9]	36.8 [36.4-37.5]	<0.001
Measured Fever ($\geq 37.5^{\circ}\text{C}$)	99 (52.9)	342 (37.2)	<0.001
Median Respiratory rate [IQR]	38.5 [34.0-42.5]	36.0 [31.5-39.5]	<0.001
Chest indrawing	4 (2.1)	9 (1.0)	0.180
Sunken eyes	177 (94.7)	848 (92.3)	0.255
Dry mouth	183 (97.9)	903 (98.3)	0.71
Skin turgor (slow/very slow)	78 (41.7)	391 (42.6)	0.833
Mental Status			
Normal	73 (39.0)	380 (41.4)	0.052
Restless/Irritable	108 (57.8)	530 (57.7)	
Lethargic/Unconscious	6 (3.2)	9 (0.9)	
Rectal prolapse	0 (0)	3 (0.3)	0.434
Bipedal edema	2 (1.1)	5 (0.5)	0.337
Abnormal hair	9 (4.8)	43 (4.7)	0.937
Under nutrition	21 (11.2)	109 (11.9)	0.807
Flaky skin	2 (1.1)	5 (0.5)	0.409
Severe Acute Malnutrition (SAM)	24 (12.8)	78 (8.5)	0.061
Wasting	13 (7.0)	39 (4.2)	0.111
Admission	27 (14.4)	87 (9.5)	0.043
Diarrhea duration (≥ 7 days)	70 (37.4)	327 (35.6)	0.631
any_antibiotic	78 (41.7)	402 (43.7)	0.609
Rotavirus vaccination doses			
0	2 (1.1)	19 (2.4)	0.385
1	8 (4.6)	25 (3.1)	
2	166 (94.3)	764 (94.5)	
ORS at facility	186 (99.5)	914 (99.9)	0.311
Zinc at facility	183 (97.9)	887 (96.9)	0.494
IV rehydration	31 (16.6)	92 (10.1)	0.01
Dehydration			
None	8 (4.3)	35 (3.8)	0.747
Some	126 (67.4)	645 (70.2)	
Severe	53 (28.3)	239 (26.0)	
Vesikari Score			
Mild	13 (7.0)	71 (7.7)	0.088
Moderate	75 (40.1)	442 (48.1)	
Severe	99 (52.9)	406 (44.2)	
Median Vesikari score [IQR]	11 [9-12]	10 [8-12]	0.0003
Diagnosis			

Dysentery	10 (5.4)	58 (6.4)	0.605
Malaria	85 (45.5)	361 (39.5)	0.131
Pneumonia	12 (6.4)	37 (4.1)	0.152
Bacterial Infection	14 (7.5)	93 (10.2)	0.258
Malnutrition	15 (8.0)	68 (7.4)	0.784

Moreover, the following variables exhibited no statistically significant association with LGF at the univariate level ($P \geq 0.05$): gender, number of children aged less than 60 months in the household, number of people sleeping in the household, total assets, refined primary fuel source, animal ownership, access to improved water and sanitation facilities, number of days with diarrhea, stool characteristics (type and frequency), presence of blood in stool, symptoms such as very thirsty, drinking poorly, unable to drink, belly pain, lethargy, unconsciousness, rectal straining, coughing, difficulty breathing, convulsions, sunken eyes, dry mouth, fast breathing, utilization of ORS or zinc at home or in medical facilities, chest indrawing, rectal prolapse, bipedal edema, abnormal hair, undernutrition indicators (such as flaky skin, severe acute malnutrition, and wasting), antibiotic usage, rotavirus vaccination status, presence of dehydration, dysentery, malaria, pneumonia, and bacterial infections.

Moreover, the synthetic dataset consisted of 8,894 observations and had a propensity score distribution that closely approximated that of the original development data (VIDA). This alignment was verified through an extensive descriptive analysis, which meticulously compared each variable. (Table 5.4).

Table 5.4: Statistical comparison of synthetic data (n = 8,894) and original training data (n=1,106).

Synthetic Data		
n=8,894		
Variables	pMSE	S_pMSE
lgf	0.0000010	0.820503
age	0.0000030	0.743468

temperature	0.0000050	1.228149
resp_rate	0.0000040	0.947305
SAM	0.0000020	1.659631
Rotavirus_vacc	0.0000020	1.131343
cur_wrinkledskin	0.0000010	0.779434

pMSE- propensity score mean-squared-error; S_pMSE- standardized ratio of propensity score mean-squared error.

Figure 5.4 presents a visualization of the comparison between the distribution of the variables in the development data and the synthetic data, affirming their alignment.

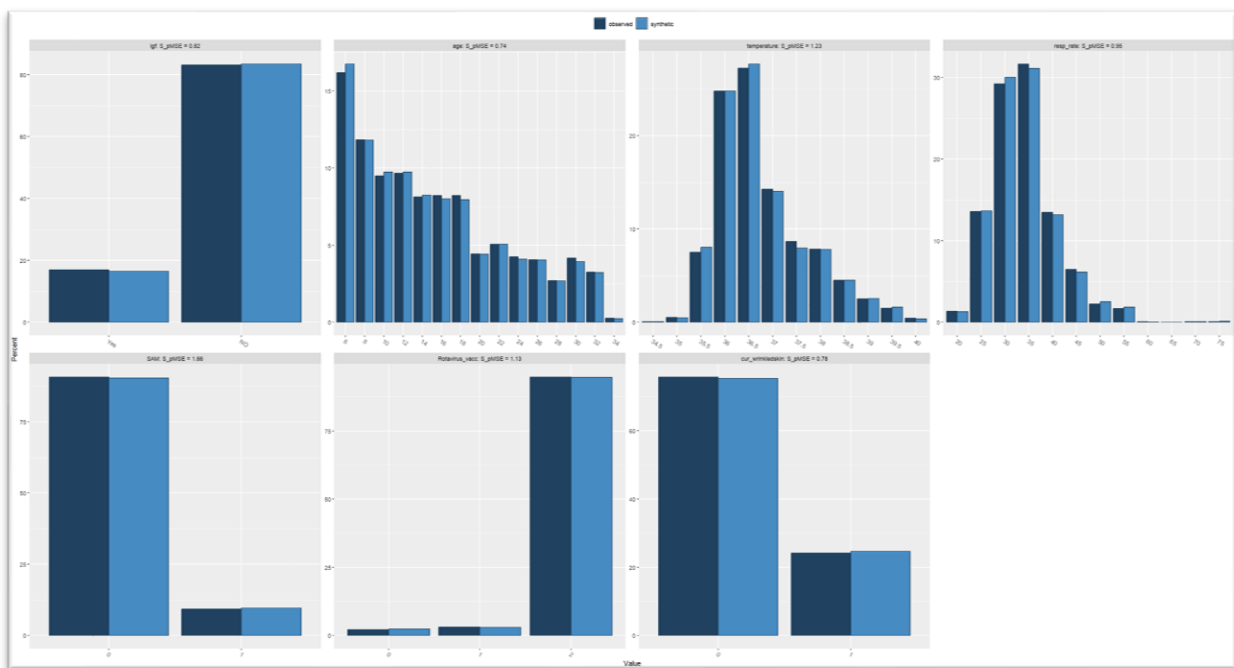
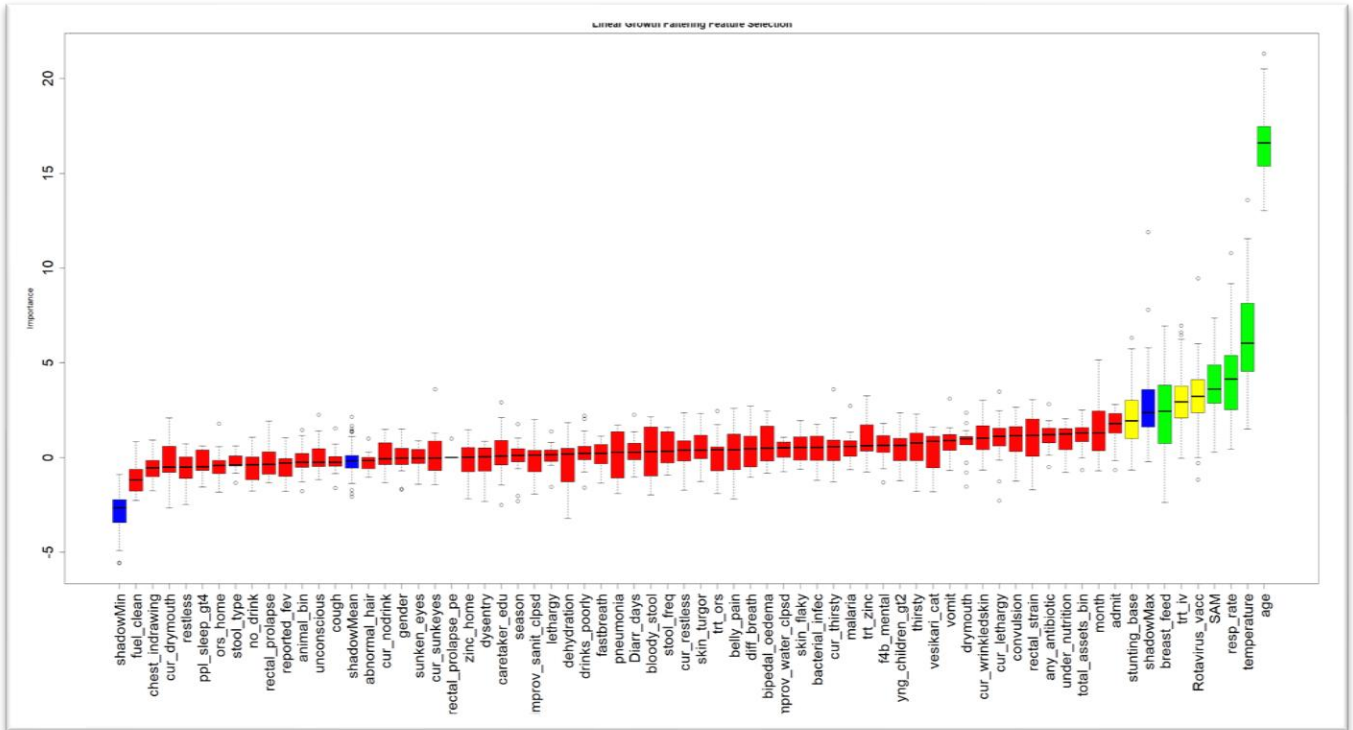


Figure 5.4: Histogram comparison of predictor variable distributions between original (n=1,106) and synthetic (n=8,894) data.

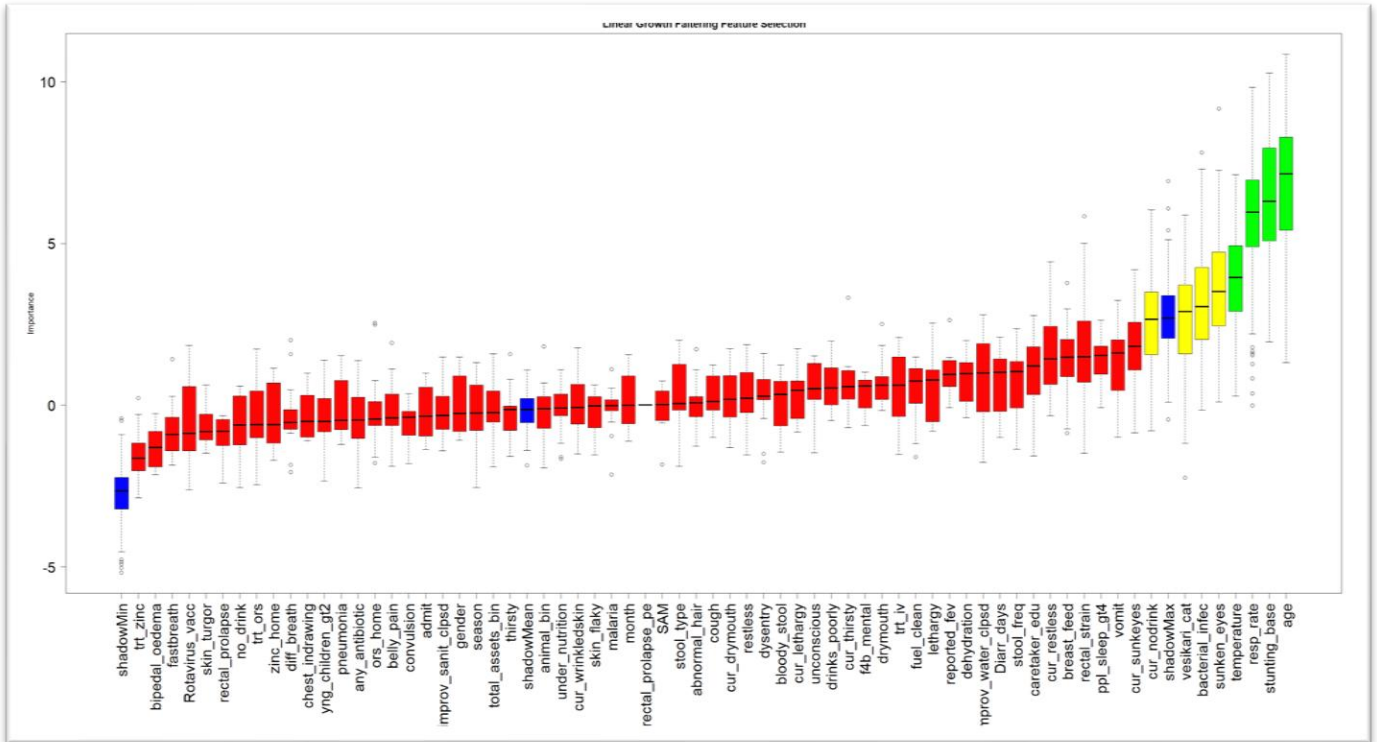
The feature selection process identified the following variables as confirmed predictors; age of child, temperature, respiratory rate and breastfeeding, having an importance of 16.6%, 6.0%, 4.1%, and 3.3%, respectively. Additionally, SAM, rotavirus vaccination, and skin turgor with importance of 3.4%, 3.3% and 2.1%, respectively, were identified as tentative features, as depicted in Figure 5.5.



Green, yellow, red and blue boxplots represent the Z scores of confirmed, tentative, rejected and shadow features, respectively.

Figure 5.5: Predictor identification for linear growth faltering ($\Delta \text{HAZ} \geq -0.5$) in development cohort (VIDA study).

The rejected features were excluded from the model building process. Additionally, breastfeeding, although a confirmed predictor was excluded, as it was not collected in the temporal validation study. In the sensitivity analysis, utilizing a negative ΔHAZ as a cut-off, several features emerged as predictors, in addition to age, respiratory rate, temperature, and breastfeeding already identified in the primary analysis, as illustrated in Figure 5.6. Confirmed predictors included stunting at baseline (5.2%), vomiting (4.0%), Vesikari score (3.7%), and sunken eyes (3.6%). Additionally, a diagnosis of bacterial infection (2.5%) was identified as a tentative predictor.



Green, yellow, red and blue boxplots represent the Z scores of confirmed, tentative, rejected and shadow features, respectively.

Figure 5.6: Predictor identification for linear growth faltering (negative Δ HAZ) in development cohort (VIDA study)

5.2.2 Discussion: Chronic Malnutrition Predictors (RQ2)

A statistically significant difference ($p=0.0042$) in the prevalence of LGF was observed between the development cohort (16.9%) and the temporal validation cohort (22.4%). A wide gap in Δ HAZ was observed between enrollment and follow-up in the temporal validation cohort compared to the development cohort: median [interquartile range] (-0.21 [-0.42- -0.01] vs -0.24 [-0.48- -0.02]). While both studies had short-term follow-up duration spanning 60-90 days and 84-127 days for the development cohort and the temporal validation cohort, respectively. The duration of follow-up is a critical factor in assessing the prevalence of stunting, as it influences the ability to capture changes in growth over time and the effectiveness of interventions aimed at addressing nutritional

deficiencies and other contributing factors. This observation merits investigation in a future research study.

From the descriptive analysis, LGF was shown to have significant association with child age, respiratory rate, temperature, Vesikari score, caregiver education level, breastfeeding, vomiting, skin turgor, restlessness, hospital admission and intravenous rehydration. But upon further feature selection analysis using boruta only the following features, which were either confirmed or tentative, were used in model development: age, temperature, respiratory rate, SAM, rotavirus vaccination and skin turgor. Despite breastfeeding status emerging as a predictor of LGF, it was not used in model development as it was not collected in the temporal validation cohort.

Although the introduction of the rotavirus vaccine has significantly influenced the epidemiology of diarrhea and the pathogen landscape, this research has revealed consistent predictors akin to those identified in prior modeling studies (Ahmed et al., 2023; Brander et al., 2019) conducted before the vaccine's introduction. These predictors include: age; breastfeeding; respiratory rate; temperature; severe acute malnutrition; and being stunted at enrolment.

The WHO conceptual framework on childhood stunting comprehensively addresses various factors including pediatric, mother, household, and ecological characteristics. The factors from this framework have been corroborated by evidence from other systematic reviews done globally (Vilcins et al., 2018) and in LMICS (Danaei et al., 2016). However, this study identified six predictors, which were all related to child characteristics. Among these predictors, five were clinical indicators (such as temperature, respiratory rate, SAM, rotavirus vaccination status, and skin turgor), while one was associated with demographic characteristics (age). Nevertheless,

further research is warranted to explore how wider socio-economic and environmental factors contribute to growth faltering, particularly over longer periods or in recurrent illness contexts.

Child age emerged as the predictor of LGF that had the highest importance. This finding is consistent with findings from past studies conducted in diverse settings in Asia (Beal et al., 2018; Blankenship et al., 2020; Chowdhury et al., 2022; Fenske et al., 2013; Ikeda et al., 2013) and Africa (Adekanmbi et al., 2013; Oluoko, 2021; Taguri et al., 2009; Takele et al., 2019). Moreover, a comprehensive study examining childhood stunting across 94 LMICs revealed that while stunting prevalence varied between different world regions, a consistent pattern emerged: the rate of stunting increased with a child's age, reaching a peak at 28 months (Karlsson et al., 2023). Following this peak, the prevalence either stabilized or declined as children grew older. This pattern suggests that prolonged exposure to malnutrition and infections during the critical first twenty-four months of life contributes significantly to being stunted. Infants and toddlers are particularly susceptible to nutritional and health issues because their bodies are still maturing physiologically. This vulnerability is amplified during episodes of diarrhea, which can significantly hinder their ability to absorb and utilize nutrients, often resulting in a higher prevalence of LGF among this group. Stunting, a key indicator of chronic malnutrition, has been documented to become largely irreversible once a child reaches 24 months of age (Victora et al., 2010). Therefore, the swift identification of vulnerable children (<24 months) empowers the use of specific preventive measures during this vital period of early childhood.

Furthermore, the developing immune system plays a crucial role in a child's susceptibility to stunting. As children grow, their immune system develops and becomes more robust, providing better protection against pathogens and reducing the incidence of infections (Maggini et al., 2018; Ygberg & Nilsson, 2012; Zhang et al., 2017). This is clearly demonstrated in the case of diarrheal

disease, a major etiology of stunting. Research conducted across various LMICs reveals a striking pattern: diarrheal episodes peak in frequency during infancy, with an average of 4.8 episodes between 6-11 months of age. As children mature, this number steadily declines, dropping to 2.6 episodes at age 2 and further diminishing to 1.4 episodes by age 4 (Karlsson et al., 2023). This decline in diarrheal frequency aligns with the observed trend of stunting prevalence stabilizing or decreasing after 28 months of age, highlighting the crucial link between LGF and the maturity of the immune system, which increases the ability of older children to fend off etiological agents of infectious disease.

Elevated baseline temperature and respiratory rate serve as physiological indicators of the metabolic and inflammatory response to diarrhea episodes. Furthermore, they are key indicators of disease severity, particularly for gastrointestinal infections like diarrhea, which can impair nutrient absorption due to intestinal damage, reduced appetite, and nutrient loss, leading to nutrient deprivation that impedes linear growth (Ahmed et al., 2023; Brander et al., 2019). Additionally, higher respiratory rates and temperatures often reflect heightened metabolic activity as the body combats infection and inflammation. Up to 50% of the total energy expenditure in mammals is dedicated to regulating a stable core temperature, underscoring the significance of this physiological process (Tusting et al., 2020). Moreover, a mere 1°C elevation in body temperature corresponds to up to 13% rise in metabolic rate. This surge in energy demand may divert energy and nutrients away from growth-related processes thereby compromising a child's linear growth potential. The associated metabolic stress, nutrient loss, inflammation, and reduced appetite can collectively contribute to LGF in children following diarrheal episodes.

The findings also showed an inverse relationship between rotavirus vaccination and LGF, aligning with the findings of a study assessing the effect of rotavirus vaccination on HAZ among children

aged 6-59 months in Peru (Loli & Carcamo, 2020). This correlation suggests that the significant reduction in rotavirus infections attributable to vaccination may mitigate the short-term effects of diarrhea on nutrient assimilation, thereby attenuating diarrhea-induced growth stunting (Loli & Carcamo, 2020). Strengthening rotavirus vaccination efforts emerges as a viable strategy for policymakers and public health professionals to combat stunting in similar settings. In addition, various caretaker characteristics have been associated with the absence of vaccination, encompassing factors such as insufficient knowledge, misconceptions, financial constraints, lack of partner support, and distrust in healthcare systems (Bangura et al., 2020). Moreover, decisions against vaccination can be influenced by additional factors like family size, lifestyle choices, parental forgetfulness, single motherhood, language barriers, and the stigma of poverty. Furthermore, the lack of access to rotavirus vaccination serves as a telling indicator of broader healthcare disparities and limited access to essential services, particularly prevalent in underserved communities and low-resource settings (Dimitrova et al., 2023; Wondimu et al., 2020). Children in these environments face barriers to vaccination due to financial limitations, deficient healthcare infrastructure, and insufficient awareness or education about vaccination's importance. These factors not only heighten the risk of childhood diarrhea but also suggest suboptimal healthcare utilization and inadequate management of diarrhea thereby exacerbating the vulnerability of these children to stunting.

In contrast to prevailing evidence (Garenne et al., 2018; Ngari et al., 2019), the study findings revealed an unexpected trend: children without SAM exhibited a heightened susceptibility to LGF. This result needs further investigation, as there could be other factors at play such as health status, genetic factors, and environmental influences. In addition to the predictors discussed above, the scenario analysis, using a cut-off of negative Δ HAZ, identified additional predictors: stunting at

baseline, vomit, Vesikari score and sunken eyes and clinical diagnosis of bacterial infection. These additional predictors are indicative of early stunting or severe illness and they have been documented as predictors of stunting in previous studies (Ahmed et al., 2023; Akombi et al., 2017; Brander et al., 2019; Nasrin et al., 2023). Changing the cut-off criteria may have led to a broader inclusion of cases or a more sensitive detection of certain predictors. The new cut-off might have been more sensitive to these short-term, severe illnesses that can contribute to a decline in HAZ.

5.3 Predictors of Diarrhea-associated Mortality

To fulfill the third study objective of “*Determine the demographic, clinical and socio-economic features that predict diarrhea-associated mortality among children aged < 60 months.*”, the third research question was formulated as follows:

- **RQ3-** What features predict diarrhea-associated mortality among inpatient children aged less than 60 months?

5.3.1 Key Findings for RQ3

The third research question (RQ3) sought to establish predictors of mortality among children aged < 60 months hospitalized with acute gastroenteritis. Specifically, in-hospital mortality was used as the outcome variable to assess this outcome. The enrolment flowchart of the study culminating in the estimation of the prevalence of mortality. Thereafter, the characteristics of children who died were compared to those who survived, and p-values were computed. This was followed by boruta feature selection, where confirmed and tentative features were used for model development.

The study flowchart is shown in Figure 5.7. A total of 12,546 pediatric patients less than five years of age admitted to SCRH were recruited in the inpatient disease surveillance.

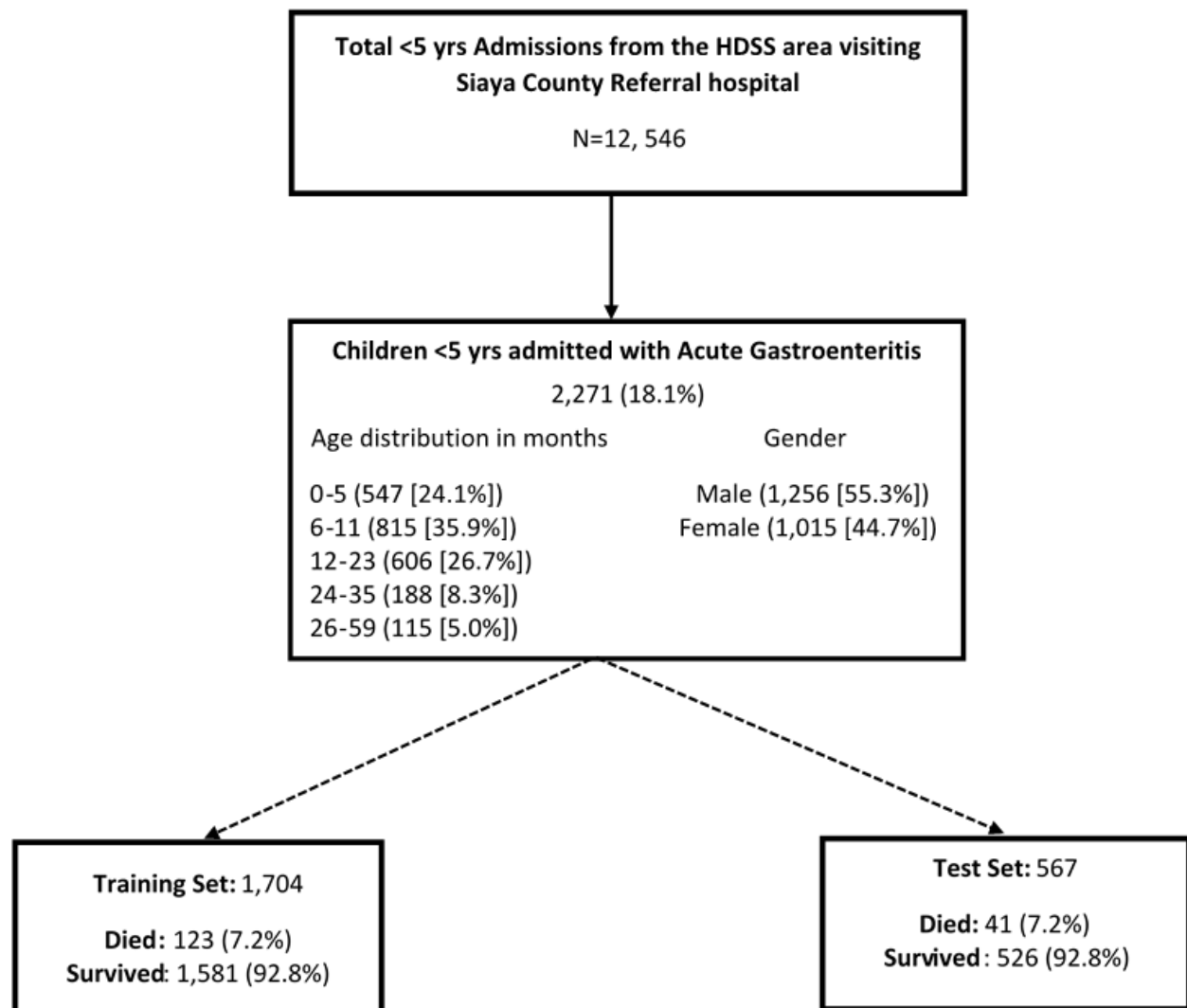


Figure 5.7: CONSORT diagram of gastroenteritis-associated deaths among children under 60 months admitted at Siaya county referral hospital, Western Kenya (2010–2020) and data segregation into training and test sets.

Specifically, a total of 2,271 children, representing 18.1% of those recruited in the inpatient disease surveillance, met the AGE case definition. Approximately 7.2% of the acute gastroenteritis cohort, totaling 164 children, died during their hospitalization. The time interval between admission and

death was a median of 2 days (IQR: 1-4 days). The median age of children diagnosed with acute gastroenteritis was 9.9 months (IQR: 6.1-17.4 months), with 1,256 cases (55.3%) being male.

The prevalence of mortality among acute gastroenteritis cases exhibited no significant variation across the years ($p = 0.3142$). The highest mortality rate was recorded in 2013, with 24 out of 167 cases (12.6%), while the lowest was observed in 2018, with no reported deaths among the 47 cases (0.0%) (Figure 5.8).

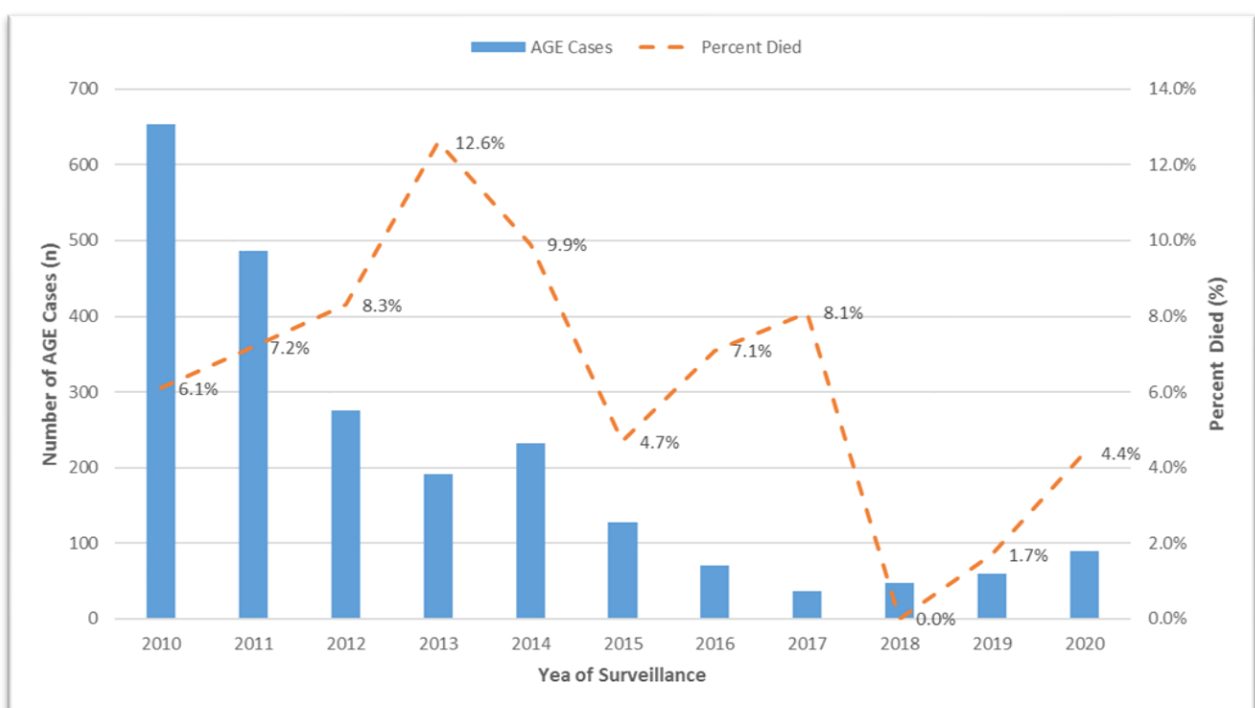


Figure 5.8: Trends of diarrhea-associated mortality among pediatric patients less than 60 months of age hospitalized for acute gastroenteritis at Siaya County Referral Hospital: 2010-2020.

The characteristics of acute gastroenteritis patients who died compared to those who survived is shown in Table 5.5. There was no significant difference in the ages of children who died and those who survived (Median age in months [IQR]: 10.4 [5.9-18.2] vs 9.9 [6.1-17.3], $p = 0.972$). However, in comparison to those who survived, the decedents had lower oxygen saturation levels

(Median [IQR]: 96 [91-98] vs 97 [94-98], $p = 0.001$), endured a longer duration of diarrhea (Median days [IQR]: 4 [3-6] vs 3 [2-4], $p < 0.001$), presented with a lower body temperature (Median [IQR]: 37.0 [36.0-38.0] vs 37.2 [36.6-38.0], $p = 0.004$) and lower number of vomiting episodes (Median [IQR]: 2 [1-4] vs 3 [1-4], $p = 0.014$).

Table 5.5: Demographic and clinical profile of pediatric patients aged less than 60 months hospitalized with acute gastroenteritis at Siaya County Referral Hospital: 2010-2020.

Characteristics	Died (n=164)	Alive (n=2,107)	p-value
Median age months [IQR]	10.35 [5.865-18.185]	9.86 [6.11-17.28]	0.972
Age			
0-5 months	43 (26.2)	504 (23.9)	0.177
6-11 months	50 (30.5)	765 (36.3)	
12-23 months	53 (32.3)	553 (26.3)	
24-35 months	14 (8.5)	174 (8.3)	
36-59 months	4 (2.4)	111 (5.3)	
Gender: Male	85 (51.8)	1,171 (55.6)	0.353
Vital signs			
Median Temperature [IQR]	37 [36-38]	37.2 [36.6-38]	0.004
Median Oxygen Saturation [IQR]	96 [91-98]	97 [94-98]	0.001
Median Respiratory Rate [IQR]	41 [32-52]	42 [34-52]	0.5648
Symptoms			
Vomiting	56 (34.4)	505 (24.0)	0.003
Convulsions	29 (18.5)	87 (19.0)	0.938
Lethargic	85 (59.9)	949 (50.9)	0.119
Unconscious	41 (26.6)	211 (10.4)	<0.001
Fever	114 (72.6)	1,786 (87.4)	<0.001
Malaria	23 (15.7)	683 (35.0)	<0.001
Lost weight	115 (78.2)	1,166 (62.0)	0.001
Unable to drink	10 (8.7)	78 (5.5)	0.234
AVPU			
Alert	130 (81.3)	2,032 (97.1)	<0.001
Voice	18 (11.3)	37 (1.8)	
Pain	7 (4.4)	20 (1.0)	
Unresponsive	5 (3.1)	3 (0.1)	
Restless	54 (33.3)	480 (22.9)	<0.001
Convulsing Now	2 (1.2)	92 (1.4)	0.863
Sunken Eyes	72 (43.9)	567 (27.1)	<0.001
Skin Pinch			
Normal	69 (42.1)	1,460 (69.7)	<0.001
Slowly	80 (48.8)	534 (25.5)	
Very Slowly	15 (9.1)	102 (4.9)	

Capillary Refill (>2 seconds)	35 (22.3)	314 (15.4)	0.024
Drink			
Not able to drink/drinks poorly	90 (55.6)	703 (33.8)	<0.001
Drinks eagerly/Very Thirsty	48 (29.6)	866 (41.7)	
Normally	24 (14.8)	509 (24.5)	
Bulging Fontanelle	5 (4.0)	24 (1.6)	0.049
Sunken Fontanelle	40 (27.4)	332 (18.2)	0.007
Red Eyes	2 (1.6)	5 (0.3)	0.035
Visible wasting	38 (31.7)	93 (6.4)	<0.001
IVF given	94 (57.7)	817 (39.2)	<0.001
Dehydration Status			
Severe dehydration	23 (14.0)	117 (5.6)	<0.001
Some dehydration	76 (46.3)	647 (30.7)	
No dehydration	65 (39.6)	1,343 (63.7)	
Vesikari Score			
Mild	25 (15.2)	294 (14.0)	0.389
Moderate	56 (34.2)	834 (39.6)	
Severe	83 (50.6)	979 (46.4)	
Stunting			
Normal	72 (44.7)	1,361 (65.1)	<0.001
Moderate	22 (13.7)	345 (16.5)	
Severe	67 (41.6)	384 (18.4)	
Wasting			
Normal	73 (45.9)	1,450 (70.6)	<0.001
Moderate	22 (13.8)	256 (12.4)	
Severe	64 (40.3)	349 (17.0)	
Underweight			
Normal	50 (30.7)	1,320 (63.9)	<0.001
Moderate	26 (16.0)	342 (16.5)	
Severe	87 (53.4)	405 (19.6)	
Prior Admission	6 (5.1)	29 (2.1)	0.041
Sought Care	109 (75.2)	1,106 (60.7)	0.001
Chest indrawing	54 (34.2)	325 (16.0)	<0.001
Stridor	16 (10.8)	84 (4.5)	0.001
Nasal flaring	94 (33.1)	346 (28.5)	<0.001
Median Diarrhea days [IQR]	4 [3-6]	3 [2-4]	<0.001
Median Diarrhea episodes [IQR]	4 [3-5]	4 [3-5]	0.651
Median Vomit episodes [IQR]	2 [1-4]	3 [1-4]	0.014
Median Vesikari Score [IQR]	11 [8-13]	10 [8-12]	0.139

IQR-Interquartile range

Moreover, the following factors were positively associated with mortality: vomiting (56 [34.4%] vs 505 [24.0%], $p=0.003$), unconsciousness (41 [26.6%] vs 211 [10.4%], $p<0.001$), weight loss

(115 [78.2%] vs 1,166 [62.0%], $p=0.001$), AVPU scale (Voice: 18 [11.3%] vs 37 [1.8%], Pain: 7 [4.4%] vs 20 [1.0%], Unresponsive: 5 [3.1%] vs 3 [0.1%]; $p<0.001$), restless (54 [33.3%] vs 480 [22.9%], $p<0.001$), sunken eyes (72 [43.9%] vs 567 [27.1%], $p<0.001$), skin pinch (Slowly: 80 [48.8%] vs 534 [25.5%], Very slowly: 15 [9.1%] vs 102 [4.9%]; $p<0.001$), capillary refill (35 [22.3%] vs 314 [15.4%], $p=0.024$), not able to drink/drinks poorly (80 [48.8%] vs 703 [33.8%], $p<0.001$), visible wasting (38 [31.7%] vs 93 [6.4%], $p<0.001$), IV given (94 [57.7%] vs 817 [39.2%], $p<0.001$), dehydration (Some dehydration: 23 [14.0%] vs 117 [5.6%], severe dehydration: 76 [46.3%] vs 647 [30.7%]; $p<0.001$), severe stunting (67 [41.6%] vs 384 [18.4%], $p<0.001$), severe wasting (64 [40.3%] vs 349 [17.0%], $p<0.001$), severe underweight (87 [53.4%] vs 405 [19.6%], $p<0.001$), prior admission (6 [5.1%] vs 29 [2.1%], $p=0.041$), chest indrawing (54 [34.2%] vs 325 [16.0%], $p<0.001$), stridor (16 [10.8%] vs 84 [4.5%], $p=0.001$) and nasal flaring (94 [33.1%] vs 346 [28.5%], $p<0.001$).

The feature selection analysis identified several key predictors, ranked in order of importance (Figure 5.9). These included the AVPU scale, Vesikari score, dehydration, sunken eyes, skin pinch, maximum number of vomits, unconsciousness, wasting, vomiting, pulse rate, fever, sunken fontanelle, restlessness, nasal flaring, duration of diarrhea, stridor, oxygen saturation below 90%, and chest indrawing. Additionally, malaria and stunting were considered as tentative features.

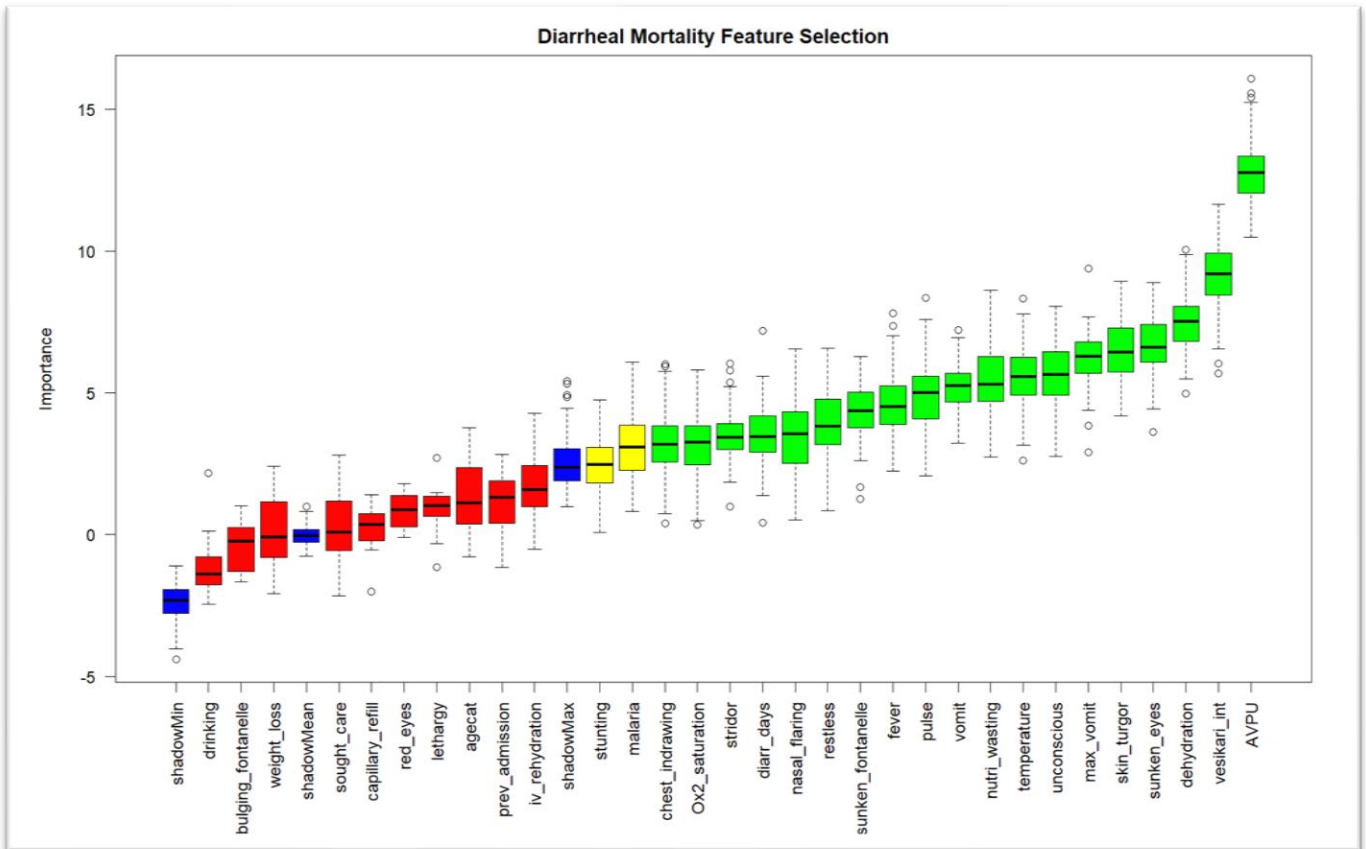


Figure 5.9: Predictor identification for gastroenteritis-associated mortality among children aged < 60 months admitted at SCRH: 2010-2020.

5.3.2 Discussion: Diarrhea-associated Mortality Predictors (RQ3)

The overall prevalence of diarrhea-associated mortality was 7.2% during the study period. Looking at the trends of mortality there were no significant variations in the prevalence of mortality over the years. This reported proportion of in-hospital mortality was higher than the estimate reported (4.6%) in Kilifi County hospital in Coastal Kenya between 2007-2015 (Talbert et al., 2019), but was lower than case fatality rate (9%) reported from two studies that assessed in-hospital mortality in Kilifi Counties respectively (Akech et al., 2018) and Siaya (O'Reilly et al., 2012). The mortality trends observed in the study deviates from the patterns of diarrheal mortality reported globally (Clark et al., 2023) and regionally (Black et al., 2019). This difference could be attributed to a

combination of regional disparities, healthcare infrastructure limitations, surveillance challenges, and variations in intervention coverage and effectiveness. Specifically, in a study looking at drivers of diarrheal mortality in 50 LMICs, a reduction in mortality rates, particularly amongst children aged less than 60 months, was observed, with deaths plummeting by over 80% from 1980 to 2015, despite a rise in the population of children in LMICs (Black et al., 2019). This remarkable decline is attributable to a range of factors, including enhanced accessibility to the rotavirus vaccine, advancements in child growth and well-being initiatives, and the establishment of universal access to safe water and adequate sanitation (Bar-Zeev et al., 2018; Black et al., 2019; Clark et al., 2023). These concerted efforts have been fundamental in significantly alleviating the burden of preventable diarrheal deaths.

In particular, the implementation of comprehensive diarrhea treatment protocols, including oral rehydration salts, zinc supplementation, antibiotics for dysentery, and management of persistent diarrhea, alongside the widespread adoption of the rotavirus vaccine, collectively contributed to a substantial 49.7% reduction in diarrhea-related mortality during this period (Black et al., 2019). Additionally, significant strides in nutrition enhancement initiatives, encompassing interventions targeting stunting, wasting, optimal breastfeeding practices, and vitamin A supplementation, accounted for another 38.8% of the mortality decline. Furthermore, advancements in water, sanitation, and handwashing practices played a crucial role, contributing 11.5% to the overall reduction in mortality attributed to diarrhea. Despite significant progress in preventive strategies and advancements in diarrhea management, diarrheal mortality persists as the second leading cause of childhood deaths globally. This burden is greatest in SSA and Southern Asia, underscoring the ongoing challenge and urgency still there. This persistence highlights the need for innovative approaches to complement ongoing strategies to address this public health problem. One such

example is the ongoing digital revolution in the healthcare sector, which can be applied in a variety of ways, including in diagnosis and treatment (Cihula, 2023).

The feature selection process identified 19 predictors, which can be organized into five distinct categories: diarrheal severity (modified Vesikari score, frequency of vomiting, vomiting, diarrhea duration, fever, dehydration status [skin turgor, restless, sunken fontanelle and sunken eyes]); IMCI general danger signs (unconsciousness, and AVPU scale); malnutrition (wasted and stunted); vital signs oxygen saturation below 90% and pulse rate); co-infections (malaria and respiratory distress [chest indrawing, stridor and nasal flaring]). The variables identified through feature selection in this study are established drivers of mortality based on prior studies (Acácio et al., 2019; Hartman et al., 2022; Lee et al., 2014; Walker et al., 2013; Wierzbica & Muhib, 2018; Zenebe et al., 2022). The consistency of these findings with prior studies further reinforces the robustness and utility of ML algorithms. In addition, these findings underscore the significance of key risk factors for diarrheal mortality in children, emphasizing the urgent necessity for targeted clinical and public health interventions aimed at addressing them effectively.

The study finding on severity of diarrhea corroborates findings from the GEMS study, which showed that children presenting with moderate-to-severe diarrhea had an 8.5 fold increase in mortality risk compared to their counterparts who were healthy (Kotloff et al., 2013; Pavlinac et al., 2018). Further affirming this finding are results from a study conducted across 13 Kenyan hospitals between 2013-2016, which showed that among children presenting with diarrhea and dehydration, those at highest risk of in-hospital mortality were those exhibiting additional signs of severe disease, rather than those with uncomplicated dehydration (Akech et al., 2018). Specifically, markers like dehydration and Vesikari score, reflect the severity of illness, are associated with a heightened risk of patient mortality and inform therapeutic management (Lee et

al., 2014; World Health Organization, 2017). Severe dehydration and fluid loss have been the leading causes of diarrheal deaths. Severe diarrhea leads to swift depletion of fluids and electrolytes, causing dehydration that can quickly become life-threatening, especially in young children. Failure to promptly treat dehydration can result in a trajectory culminating in shock, organ dysfunction, and subsequent mortality. Moreover, severe diarrhea can result in electrolyte imbalances, such as hyponatremia, hypokalemia or metabolic acidosis, which can disrupt normal physiological processes and lead to cardiac arrhythmias, seizures, or other serious complications (Shah et al., 2007). Furthermore, severe diarrhea may precipitate a systemic inflammatory response, culminating in septicemia, a critical syndrome marked by widespread inflammation and associated organ failure (Chakraborty & Burns, 2023; Ye-ting & Dao-ming, 2018).

Additionally, a systematic review investigating the recognition of childhood illnesses and care-seeking behavior in developing nations revealed a concerning trend: caregivers exhibited a median sensitivity of only 26.0% in recognizing diarrhea (Geldsetzer et al., 2014). This finding resonates with research conducted in Kenya, where a significant proportion of caretakers perceived diarrhea as a mild ailment not requiring medical attention beyond home remedies (Nasrin et al., 2013; Wambui et al., 2018). These findings exacerbates an already grievous public health problem, as studies have consistently shown that delays in seeking care by caregivers in LMICs expose children to more severe manifestations of disease, leading to complications and even death (Zenebe et al., 2022). The above evidence highlights the need for educating caregivers on diarrheal illness, its severity and appropriate management strategies to promote timely and effective care-seeking behavior. Linking severity of disease to etiology of diarrhea, a systematic review aimed at determining the global causes of childhood diarrheal mortality identified 13 pathogens contributing to severe diarrheal illness (Lanata, Fischer-Walker, Olascoaga, Torres, Aryee, Black,

& Unicef, 2013). These include 4 viruses (rotavirus; calicivirus; astrovirus; and adenovirus), 6 bacteria (EPEC; ETEC; *Shigella*; *Campylobacter*; *Salmonella*; and *Vibrio cholerae*), and 3 parasites (*Cryptosporidium*; *Giardia*; and *Entamoeba histolytica*). These pathogens align closely with those identified in the GEMS as posing heightened risks of mortality (Levine et al., 2020). While the primary study only tested for rotavirus, it is hypothesized that the etiological agents of diarrheal illness in children who died were likely those associated with severe disease and mortality.

The documentation of general danger signs as indicators of increased risk of mortality is well-established in medical literature and clinical practice (Acácio et al., 2019; Akech et al., 2018; Basaleem et al., 2021; Mehal et al., 2012; Mitra et al., 2000; Uysal et al., 2000; World Health Organization & UNICEF, 2005). These signs reflect a breakdown of vital physiological and neurological functions, and imminent risk of death in the absence of immediate medical attention. In recognition of the critical significance of these danger signs, WHO has formulated the IMCI guidelines emphasizing the essential role of the assessment general danger signs in the process of evaluating pediatric patients presenting with illness at healthcare facilities (World Health Organization & UNICEF, 2005). These guidelines recommend initiating urgent pre-referral treatment while simultaneously arranging for prompt referral to a higher level of care for further evaluation and management. Specifically, the manifestation of unconsciousness and assessment through the AVPU scale, as observed in this study, highlights an alarming escalation of severity and imminent danger. In the setting of diarrheal illness, unconsciousness may signal severe dehydration, electrolyte imbalances, metabolic acidosis or systemic inflammatory response, leading to neurological compromise and potential organ failure including brain dysfunction and loss of consciousness. Likewise, an abnormal AVPU scale, particularly a score of "Pain" or

"Unresponsive," indicates a significant decline in neurological function. The presence of unconsciousness or altered consciousness levels according to the AVPU scale reflects an urgent need for immediate medical intervention to address underlying complications, which is crucial for improving the chances of survival and preventing potentially fatal complications associated with severe diarrheal illness.

The relationship between malnutrition and diarrheal mortality is bidirectional and well-documented over decades of research. Malnutrition exacerbates the risk of mortality among children with diarrhea through various mechanisms, including weakened immune response, inappetence/anorexia, reduced absorptive capacity, damage to the gastrointestinal mucosa, and nutrient depletion (Acácio et al., 2019; Wierzba & Muhib, 2018). In addition, disruption of gut function and an imbalance in the gut microbiota are closely associated with being acutely malnourished. These conditions may create an environment in which enteropathogens are more prone to triggering severe and potentially fatal diseases (Tickell et al., 2020). Furthermore, this vicious cycle is perpetuated as being malnourished compromises the immunological defenses of the body against future diarrheal illness, leading to increased frequency, prolonged, and severe illness. Previous research has shown that children suffering from severe acute malnutrition face a 10-fold increase in mortality risk compared to their healthy peers (Roy et al., 2011; Uauy et al., 2012). Moreover, the GEMS study has highlighted a significant association between stunting and mortality, particularly during or following discharge from healthcare facilities among children with diarrheal disease (Kotloff et al., 2013). Moreover, a post-hoc analysis of GEMS showed that children experiencing acute malnutrition face a significantly elevated likelihood of death after experiencing an episode of diarrheal disease, comprising approximately 67% of all deaths reported in this population (Tickell et al., 2020). Importantly, this association between being acutely

malnourished and death during episodes of MSD was not confined to specific infections; rather, it spanned across a diverse range of enteric pathogens.

Children with malnutrition exhibit a distinctive physiological characteristic: a disproportionately larger gut mucosal surface area relative to their body weight compared to well-nourished counterparts (Nuzhat et al., 2020). This physiological difference heightens their susceptibility to more frequent and severe purging episodes during diarrheal illness, leading to increased stool output. Moreover, malnourished individuals often exhibit host characteristics that exacerbate disease severity, including compromised immune function and alterations in gut structure and function (Rytter et al., 2014; Tickell et al., 2017). These factors collectively contribute to the impairment of the intestinal barrier, immaturity of the gut microbiota, and heightened gut permeability. Consequently, pathogens are more likely to breach the gut barrier, entering the bloodstream and potentially causing serious complications such as gram-negative bacteremia or septicemia. This intricate interplay accentuates the elevated risk of mortality among malnourished children during diarrheal episodes.

Moreover, co-morbidities have been demonstrated to exacerbate the complexity and severity of the ailment, elevating the mortality risk in children suffering from diarrhea (Acácio et al., 2019; Black et al., 2003; Walker et al., 2013). This phenomenon might be elucidated by heightened physiological strain, resource competition, and strain on organ systems. Additionally, there is a possibility that concurrent illnesses or the immune response to them could modify gut flora, consequently augmenting susceptibility to, or severity of, intestinal infections (Walker et al., 2013). Co-infections can overwhelm the immune system, leading to a compromised ability to fight off infections effectively (Mulatya & Mutuku, 2020). This weakened immune response may exacerbate the severity of both diarrhea and co-infecting illnesses, increasing the risk of mortality.

Among issues, co-infections can also put additional strain on various organ systems, particularly the respiratory and circulatory systems further compromising organ function. Delayed recognition and treatment of both diarrhea and co-infections can result in worsening symptoms and progression to severe illness, increasing the risk of mortality.

Looking at the specific co-infections identified as predictors in this study, respiratory distress, such as chest indrawing, stridor, and nasal flaring, can impair lung function and gas exchange, leading to hypoxemia (Salah et al., 2015). Hypoxemia, combined with the metabolic demands of diarrheal illness, can result in respiratory failure and contribute to mortality. Respiratory distress can lead to hypoxia, which may further compromise the body's ability to cope with diarrheal illness. In addition, malaria, a parasitic infection, can cause severe anemia, which reduces oxygen-carrying capacity in the blood and weakens the immune response (Sey et al., 2020; White, 2018). Malaria presents a significant risk factor for invasive bacterial diseases, predominantly attributed to invasive enteropathogens, specifically non-Typhoidal Salmonella (NTS). Bacteremia accompanies roughly 6.5% of severe malaria cases and carries a high fatality rate (Sey et al., 2020). Future modeling studies could benefit from the inclusion of socio-economic factors, which were not captured in the primary dataset used for this analysis.

While the predictors for each outcome has been discussed in detail in Sections 5.1.2, 5.2.2, and 5.3.2, a comparison across the three outcomes reveals both overlap and distinct divergence in the predictors, highlighting shared pathways of illness alongside outcome-specific mechanisms. Several features were common across outcomes, particularly age, breastfeeding status, respiratory rate, skin turgor, and rotavirus vaccination, highlighting the central role of early-life vulnerability, nutrition, hydration status, and infection prevention in shaping pediatric outcomes. Indicators of dehydration and illness severity including vomiting, stool frequency, and Vesikari score were

especially prominent for LDD and mortality, highlighting the importance of acute gastrointestinal burden and fluid loss in driving short-term complications and fatal outcomes. In contrast, LGF was characterized by fewer predictors and was more strongly associated with chronic nutritional and systemic factors, such as severe acute malnutrition and temperature, rather than acute severity markers. This likely reflects the cumulative and longer-term biological processes underlying growth impairment, which are less sensitive to transient illness features. Inpatient mortality, conversely, was associated with the largest and most complex set of predictors, dominated by neurological, respiratory, and circulatory, as well as comorbid infections like malaria, reflecting multisystem failure and advanced disease severity. Overall, these similarities and differences likely reflect the temporal and physiological scales of the outcomes. Early nutritional status and basic clinical signs influence all three, but mortality represents an acute, high-risk event driven by systemic decompensation and comorbidities, requiring indicators of immediate organ compromise. LGF, in contrast, is a sub-acute process shaped by a child's nutritional reserves, such as severe acute malnutrition, and protective factors like breastfeeding. LDD lies between these extremes, with illness duration determined by the interplay between pathogen virulence, captured by the Vesikari score, and the child's local mucosal immune response.

5.4 Model Development and Evaluation for Poor Diarrheal Outcomes

To fulfill the study objective of “*Develop and evaluate (internal and temporal validation) the performance of predictive models for each poor diarrheal outcome, compared to logistic regression, among children aged <60 months using machine learning.*”, three distinct research questions were formulated for each poor diarrheal outcome (LDD, chronic malnutrition and mortality) under investigation.

- **RQ4-** What machine learning model best predicts longer duration diarrhea among children aged less than 60 months?
- **RQ5-** What machine learning model best predicts chronic malnutrition among children aged less than 60 months following a diarrheal illness?
- **RQ6-** What machine learning model best predicts diarrhea-associated mortality among inpatient children aged less than 60 months?

5.4.1 Assessing Model Performance for Longer Duration Diarrhea Prediction

5.4.1.1 Key Findings for RQ4

To address the fourth research question (RQ4), the model performance was evaluated using internal validation. Subsequently, a comparative evaluation was conducted of the predictive capability of the models, with logistic regression as a reference, using the Friedman test. Next, SHAP attributions from explanatory model analysis were utilized to enhance interpretability and explainability. Moreover, an assessment of how well the models were calibrated, was done using Brier score. Additionally, business value evaluation charts for the champion model were included. Finally, model performance was also assessed through temporal validation.

The primary use case modelled for this research question was the likelihood of acute diarrhea evolving to LDD. The results presented in Table 5.6 show the model performance when oversampling technique was employed as a component of the resampling methodology to mitigate class imbalance. Among the developed models, the RF, LR and ANN models exhibited the highest sensitivity at 80.7%, 76.5% and 75.6%, respectively. The GBM, SVM, and ANN models demonstrated the highest specificity at 76.5%, 76.5 % and 74.9%, respectively. The PPV ranged from 55.4% to 59.9%, while the NPV ranged from 83.8% to 89.0%. In terms of the discriminatory

ability, the RF model performed the best with an AUC of 83.0%, followed by SVM (82.0%), ANN (81.5%), GBM (81.1%), LR (80.5%), KNN (79.7%), and NB (77.3%).

Table 5.6: Performance of LDD predictive models utilizing over-sampling approach within the resampling procedure

Comparative analysis of the performance of the 7 models was done using Friedman's test

Algorithm	LDD Prediction with over-sampling technique						
	Sensitivity % [95% CI]	Specificity % [95% CI]	PPV % [95% CI]	NPV % [95% CI]	F1-Score % [95% CI]	AUC % [95% CI]	PRAUC % [95% CI]
RF	80.7 [72.4-87.3]	74.1 [68.2-79.4]	59.6 [51.6-67.3]	89.0 [83.9-92.9]	68.6 [56.3-75.3]	83.0 [78.6-87.5]	90.0 [85.9-93.8]
GBM	72.3 [63.3-80.1]	76.5 [70.8-81.1]	59.3 [50.8-67.4]	85.3 [80.0-89.7]	65.2 [41.7-73.1]	81.1 [76.3-86.0]	87.8 [83.3-93.0]
NB	69.7 [60.7-77.8]	74.1 [68.2-79.4]	56.1 [47.7-64.2]	83.8 [78.3-88.4]	62.2 [36.4-68.9]	77.3 [72.1-82.6]	85.7 [82.2-89.0]
LR	76.5 [67.8-83.8]	74.5 [68.6-79.8]	58.7 [50.5-66.5]	87.0 [81.7-91.2]	66.4 [44.1-69.4]	80.5 [75.6-85.4]	88.7 [84.4-92.7]
SVM	73.9 [65.1-81.6]	76.5 [70.8-81.6]	59.9 [51.5-67.9]	86.1 [80.9-90.4]	66.2 [42.2-69.6]	82.0 [77.3-86.7]	89.3 [85.1-93.1]
KNN	73.1 [64.2-80.8]	72.1 [66.1-77.6]	55.4 [47.3-63.3]	85.0 [79.5-89.5]	63.0 [40.4-70.7]	79.7 [74.9-84.5]	87.0 [83.9-91.0]
ANN	75.6 [66.9-83.0]	74.9 [69.1-80.1]	58.8 [50.6-66.7]	86.6 [81.4-90.9]	66.1 [46.3-74.1]	81.5 [76.8-86.2]	89.1 [84.8-93.6]

(Friedman chi-squared = 16.71, df = 6, p-value=0.01041), which revealed that the difference in the performance of the models across the different metrics was statistically significant. While the Friedman test showed that at least one model differed significantly in performance compared to the others, it did not give insights about which specific model. Consequently, post-hoc tests were conducted to determine the pairwise difference between the models using Nemenyi test, as shown in Table 5.7.

Table 5.7: Post-hoc analysis of LDD model performance using Nemenyi Test

	ANN	GBM	KNN	LR	NB	RF
GBM	0.9943	-	-	-	-	-
KNN	0.1961	0.5908	-	-	-	-
LR	1.0000	0.9943	0.1961	-	-	-
NB	0.6360	0.9537	0.9904	0.6360	-	-
RF	0.9171	0.5453	0.0073	0.9171	0.0743	-

Friedman chi-squared = 16.71, df = 6, p-value=0.01041

The results of the pairwise differences indicate that the RF model had a significantly better performance than the KNN model ($p < 0.05$). Additionally, Nemenyi test results were plotted to observe patterns in model performance among the 7 models. Groups that share a common letter are not significantly different from each other, while those with different letters are considered to have statistically significant differences in their performance. These results showed three statistically distinct groups based on Nemenyi test results as shown in Figure 5.10. The RF model (group “b”) had the highest performance and was significantly different from KNN (group “a”). KNN (group “a”) had the lowest performance and was significantly different from both “b” and “ab” groups. Group “ab” (ANN, GBM, LR, NB and SVM) had intermediate performance. The models in these group were not significantly different from each other.

For this primary use case, a sensitivity analysis was conducted using the data without correcting for the class imbalance. The results of the sensitivity analysis are shown in Table 5.8. Model performance ranged from 43.7% to 52.9%, 84.9% to 90.8%, 60.0% to 72.6% and 76.9% to 79.9% for sensitivity, specificity, PPV and NPV, respectively. In terms of discriminative ability assessed by model AUC, ANN had the highest AUC (81.2%), followed by LR (81.1%), SVM (80.5%), RF (80.3%), NB (77.6%), GBM (76.6%) and KNN (75.1%).

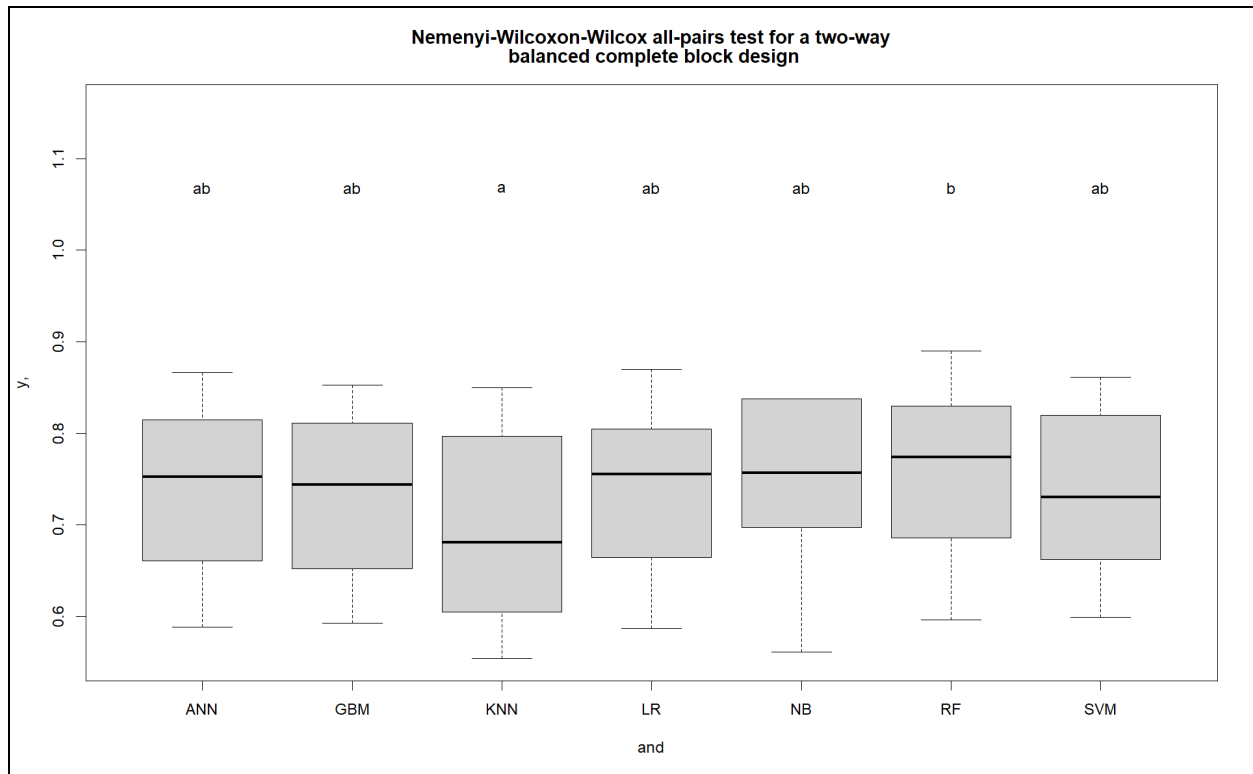


Figure 5.10: Nemenyi Post-hoc Test for Performance Differences among LDD Prediction Models

Table 5.8: Performance of LDD Predictive Models without correcting for class imbalance.

LDD Prediction with no sub-sampling technique used in the resampling procedure							
Algorithm	Sensitivity % [95% CI]	Specificity % [95% CI]	PPV % [95% CI]	NPV % [95% CI]	F1-Score % [95% CI]	AUC % [95% CI]	PRAUC % [95% CI]
RF	47.9 [38.7-57.2]	84.9 [79.8-89.1]	60.0 [49.4-69.9]	77.5 [72.1-82.3]	53.3 [25.6-69.9]	80.3 [75.6-85.0]	89.0 [86.4-92.1]
GBM	52.9 [43.6-62.2]	85.7 [80.7-89.7]	63.6 [53.4-73.1]	79.3 [74.0-84.0]	57.8 [33.2-71.9]	76.6 [71.4-81.7]	86.7 [84.1-89.8]
NB	49.6 [40.3-58.9]	87.3 [82.5-91.1]	64.8 [54.1-74.6]	78.5 [73.2-83.2]	56.2 [29.7-69.6]	77.6 [72.4-82.8]	86.0 [82.9-89.2]
LR	51.3 [41.9-60.5]	90.8 [86.6-94.1]	72.6 [61.8-81.8]	79.7 [74.6-84.2]	60.1 [35.3-66.9]	81.1 [76.4-85.9]	89.0 [84.9-92.8]
SVM	48.7 [39.5-58.1]	90.0 [85.6-93.5]	69.9 [58.8-79.5]	78.8 [73.6-83.3]	57.4 [30.8-72.5]	80.5 [75.6-85.4]	88.2 [84.3-92.6]
KNN	43.7 [34.6-53.1]	88.8 [84.3-92.5]	65.0 [53.5-75.3]	76.9 [71.6-81.6]	52.3 [22.1-66.7]	75.1 [69.7-80.5]	84.7 [82.4-87.3]
ANN	52.1 [42.8-61.3]	90.4 [86.1-93.8]	72.1 [61.4-81.2]	79.9 [74.8-84.4]	60.5 [36.1-67.1]	81.2 [76.4-86.0]	89.0 [85.1-92.9]

Figure 5.11 illustrates the ROC curves for the 7 models in the prediction of LDD for the primary use case. The results of the calibration assessment of the models are shown in Table 5.9. Despite generally low Brier scores ranging from 0.17 to 0.21 overall, the automated algorithm revealed poor calibration for the NB and KNN models, as indicated by Spiegelhalter's p-value ($p < 0.05$).

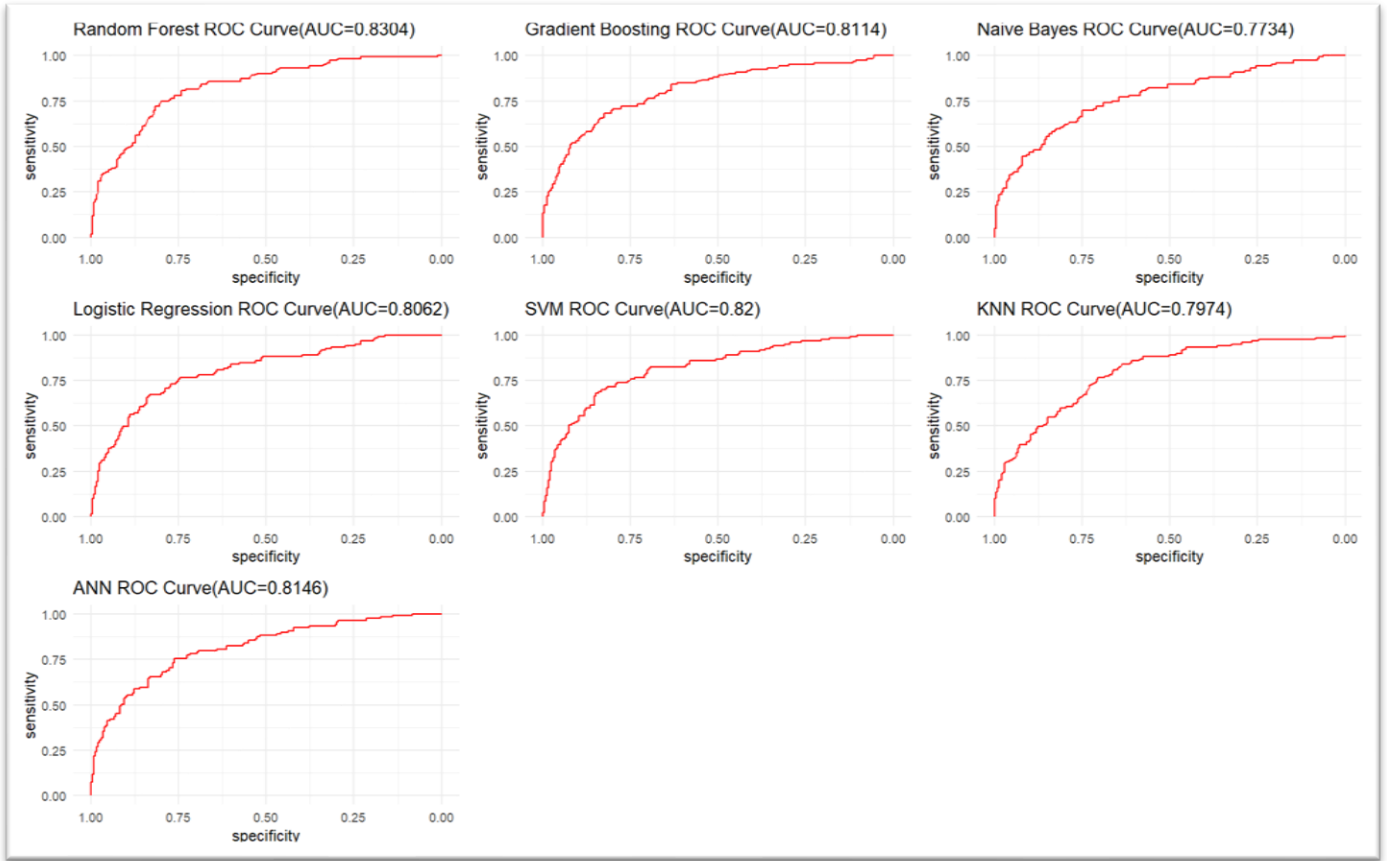


Figure 5.11: ROC analysis of the 7 models in the prediction of LDD in the primary use case

Table 5.9: Calibration results for LDD prediction models.

Algorithm	Brier Score	Spiegelhalter Z-score	Spiegelhalter p-value
RF	0.171	-1.230	0.219
GBM	0.178	-0.952	0.341
NB	0.201	4.819	0.000
LR	0.178	-0.576	0.565
SVM	0.177	-1.450	0.147
KNN	0.194	-3.859	<0.001
ANN	0.176	-0.182	0.855

In the explanatory model analysis for predicting LDD among the top 4 models for the primary use case, a consistent pattern was observed (Figure 5.12). While the direction of association among

the predictors remained similar across models, variations were observed in the magnitude of the associations. Specifically, pre-enrollment diarrhea days, severity based on the modified Vesikari score, lack of rotavirus vaccination, absence of skin turgor, and older age were consistently associated with an increased likelihood of developing LDD. On the contrary, the likelihood of progressing to LDD decreased with the absence of vomiting (vomit=0, no vomits in the last 24 hours, and no vomit days) and a lower number of loose stools in the last 24 hours (< 6).

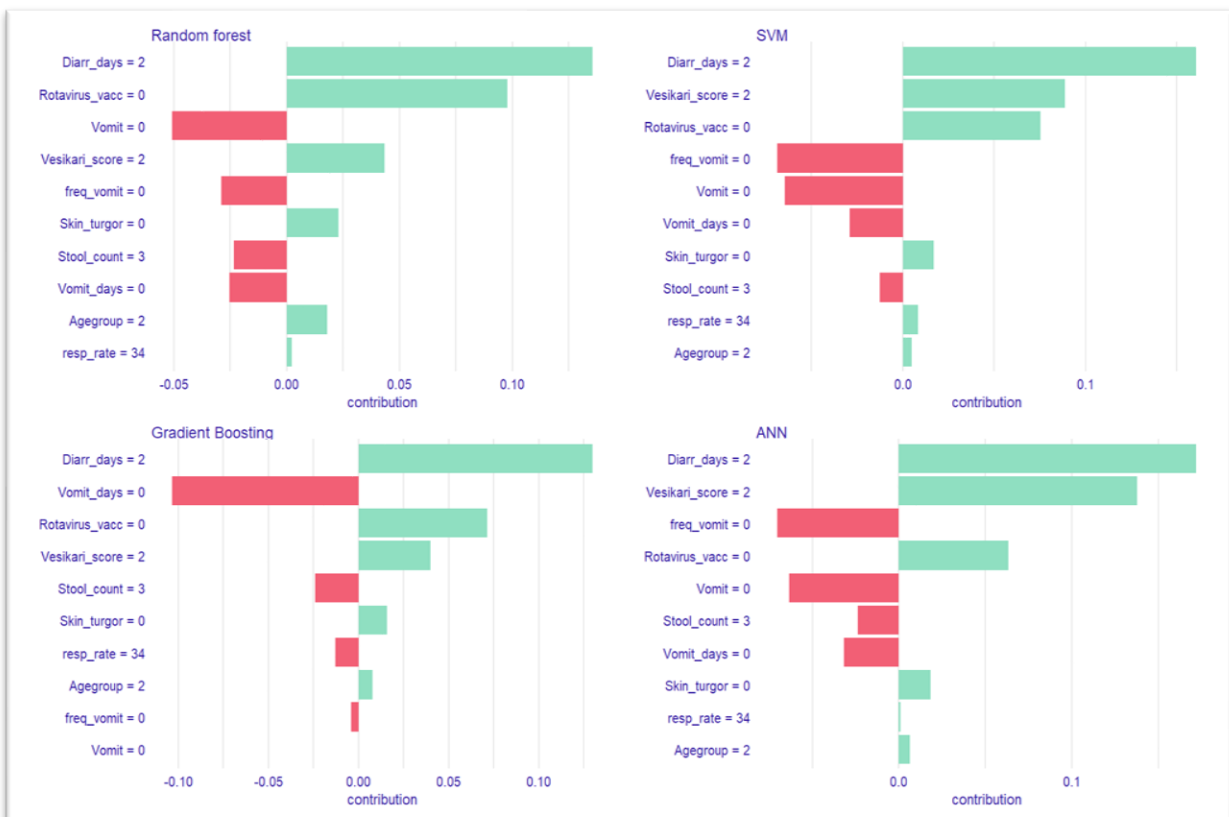


Figure 5.12: Explanatory profile for the top-performing 4 LDD models for the primary use case

In the primary use case, RF emerged as the champion model based on its highest discriminative ability (AUC=83.0%). The business value of the RF model that performed best was evaluated further, results of which are shown in Figure 5.13. The cumulative gains plot indicates that by

picking the top 20% of cases with the highest predicted probabilities, the champion model can identify 46% of the LDD instances. Furthermore, according to the cumulative lift plot, the RF model demonstrates a 2.6-fold increase in the identification of LDD cases compared to a random approach when a 20% selection threshold based on its predicted probabilities is applied. Finally, the cumulative response plot indicates that 72% of the observations within the top 20% quantile ranked by their predicted probability scores, are members of the target class.

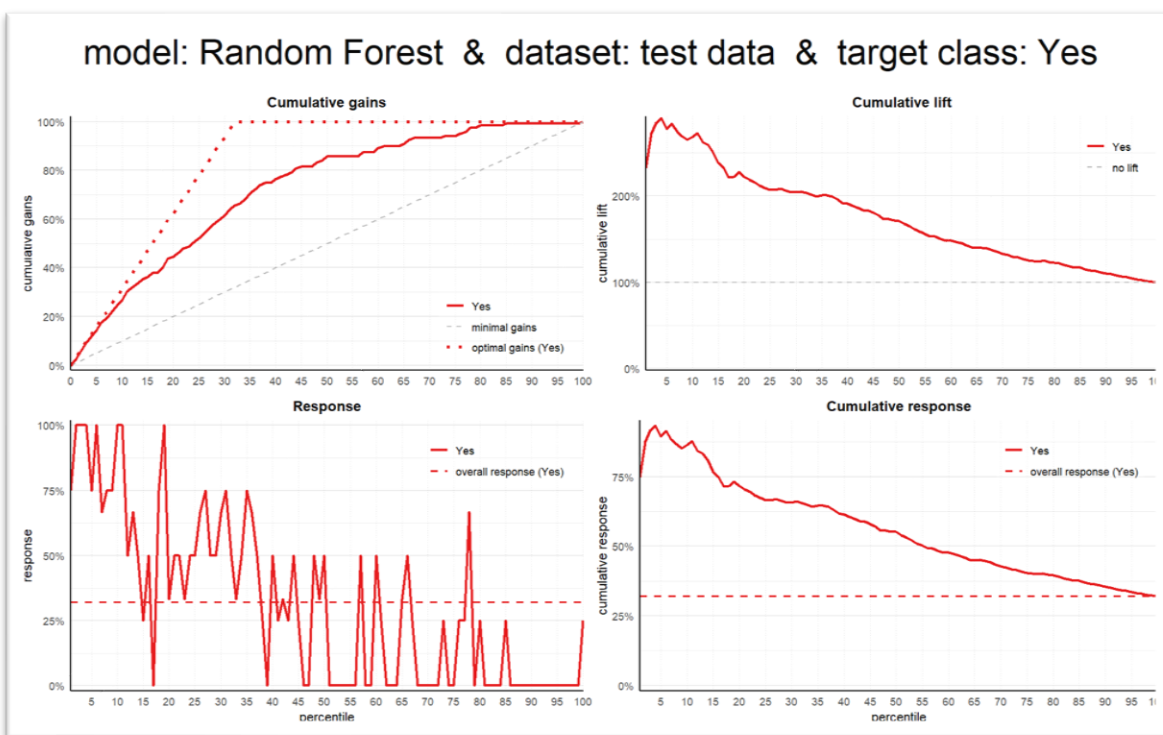


Figure 5.13: Evaluation of the business value of the champion model in LDD prediction

The champion RF model was also temporally validated using data from the EFGH study. The results of this evaluation are shown in Figure 5.14. A decline in predictive power of the model was observed when validating on the temporal cohort. Specifically, the RF model achieved 37.7% sensitivity, 86.0% specificity, 23.2% PPV, 92.5% NPV, 27.9% F1-score, 68.4% AUC, and 94.4%

PRAUC. In a sensitivity analysis focusing on EFGH enrollees meeting the VIDA diarrheal disease case definition, the model registered a slight improvement in its performance. Here, the RF model achieved 47.5% sensitivity, 80.5% specificity, 25.7% PPV, 91.5% NPV, 33.3% F1-score, 71.0% AUC, and 93.8% PRAUC.

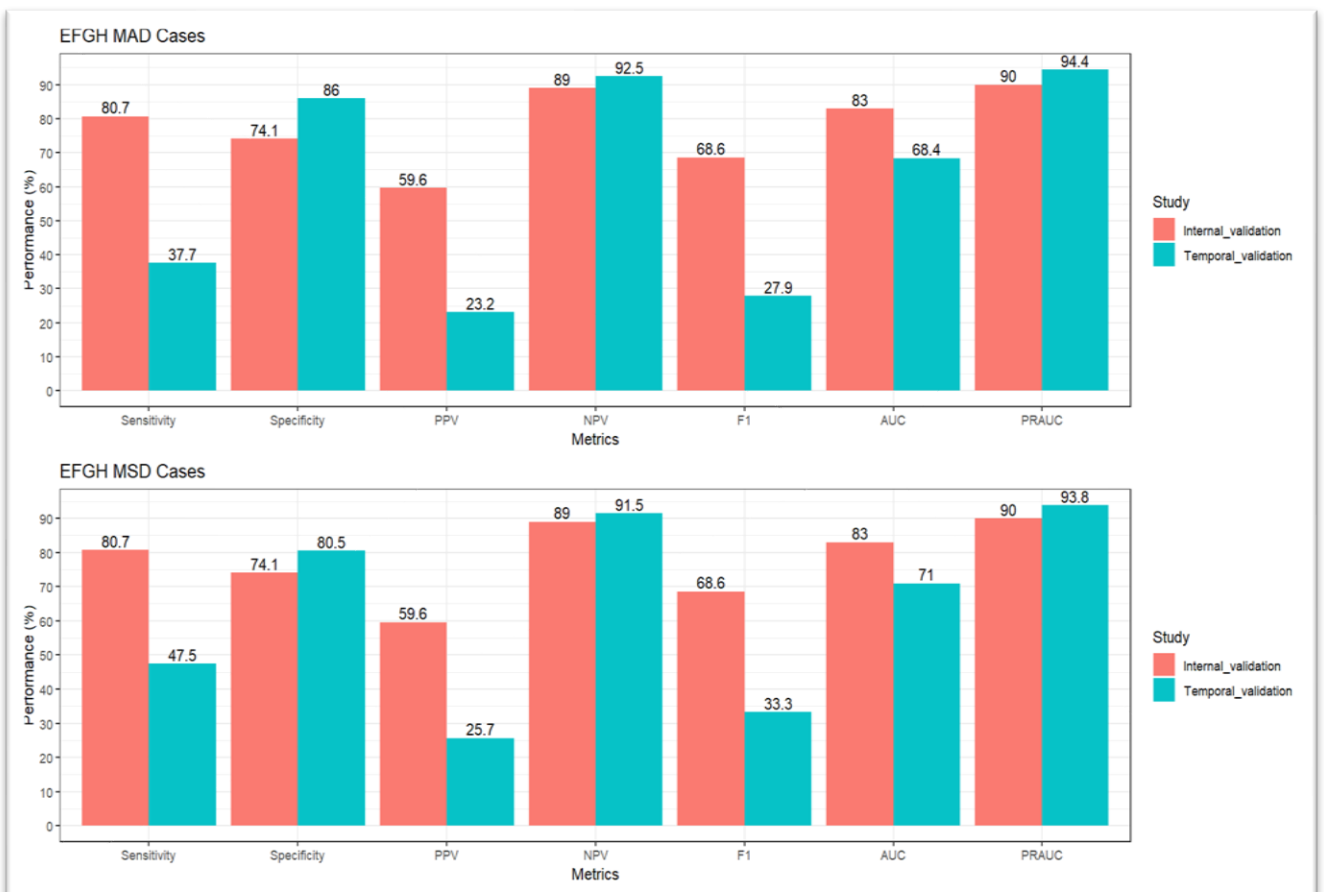


Figure 5.14: Comparative evaluation of model performance in the prediction of LDD between the development (2015-2018) and temporal validation (2022-2023) cohorts.

The second use case that was modelled for this diarrheal outcome was the likelihood of a diarrheal illness persisting for another seven days or more following the child's hospital presentation. The performance of the models in this scenario are shown in Table 5.10. The performance of the models varied across metrics, with sensitivity ranging from 42.3% to 78.8%, specificity from 45.3% to

72.3%, PPV from 16.8% to 22.1%, NPV from 88.3% to 90.9%, F1-score from 26.5% to 30.7%, AUC from 52.9% to 64.4%, and PRAUC from 86.9% to 92.0%.

The explanatory model analysis for predicting ≥ 7 days of post-enrollment diarrhea revealed similar patterns to the first use case. However, one key distinction emerged: the number of pre-enrollment diarrhea days. In this model, a higher number of pre-enrollment diarrhea days was associated with a decreased likelihood of experiencing prolonged diarrhea after enrollment (Figure 5.15).

Table 5.10: Predictive capability of models in predicting post-enrolment duration (≥ 7 days)

Post-enrolment Duration Prediction (≥ 7 days)							
Algorithm	Sensitivity % [95% CI]	Specificity % [95% CI]	PPV % [95% CI]	NPV % [95% CI]	F1-Score % [95% CI]	AUC % [95% CI]	PRAUC % [95% CI]
RF	48.1 [34.0-62.4]	72.3 [67.1-77.2]	22.1 [14.9-30.9]	89.5 [85.1-93.0]	30.3 [-19.4-56.0]	63.3 [55.9-70.7]	91.6 [88.1-94.4]
GBM	42.3 [28.7-56.8]	71.1 [65.7-76.0]	19.3 [12.5-27.7]	88.3 [83.7-92.0]	26.5 [-25.7-55.3]	61.1 [54.1-68.1]	91.7 [87.7-94.8]
NB	78.8 [65.3-88.9]	45.3 [39.7-50.9]	19.1 [14.0-25.0]	92.9 [87.7-96.4]	30.7 [0.5-45.2]	62.3 [54.6-70.0]	90.6 [87.6-92.9]
LR	61.5 [47.0-74.7]	59.1 [53.5-64.6]	19.8 [13.9-26.7]	90.4 [85.5-94.0]	29.9 [-11.7-39.7]	64.2 [57.1-71.3]	91.8 [88.5-95.2]
SVM	59.6 [45.1-73.0]	62.6 [57.0-67.9]	20.7 [14.5-28.0]	90.5 [85.8-94.0]	30.7 [-12.6-50.1]	62.7 [55.7-69.7]	91.2 [87.3-95.3]
KNN	59.6 [45.1-73.0]	51.6 [45.9-57.2]	16.8 [11.7-22.9]	88.6 [83.2-92.8]	26.2 [-12.9-43.5]	52.9 [44.5-61.3]	86.9 [83.7-89.9]
ANN	67.3 [52.9-79.7]	53.5 [47.8-59.0]	19.1 [13.7-25.6]	90.9 [85.8-94.6]	29.8 [-7.8-50.2]	64.4 [57.2-71.6]	92.0 [88.8-95.3]

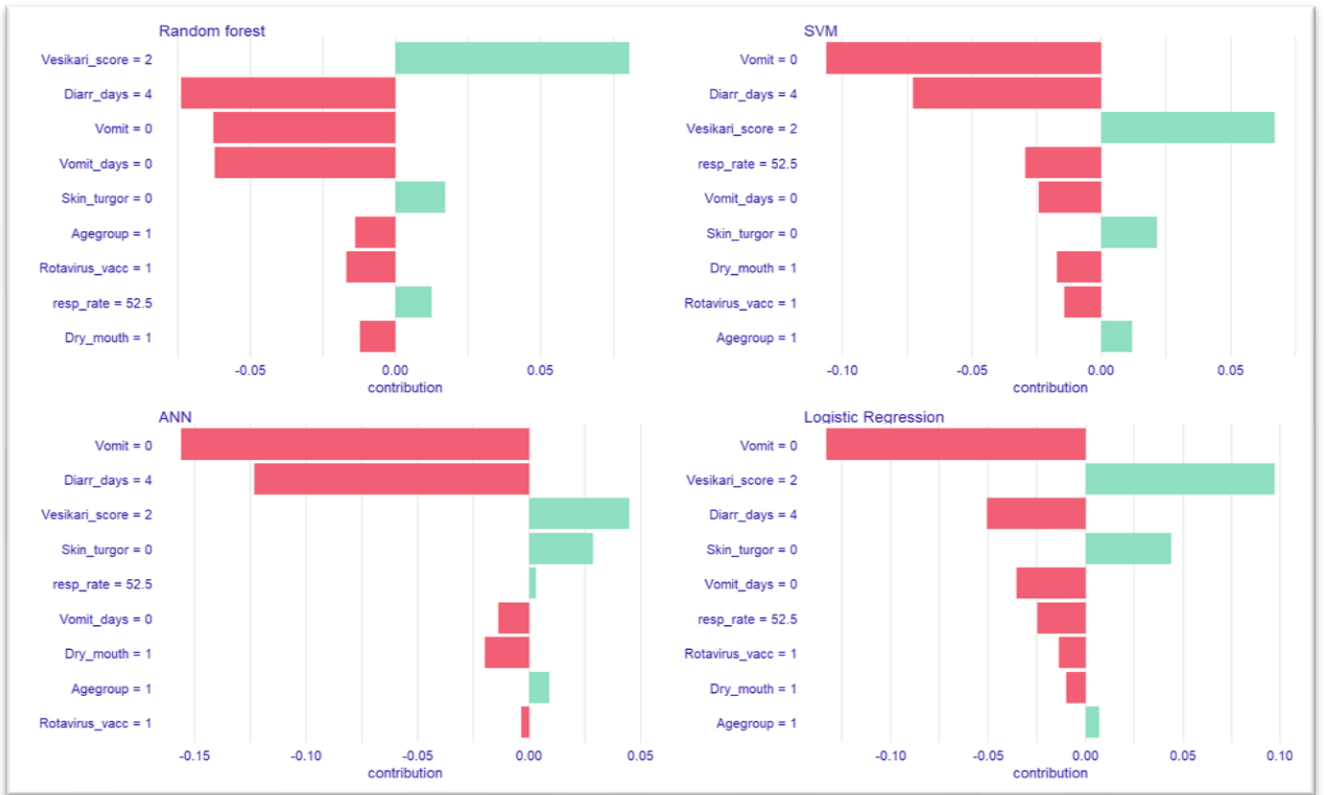


Figure 5.15: Explanatory profile for the top-performing 4 models in the prediction of post-enrolment duration (≥ 7 days).

5.4.1.2 Discussion: Model Performance in Prediction of Longer Duration Diarrhea

An estimated 32 out of 100 children with moderate-to-severe, and 11 out of 100 children with medically-attended diarrhea, will develop LDD. The confluence of this significant burden, limited diagnostic infrastructure (Bahati et al., 2022), and the strained healthcare personnel (Kokonya, 2014), highlights the critical need for new approaches, including risk assessment through clinical prediction models, to triage resources and optimize care for vulnerable children while facilitating the earlier discharge of low-risk children. Despite the different algorithms used, wide disparities in model performance were not observed, highlighting the feasibility and efficacy of these

algorithms. The RF model emerged as the best model in the prediction of LDD achieving an AUC of 83.0%. Although the review of literature did not uncover any previous attempts to predict LDD, it did provide insights into predicting illness durations for a range of other diseases. Remarkably, the model exhibited superior performance compared to a majority of these studies, namely: Duration of hospitalization for COVID-19 in USA with an AUC of 81.9% (Ebinger et al., 2021); Duration of stay among outpatient surgery patients in USA with an AUC of 73.5% (Gabriel et al., 2017); All-cause duration of hospitalization in USA with AUC ranging between 70.0% to 80.0% (Levin et al., 2021); Duration of stay among critically ill patients with organ failure in Belgium with an AUC of 75.0% (Houthoofd et al., 2015); Duration of sickness absence for knee osteoarthritis in Sweden with an AUC of 75.0% (Holm et al., 2021); Duration of hospitalization for cardiac surgery patients in Netherlands with an AUC of 75.0% (Ettema et al., 2010); All-cause duration of hospitalization in Australia with an AUC of 82.0% (Cai et al., 2016). Among the myriad studies examined, only three reported an AUC greater than 83%, a benchmark achieved by the model: Duration of hospitalization for lung cancer patients in USA with an AUC of 98.0% (Alsinglawi et al., 2022); Duration of hospitalization for septic patients in USA with an AUC of 86.0% (Ling et al., 2018); Duration of hospitalization for cardiac arrest in Saudi Arabia with an AUC of 94% (Daghistani et al., 2019). These findings highlight the good predictive capability of the model in predicting LDD duration, compared to a broad spectrum of illness duration prediction studies. Furthermore, comparative analysis using the Friedman test and Nemenyi Post-hoc test highlighted the RF model to outperform the KNN models in prediction accuracy and discriminative ability, having a statistically significant difference in the pairwise comparison. Additionally, while its performance over the ANN, GBM, LR, NB and SVM models was not statistically significant, it was distinctively higher.

These results can be explained by the documented success of RF as a general-purpose classification approach, which amalgamates several randomized decision trees aggregating their predictions through a bagging approach (Biau & Scornet, 2016; Khalilia et al., 2011). Furthermore, RF has shown excellent performance in the modelling of high dimensional data using many trees in the ensemble. It has also been shown that RF models are capable of capturing non-linear relationships and interactions between predictor variables, which may be present in complex epidemiological (Auret & Aldrich, 2012; Ryo & Rillig, 2017; Schonlau & Zou, 2020). This flexibility enhances the model's capability to uncover the fundamental structures present in the data, leading to improved predictive performance. Moreover, RF models are less prone to overfitting compared to other complex models, especially when the dataset is relatively small or noisy (Pantic et al., 2025; Schonlau & Zou, 2020). Its parallel approach facilitates learning from heterogeneous subsets of data, thereby mitigating the risk of overfitting. This robustness helps reduce the susceptibility of the model to learn spurious correlations from noisy data and enables it to generalize well to unseen data, resulting in better performance. In addition, RF has an intrinsic capability of estimating variable importance as well as a method for handling class imbalance and estimating missing data (Khalilia et al., 2011). This finding is consistent with findings from other studies that have shown RF algorithms to be superior to other supervised learning algorithms such as: Sarica et al. in the prediction of Alzheimer's disease (Sarica et al., 2017); Rigatti et al. in the prediction of colon cancer (Rigatti, 2017); Dai et al. in the diagnosis of breast cancer (Dai et al., 2018); VijiyaKumar et al. in the prediction of diabetes (VijiyaKumar et al., 2019).

The sensitivity analysis of the primary use case that modelled LDD without subsampling showed that all models except LR had a lower discriminative ability. The performance of most models except LR were negatively impacted, with the models having lower sensitivities and higher

specificities compared to when oversampling was used. This suggests that oversampling technique played an essential part in improving the predictive power of these models by addressing class imbalance challenges. Class imbalance can result in the computation of performance metrics that are susceptible to bias, particularly when evaluating sensitivity and specificity resulting in a higher count of misclassifications as positive and a lower count of valid positive classifications (Kong et al., 2020; Sharma et al., 2022; Werner de Vargas et al., 2023). In this scenario, the models might prioritize predicting the majority class (non-LDD) at the expense of the minority class (LDD), resulting in lower sensitivity and higher specificity. Models trained without subsampling may struggle to generalize unseen data, particularly when faced with imbalanced datasets. By not adequately representing the minority class during training, the models may fail to learn the underlying patterns associated with LDD, resulting in poorer discriminative ability and sensitivity. While LR does not explicitly address class imbalance during training like some ensemble methods, its probabilistic nature allows for direct estimation of the probability of the outcome variable, rather than making discrete predictions thereby potentially handling class imbalance naturally by adjusting the decision threshold. Moreover, it is a relatively simple and interpretable model, whose simplicity allows it to generalize well to imbalanced datasets without overfitting.

The calibration results conform to what has been documented about the calibration of various algorithms and can possibly be explained by the diversity in intrinsic characteristics and approaches used by various algorithms in the implementation of the classification task (Huang et al., 2020; Pavlovic, 2025). It has been shown that algorithms that use maximum margin methods move the real posterior probability away from 0 and 1 while other methods like NB tend to move the probabilities close to 0 and 1 due to their underlying assumptions on feature independence. Additionally, other algorithms such as logistic regression have been shown to be well calibrated

due to the loss function employed in its implementation while for neural networks calibration of the model decreases as the architecture complexity increases. Moreover, NB and KNN are both sensitive to outliers and noise in the data (Md. S. Ahmed et al., 2017; Halder et al., 2024; Jannah & S, 2025). Outliers can disproportionately influence the model's decision boundaries and nearest neighbor calculations, leading to erratic probability estimates and poor calibration. In addition, NB and KNN models may require a large amount of data to produce well-calibrated predictions, especially when dealing with high-dimensional feature spaces or complex decision boundaries. NB and KNN have also been shown to be relatively simple models that may exhibit limitations in modeling intricate dependencies and nonlinearities within the data. NB specifically, posits that features exhibit conditional independence within the strata defined by class membership (Jiang et al., 2013; Khan et al., 2024). However, in real-world datasets, this assumption is often violated, leading to poor calibration. It is also associated with zero-frequency problem, which occurs when a feature value fails to appear with a particular class in the training data, which potentially leads to underestimation of probabilities for certain instances by NB (Pajila et al., 2023; Phatcharathada & Srisuradetchai, 2025).

In the realm of ML, the increasing complexity of models often comes at the expense of interpretability, reducing the transparency of the model's decision-making process. Lack of transparent or trustworthy AI has been identified as one of the main barriers to implementation particularly in high-stakes domains like healthcare (González-Gonzalo et al., 2022; Markus et al., 2021). Explanatory Model Analysis using SHAP attribution empowers users to understand and trust ML models by offering explanations characterized by transparency and interpretability explanations. The explanatory model analysis across the top 4 models for the primary use case, showed that there was a consistent direction of association among the predictors and LDD. This

consistency suggests these factors consistently influence the likelihood of developing LDD, regardless of the specific predictive model used. This provides robust evidence of the relationship between these predictors and LDD. While the direction of association was consistent, there were variations in the magnitude of the associations across different models. This variation indicates that the strength of the relationship between predictors and the likelihood of LDD vary depending on the specific model used. Model complexity and algorithmic differences could contribute to these variations.

Several predictors emerged as consistently associated with an increased probability of developing LDD. These include: pre-enrollment diarrhea days; severity based on the modified Vesikari score; lack of rotavirus vaccination; absence of skin turgor; and older age. These findings align with existing knowledge in the field and have been extensively discussed in section 5.1.2, highlighting the importance of these factors in predicting the likelihood of LDD. Conversely, certain factors were associated with a decreased likelihood of progressing to LDD. These included the absence of vomiting and a lower stool frequency. Identifying these protective factors provides insights into potential interventions or preventive measures that may help reduce the risk of developing LDD among pediatric patients.

The business value evaluation further strengthens the case for RF by demonstrating its practical utility in identifying LDD cases efficiently. The cumulative gains plot estimated that 46% of LDD cases can be captured by selecting the top 20% of cases with the highest predicted probabilities. This highlights the model's effectiveness in prioritizing the most vulnerable children through heightened monitoring and more tailored support optimizing resource allocation and maximizing impact of intervention strategies like diagnostic testing and treatment. Additionally, the cumulative lift plot illustrated that the RF model outperforms random selection by identifying a 2.6-fold

increase of LDD cases when a 20% selection threshold based on its predicted probabilities is applied. This demonstrates the model's ability to significantly enhance the efficiency of case identification, enabling more targeted and effective interventions for individuals at risk of prolonged diarrhea. Lastly, the cumulative response plot shows that 72% of observations within the top 20% quantile ranked by their predicted probability scores, are members of the target class (LDD). This indicates a high precision rate, meaning that when the model predicts a high probability of LDD, it is likely to be correct. This high precision minimizes the risk of false positives, ensuring that resources are directed towards children who truly require enhanced attention. Overall, these results underscore the utility and value of the RF model in predicting LDD, offering both high discriminative ability and substantial business value in terms of identifying at-risk children and optimizing intervention strategies.

However, when temporally validated, the RF champion model exhibited suboptimal performance, with a 12.0% drop in the discriminatory ability of the model, largely explained by a reduction in the model's recall (sensitivity). This decline in model performance during temporal validation is likely attributed to heterogeneous study populations (Bayram et al., 2022; Rahmani et al., 2023). The temporal validation cohort comprised fewer instances of LDD and included children with less intense diarrheal episodes in terms of severity compared to the development cohort. While the development cohort encompassed the full age range (1–59 months) and primarily moderate-to-severe diarrhea, the temporal validation cohort was limited to children aged 6–35 months seeking care for any diarrhea. These differences in age distribution and illness severity were clearly evident between the two cohorts, highlighting potential influences on model performance. Although these differences in population characteristics, differences in case-mix characteristics, variations in the predictive influence of model features across the two cohorts, or a synergistic interplay of these

factors could potentially explain the drop in performance (Agrahari & Singh, 2022; Ali Abdu & Basulaim, 2023; Bayram et al., 2022; Steyerberg, 2009), the possibility of model overfitting contributing to the performance decrement in the validation dataset cannot be precluded. These findings also indicate that LDD prediction methodologies may need to be calibrated based on severity in each population. The observed decline in model generalizability during temporal or external validation, while consistent with prior research (Ukah et al., 2019), necessitates important caveats. Heterogeneous diarrhea severity profiles between the two study groups, compounded by potential population drift, impede the maintenance of robust predictive performance. These findings highlight the challenges of maintaining consistent predictive accuracy over time and emphasize the importance of ongoing model monitoring, periodic retraining, and the integration of dynamic, agentic AI approaches to adaptively sustain optimal performance (Heymann et al., 2022; Hinostroza Fuentes et al., 2025; Mahadevan & Mathioudakis, 2023).

Predicting the duration of post-enrolment diarrhea lasting seven or more days proved challenging compared to overall diarrheal duration (primary use case). The analysis revealed a significant reduction in model AUCs, with performance degradation of up to 26.8%. Various factors identified in the literature contribute to suboptimal predictive capability of ML algorithms, including atypical data points, class imbalances, poor model fit (over or under), suboptimal performance metrics, and data that lacks a discernible pattern (Hakami, 2024; Xie et al., 2025). The model development approach, which incorporates data partitioning, 10-fold cross-validation, and oversampling techniques, exhibits a robust capacity to mitigate most of these limitations. However, the persistence of unpredictable data patterns remains a potential explanation for the observed suboptimal results. EMA findings regarding post-enrolment duration prediction closely mirrored those for overall diarrheal duration in the primary use case. These findings highlighted that factors

such as severity based on the modified Vesikari score, lack of rotavirus vaccination, absence of skin turgor, and age positively influenced outcome prediction. Conversely, the absence of vomiting and the presence of six or more loose stools had a negative impact on the outcome prediction. The main disparity between the two outcomes lay in the inclusion of pre-enrolment diarrhea days, information readily available to both models and clinicians upon presentation.

Future research should aim to investigate the observed decline in sensitivity during temporal validation, which may be attributed to baseline differences in disease severity or potential changes in the study population over time. Additionally, since this analysis relied on secondary data from prior enteric studies that were not specifically designed to assess LDD, important predictors—such as history of LDD and HIV status—may have been overlooked. Purpose-built studies are needed to capture these variables and further enhance model accuracy.

5.4.2 Assessing Model Performance for Chronic Malnutrition Prediction

5.4.2.1 Key Findings for RQ5

The predictive capability of the models in the first scenario are shown in Table 5.11. The performance evaluation of the models revealed wide variations across different metrics. Sensitivity, a metric that evaluates the true positive rate of a model, was highest in the RF model at 80.7%, closely followed by the ANN at 79.5%. Other models ranked in decreasing order of sensitivity: SVM at 77.3%, NB at 76.5%, GBM at 75.6%, LR at 75.4%, and the least sensitivity observed in the KNN model at 72.4%. Specificity, which gauges a model's ability to accurately identify negative cases, ranged from 58.2% to 71.8%. The GBM model exhibited the highest specificity at 71.8%, followed closely by RF at 70.1%. LR showed a specificity of 61.9%, while NB and SVM performed similarly at 61.6%, KNN at 61.4%, and ANN had the lowest specificity

at 58.2%. PPV ranged between 27.4% and 34.9%, whereas NPV ranged from 92.3% to 94.8%.

The AUC values, indicative of overall model performance, ranged from 73.4% to 83.5%.

Table 5.11: Model performance in the prediction of linear growth faltering^β using composite data (Original/simulated data).

Algorithm	Sensitivity % [95% CI]	Specificity % [95% CI]	PPV % [95% CI]	NPV % [95% CI]	F1-Score [95% CI]	AUC % [95% CI]	PRAUC % [95% CI]
RF	80.7 [76.5-84.4]	70.1 [68.1-72.1]	34.9 [31.9-38.0]	94.8 [93.6-95.9]	48.7 [16.8-59.7]	82.8 [80.8-84.8]	96.0 [93.8-96.2]
GBM	75.6 [71.2-79.7]	71.8 [69.8-73.7]	34.7 [31.6-37.9]	93.7 [92.4-94.8]	47.6 [13.9-74.5]	83.5 [81.6-85.4]	96.2 [94.9-96.5]
NB	76.1 [71.7-80.1]	61.6 [59.5-63.7]	28.2 [25.6-31.0]	92.8 [91.4-94.1]	40.2 [12.0-42.4]	75.6 [73.3-77.9]	94.0 [92.1-95.0]
LR	75.4 [70.9-79.4]	61.9 [59.7-64.0]	28.2 [25.5-30.9]	92.7 [91.2-94.0]	38.2 [3.1-64.2]	73.7 [71.3-76.1]	93.0 [91.1-94.0]
SVM	77.3 [73.0-81.2]	61.6 [59.5-63.7]	28.6 [25.9-31.3]	93.2 [91.7-94.5]	41.7 [9.3-56.8]	73.4 [71.0-75.8]	93.0 [91.6-94.1]
KNN	72.4 [69.7-75.0]	61.4 [59.3-63.5]	27.6 [25.0-30.3]	92.3 [90.8-93.6]	40.2 [6.7-67.0]	74.8 [72.3-77.2]	93.0 [90.8-93.6]
ANN	79.5 [75.3-83.3]	58.2 [56.1-60.4]	27.4 [24.9-30.0]	93.5 [92.0-94.7]	40.8 [9.8-58.5]	73.6 [71.3-76.0]	93.0 [90.9-94.1]

^β LGF definition: $\Delta \text{HAZ} \geq -0.5$

The GBM model demonstrated the highest AUC at 83.5%, with a 95% CI ranging from 81.6% to 85.4%. The GBM model emerged as the champion model based on its highest discriminative ability (AUC=83.5%), demonstrating robustness and reliability in predicting the target outcome. Comparative analysis of the performance of the 7 models was done using Friedman's test (*Friedman chi-squared* = 20.71, *df* = 6, *p-value*=0.0021), which revealed that the difference in the performance of the models across the different metrics had statistical significance. While the Friedman test showed that at least one model differed significantly in performance compared to the others, it did not identify the specific model. Consequently, post-hoc tests were conducted to determine the pairwise difference between the models using Nemenyi test as shown in Table 5.12.

Table 5.12: Post-hoc analysis of LGF model performance using Nemenyi Test.

	ANN	GBM	KNN	LR	NB	RF
GBM	0.074	-	-	-	-	-
KNN	1.000	0.169				
LR	1.000	0.028	0.994	-	-	-
NB	1.000	0.196	1.000	0.990	-	-
RF	0.074	1.000	0.169	0.028	0.196	-
SVM	0.998	0.258	1.000	0.977	1.000	0.258

* *Friedman chi-squared* = 20.71, *df* = 6, *p-value* <=0.0021

The results of the pairwise differences indicate that the GBM and RF model had a significantly better performance than the LR model ($p < 0.05$). The Nemenyi test results were then plotted to observe patterns in model performance among the 7 models. Groups that share a common letter are not significantly different from each other, while those with different letters are considered to have statistically significant differences in their performance. These results showed three statistically distinct groups based on Nemenyi test results as shown in Figure 5.16. The models in group “a” (GBM and RF) had the highest performance and were significantly different from group “b” (LR). Group “b” (LR) had the lowest performance and was significantly different from both “a” and “ab” groups. Group “ab” (ANN, KNN, NB and SVM) have intermediate performance. They are not significantly different from each other.

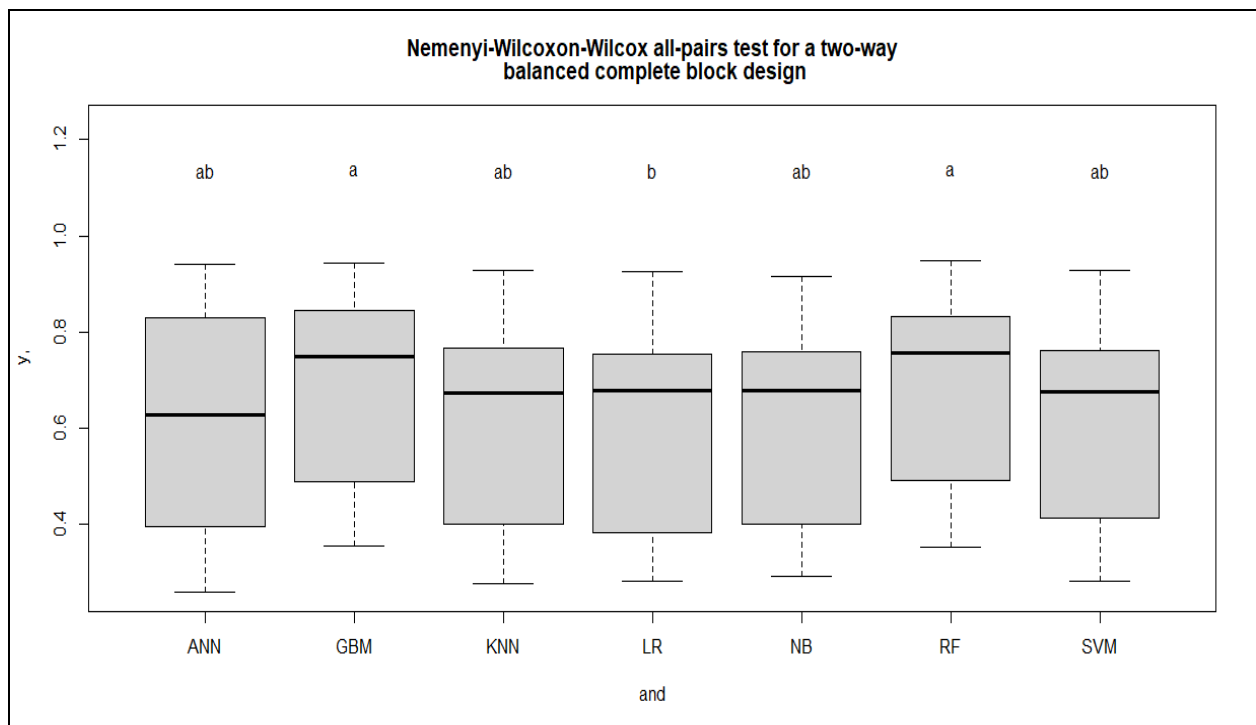


Figure 5.16: Nemenyi Post-hoc Test for Performance Differences among LGF Prediction Models.

The ROC curves for LGF prediction using the primary use case are shown in Figure 5.17. Moreover, in the sensitivity analysis that omitted synthetic data and exclusively utilized the VIDA dataset for development, the model's performance spanned a range of metrics as shown in Table 5.13. Sensitivity ranged from 52.2% to 82.6%, specificity from 55.9% to 78.6%, PPV from 27.3% to 33.7%, NPV from 91.0% to 94.2%, F1-score from 40.3% to 44.3%, AUC from 68.0% to 75.5%, and PRAUC from 90.6% to 94.4%. This sensitivity analysis demonstrated a degradation in the predictive capability of all models, with the exception of the SVM model, which exhibited a marginally increased efficacy.

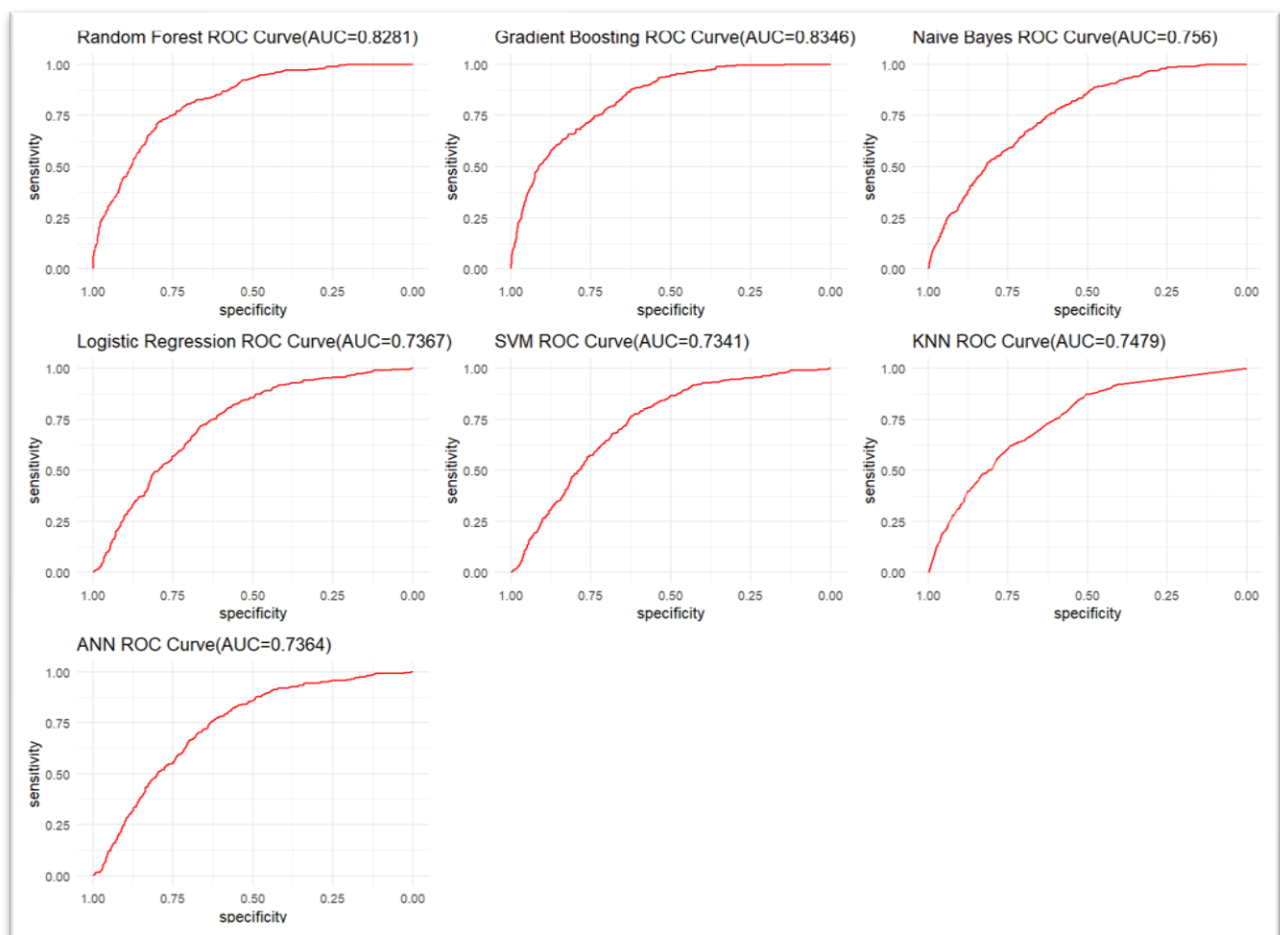


Figure 5.17: Receiver operating characteristic curve for linear growth filtering prediction models

Table 5.13: Model performance in the prediction of linear growth faltering^β using original data only

Algorithm	Sensitivity % [95% CI]	Specificity % [95% CI]	PPV % [95% CI]	NPV % [95% CI]	F1-Score [95% CI]	AUC % [95% CI]	PRAUC % [95% CI]
RF	52.2 [36.9-67.1]	78.6 [72.7-83.7]	32.9 [22.3-44.9]	89.1 [84.0-93.0]	40.3 [10.9-51.4]	70.3 [61.8-78.7]	90.6 [88.4-90.8]
GBM	80.4 [66.1-90.6]	63.3 [56.7-69.6]	30.6 [22.5-39.6]	94.2 [89.2-97.3]	44.3 [12.9-55.8]	75.5 [68.2-82.8]	93.6 [92.3-93.9]
NB	63.0 [47.5-76.8]	75.1 [69.0-80.6]	33.7 [23.9-44.7]	91.0 [86.0-94.7]	43.9 [5.6-60.3]	73.6 [66.1-81.2]	93.0 [91.1-94.0]
LR	73.9 [58.9-85.7]	63.3 [56.7-69.6]	28.8 [20.8-37.9]	92.4 [87.0-96.0]	41.5 [7.8-52.1]	73.8 [67.0-80.5]	93.9 [92.0-94.9]
SVM	71.7 [56.5-84.0]	65.9 [59.4-72.1]	29.7 [21.4-39.1]	92.1 [86.8-95.7]	42.0 [7.2-51.9]	75.2 [68.9-81.5]	94.4 [93.0-95.5]
KNN	82.6 [68.6-92.2]	56.8 [50.1-63.3]	27.7 [20.4-36.0]	94.2 [88.9-97.5]	41.5 [12.2-51.8]	73.1 [66.3-79.9]	93.6 [91.4-94.2]
ANN	82.6 [68.6-92.2]	55.9 [49.2-62.4]	27.3 [20.1-35.5]	94.1 [88.7-97.4]	41.1 [12.0-57.9]	68.0 [60.5-75.6]	91.4 [89.3-92.5]

^β- LGF definition: $\Delta \text{HAZ} \geq -0.5$

During sensitivity assessment applying the alternate definition of LGF (negative ΔHAZ per month of follow-up), the performance of the models varied across different evaluation metrics. Sensitivity ranged from 45.8% to 73.1%, specificity from 53.2% to 76.6%, PPV from 79.0% to 90.5%, NPV from 28.6% to 48.5%, F1-score from 58.3% to 80.9%, AUC from 58.0% to 82.4%, and PRAUC from 29.0% to 62.6%, as presented in Table 5.14 This scenario revealed a decline in model performance across the board, with the exceptions being that KNN exhibited marginal performance gains, and RF maintained the same performance characteristics as seen previously.

Table 5.14: Model performance in the prediction of linear growth faltering^β using composite data.

Algorithm	Sensitivity % [95% CI]	Specificity % [95% CI]	PPV % [95% CI]	NPV % [95% CI]	F1-Score % [95% CI]	AUC % [95% CI]	PRAUC % [95% CI]
RF	73.1 [71.1-75.1]	76.6 [73.1-79.9]	90.5 [88.9-91.9]	48.5 [45.3-51.6]	80.9 [65.2-97.0]	82.4 [80.5-84.2]	62.6 [59.4-65.8]
GBM	71.2 [69.0-73.2]	72.4 [68.7-75.9]	88.7 [87.0-90.2]	45.3 [42.2-48.5]	80.4 [76.6-96.4]	78.0 [76.0-80.1]	53.0 [50.2-55.5]
NB	66.0 [63.9-68.2]	53.2 [49.2-57.2]	81.1 [79.0-83.0]	34.1 [31.1-37.2]	72.8 [58.3-85.8]	62.2 [59.7-64.7]	34.3 [32.8-36.6]
LR	56.9 [54.7-59.2]	54.0 [50.0-58.0]	79.0 [76.7-81.1]	29.3 [26.7-32.0]	64.9 [50.1-89.9]	56.9 [54.3-59.5]	29.0 [26.5-31.9]
SVM	58.5 [56.3-60.8]	60.0 [56.0-63.9]	81.6 [79.4-83.6]	32.3 [29.6-35.1]	68.2 [62.2-94.3]	62.9 [60.4-65.4]	35.6 [32.3-39.0]
KNN	65.0 [62.8-67.1]	73.7 [70.1-77.1]	88.2 [86.4-89.9]	41.0 [38.1-43.9]	74.8 [59.9-97.0]	77.3 [75.2-79.4]	54.0 [51.2-57.3]
ANN	45.8 [43.6-48.1]	65.8 [61.9-69.5]	80.2 [77.7-82.6]	28.6 [26.3-31.0]	58.3 [43.6-85.4]	58.0 [55.5-60.6]	31.0 [29.1-32.8]

LGF definition: negative ΔHAZ per month of follow-up

Across all models, the Brier scores exhibited a range from 0.163 to 0.230, as shown in Table 5.15.

However, the results of Spiegelhalter's p-value test indicate that only the SVM model achieved

proper calibration ($p=0.837$). The calibrated GBM model performed better than its uncalibrated counterpart, achieving an AUC of 84.6%.

Table 5.15: Model calibration in the prediction of linear growth faltering.

Algorithm	Brier Score	Spiegelhalter Z-score	Spiegelhalter p-value
RF	0.190	6.610	<0.001
GBM	0.163	4.624	<0.001
NB	0.190	3.577	<0.001
LR	0.214	4.093	<0.001
SVM	0.213	-0.205	0.837
KNN	0.230	2.258	<0.001
ANN	0.218	6.312	<0.001

The two top-performing models from the primary analysis showed similar patterns in EMA findings, although their predictors showed varying levels of statistical importance across the models (Figure 5.18). Factors such as absence of SAM, absence of skin turgor, lack of rotavirus vaccine, age, high temperature, and respiratory rate emerged as predictors of LGF, as illustrated in Figure 5.18. Likewise, in the sensitivity analysis utilizing the alternative LGF definition, both models displayed a consistent direction of association, albeit with variations in the magnitude of importance. Beyond age, respiratory rate, and temperature, additional predictors of LGF included disease severity, absence of vomiting, baseline stunting, bacterial infection, and normal eyes, as depicted in Figure 5.18.



Figure 5.18: Explanatory profile for the two top-performing linear growth faltering models.

In the business value assessment of the champion model (GBM), the cumulative gains plot demonstrates that by selecting the top 20% of cases according to the model, approximately 60% of the target class (LGF) can be captured (Figure 5.19).

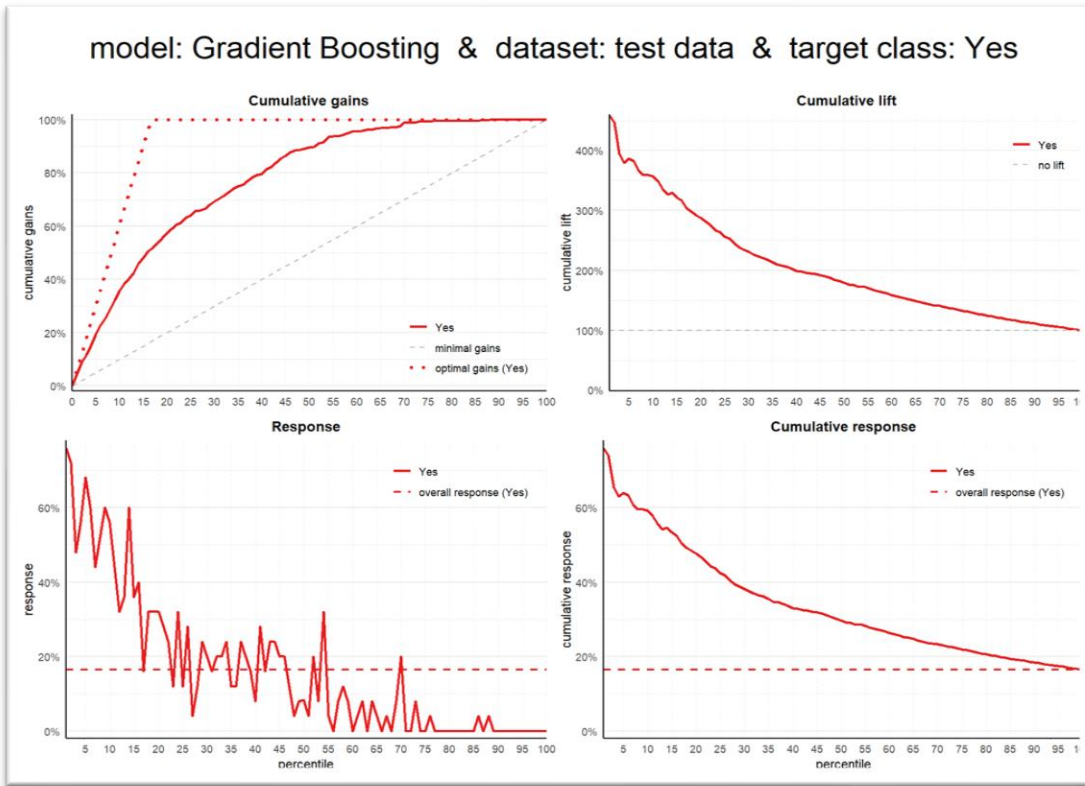


Figure 5.19: Evaluation of the business value of the champion model in linear growth faltering prediction

Moreover, the cumulative lift plot reveals that the champion model is approximately three times more effective at identifying the target class than random selection when considering the top 20% of cases based on predicted probabilities. Finally, the cumulative response plot demonstrates that within the top 20% of observations ranked by model-generated probabilities, 48% belong to the target class, as depicted in Figure 5.19.

The model's performance significantly decreased when temporally validated, evidenced by an approximate 18% drop in the AUC. A similar downward trend was observed across all metrics. Specifically, the GBM model demonstrated the following performance characteristics: a sensitivity of 53.7%, a specificity of 67.7%, a PPV of 32.5%, an NPV of 83.5%, an F1-score of 40.5%, an AUC of 65.6%, and a PRAUC of 86.4% as illustrated in Figure 5.20.

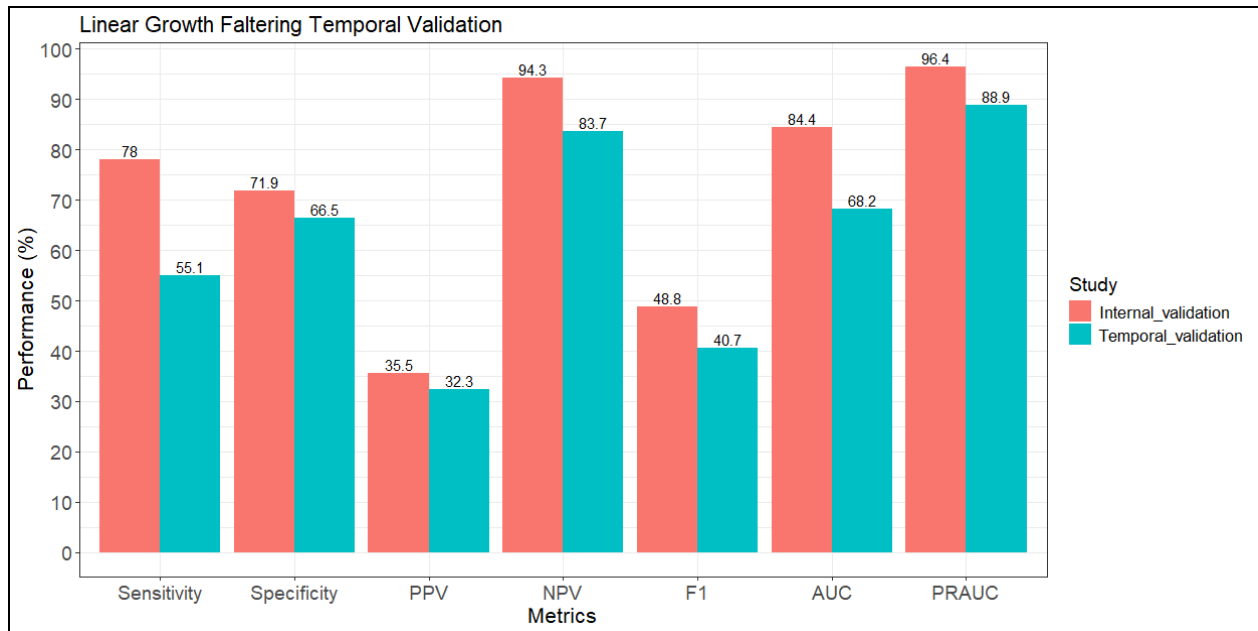


Figure 5.20: Comparative evaluation of model performance in the prediction of linear growth faltering between the development (2015-2018) and temporal validation (2022-2023) cohorts.

5.4.2.2 Discussion: Model Performance in Prediction of Chronic Malnutrition

Approximately, 17 out of every 100 children with MSD, and 22 out of every 100 children with MAD, experience LGF within 2-3 months following a diarrheal episode, confirming a previously established fact that a higher cumulative burden of diarrhea detrimentally impacts the nutritional status of children during early childhood. The tree-based ensemble models demonstrated robust predictive capabilities for LGF, with the GBM model (AUC=83.5%) exhibiting a marginally superior performance to the RF model (AUC=82.8%). This finding was further substantiated by comparative analysis using the Friedman test and Nemenyi Post-hoc test, which identified three distinct clusters based on model performance. The first cluster, comprising RF and GBM, demonstrated the highest performance, followed by ANN, KNN, NB and SVM in the second cluster, having intermediate performance. Logistic regression LR was in the last cluster, exhibiting

the lowest performance. The champion model exhibited superior performance compared to models published by Brander et al. and Ahmed et al., that had AUCs of 67.0% and 75.0%, respectively (Ahmed et al., 2023; Brander et al., 2019). The enhanced model performance stems from the synergistic combination of a robust modeling framework and the inherent predictive strengths of tree-based ensemble algorithms. The champion model's robust discriminatory capacity holds implications for public health interventions, affirming that ML algorithms are feasible and effective in rapidly identifying pediatric populations with elevated likelihood of LGF. This capability facilitates timely deployment of nutritional and healthcare, thereby potentially mitigating adverse outcomes. By enabling targeted screening and offering healthcare personnel with a useful decision-making tool, the model enhances the efficiency of resource allocation. Moreover, it empowers healthcare professionals to provide personalized treatment strategies based on risk assessment profiles, thereby optimizing the effectiveness of interventions and ultimately improving health outcomes for vulnerable populations.

The reason the tree-based ensembles outperformed other models lies in their purposeful design and inherent ability to mitigate the issues of high variance and suboptimal performance often encountered in other ML algorithms. By employing an iterative learning process, these ensembles create multiple individual models and aggregate their predictions into a final meta-model. This approach effectively reduces the impact of random artifacts present in the training data or introduced during the learning process, leading to more robust and generalizable models. Extensive empirical evidence has consistently demonstrated the superior predictive accuracy of tree-based ensembles compared to alternative methods, solidifying their position as powerful and reliable tools for ML tasks (Blockeel et al., 2023). In the healthcare domain, their superior performance has been demonstrated across a wide range of predictive modeling tasks. For instance,

they have achieved remarkable results in breast cancer classification (Ghiasi & Zendehboudi, 2021), COVID-19 death probability profiling (Gonçalves & Rouco, 2020), and both ischemic heart disease classification and adverse event prediction in patients with high blood pressure (Banerjee et al., 2019). These studies provide compelling evidence of the efficacy and versatility of tree-based ensembles as powerful tools for tackling complex prediction problems.

While tree-based ensembles encompass various algorithms, they can be broadly categorized into those built through randomization methods, such as Random Forest, and those constructed using boosting methods, like Gradient Boosting Machine. The strengths and potential reasons behind the discriminative ability of RF were previously discussed in section 5.4.1.2, where it emerged as the champion model for predicting LDD. In this section, focus shifted to the GBM model, which exhibited slightly superior performance compared to RF and serves as the champion model for LGF prediction. GBM has demonstrated superior performance in classification problems because they are effective at identifying intricate relationships within the data, robustness against outliers, and versatility in processing various data formats through various loss functions (Alwanin et al., 2025; Kyrychek, 2025; Natekin & Knoll, 2013). This combination of strengths often leads to higher accuracy compared to other ML algorithms, further enhanced by GBM's minimal pre-processing requirements and extensive hyper-parameter tuning options. Moreover, GBM has demonstrated significant success in its application in the healthcare sector to predict various diseases. Specifically, it has been used in the prediction of: risk of sepsis (Delahanty et al., 2019); readmissions following lumbar laminectomy (Kalagara et al., 2018); delirium risk (Wong et al., 2018); risk of cardiovascular disease (Theerthagiri, 2022; Zhao et al., 2019); length of Stay among inpatients (Zelege et al., 2023); risk stratification on day 5 of COVID-19 patients admitted to the

ICU (Montomoli et al., 2021); mortality prediction in an emergency department (Klug et al., 2020); COVID-19 mortality (Wichmann et al., 2023); Type-2 diabetes risk (L. Wang et al., 2020).

The sensitivity analysis conducted to assess the impact of excluding synthetic data from the model training process for the primary use case of LGF prediction showed that the models' ability to predict LGF weakened across the board, with the exception of the SVM model, which experienced a minor improvement. This shows that the inclusion of synthetic data in model building generally enhanced the performance of most models, while the SVM model demonstrated robustness and effectiveness even without the additional synthetic data points. Synthetic data proves its efficacy in addressing the limitations of class imbalance, scarce, or privacy-constrained real-world data by extending and diversifying the available data resources. (Giuffrè & Shung, 2023; Gonzales et al., 2023). This approach effectively addresses the problem of limited data, enabling more extensive model development and enhancing generalization capabilities. By expanding the data distribution, synthetic data helps overcome biases and ensures greater model fairness, allowing for the accommodation of complex risk factors influencing health outcomes. When employed strategically, synthetic data offers a valuable pathway to strengthening the veracity, applicability, and ethical underpinnings of predictive models utilized in healthcare. This approach lays the groundwork for more targeted and individualized interventions, ultimately contributing to enhanced patient care standards and favorable health outcomes. However, it is crucial to acknowledge the potential risks associated with synthetic data. If biases exist within the primary data, they can be inadvertently propagated and even amplified in the generated synthetic data (Marwala et al., 2023). This underscores the importance of carefully evaluating and mitigating biases in the primary data before utilizing synthetic data generation techniques.

The divergent result observed in the SVM model in this sensitivity analysis could potentially be explained by the following factors; SVMs focus on finding the largest separation between different classes as well as employ regularization techniques, which makes them less susceptible to overfitting on the training data making them robust even with smaller datasets (Bui, 2021; Yousefi et al., 2022). Additionally, SVMs utilize the kernel trick to implicitly map data to higher dimensional spaces, which allows the inner product of mapping function instead of the data points, potentially capturing complex relationships even with smaller datasets (Chang et al., 2025; Maggioni & Spinelli, 2025). However, possibility that the observed improvement in the SVM's performance may be due to random chance or statistical fluctuations cannot be ruled out. Further investigation is needed to better understand this divergent finding.

In the second sensitivity analysis that employed a less stringent cutoff of a negative ΔHAZ to define LGF, saw the prevalence go up from ~17% in the primary analysis to ~29% in this sensitivity analysis. Moreover, there was a significant reduction in the predictive capability of models compared to the primary analysis, which used a decrease of 0.5 HAZ or more as the threshold. This outcome highlights the impact of the chosen LGF definition on model performance. Possible explanations for this observation are that the models are exposed to more noise and minor fluctuations in growth that might not be indicative of true LGF when negative change in HAZ cutoff was used. This increased variability can make it more difficult for models to identify meaningful patterns and distinguish between true cases of LGF and normal growth fluctuations. Moreover, including minor growth dips can dilute the signal of true LGF cases within the data, making it harder for the models to learn the distinctive characteristics associated with actual growth faltering. Interestingly, amidst the general decline, the KNN model demonstrated a slight improvement in performance, while the RF model maintained its performance level from the

primary analysis. This suggests potential differences in the sensitivity of these algorithms to variations in the LGF definition. The KNN model's slight improvement and RF's consistent performance suggest that these models may be more robust to the inclusion of minor growth fluctuations compared to other models in the analysis.

The public health implications of this finding is that using a less stringent LGF definition could lead to over-diagnosis, where children experiencing normal growth fluctuations are incorrectly identified as having LGF. This might result in unnecessary interventions and inefficient allocation of resources. The findings highlight the critical role of the LGF definition in model performance and the need for careful consideration when establishing thresholds for identifying growth faltering. The suitability of cutoff criteria varies depending on the context, and the selection should be guided by clinical knowledge and practicality, considering the specific healthcare environment, the population being studied (including age variations), and the clinical importance of changes in HAZ. Furthermore, from a modelling perspective, it's essential to evaluate models under different LGF definitions and robust models that perform well across various definitions are more desirable for real-world applications. Moreover, complementary strategies that include using clinical predictive tools alongside clinical judgment and additional growth monitoring data can help mitigate the risk of over-diagnosis and ensure appropriate interventions for children truly at increased risk of experiencing LGF.

Despite the documented literature on the calibration of various ML models (Huang et al., 2020; Pavlovic, 2025), relatively high brier scores were observed and only the SVM model achieved proper calibration. The calibration of specific ML models have previously been discussed in section 5.4.1.2, which showed all models had good calibration save for NB and KNN aligning well with literature. Here, the finding of poor calibration among the remaining six models that is

discordant with existing literature is discussed. Possible explanations for this finding is a potentially complex relationships in the LGF prediction task and the underlying data that may not be well-captured by standard model assumptions, leading to poor calibration (Van Calster et al., 2019). Additionally, there was a relatively high-class imbalance in this outcome (1:5), which can hinder calibration, especially for models sensitive to class imbalance. Data quality related issues such as outliers, measurement errors or missing data may also contribute to poor calibration. The finding shows that calibration is not guaranteed as even models known for good calibration might require explicit calibration techniques depending on the dataset and the prediction task. The observed miscalibration highlights the importance of calibrating models before using their probability estimates for decision-making as well as the need to evaluate both discrimination and calibration performance when selecting models for real-world applications. This further reinforces the need to adhere to the TRIPOD guidelines that recommend the reporting of calibration performance particularly in the existing contexts where calibration is assessed far less often than discrimination. It was also observed that the performance of the calibrated GBM model surpassed that of its uncalibrated form yielding an AUC of 84.6%. The improvement achieved through Platt scaling highlights the potential of calibration techniques to enhance model reliability. However, it's crucial to recognize that calibration is not a one-size-fits-all process and requires careful consideration of dataset characteristics, model choices, and calibration methods to achieve optimal results.

The EMA analysis showed several factors consistently linked with a higher probability of LGF subsequent to a diarrheal episode, including the absence of SAM and normal skin turgor. Conversely, certain factors appeared to have a protective effect, reducing the likelihood of LGF development. These included the lack of rotavirus vaccination, older age, elevated temperature,

and increased respiratory rate. Identifying these protective factors offers valuable insights into potential preventive measures or interventions that could mitigate the risk of LGF in pediatric patients. A comprehensive discussion of these predictors and their alignment with existing knowledge in the field is presented in section 5.2.2, emphasizing the importance of these factors in predicting LGF risk following diarrheal illness. Understanding these associations can inform targeted interventions and guide the development of strategies to prevent LGF and improve child health outcomes.

The business value evaluation further underscored the practical utility of the GBM model in identifying children at heightened risk of LGF following a diarrheal episode. The cumulative gains plot reveals that by targeting the top 20% of cases as predicted by the model, 60% of actual LGF cases can be effectively captured. This highlights the model's effectiveness in prioritizing children with high-risk profiles for intensive monitoring and more effective management strategies optimizing resource allocation and maximizing impact of intervention strategies like early nutritional and healthcare interventions. Furthermore, the cumulative lift plot highlights the GBM model's superior performance compared to random selection. Within the top 20% of cases identified by the model's predicted probabilities, it successfully captures three times more LGF cases than random selection alone. This demonstrates the model's effectiveness in pinpointing high-risk individuals, leading to a substantial improvement in case identification efficiency. The cumulative response plot further reinforces the model's precision by revealing that 48% of cases within the top 20% of predicted probabilities actually belong to the target class (LGF). This high precision rate signifies that the model's predictions of high LGF risk are highly reliable, minimizing the occurrence of false positives. Consequently, resources can be confidently directed towards children who genuinely require intervention and support. Collectively, these findings

underscore the exceptional utility and value of the GBM model in predicting LGF. It demonstrates both strong discriminatory power and substantial practical value by effectively identifying at-risk children and enabling the optimization of intervention strategies. This ultimately contributes to improved resource allocation, targeted interventions, and better child health outcomes.

Temporal validation, however, revealed a significant decline in model performance, with an approximate 18% reduction in AUC. This emphasizes the ever-evolving nature of health data and the crucial requirement for ongoing model assessment and adjustment. While this decline aligns with findings reported by Ahmed et al. (Ahmed et al., 2023), It highlights crucial aspects that must be taken into account regarding the long-term reliability of predictive models in evolving healthcare landscapes. Several factors may have contributed to the observed performance decline during temporal validation as well as the proposed solutions, as discussed in section 5.2.1.2. These include temporal shifts in the study population characteristics, variations in diarrhea severity between the development and validation cohorts, and potential differences in the case-mix. Furthermore, variations in the predictive influence of model features across the two cohorts, or a complex interplay of these factors, could also have played a role in the reduced performance. These findings highlight the challenges associated with maintaining consistent predictive accuracy over time and emphasize the importance of ongoing model monitoring, periodic retraining and the integration of dynamic, agentic AI approaches to adaptively sustain optimal performance. The dynamic nature of health data necessitates a proactive approach to model maintenance, ensuring that predictive models remain relevant and effective in addressing evolving healthcare needs.

5.4.3 Evaluating Model Performance for In-hospital Mortality Prediction

5.4.3.1 Key Findings for RQ6

Existing diarrhea scores showed limited ability to predict mortality; dehydration score (AUC: 67.9), Vesikari score (AUC: 50.3), and IMCI danger signs (AUC: 61.3) as shown in Table 5.16.

Table 5.16: Predictive capability of existing diarrheal scores in the prediction of in-hospital mortality among children aged < 60 months hospitalized with AGE.

Mortality Prediction					
Score	Sensitivity % [95% CI]	Specificity % [95% CI]	PPV % [95% CI]	NPV % [95% CI]	AUC % [95% CI]
Dehydration	70.7 [54.5-83.9]	65.0 [60.8-69.1]	13.6 [9.3-19.0]	96.6 [94.2-98.2]	67.9 [60.5-75.2]
Vesikari score	61.0 [44.5-75.8]	37.5 [33.3-41.7]	7.1 [4.6-10.2]	92.5 [88.1-95.6]	50.3 [41.6-58.9]
IMCI danger signs	46.3 [30.7-62.6]	69.2 [65.1-73.1]	10.5 [6.4-15.9]	94.3 [91.5-96.4]	61.3 [53.2-69.4]

Evaluation of the 7 models built using the selected and tentative features revealed variations in sensitivity and specificity across different algorithms as shown in Table 5.17. The RF and NB models demonstrated the highest sensitivity, both at 78.0%, followed by the ANN at 75.6%, LR at 73.2%, GBM at 61.3%, and lastly, KNN at 46.3%. Specificity ranged from 71.7% to 78.7%, with the ANN model exhibiting the highest value at 78.7%, followed by RF (76.6%), SVM at 74.1%, GBM at 74.0%, NB at 73.4%, KNN at 72.8%, and LR with the lowest at 71.7%. PPV ranged from 11.0% to 21.7%, while NPV displayed a range of 94.6% to 97.8%. AUC scores followed a similar pattern, with RF and ANN achieving the highest values (82.6% and 82.6%, respectively), followed by SVM (80.2%), NB (79.3%), LR (74.3%), and lastly, GBM and KNN with 56.5%.

Table 5.17: Model Performance in the prediction of in-hospital mortality among children aged < 60 months hospitalized with AGE.

In-hospital Mortality					
Algorithm	Sensitivity % [95% CI]	Specificity % [95% CI]	PPV % [95% CI]	NPV % [95% CI]	AUC % [95% CI]
RF	78.0 [62.4-89.4]	76.6 [72.8-80.2]	20.6 [14.6-27.9]	97.8 [95.9-99.0]	82.6 [77.1-88.1]
GBM	61.0 [44.5-75.8]	74.0 [70.0-77.0]	15.4 [10.2-21.9]	96.0 [93.7-97.7]	74.3 [66.9-81.7]

NB	78.0 [62.4-89.4]	73.4 [69.4-77.1]	18.6 [13.11-25.2]	97.7 [95.7-99.0]	80.2 [73.8-86.5]
LR	73.2 [57.1-85.8]	71.7 [67.6-75.5]	16.8 [11.6-23.1]	97.2 [95.0-98.6]	79.3 [72.3-86.4]
SVM	78.0 [62.4-89.4]	74.1 [70.2-77.8]	19.0 [13.4-25.8]	97.7 [95.8-99.0]	82.0 [76.8-87.2]
KNN	46.3 [30.7-62.6]	72.8 [68.8-76.6]	11.7 [7.2-17.7]	94.6 [91.9-96.6]	56.5 [47.4-65.7]
ANN	75.6 [59.7-87.6]	78.7 (75.0-82.1)	21.7 (15.2-29.3)	97.6 [95.7-98.9]	82.6 [77.2-88.0]

Comparative analysis of the performance of the 7 models was done using Friedman's test (*Friedman chi-squared* = 24.61, *df* = 6, *p-value*=0.0004026), which revealed that the difference in the performance of the models across the different metrics had statistical significance. While the Friedman's test showed that at least one model differed significantly in performance compared to the others, it did not give insights about which specific model. Consequently, post-hoc tests to determine the pairwise difference between the models using Nemenyi test as shown in Table 5.18.

Table 5.18: Post-hoc analysis of mortality model performance using Nemenyi Test.

	ANN	GBM	KNN	LR	NB	RF
GBM	0.1919	-	-	-	-	-
KNN	0.0172	0.9758	-	-	-	-
LR	0.2592	1.00000	0.9485	-	-	-
NB	0.9758	0.7224	0.1919	0.8071	-	-
RF	0.9995	0.0651	0.0036	0.0961	0.844	-
SVM	0.9999	0.3391	0.0429	0.4299	0.9972	0.9906

Friedman chi-squared=24.61, *df*=6, *p-value*=0.0004026

The results of the pairwise differences indicate that the RF and ANN models had significantly better performance than the KNN model ($p < 0.05$). Nemenyi test results were then plotted to observe patterns in model performance among the 7 models. Groups that share a common letter are not significantly different from each other, while those with different letters are considered to have statistically significant differences in their performance. These results showed three statistically distinct groups based on Nemenyi test results as shown in Figure 5.21. The RF, ANN and SVM models (group "a") had the highest performance and were significantly different from KNN (group "b"). KNN (group "b") had the lowest performance and was significantly different

from both "a" and "ab" groups. Group “ab” (GBM, LR and NB) had intermediate performance. The models in these group were not significantly different from each other.

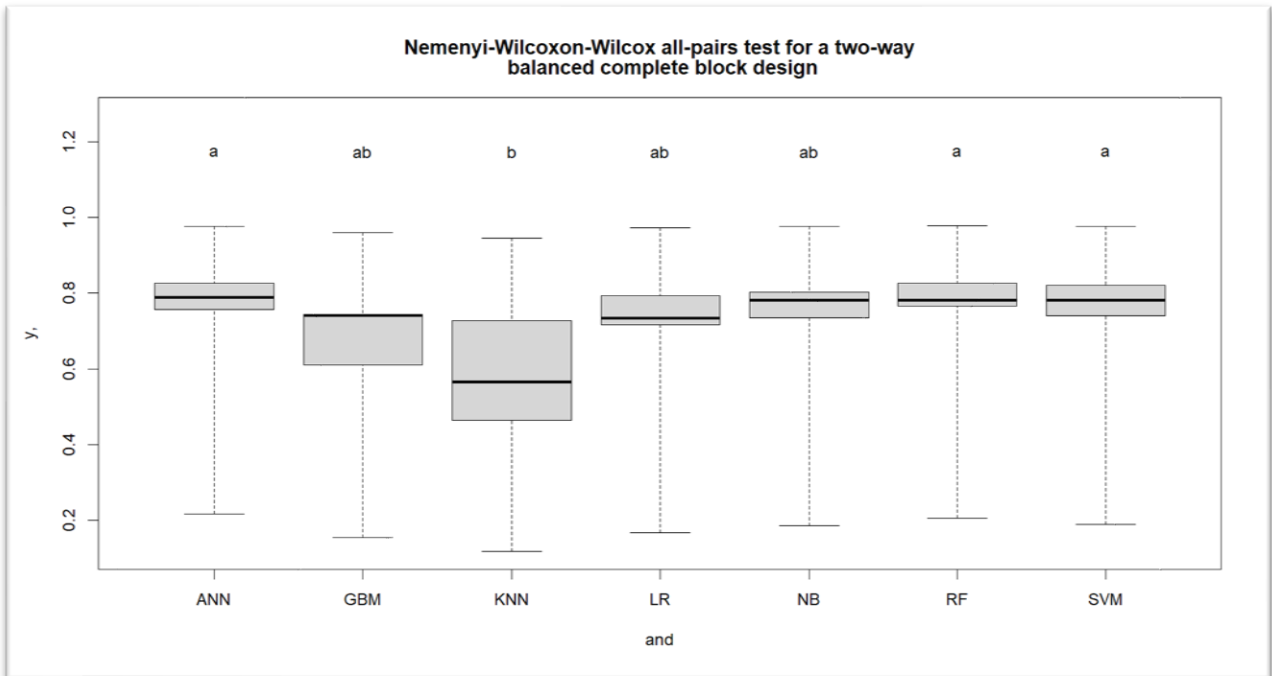


Figure 5.21: Nemenyi Post-hoc Test for Performance Differences among Mortality Prediction Models.

The ROC analysis of the prediction models is shown in Figure 5.22. The stacked ensemble demonstrated a sensitivity of 80.5%, a specificity of 73.6%, a PPV of 19.2%, and an NPV of 98.0%. Among the individual models, both RF and ANN models demonstrated the highest AUC values, indicating superior overall performance in discriminating between the classes.

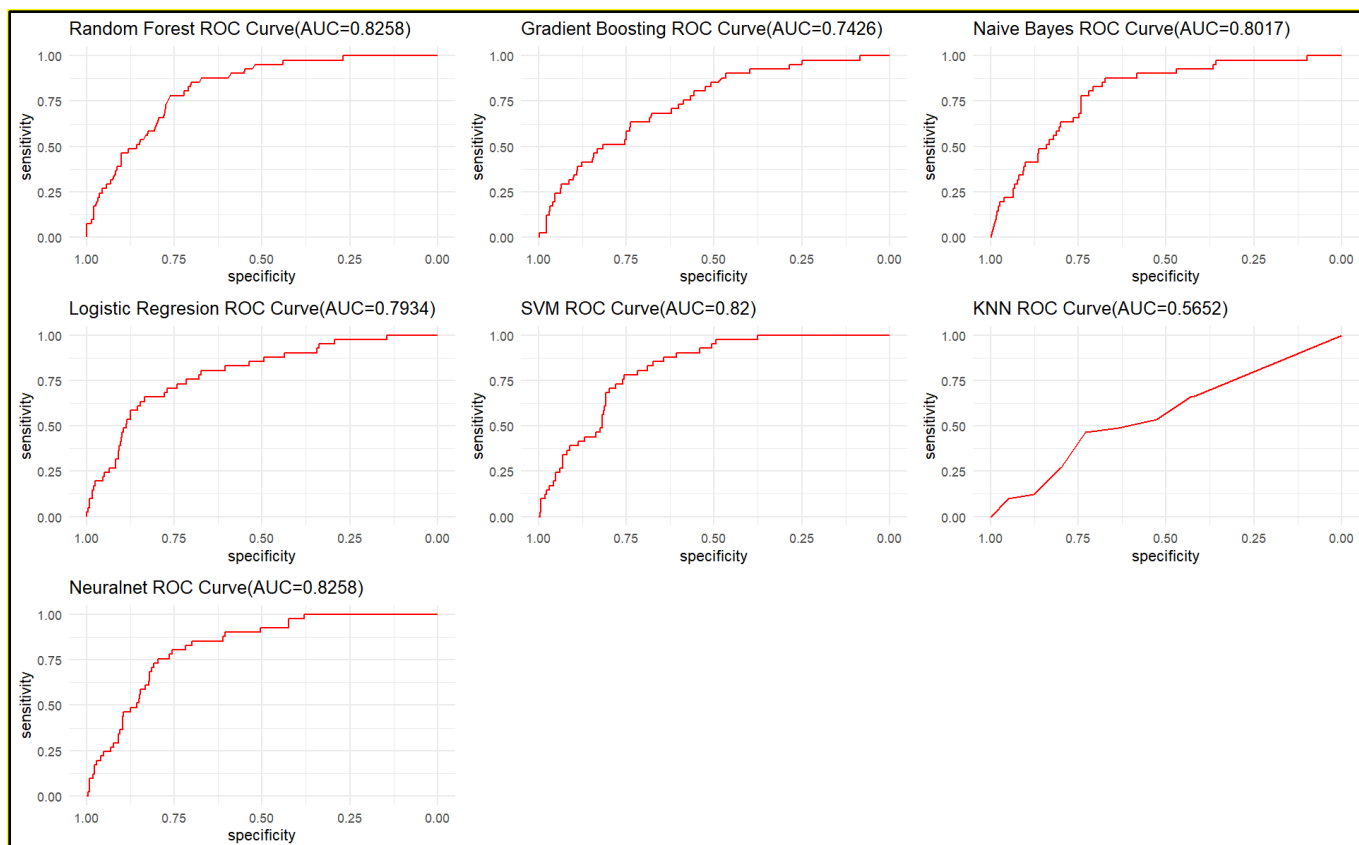


Figure 5.22: ROC analysis of the 7 models in the prediction of in-hospital mortality prediction

Model calibration was assessed using Brier scores and Spiegelhalter's z-test as shown in Table 5.19. While the Brier scores, ranging from 0.16 to 0.437, suggested acceptable calibration, the Spiegelhalter's test revealed statistically significant poor calibration ($p < 0.05$) for six of the models: RF, NB, LR, SVM, KNN, and ANN. Calibration of the RF model resulted in consistent sensitivity and NPV, both remaining at 78.0% and 97.8%, respectively. However, a slight decrease was observed in specificity (75.9%) and PPV (20.1%).

Table 5.19: Calibration results of prediction models applied to test data.

Algorithm	Brier Score	Spiegelhalter Z-score	Spiegelhalter p-value
RF	0.160	-6.623	<0.001
GBM	0.184	0.023	0.981
NB	0.216	19.594	<0.001
LR	0.199	3.893	<0.001
SVM	0.180	-3.358	0.001
KNN	0.217	20.287	<0.001
ANN	0.437	30.971	<0.001

Explanatory model analysis (EMA) was conducted for four top-performing models to understand the influence of features on mortality prediction. The analysis revealed that chest indrawing, nasal flaring, stunting, wasting, and a negative malaria test result consistently exhibited a negative association with mortality across all four high-performing models (RF, ANN, SVM, and NB) (Figure 5.23). Furthermore, the Vesikari score and dehydration were negatively associated with mortality in the ANN model, while the pulse rate showed a negative association in the SVM model.

Focusing on the champion RF model, SHAP attributions provided more granular insights into feature importance. The top contributors, in descending order of impact on reducing the predicted probability of mortality: chest indrawing (-0.08), nasal flaring (-0.07), wasting (0.03), stunting (-0.03), malaria negative result (-0.02), sunken eyes (0.02), dehydration (0.02), oxygen saturation (0.02), restless (0.01) and skin turgor (0.01).

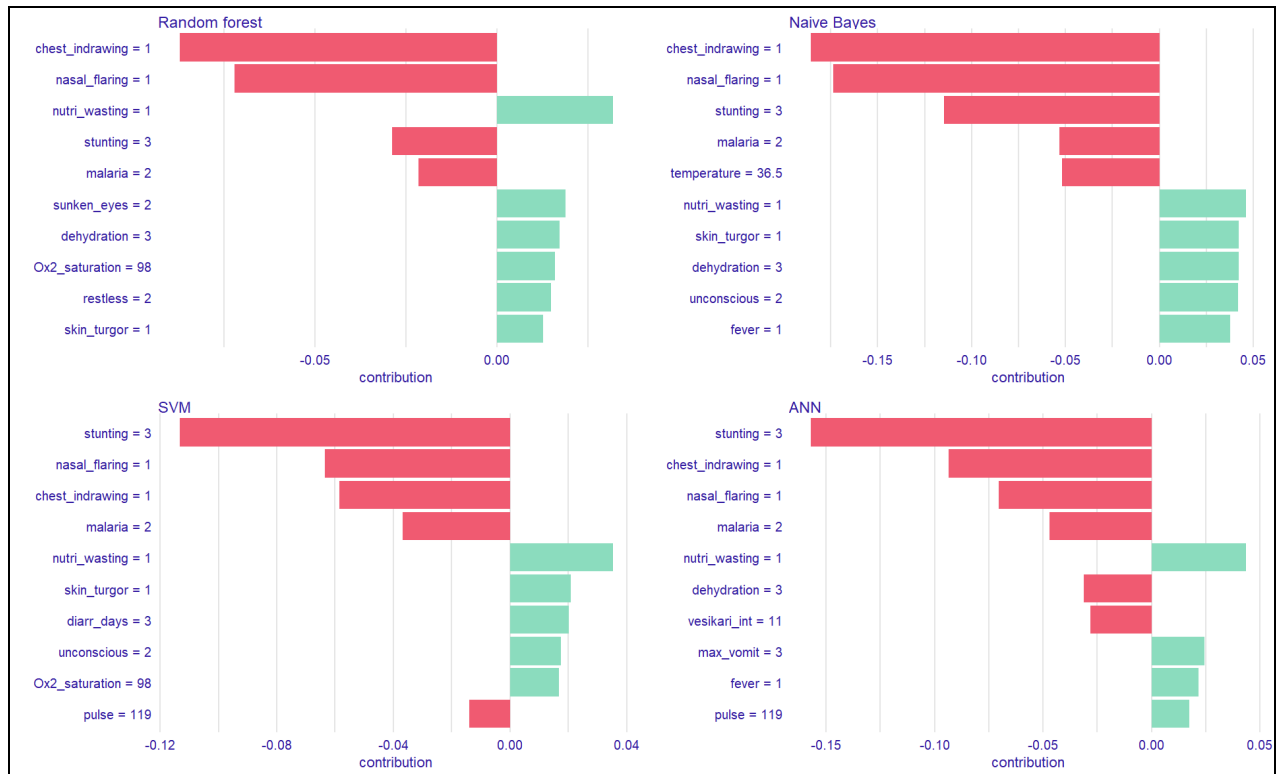


Figure 5.23: Explanatory profile for the top-performing 4 models in the prediction of in-hospital-mortality.

Despite achieving comparable AUC values, the ANN model underperformed in the business value evaluation compared to the RF model, which was ultimately selected as the champion model. The cumulative gains plot for the RF model demonstrates its ability to capture 67% of actual mortality cases when focusing on the top 25% of cases as predicted by the model (Figure 5.24). Furthermore, the cumulative lift plot reveals that the RF model identifies 2.8 times more true mortality cases within the top 25% of predicted probabilities compared to random selection. Finally, the cumulative response plot indicates that 20% of cases within the top 25% of predicted probabilities belong to the target class (mortality).

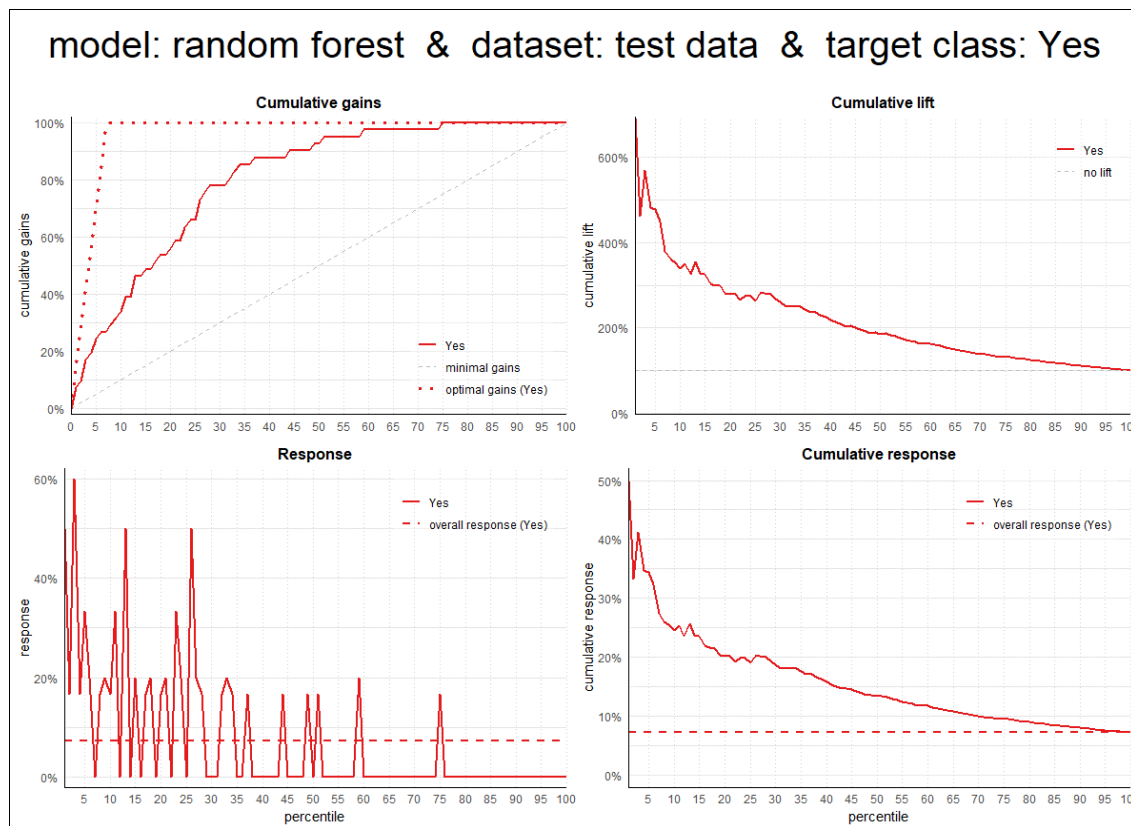


Figure 5.24: Evaluation of the business value of the champion model in in hospital-mortality prediction

5.4.3.2 Discussion: Model Performance in Mortality Prediction

The observed in-hospital mortality rate in this study was approximately 7% among children aged less than 60 months admitted with AGE. While the ML models developed to predict mortality risk demonstrated varying performance across different algorithms, the results highlight the potential utility of such models, particularly the RF algorithm that achieved an AUC of 82.6%, in the identification of pediatric patients with increased likelihood of mortality. The observed variation in performance likely stems from the inherent differences in the algorithms' approaches to classification tasks. Each algorithm employs distinct strategies for pattern recognition and trend identification, influenced by its underlying principles and purpose (IBM, 2021). Despite these

variations, the promising performance of the RF model, consistent with findings from other studies (Uddin et al., 2019), reinforces the capabilities of ML to strengthen clinical decision [*making*] and enhance patient outcomes in pediatric AGE. Furthermore, comparative analysis using Friedman test and Nemenyi Post-hoc test highlighted RF model to outperform the KNN models in prediction accuracy and discriminative ability having a statistically significant difference in the pairwise comparison. Additionally, while its performance over the GBM, LR and NB models was not statistically significant, it was distinctively higher.

Although the literature review did not reveal any prior attempts to predict in-hospital mortality specifically among pediatric patients with acute gastroenteritis (AGE), it did provide valuable insights into mortality prediction for various other diseases. Importantly, the top-performing model (RF), exhibited superior performance compared to a majority of these studies: Mortality among patients with bloodstream infections with an AUC of 80.0% (Gutiérrez-Gutiérrez et al., 2016); Mortality among patients with infective endocarditis with an AUC of 79.0% (Fernandez-Felix et al., 2021); All-cause mortality among veterans in the USA with AUC ranging between 73.0% to 74.0% (DeSalvo et al., 2005); Mortality among COVID-19 patients in the Netherlands with AUC ranging between 81.0% to 82.0% (Ottenhoff et al., 2021); Mortality among older patients in Ireland with an AUC of 78.0% (Matthews et al., 2022); Mortality among patients with ischemic stroke in Spain with an AUC of 74.2% (Lea-Pereira et al., 2022); Mortality among Mexican patients with COVID-19 with an AUC of 82.3% (Bello-Chavolla et al., 2020). Among the myriad studies examined, only two reported an AUC greater than 82.6%, a benchmark achieved by the model, namely: Mortality in ICU patients in USA with an AUC of 92.8% (Chiu et al., 2022); Mortality among sepsis patients in USA with an AUC of 85.3% (Alnsour et al., 2019).

These findings highlight the good predictive capability of our model in predicting in-hospital mortality compared to a broad spectrum of mortality prediction studies, positioning it among the top performers in this domain. The superior performance of the RF model in predicting in-hospital mortality among pediatric AGE patients shows that advanced ML algorithms can provide highly accurate risk assessments early enough, allowing for timely and targeted interventions. Furthermore, by accurately predicting which children are at higher risk of mortality, clinicians can allocate resources more efficiently. This includes prioritizing critical care resources, optimizing staffing, and ensuring that high-risk patients receive the necessary medical attention and support. Moreover, the finding that the RF model outperforms many existing models in different contexts provides a benchmark for other studies and encourages continuous improvement in predictive modeling. Additionally, the RF model if validated could provide policymakers with insights to develop guidelines and protocols that incorporate ML models for risk prediction enabling standardized care practices and improvement in overall patient outcomes at a systemic level.

The robust predictive power of the RF model and its successful applications in various healthcare contexts have been extensively discussed in Section 5.4.1.2. The findings for this diarrheal outcome provide further evidence supporting the efficacy and versatility of RF models in addressing complex medical prediction challenges.

Despite extensive literature on calibrating various ML models (Davis et al., 2017; Huang et al., 2020; Pavlovic, 2025), this study observed relatively low Brier scores (0.160-0.437), indicating acceptable calibration. However, Spiegelhalter's test revealed statistically significant poor calibration ($p < 0.05$) for six models: RF, NB, LR, SVM KNN, and ANN. These findings highlight the necessity of using multiple metrics for comprehensive calibration evaluation and suggest the need for calibration refinement in the identified models. Several possible explanations for these

findings have been advanced as detailed in section 5.4.2.2, including the potentially complex relationships within the data, a high class imbalance (1:13), and data quality issues. These observations reaffirm that calibration is not inherently guaranteed; even models traditionally known for good calibration may require explicit techniques depending on the dataset and prediction task. This highlights the critical importance of calibrating models before utilizing their probability estimates for decision-making and emphasizes the need to evaluate both discrimination and calibration performance when selecting models for real-world applications.

The EMA analysis identified similar predictors to the feature selection even though the importance of the features and direction of the associations were different. The factors can be broadly categorized into: diarrhea-specific severity features, danger signs, vital signs and co-infections. Prior research has established these as determinants of mortality (Acácio et al., 2019; Hartman et al., 2022; Lee et al., 2014; Walker et al., 2013; Wierzba & Muhib, 2018; Zenebe et al., 2022), and have been elaborately discussed in section 5.3.2.

The business value evaluation demonstrates the significant clinical value of the RF model in the identification of children with a heightened risk of death. The cumulative gains plot reveals that by targeting interventions on the top 25% of cases predicted to be at highest risk by the model, 67% of actual deaths can effectively be captured. This finding highlights the model's potential to improve patient outcomes by enabling targeted interventions including critical care and intensive monitoring for those most in need potentially reducing mortality rates and improving patient survival. Moreover, the cumulative lift plot emphasizes the significant advantage of the RF model over random selection. By focusing on the top 25% of cases predicted to be at highest risk, the model identifies 2.8 times more actual deaths than would be expected by chance alone. This demonstrably superior performance highlights the model's ability to effectively identify high-risk

individuals and offers a compelling argument for its use in clinical practice. The cumulative response plot further reinforces the model's precision by revealing that 20% of cases within the top 25% of predicted probabilities, actually belong to the target class (death). Collectively, these findings underscore the exceptional utility and value of the RF model in predicting in-hospital mortality. It demonstrates both strong discriminatory power and substantial practical value by effectively identifying at-risk children and enabling the optimization of intervention strategies. This ultimately contributes to improved resource allocation, targeted interventions, and better child health outcomes. Further research and implementation of the model in clinical settings are warranted to fully realize its potential.

5.5 Product Development and Deployment

5.5.1 Key Findings for RQ7

To address the final research question (RQ7), a proof-of-concept online dashboard that incorporated the champion models for each of the poor diarrheal outcomes was developed. The R Shiny package was leveraged, complemented by *Shinydashboard* in creating an interactive web-accessible dashboard with minimal web development. The dashboard accepts user input and provides prediction outputs visualized as risk gauges indicating the risk profile of the child and SHAP plots, that helped to improve the clarity of how the models arrive at their decisions.

The summary of the methodology and the performance of the champion models are summarized in Table 5.20. The champion models for the three diarrheal outcomes achieved good discriminative ability achieving AUCs ranging between 82.6%-84.3%. The robust predictive performance observed, particularly with the RF algorithm excelling in predicting LDD and in-hospital

mortality, as well as the GBM model demonstrating superiority in predicting LGF, suggests the potential utility of ML in this context.

Table 5.20: Summary of Methodology and Model Performance in the Prediction of Poor Diarrheal Outcomes.

	Diarrheal Outcome		
	LDD	LGF	Mortality
Development Cohort: Year of study	VIDA study (n=1,482): 2015-2018	VIDA study + Synthetic data (n=10,000): 2015-2018	Rotavirus Inpatient Surveillance (n= 2,271): 2010-2020
Temporal validation Cohort: Year of study	EFGH Study (n=685): 2022-2023	EFGH Study (n=655):2022-2023	-
Age Category	0-59 months	6-35 months	0-59 months
Missing Values	Imputed using MICE package		
Split-sampling	75%:25%		
Sub-sampling for Class Imbalance	Over-sampling technique		
Feature Selection	Conducted using Boruta Package		
Number of Potential Predictors	68	68	39
Number of Predictors Selected	11	7	19
Features Selected	Diarrhea days; Vesikari score; Age group; Vomit days; breastfeed ; respiratory rate; Vomit; Frequency vomiting; Rotavirus vaccination; Rectal straining ; Stool frequency; Skin turgor	Age; Temperature; Respiratory rate; Severe acute malnutrition (SAM); Rotavirus vaccination; Breastfeeding ; Skin turgor	AVPU scale; Vesikari score; Dehydration; Sunken eyes; Skin pinch; Frequency vomiting; Unconsciousness; Wasting; Vomiting; Pulse; Fever; Sunken fontanelle; Restless; Nasal flaring; Diarrhea days; Stridor; <90% Oxygen saturation; Chest indrawing; Malaria and Stunting
Algorithms Used	Random Forest; Gradient Boosting; Naïve Bayes; Logistic Regression; Support vector machine; K-nearest neighbors; Artificial Neural Networks		
Performance Metrics Reported	Sensitivity; Specificity; Positive Predictive Value; Negative Predictive Value; F1-score; Area under the Curve; Precision Recall Area under the Curve		
Champion Model (AUC%[95%CI])	Random Forest (83.0 [78.6-87.5])	Gradient Boosting (83.5 [81.6-85.4])	Random Forest (82.6 [77.1-88.1])
Explanatory Model Analysis	SHAP attribution estimated using DALEX package for models with AUC $\geq$ 80%		
Business Value Evaluation	2.6 times better than random chance	3 times better than random chance	2.8 times better than random chance

Despite ML demonstrating exceptional predictive capabilities as was witnessed in the study, a substantial disparity exists between the development of these models and their seamless integration into clinical practice (González-Gonzalo et al., 2022). The primary challenge facing ML in healthcare has moved from developing models to effectively deploying model (Zhang et al., 2022). Specifically, lack of transparency or trustworthy AI has been identified as one of the main barriers to implementation (González-Gonzalo et al., 2022; Markus et al., 2021).

To bridge these gaps within the context of childhood diarrhea, champion ML models for each poor diarrheal outcome were deployed in a proof-of-concept online prediction dashboard that not only delivers predictions but also incorporates explanatory model analysis (EMA) to foster trust and facilitate the practical application of ML insights in clinical decision-making.

The online dashboard's architecture, illustrated in Figure 5.25, is designed to streamline the prediction process for diarrheal outcomes. User data is inputted through a consolidated interface and preprocessed for compatibility with three pre-trained champion models. Each model, utilizing relevant predictors, generates predictions that are then analyzed by saved explainer objects to offer understanding into the model's reasoning. Finally, the risk profile of the patient together with the SHAP attribution plot for each diarrheal outcome are sent back to the user interface.

R Shiny package was utilized to develop a web-accessible dashboard (https://bogwel.shinyapps.io/Poor_diarrheal_outcomes/), which allowed users to input data into the designated fields and dynamically displayed results upon submission. To streamline data entry and avoid duplication, the input fields for the three diarrheal outcomes under investigation were consolidated into a single interface, as shown in Figure 5.26.

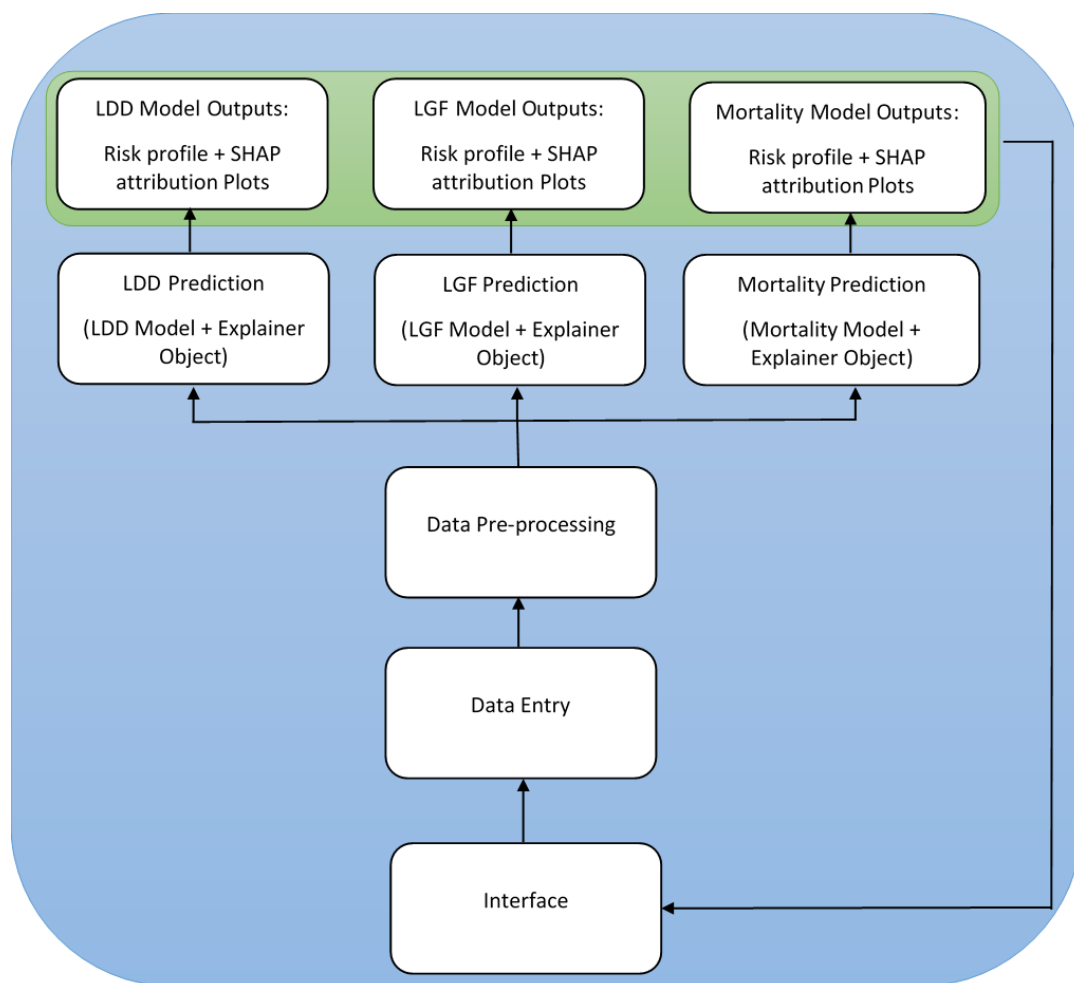


Figure 5.25: Online Dashboard Architecture

The LDD, LGF, and mortality models utilized 10, 6, and 19 predictors, respectively. In contrast, the consolidated interface incorporated a total of 31 predictors. Additionally, the severity of diarrheal illness was calculated using the Vesikari score, which was computed on the backend based on the inputted data. After submitting user inputs (Figure 5.27), pre-trained models for each outcome generated predictions, displayed as probabilities on a risk gauge, along with a risk categorization. SHAP attribution plots, generated from saved explainer objects, provided insights into each models' decision-making processes. Figures 5.28, 5.29, and 5.30 illustrate example prediction outputs for LDD, LGF, and mortality outcomes, respectively.

Clinical Prediction Tool for Poor Diarrheal Outcomes

Data Entry and Output Interface

Enter the Patient's age in months: 0

Number of loose stool: 3

Diarrhea days: 3

Vomiting: No

Maximum number of vomits in last 24 hours: 0

Vomiting days: 0

Temperature: 38.5

Individual Risk Profile 1: Longer Duration Diarrhea

Explanatory Model Analysis 1: Longer Duration Diarrhea

Respiratory Rate: 0

Oxygen Saturation: 0

Pulse: 0

Rotavirus Vaccination: Not vaccinated

Skin Turgor: Normal

Wrinkled Skin: No

Unconscious: Yes

Fever: Yes

Malaria: Yes

Individual Risk Profile 2: Linear Growth Faltering

Explanatory Model Analysis 2: Linear Growth Faltering

Straining: Normal

Wasting: Normal

Chest Indrawing: Yes

Stridor: Yes

Nasal flaring: Yes

Admission recommended: No

Oral rehydration given: No

IV rehydration given: No

Explanatory Model Analysis 3: Mortality

Figure 5.26: Consolidated User Interface for Poor Diarrheal Outcomes Dashboard.

The `reactive{ }` functions were incorporated into the dashboard to ensure the application remained dynamic, automatically refreshing its display whenever user inputs changed.

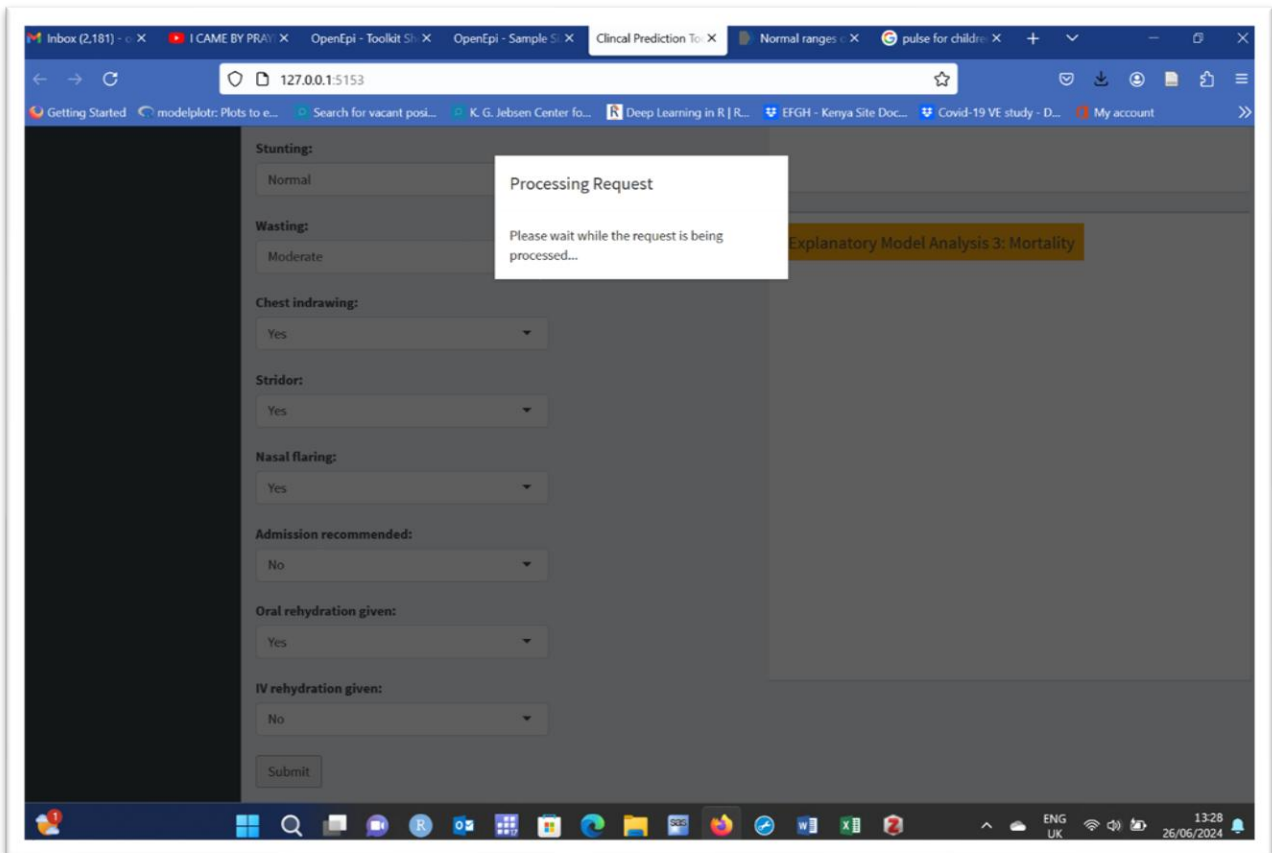


Figure 5.27: Processing of Request upon Submission of User Inputs

This dashboard, developed as a proof of concept, demonstrates the potential of a web-based application to streamline data entry, facilitate prediction modeling, and provide insights into diarrheal outcomes. While currently limited in scope and functionality, it serves as a foundational prototype for future development. Further iterations will focus on expanding the dataset, incorporating user feedback, and refining the model performance to enhance the application's utility and impact. While a formal evaluation of the dashboard's usability has not yet been

conducted, future development will prioritize incorporating user feedback to make enhancements to the application's interface and the user experience as a whole. More specifically, future efforts will evaluate usability, acceptability, and integration among healthcare workers; assess the clinical impact on decision-making and patient outcomes; and examine health system and societal economic impacts through a mixed-methods approach, including a prospective quasi-experimental before–after study, quantitative surveys, and focus group discussions with healthcare workers. Complementing this approach, decision-analytic modeling will also be used to estimate the clinical and health economic impact of implementing the mortality prediction model given the relative rarity of mortality events.

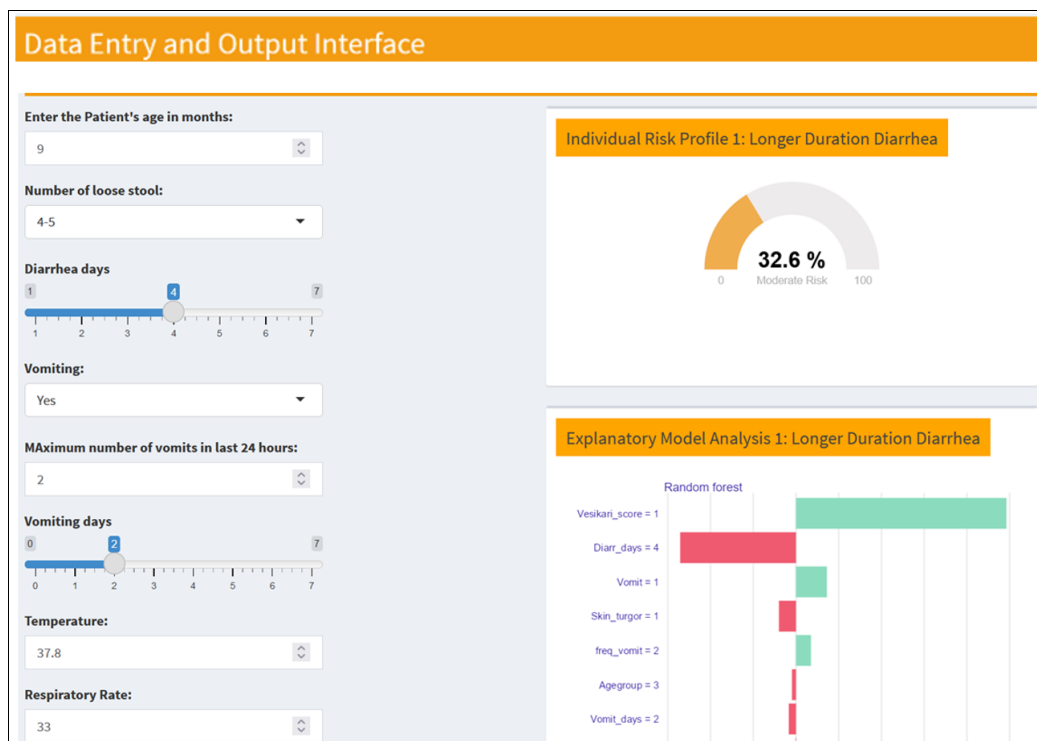


Figure 5.28: Prediction Outputs for Longer Duration Diarrhea.

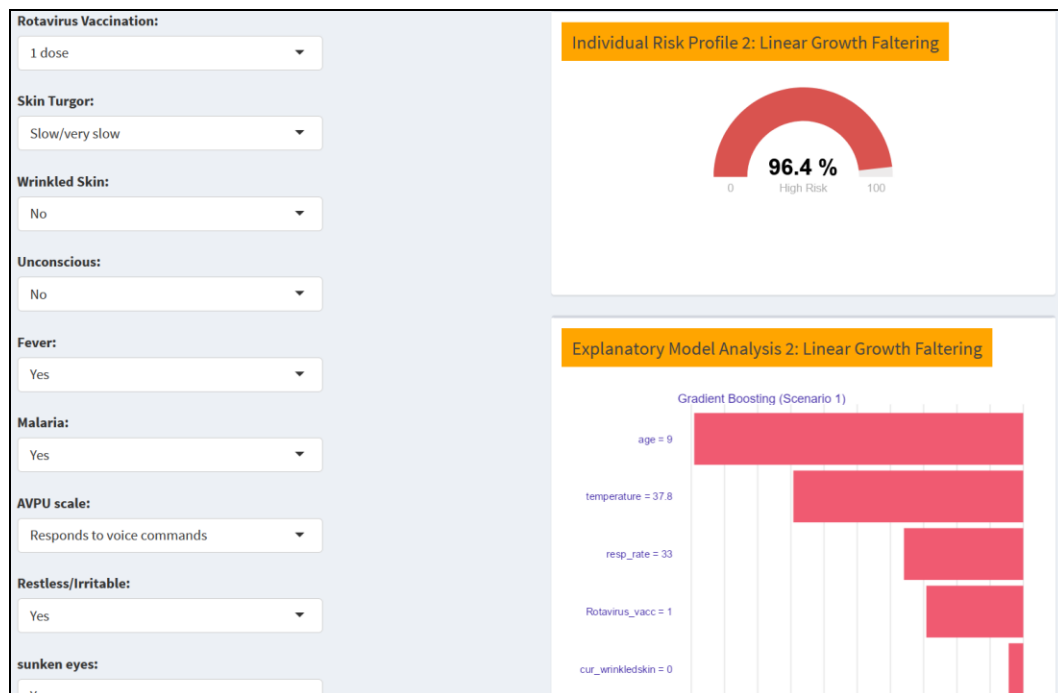


Figure 5.29: Prediction Outputs for Linear Growth Faltering.

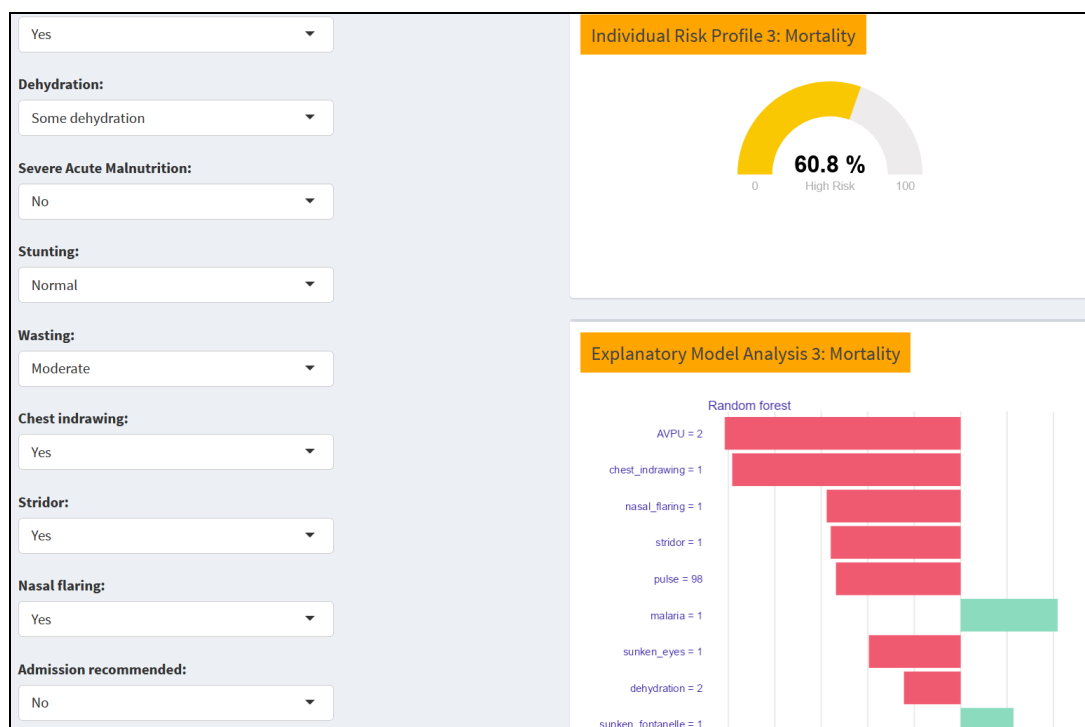


Figure 5.30: Prediction Outputs for Diarrhea-associated Mortality.

Moreover, the effectiveness and user engagement of the dashboard could be analyzed through web analytics utilizing a tool like Google Analytics. Specifically, user engagement could be evaluated by metrics that include: Total number of distinct users visiting the dashboard (Unique Visitors); Average time spent on the dashboard during a single visit (Session Duration); Percentage of users who revisit the dashboard (Returning Visitors). This evaluation could provide valuable insights into the effectiveness of the diarrheal outcome dashboard and guide further development and optimization efforts. Specifically, it could help in interface improvement by optimizing the dashboard's design and layout to enhance user experience, update the content based on user needs and feedback and model refinement to improve accuracy and robustness.

5.5.2 Discussion: Development of Decision Support Dashboard

Enhancing the timely identification and management of children at heightened risk of poor diarrheal outcomes is imperative, particularly in resource-constrained settings where diarrhea burden is disproportionately high, yet diagnostic capabilities remain inadequate, and the healthcare providers overburdened. To empower informed and rapid clinical decision-making in the management of childhood diarrhea, the study leveraged the champion models to develop an R-shiny intuitive online dashboard designed to augment decision making, by predicting the risk profile of pediatric patients and providing SHAP estimates from an explainable model agnostic approach to help understand the reasoning behind the predictions.

The successful development and deployment of an online dashboard incorporating explanatory model analysis using R Shiny reinforces the platform's versatility for such applications (Beckman et al., 2020; Melfi et al., 2020; Zhu, 2016). This is further evidenced by a comprehensive review

of literature from 2012 to 2018, which identified over 5,251 articles related to Shiny applications in academia (University of Adelaide search engine) (Kasprzak et al., 2020). While a significant portion of these publications were found in bioinformatics journals, the diversity of applications demonstrates Shiny's growing adoption across various disciplines. Specifically, R Shiny's intuitive framework simplifies the process of building interactive web applications. Its simplicity allows for rapid prototyping and iterative development, while its reactive programming paradigm allows for dynamic updates based on user interactions, making it ideal for creating dashboards that respond to user input. While the current dashboard was deployed using shinyapps.io, R Shiny can also be integrated with other hosting platforms for greater scalability to allow for expanding the user base and handling increasing data volume. R Shiny is an open-source tool, making it accessible and affordable for researchers and institutions with limited resources. Its open-source nature also facilitates collaboration and community development.

The clinical relevance of the online dashboard tailored for predicting poor diarrheal outcomes, lies in its ability to offer crucial insights into patient prognosis, thereby profoundly impacting healthcare decision-making. By harnessing predictive analytics, this tool has the potential to facilitates early identification of individuals at heightened risk for adverse outcomes associated with diarrheal illnesses. This predictive ability offers significant opportunities to improve healthcare interventions, with capability to enable clinicians to make better-informed decisions and tailor treatment plans to individual patients. The integration of EMA results alongside risk profile predictions not only enhances the interpretability of the predictions but also potentially enables healthcare providers to gain deeper insights into the disease mechanisms and patient-specific factors driving poor outcomes. Furthermore, the platform developed for predicting poor outcomes in diarrheal illnesses can be adapted and applied to the management of other pediatric

diseases. Once successfully implemented the platform could be expanded to other pediatric diseases with similar predictive modeling needs. This iterative approach would leverage the infrastructure and expertise developed for the initial disease platform while tailoring the models and dashboard to the specific characteristics of each new disease.

The implementation of a tool like this dashboard, which predicts poor diarrheal outcomes and aids in healthcare decision-making, carries significant public health implications. Specifically, by the potential to rapidly identify individuals with elevated risk, public health officials could implement targeted interventions such as community-based education campaigns on hygiene practices, vaccination programs, or early detection and treatment initiatives to avert poor outcomes. Moreover, such tools could potentially help to optimize allocation of scarce resources by prioritizing interventions for high-risk populations and thereby maximize their impact on overall health outcomes. Additionally, insights from predictive tools can inform long-term healthcare system planning and policy development by anticipating future healthcare needs and designing interventions that address the specific needs of different populations. Lastly, by identifying populations at heightened risk of poor diarrheal outcomes, public health interventions can be tailored to tackle root societal drivers of health, which influence health inequalities helping to mitigate inequalities in access to healthcare services.

Integrating the online dashboard into clinical practice requires collaboration between healthcare providers, IT professionals, and developers to ensure seamless integration, effective training, and ongoing support for its use in improving patient care and outcomes. While simple and flexible deployment methodologies such as R-shiny or plumber web-based applications (Eddington et al., 2022; Murphree et al., 2018) employed in the study offer the easiest option for quick deployment, debugging, and user testing of the application. The ultimate integration of such tools into existing

clinical systems, such as electronic health records (Khalilia et al., 2015) or clinical decision support systems could allow clinicians to access the outputs within their workflow, without needing to switch between different systems or interfaces. Future studies need to comprehensively evaluate this tool's acceptability, its effects on clinical practice and patient outcomes, and the cost-effectiveness and long-term viability of implementing such models within healthcare systems along with external validation in cohorts different from study population to ensure robustness and generalizability.

6. CHAPTER SIX: REFLECTIONS ON STUDY CONTRIBUTION IN MACHINE LEARNING AND PEDIATRIC ENTERIC DISEASES

6.1 Introduction

This thesis embarked on a critical inquiry at the intersection of computational science and global public health, setting out to develop and evaluate a ML approach for predicting the poor outcomes of pediatric diarrhea: LDD, chronic malnutrition, and mortality. The preceding chapters have laid the groundwork, presenting the methodology, empirical results and discussion of this investigation. This chapter moves beyond empirical findings to offer a broader reflection on the study's contributions and future potential. It pivots from what was found to why it matters. While the overarching goal was to create three models as well as the prototype digital dashboard, the journey of their creation yielded many insights.

The primary objective here is two-pronged. First, it aims to articulate the study's contribution to the knowledge base on using ML in pediatric enteric diseases, dissecting this contribution through a dual-dimensional lens: i) The methodological dimension, which explores how this work advanced the technical rigor, interpretability, and practical applicability of ML in low-resource clinical settings; and ii) The health services dimension. Table 6.1 summarizes these contributions under these two distinct yet complementary perspectives. These dimensions are underpinned by a triadic theoretical framework, integrating the design science approach, socio-technical framework, and the PROGRESS prognostic research framework, which together ensured the research was technically robust, clinically grounded, and contextually relevant.

The reflections in this chapter have been framed to address objective six and seven, which have been formulated into research questions eight (**RQ8**: What are the contributions of this study to the knowledge base on predictive modelling in pediatric enteric diseases from methodological and health services dimensions?) and nine (**RQ9**: What are the future research directions on predictive modelling in pediatric enteric diseases?).

Table 6.1: Contributions of the Thesis to the Knowledge Base

Dimension	Contribution to Knowledge
Methodological	i. Robust handling of missing data; ii. Hybrid Feature Selection; iii. Handling Class imbalance; iv. Leveraging synthetic Data; v. Interpretability Focus; vi. Rigorous Evaluation (Friedman Test and Nemenyi Post-hoc, model calibration, and Temporal validation); vii. Non-technical evaluation
Health Services	i. Potential for early Risk Stratification and Proactive Care; ii. Blueprint for a Low-Cost Clinical Decision Support Tool; iii. Informing Targeted Public Health Interventions; iv. Informing national/county child health guidelines

6.2 Contributions to the Methodological Dimension

This thesis advanced the application of ML in pediatric health research through seven-pronged methodological approach designed to address the unique challenges of health data modeling in low-resource settings. The following headings are elaboration on the points in the table of contributions, Table 6.1.

i.) Robust Handling of Missing Data

The prevalent challenge of missing data in clinical surveillance datasets threatens statistical power and model validity (Liu et al., 2023). Since naive methods like listwise deletion or mean imputation

introduce bias and reduce sample integrity, the study employed Multiple Imputation by Chained Equations (MICE). MICE was chosen for its ability to handle mixed data types and preserve the complex, multivariate relationships inherent in the data. This approach provided a more robust and statistically sound foundation for the models by creating plausible, complete datasets that account for imputation uncertainty, thereby ensuring the study findings are based on the maximal use of all collected information.

ii.) A Hybrid Approach to Feature Selection

To navigate the "curse of dimensionality" in datasets with numerous clinical and demographic variables but limited sample sizes, a two-stage hybrid feature selection strategy was designed. This approach synergized the computational efficiency of filter methods with the predictive precision of wrapper methods. An initial, lenient statistical filter ($p < 0.2$) reduced the feature space, followed by the Boruta algorithm, a robust Random Forest-based wrapper. This hybrid method proved highly effective, balancing the need to identify complex, interactive predictors with the practical necessity of mitigating overfitting and maintaining computational feasibility in a resource-constrained context.

iii.) Principled Handling of Class Imbalance

The target outcomes (LDD, LGF, and mortality) exhibited significant class imbalance (up to 1:13 for mortality). Standard algorithms are inherently biased toward the majority class, risking low sensitivity for the most critical, albeit rare, outcomes. To mitigate this, data-centric over-sampling (upsampling) was integrated directly into the cross-validation resampling framework. This strategic choice, favored over under-sampling to avoid data loss in the limited sample, ensured the models were sufficiently exposed to minority class instances, forcing them to learn the subtle patterns of high-risk cases and thereby enhancing predictive equity.

iv.) Leveraging Synthetic Data for Model Stability

For the LGF use case, the primary dataset's sample size ($n=1,106$) was a potential constraint on model stability and generalizability. This was addressed by generating a high-fidelity synthetic dataset ($n=8,894$) using the `synthpop` package. After rigorously validating that the synthetic data preserved the statistical distributions and propensity scores of the original data, the models were trained on the combined dataset. The sensitivity analysis confirmed that this high-ratio synthetic augmentation significantly improved the performance of most algorithms, demonstrating a validated and effective method for enhancing model robustness when primary data is scarce.

v.) Model Interpretability

A primary barrier to the clinical adoption of ML is the "black box" problem. To transform the models from opaque predictors into transparent clinical partners, Explanatory Model Analysis (EMA) was embedded as a core component of the workflow. Using a model-agnostic SHapley Additive exPlanations (SHAP), the opacity of each prediction was demystified, visualizing the specific contribution of every feature for individual patients. This provided not just a risk score, but a justification, enabling clinicians to understand the direction and magnitude of each feature's contribution to individual predictions.

vi.) Tripartite Model Evaluation

Recognizing that a single performance metric is insufficient, a tripartite evaluation approach was established to ensure the models were accurate, reliable, and robust:

- i. Statistical comparison: The Friedman and Nemenyi post-hoc tests were used for a statistically rigorous comparison of algorithms, moving beyond simple point-estimate rankings.

- ii. Calibration assessment and correction: Probabilistic reliability was assessed using the Brier score and Spiegelhalter's z-test. Crucially, for miscalibrated models, Platt scaling was applied to correct the probability outputs, ensuring they were clinically trustworthy.
- iii. Temporal validation: The most stringent test involved validating the models on unseen, future data from the EFGH study (2022-2024), providing strong evidence of their generalizability and robustness against concept drift.

vii.) Translating Performance into Health System Value

To bridge the gap between technical performance and operational utility, a "public health value" evaluation was conducted using the `modelplotr` package. This analysis translated model performance into actionable, resource-oriented visualizations like cumulative gains and lift charts. This allowed us to answer the critical stakeholder question: "How best can this model be used to target the limited resources?". This final evaluation step framed the model not as a statistical artifact but as a strategic tool for health system planning and impact optimization.

6.2.1 Discussion of Methodological Contributions

The methodological approaches detailed above coalesce into three overarching contributions to the knowledge base, advancing the application of machine learning in pediatric health from technical feasibility toward clinical and health system reality.

Firstly, the thesis establishes a new comprehensive standard for the development and validation of predictive models in pediatric health. This work demonstrated that methodological rigor is not a single action but a complete workflow. It begins with a principled foundation, robustly handling missing data (MICE) and systematically selecting features (hybrid approach). It culminates in a tripartite evaluation that moves beyond accuracy metrics. By integrating a triad of statistically

sound model comparison (Friedman/Nemenyi), probabilistic calibration and correction (Brier/Spiegelhalter/Platt), and temporal validation, this work provides a replicable blueprint for building models that are not just accurate, but demonstrably robust, reliable, and safe for clinical consideration. Additionally, the study through the chosen approaches and their implementation offers a computationally feasible validated pipeline architecture that enables the development and deployment of sophisticated ML models within low-infrastructure settings.

Next, a core contribution of this work lies in providing effective and validated solutions to the widespread challenge of data scarcity in global health research. This was addressed on two fronts. First, for rare but critical events like mortality, the study demonstrated how programmatic oversampling can ensure predictive equity without discarding precious data. Secondly, for outcomes where overall sample size was a constraint (LGF), the study provided a framework for high-ratio synthetic data generation, proving it is possible to significantly enhance model performance while rigorously maintaining the statistical fidelity of the original dataset. Together, these strategies offer a pragmatic toolkit for researchers seeking to apply advanced ML techniques in data-limited, high-stakes environments.

Finally, perhaps the most significant contribution of this thesis is its deliberate focus on bridging the gap between algorithmic development and real-world implementation. The study advocates that a model's true value is measured by its utility and trustworthiness achieved through a dual strategy. First, by prioritizing model interpretability through SHAP, the study transformed the ML model opacity into a transparent diagnostic partner, providing the explanations necessary to build clinician trust. Second, through the value-based evaluation (modelplotr), predictive accuracy was translated into the language of resource allocation and strategic impact, demonstrating clear health system value to administrators and policymakers. This end-to-end focus, from creating a

technically sound and explainable algorithm to proving its practical utility, provides a holistic model for translation of data science results in the global public health space.

6.3. Contribution in the Health Services Dimension

Beyond its methodological contributions, this thesis makes four pivotal contributions to the knowledge base from a health services perspective, advancing the practical application of ML in pediatric health within low-resource settings. These contributions reflect a translational focus, bridging the gap between computational research and clinical application, health system efficiency, and public health strategy, thereby enhancing the relevance, accessibility, and impact of ML-driven tools in pediatric care. Here is the elaboration on the points from the table of contributions, Table 6.1.

6.3.1 Enabling Early Risk Stratification and Proactive Care

This thesis delivers validated, high-performing ML models (AUCs: 82.6–83.5%) capable of early risk stratification for LDD, LGF, and diarrhea-associated mortality. By operationalizing routinely collected clinical data, the models convert static observations into actionable, dynamic risk scores. This enables scalable, point-of-care triaging of the risk profile of pediatric diarrhea cases to support individualized and timely clinical decision-making. High-risk children flagged by the models can receive intensified monitoring, timely interventions, and closer follow-up, potentially preventing LDD, mitigating growth faltering, or averting death.

6.3.2 Blueprint for a Low-Cost Clinical Decision Support Tool

The thesis demonstrates the technical feasibility of developing a low-cost, clinician-facing Clinical Decision Support Tool optimized for resource-limited settings. This proof-of-concept dashboard provides a functional blueprint for implementation, advancing from theoretical modeling to a

deployable prototype. The R Shiny-based dashboard integrates champion models for LDD, LGF, and mortality into a single, streamlined interface, minimizing data entry burden while offering actionable outputs. Built entirely on open-source infrastructure, the tool ensures accessibility without licensing barriers. It features intuitive risk gauges for rapid clinical assessment of pediatric diarrhea cases and SHAP attributions with prospects of enhancing interpretability and fostering clinician trust. This prototype lays the foundation for scalable deployment, further development and potential integration into electronic medical records or mHealth platforms to support overburdened clinicians and reduce diagnostic uncertainty.

6.3.3 Informing Targeted Public Health Interventions

This research identified distinct, modifiable predictors for each diarrheal outcome, providing a granular evidence base enabling a shift from a one-size-fits-all approach to targeted, outcome-specific public health interventions. The strong, consistent protective effect of rotavirus vaccination against LDD and LGF provides locally validated evidence to support scaling up immunization coverage as a primary preventive strategy. Malnutrition-related indicators (SAM, wasting, stunting) emerged as dominant risk factors for mortality, reinforcing the need to embed nutritional screening and interventions (including therapeutic feeding, supplementation and breastfeeding support) within diarrhea management protocols. Although excluded from model training due to unavailability in the temporal validation cohort, the identification of breastfeeding as a significant predictor underscores the continued need to promote and support optimal breastfeeding practices as a modifiable risk factor. Moreover, pre-enrollment diarrhea duration points to delayed care-seeking or suboptimal home management, highlighting the need for caregiver education on early danger signs. While socioeconomic and WASH variables were less

predictive in the final models, their initial associations indicate structural determinants that warrant parallel sanitation and hygiene interventions in high-risk settings.

6.3.4 Informing National and County-Level Child Health Guidelines

The thesis delivers concrete, evidence-based guidance that could be incorporated into national and county-level child health policies and clinical guidelines on diarrheal management. The validated predictive models, operationalized through a functional online R-Shiny dashboard, offer a clear policy pathway for integrating ML-driven risk stratification tools or simplified risk scores into clinical protocols and digital health platforms. These tools have the potential to enhance decision-making efficiency, reduce health system burden, and improve care quality. The study findings also support the development of risk-based clinical pathways, such as mandatory admission for high predicted mortality risk and routine nutritional assessment and intervention for children with diarrhea and SAM/wasting. Combined with the identified clinical predictors, this work provides objective, context-specific criteria that could potentially inform triage, escalation of care, and additions to the knowledge base that can benefit policy and practice, as well as clinical guidelines for management of diarrheal diseases.

6.4. Reflections on Triadic Conceptual Framework

The strength of this study lies in its integration of three complementary frameworks that together bridge the gap between theory and practice. The first part of the triadic conceptual framework is the design science approach. This framework positioned the predictive models and R-Shiny dashboard not as analytical byproducts, but as purpose-built artifacts designed to address a clearly defined clinical challenge: the early identification of at-risk children. Design science emphasized an iterative, problem-solving orientation, ensuring that the outputs were tangible, context-sensitive, and solution-focused, rather than abstract or purely theoretical. Within this paradigm,

the models and dashboard developed in the study meet the threshold of legitimate research artifacts by fulfilling five critical dimensions (Gerber & Kotze, 2015). First, each model demonstrated clinical utility and actionability, directly addressing a real-world decision-making need in pediatric care by producing risk scores for poor diarrheal outcomes (LDD, LGF and mortality) that inform triage, monitoring, and targeted interventions. Second, the modeling process embodied methodological rigor, incorporating hybrid feature selection, multi-algorithm benchmarking, calibration checks, and both internal and temporal validation—safeguarding against overfitting and ensuring external generalizability. Third, the models delivered performance benchmarking, consistently outperforming random allocation by capturing two to three times more high-risk cases within the top-risk deciles, thereby demonstrating clear clinical lift. Fourth, the inclusion of SHAP explanations introduced transparency and interpretability, fostering clinician trust, enhancing regulatory acceptability, and aligning with ethical requirements for AI deployment in health settings. Finally, the artifact instantiation of these models in a web-based R-Shiny dashboard ensured their potential for usability and testability, particularly in low-resource environments. Together, these elements exemplify how predictive modeling, when grounded in design thinking, can yield scientifically robust, operationally relevant, and ethically deployable solution.

The second part of the triadic framework is the socio-technical approach. The socio-technical framework provided important context, acknowledging that the artifacts do not operate in a vacuum. It forced consideration of the interplay between the technology (the models), the people (clinicians, patients), and the processes (hospital workflows) addressing usability, equity, and trust challenges (Bauer & Herder, 2009; Patnayakuni & Ruppel, 2010). This lens was critical for designing an interpretable, clinician-centric dashboard (consolidated input interface, SHAP explanations) and for understanding why adoption hinges on more than just algorithmic accuracy

The third part of the triadic framework is the PROGRESS framework, which supplied the clinical and epidemiological rigor. The PROGRESS framework ensured the research question was well-posed, the patient populations clearly defined, and the outcomes clinically meaningful (Kent et al., 2020). It grounded the design science approach and socio-technical elements in established best practices for prognostic modeling, ensuring the artifacts were not just technically functional but clinically valid and relevant. Together, this triadic framework created a robust methodology for applied eHealth research, ensuring the solution was technically sound (design science approach), clinically relevant (PROGRESS), and organizationally/health system aware (Socio-technical).

Further reinforcing its contributions, this research is strongly aligned with the WHO's ethical guidelines for artificial intelligence in health (World Health Organization, 2021a) as well as core principles of Health Policy and Systems Research (HPSR) (Gilson et al., 2012). Specifically, for the WHO AI guidelines, the R-Shiny dashboard developed in this study, which integrates validated predictive models, is explicitly designed as a clinical decision support tool to augment, not replace, clinician judgment, thereby with capacity to uphold the principle of autonomy. The rigorous tripartite evaluation framework, which includes statistical comparison, calibration assessment and correction, and temporal validation, reflects a deep commitment to safety, efficacy, and accountability, ensuring that model outputs are reliable and context-sensitive. The incorporation of SHAP-based explanations within the dashboard supports the principle of transparency and explainability, potentially enabling frontline users to understand the rationale behind individual predictions: an essential step toward ethical and trustworthy AI deployment. Moreover, the study's use of a socio-technical lens and its reflections on model decay foreground the importance of ongoing human oversight and governance structures, aligning with WHO's call for continuous monitoring of AI systems in health environments. By prioritizing a disease that disproportionately

affects children in low- and middle-income countries, the study further advances the principle of equity, ensuring that AI innovation addresses global health disparities.

In parallel, the research embodies the vision of HPSR by generating context-specific, problem-driven, and policy-relevant evidence (Gilson et al., 2012). It offers a functional blueprint for embedding predictive models into routine pediatric care in resource-limited settings, thus transforming rehabilitation and diarrhea management services from reactive to proactive. By holding the potential to strengthen system responsiveness, improve resource allocation, and facilitate evidence-informed decision-making, the study makes a latent contribution to health system performance. In doing so, it could potentially support global goals such as Universal Health Coverage and the Sustainable Development Goals, while offering a replicable model for integrating neglected yet high-impact digital innovations into national health systems.

6.5 Future Research Directions: Pathways for Implementation, Adaptation, and Integration

Building on this thesis' contributions, several key avenues emerge to transition the validated tools into practice, enhance model adaptability, and further elucidate the drivers of adverse diarrheal outcomes. These directions span three core domains: clinical implementation and impact evaluation, addressing concept drift and data complexity, integrating multi-level determinants.

6.5.1 Clinical Implementation and Impact Evaluation

The most critical next step is to evaluate the real-world impact of the developed Clinical Decision Support Tool. This involves moving beyond the current proof-of-concept to rigorous implementation science research:

- i. Prospective validation and clinical trials: Conduct a multi-site, cluster-randomized controlled trial or a before-after study to evaluate the tool's impact on patient outcomes and clinical practice. Such studies would aim to answer key questions: Does the use of the tool lead to a measurable reduction in mortality, duration of diarrhea, or incidence of growth faltering? Does it improve the timeliness of critical interventions?
- ii. Usability and human-computer interaction: Before large-scale deployment, targeted usability and acceptance studies with frontline clinicians are essential. This research would refine the dashboard's interface, workflow integration, and the presentation of SHAP explanations to maximize adoption and minimize cognitive load in high-pressure environments.
- iii. Health system integration and interoperability: Future work should also look into the integration of the predictive models into existing health information systems, such as Electronic Medical Records or District Health Information Software 2. Research into lightweight, interoperable solutions that can push risk alerts directly into routine clinical workflows would be a significant advancement.

6.5.2 Addressing Concept Drift and Data Complexity

The findings on model decay and the deliberate exclusion of certain data types open new avenues for advanced methodological research:

- i. Operationalizing model governance and adaptive learning: The observed temporal decay necessitates research into adaptive AI systems. This includes developing automated pipelines for continuous model monitoring, anomaly detection, and periodic recalibration. Furthermore, exploring online learning or federated learning models that can adapt in near

real-time across different facilities, without centralizing sensitive patient data, would represent a major leap toward creating truly resilient and trustworthy clinical AI.

- ii. Integrating multi-modal data: This study intentionally focused on readily available data to ensure applicability. Future research, particularly in settings with growing diagnostic capacity, should explore the integration of multi-modal data. This could involve incorporating etiological data (viral, bacterial pathogens), host inflammatory biomarkers, or even microbiome data to create more mechanistically informed and accurate predictive models.

6.5.3 Integrating Multi-Level Determinants

The unexpected dominance of clinical predictors over socio-economic ones calls for deeper exploration of structural and contextual factors:

- a) Longitudinal and life-course studies: While clinical variables predict acute outcomes, socio-economic factors are critical drivers of long-term health. Future longitudinal studies should follow cohorts post-discharge to determine how household wealth, maternal education, and food security, influence recovery, recurrence of illness, and long-term growth trajectories.
- b) Multi-level and geospatial modeling: Although socio-economic variables did not emerge as key predictors in acute settings, future studies should move beyond individual-level data to incorporate household, community, and environmental, WASH, and contextual economic indicators to understand upstream determinants of repeated illness, care-seeking behavior, and recovery trajectories. Using multi-level models that integrate geospatial data on climate, sanitation infrastructure (WASH), and proximity to health facilities could

uncover structural drivers of risk and inform geographically targeted public health interventions. This would help clarify the context in which individual clinical risk factors become most potent.

These future directions aim to enhance the clinical relevance, technical robustness, and contextual sensitivity of predictive tools for pediatric diarrheal illness, advancing the field toward real-world implementation, and equity-oriented decision support systems.

CHAPTER SEVEN: CONCLUSION AND RECOMMENDATIONS

7.1 Research Summary

The integration of digital technologies in healthcare has led to a generation of diverse and complex data, fueling the creation of data-driven predictive models. These models offer significant advantages to both patients and the healthcare system, particularly in the face of insufficient diagnostic capacity and overworked healthcare professionals. A prime example of this transformative potential is the prediction of poor diarrheal outcomes, marking a key advancement in pediatric health predictive modeling. This study developed seven predictive models for each diarrheal outcome (LDD, LGF and mortality), rigorously evaluating their performance through internal validation, temporal validation, and business value assessment. The top-performing models were then subjected to explanatory analysis to improve the interpretability of their decision-making processes. These models and their explainer objects were subsequently integrated into a proof-of-concept online prediction dashboard.

In the research, the literature review grounded the research in theory, and provided the evidence base for the study's findings, was presented across two chapters (Chapter Two and Chapter Three). Chapter Two explored the public health context, discussing the burden, risk factors, pathogenesis, diagnosis, and management of each diarrheal outcome. Chapter Three delved into the informatics context, demystifying the concept of ML and detailing the seven algorithms used in the study. This chapter also offered a comprehensive overview of the current landscape of predictive modeling for the diarrheal outcomes and critically evaluated key considerations for selecting a deployment framework. It provided valuable insights into the diverse frameworks available for deploying ML models, particularly within the R programming environment. Additionally, it outlined the

theoretical and conceptual framework guiding the study, establishing the evidence foundation for the findings presented.

Chapter Four discussed the research design and methodology employed in the study. It provided an in-depth explanation of the correlational research design, along with the rationale for its selection. The chapter outlined the study area and the research process, offering justifications for the chosen methods. It also detailed the data used in the study, including collection procedures and modeling techniques. Furthermore, the chapter documented the approach taken for the development and deployment of the proof-of-concept online prediction dashboard.

Chapter Five presented the study's results, organized according to the study objectives. This chapter documented key findings through tables and charts, utilizing a blend of descriptive and inferential statistics. Each study objective was aligned with logically formulated research questions, carefully designed to guide the study towards its goals. The results formed the foundation for a comprehensive discussion of the outcomes, providing a detailed analysis of each objective.

Chapter Six transitions from the empirical results of the preceding chapters to a broader reflection on the study's significance and impact. It provides a multi-faceted analysis of the thesis's core contributions, dissected through two complementary lenses: the methodological dimension, and the health services dimension. Grounding these insights in the study's unique triadic theoretical framework. The triadic framework has three parts: design science approach, socio-technical framework, and the PROGRESS framework. The reflection also views the study under the lens of the WHO guidelines on ethical AI in healthcare. This chapter also discussed future research directions emanating from the study.

7.2 Accomplishment of the Research Objectives

This section summarizes the study's accomplishment, of its objectives and research questions. The study had seven objectives, each with one specific research question, except for the fourth objective, which had three. Objectives six and seven were the reflection on the study's contribution to the knowledge base, as well as delineation of future research directions emanating from the study. Successfully answering these questions would be considered fulfillment of the study's objectives.

7.2.1 Objective One: *Determine the demographic, clinical and socio-economic features that predict longer duration diarrhea, among children aged < 60 months.*

This objective explored the relationships between demographic, clinical, and socioeconomic variables and LDD. To achieve this, a hybrid feature selection strategy was employed, combining a filtering approach with a wrapper algorithm.

7.2.1.1 RQ1: *What features predict longer duration diarrhea among children aged < 60 months?*

Eleven variables were identified as predictors of LDD, with ten of them being clinical characteristics. These predictors included pre-enrollment diarrhea days, Vesikari score, age category, number of vomiting days, breastfeeding status, respiratory rate, frequency of vomiting, rotavirus vaccination status, rectal straining, stool frequency, and skin turgor. The EMA results indicated positive associations for variables such as longer pre-enrollment diarrhea duration, higher severity according to the modified Vesikari score, lack of rotavirus vaccination, normal skin turgor, and older age. Conversely, the absence of vomiting and fewer loose stools in the last 24 hours (< 6) were negatively associated with LDD.

7.2.2 Objective Two: *Determine the demographic, clinical and socio-economic features that predict chronic malnutrition following a diarrheal illness among children aged < 60 months.*

This objective explored the relationships between demographic, clinical, and socioeconomic variables and linear growth faltering a precursor of stunting. To achieve this, a similar approach to that used in objective one was employed.

7.2.2.1 RQ2: *What features predict chronic malnutrition following a diarrheal illness among children aged <60 months?*

Seven variables were identified as predictors of LGF, a precursor to stunting, with six of these being clinical characteristics. The predictors included age, temperature, respiratory rate, severe acute malnutrition (SAM), rotavirus vaccination status, breastfeeding status, and skin turgor. EMA analysis revealed that the absence of SAM and normal skin turgor were positively associated with LGF. In contrast, lack of rotavirus vaccination, older age, elevated temperature, and increased respiratory rate were negatively associated with LGF.

7.2.3 Objective Three: *Determine the demographic, clinical and socio-economic features that predict diarrhea-associated mortality following a diarrheal illness among children aged < 60 months.*

This objective explored the relationships between demographic, clinical, and socio-economic variables and in-hospital mortality. To achieve this, a similar approach to that used in objective one and two was employed.

7.2.3.1 RQ3: *What features predict diarrhea-associated mortality among inpatient children aged < 60 months?*

Nineteen variables were identified as predictors of mortality among children hospitalized with diarrhea. These variables included the AVPU scale, Vesikari score, dehydration, sunken eyes, skin pinch, frequency of vomiting, unconsciousness, wasting, vomiting, pulse, fever, sunken fontanelle, restlessness, nasal flaring, duration of diarrhea, stridor, oxygen saturation below 90%, chest indrawing, malaria, and stunting. EMA analysis highlighted the top 10 predictors, where chest indrawing, nasal flaring, wasting, stunting, and a negative malaria result were negatively associated with mortality. In contrast, sunken eyes, dehydration, low oxygen saturation, restlessness, and poor skin turgor showed positive associations with mortality.

7.2.4 Objective Four: *Develop and evaluate (internal and temporal validation) the performance of predictive models for each poor diarrheal outcome among children aged <5 years using machine learning.*

The study built a set of seven models for each diarrheal outcome, which were then evaluated and a champion model chosen. The achievements of the three research questions are as outlined in 7.2.4.1, 7.2.4.2 and 7.2.4.3, respectively.

7.2.4.1 RQ4: *What machine learning model best predicts longer duration diarrhea among children aged < 60 months?*

The study built 7 models using RF, GBM, NB, LR, SVM, KNN and ANN algorithms. The RF model emerged as the champion model (AUC of 83.0% [95%CI: 78.6-87.5]). The champion model also underwent temporal validation and registered a decline in model performance (AUC=68.4%). The model was well calibrated and showed good business value outperforming random selection

by identifying 2.6 times more LDD cases within the top 20% of observations based on predicted probabilities.

7.2.4.2 RQ5: *What machine learning model best predicts chronic malnutrition among children aged < 60 months following a diarrheal illness?*

Similarly, seven models were developed to predict LGF for this research question. The GBM model demonstrated the highest discriminative ability with an AUC of 83.5% (95% CI: 81.6-85.4). However, it exhibited poor calibration, which was improved using Platt scaling, resulting in an AUC of 84.6%. Temporal validation revealed an 18% decline in model AUC. In terms of business value, the champion model identified three times as many instances of LGF compared to random selection when the top 20% of observations were chosen based on predicted probabilities.

7.2.4.3 RQ6: *What machine learning model best predicts chronic malnutrition among children aged < 60 months following a diarrheal illness?*

For this research question, the RF model exhibited strong discriminatory power compared to the other algorithms, with an AUC of 82.6% (95% CI: 77.1-88.1). It also demonstrated substantial practical value by identifying 2.8 times more true mortality cases within the top 25% of predicted probabilities compared to random selection, effectively pinpointing children at heightened risk of death. This champion model was not subjected to temporal validation due to the unavailability of an appropriate dataset.

7.2.5 Objective Five: *Design an R-shiny product suite for predictive models of poor diarrheal outcomes.*

This objective involved developing a proof-of-concept clinical predictive tool that integrated the champion models for each diarrhea outcome, with the potential to help clinicians identify at-risk

children in a timelier manner and put them on better management, leveraging the R-shiny framework.

7.2.5.1 RQ7: *Is it possible for R-shiny to adequately support the design of a product suite for predictive models of poor diarrheal outcomes?*

R Shiny package was utilized to develop a proof-of-concept web-accessible dashboard, which allowed users to input data into the designated fields and dynamically displayed results upon submission, as a proof-of-concept. To streamline data entry and avoid duplication, input fields were consolidated for the three diarrheal outcomes under investigation into a single interface. Deployment used shinyapps.io that provided the easiest option for quick deployment, debugging, and user testing of the application. The dashboard accepted user input and provided prediction outputs visualized as risk gauges, indicating the risk profile of the child and SHAP plots that helped to enhance the interpretability of the models' decision-making process. The successful development and deployment of an online dashboard incorporating explanatory model analysis using R Shiny reinforces the platform's versatility for such applications. While currently limited in scope and functionality, it serves as a foundational prototype for future development.

7.3 Reflection on Study Contributions to the Knowledge Base and Future Research

Directions

The findings of the study provided theoretical, methodological and policy implications, this relates to objective six. Objective seven was about reflection on future research directions emanating from the study.

Firstly, the study identified a set of demographic and clinical variables significantly associated with poor diarrheal outcomes in children under 5 years of age. These findings not only align with

existing literature that employed standard statistical methods but also underscore the enhanced robustness and reliability in the application of ML approaches in epidemiological studies. The study revealed that clinical variables were the primary predictors of poor diarrheal outcomes. However, different sets of variables were associated with distinct outcomes—longer duration of diarrhea, chronic malnutrition, and mortality—indicating that interventions should be tailored to address specific risk factors for each outcome. This differentiation emphasizes the importance of targeted strategies in managing pediatric diarrheal illness. Moreover, the study contributes to and extends existing theoretical frameworks by providing a deeper understanding of the determinants of pediatric diarrheal illness and its associated outcomes, paving the way for more focused and effective public health interventions.

Secondly, the study introduced an innovative hybrid feature selection strategy that effectively combined filtering and wrapper algorithms to identify key predictors of poor diarrheal outcomes. This approach demonstrated significant efficacy in identifying relevant variables, marking a methodological advancement in pediatric enteric research. By leveraging ML models to predict diarrheal outcomes, the study contributes to the growing body of research that applies ML in healthcare, offering new insights into predictive modeling for pediatric populations. A key highlight of the study was the use of model-agnostic explanatory model analysis, which helped demystify the "black-box" nature of ML models. This analysis provided transparency by offering insights into how model predictions were made, thereby enhancing interpretability and fostering trust in the model's outcomes. Additionally, the study addressed crucial aspects of model calibration and temporal validation, ensuring that the models are not only reliable but also applicable in real-world settings. The integration of an R-Shiny dashboard with the potential for clinical decision support further exemplifies the study's demonstrated practical value, which holds

prospects for timely identification and better management of at-risk children. This novel application of digital tools in healthcare demonstrates the feasibility of translating complex predictive models into accessible, user-friendly platforms. By creating an interactive interface, the study could potentially pave the way for clinicians to leverage advanced predictive analytics in their decision-making processes, ultimately improving patient outcomes.

Thirdly the findings provide valuable insights for healthcare providers, with the capability to enable more accurate and timely identification of children at risk of severe diarrheal outcomes. This could lead to more targeted interventions, potentially improving child health outcomes. By identifying children at heightened risk, the models can help optimize resource allocation, ensuring that limited resources are used more efficiently by focusing on the most vulnerable populations. The study's development of predictive models and the R-Shiny dashboard offers a practical tool for clinicians that could support their decision making in the backdrop of a demanding work environment, burnout and pressure of severely ill patients, enhancing the timeliness and accuracy of treatment decisions. This work highlights the feasibility of integrating data-driven approaches into clinical practice, with significant policy implications for health systems aiming to reduce childhood mortality and malnutrition. The national and county government have the opportunity to incorporate predictive modelling in the policies and strategies on diarrhea management. Materials such as guidelines and protocols, as well as public health strategies could benefit from the notion of predictive modelling. Such an opportunity offers the incorporation of digital technologies in the management of diarrhea which could potentially reduce the burden on the healthcare system and improve quality of care. Additionally, the identified predictors can guide the development of targeted public health interventions that address the root causes of poor diarrheal outcomes. These interventions could include promoting breastfeeding, improving

sanitation, and increasing access to vaccinations, all of which are critical steps in improving child health on a broader scale.

Lastly, several future research directions were identified in the study. These were discussed in chapter six, and include the matter of integrating the system into clinical care, as well as the matter of socio-economic variables, how they could be used in improving the models.

7.4 Limitations

This study had a number of limitations that provide opportunities for future research. Etiological laboratory data were intentionally excluded from model development due to limited diagnostic capacity in most LMICs, which would hinder the practical application of models reliant on such data for healthcare decision-making. Instead, models were built using only readily available clinical, anthropometric, and socioeconomic data to maximize their utility in resource-constrained settings. While this approach prioritizes immediate applicability, future studies should explore integrating biological data into predictive frameworks, especially in settings with strengthened laboratory capacity, as an alternative pathway to improving the accuracy of clinical judgments.

Additionally, this study relied on data from previous enteric studies that were not specifically designed to evaluate the poor diarrheal outcomes investigated. As a result, certain important predictors—such as HIV, which is highly prevalent in the study setting—were not captured. The absence of such predictors could limit the models' ability to account for critical contextual factors. However, despite this challenge, all models achieved good predictive power, with AUCs exceeding 80%. To further enhance the robustness of future models, the incorporation of data collection frameworks that explicitly target known risk factors, such as HIV status, and other context-specific predictors identified through expert opinion or literature review is recommended.

Moreover, the long-term utility of the models may be constrained by the observed 12–18% decay in predictive accuracy over time, highlighting the necessity for continuous recalibration to match evolving health trends. Furthermore, the geographic focus on Kenya and the reliance on facility-based data may limit the models' generalizability to broader global contexts or community-level care where clinical profiles differ. While the models' generalizability beyond Kenya has not been quantified, the findings suggest that their performance could be replicated in other settings with a similar epidemiology of childhood diarrhea.

7.5 Recommendations

The findings of this study advocate for the adoption of digital tools in pediatric enteric healthcare. To facilitate this transition and improve patient outcomes, the following recommendations are proposed for the Ministry of Health (MoH) in close collaboration with county governments and other partners, outlined in the sub-sections that follow.

7.5.1 Features predicting poor diarrheal outcomes

Clinical features emerged as the primary predictors of all the poor diarrheal outcomes. This finding suggests that the MoH should implement targeted interventions based on these identified risk factors and optimize resource allocation to support the most vulnerable children. A key recommendation is to enhance vaccination campaigns to increase rotavirus vaccine coverage, as the lack of vaccination was a significant predictor of poor outcomes. The study also highlights a strong association between malnutrition (stunting and wasting) and poor diarrheal outcomes. To address this, the MoH should reinforce existing programs aimed at combating malnutrition, such as promoting breastfeeding, providing essential nutritional supplements, and improving access to

quality food. Moreover, the Ministry should integrate these findings into national and county-level health policies aligning health strategies with evidence-based insights.

7.5.2 Develop and evaluate the performance of predictive models for poor diarrheal outcomes

The study successfully developed predictive models with strong discriminative ability for the three diarrheal outcomes. To build on this achievement, the MoH should implement a national strategy to enhance data collection and epidemiological surveillance. Additionally, they should develop national data-sharing protocols and platforms to enable access to anonymized health data for model development and validation. This will support more accurate predictive modeling and better monitoring of health outcomes. The Ministry should also develop a plan of how to effectively integrate validated models into existing national health strategies to support decision-making and resource allocation.

7.5.3 Develop an R-shiny product suite for predictive models of poor diarrheal outcomes.

A proof-of-concept online dashboard using the R Shiny platform, integrating the champion models for each of the poor diarrheal outcomes was developed. Based on these promising results, it is recommended that the Ministry of Health working together with other stakeholders initiate a pilot program to evaluate the tool's user acceptability, as well as its impact on clinical practice and patient outcomes. This pilot will provide valuable insights into the tool's effectiveness and guide its broader implementation in healthcare settings.

REFERENCES

- Abdulkareem, N. M., & Abdulazeez, A. M. (2021). *Machine Learning Classification Based on Radom Forest Algorithm: A Review*. <https://doi.org/10.5281/ZENODO.4471118>
- Acácio, S., Mandomando, I., Nhampossa, T., Quintó, L., Vubil, D., Sacoór, C., Kotloff, K., Farag, T., Nasrin, D., Macete, E., Levine, M. M., Alonso, P., & Bassat, Q. (2019). Risk factors for death among children 0–59 months of age with moderate-to-severe diarrhea in Manhica district, southern Mozambique. *BMC Infectious Diseases*, 19(1), 322. <https://doi.org/10.1186/s12879-019-3948-9>
- Adekanmbi, V. T., Kayode, G. A., & Uthman, O. A. (2013). Individual and contextual factors associated with childhood stunting in Nigeria: A multilevel analysis. *Maternal & Child Nutrition*, 9(2), 244–259. <https://doi.org/10.1111/j.1740-8709.2011.00361.x>
- Adomavicius, G., Bockstedt, J. C., Gupta, A., & Kauffman, R. J. (2008). Making Sense of Technology Trends in the Information Technology Landscape: A Design Science Approach. *MIS Quarterly*, 32(4), 779–809. <https://doi.org/10.2307/25148872>
- Agrahari, S., & Singh, A. K. (2022). Concept Drift Detection in Data Stream Mining: A literature review. *Journal of King Saud University - Computer and Information Sciences*, 34(10, Part B), 9523–9540. <https://doi.org/10.1016/j.jksuci.2021.11.006>
- Aguiar, F. S., Torres, R. C., Pinto, J. V. F., Kritski, A. L., Seixas, J. M., & Mello, F. C. Q. (2016). Development of two artificial neural network models to support the diagnosis of pulmonary tuberculosis in hospitalized patients in Rio de Janeiro, Brazil. *Medical & Biological Engineering & Computing*, 54(11), 1751–1759. <https://doi.org/10.1007/s11517-016-1465-1>
- Ahmed, M., Naser Mahmood, A., & Hu, J. (2016). A survey of network anomaly detection techniques. *Journal of Network and Computer Applications*, 60, 19–31. <https://doi.org/10.1016/j.jnca.2015.11.016>

- Ahmed, Md. S., Shahjaman, Md., Rana, Md. M., & Mollah, Md. N. H. (2017). Robustification of Naïve Bayes Classifier and Its Application for Microarray Gene Expression Data Analysis. *BioMed Research International*, 2017, 3020627. <https://doi.org/10.1155/2017/3020627>
- Ahmed, S. M., Brintz, B. J., Pavlinac, P. B., Shahrin, L., Huq, S., Levine, A. C., Nelson, E. J., Platts-Mills, J. A., Kotloff, K. L., & Leung, D. T. (2023). Derivation and external validation of clinical prediction rules identifying children at risk of linear growth faltering. *eLife*, 12, e78491. <https://doi.org/10.7554/eLife.78491>
- Ahmed, S., N, D., As, A., D, S.-P., D, B.-T., Ve, B., & Jm, T. (2017). Effect of fuzzy partitioning in Crohn's disease classification: A neuro-fuzzy-based approach. *Medical & Biological Engineering & Computing*, 55(1). <https://doi.org/10.1007/s11517-016-1508-7>
- Akech, S., Ayieko, P., Gathara, D., Agweyu, A., Irimu, G., Stepniewska, K., English, M., Ngarngar, S., Aduro, N., Mutai, L., Kimutai, D., Emadau, C., Mutiso, C., Muturi, C., Nzioki, C., Kanyingi, F., Mithamo, A., Kuria, M., Otido, S., ... Wafula, J. (2018). Risk factors for mortality and effect of correct fluid prescription in children with diarrhoea and dehydration without severe acute malnutrition admitted to Kenyan hospitals: An observational, association study. *The Lancet Child & Adolescent Health*, 2(7), 516–524. [https://doi.org/10.1016/S2352-4642\(18\)30130-5](https://doi.org/10.1016/S2352-4642(18)30130-5)
- Akkaya, B., & Çolakoğlu, N. (2019). *Comparison of Multi-class Classification Algorithms on Early Diagnosis of Heart Diseases*.
- Akombi, B. J., Agho, K. E., Hall, J. J., Merom, D., Astell-Burt, T., & Renzaho, A. M. N. (2017). Stunting and severe stunting among children under-5 years in Nigeria: A multilevel analysis. *BMC Pediatrics*, 17(1), 15. <https://doi.org/10.1186/s12887-016-0770-z>
- Akram, R., Sultana, M., Ali, N., Sheikh, N., & Sarker, A. R. (2018). Prevalence and Determinants of Stunting Among Preschool Children and Its Urban–Rural Disparities in Bangladesh. *Food and Nutrition Bulletin*, 39(4), 521–535. <https://doi.org/10.1177/0379572118794770>

- Akseer, N., Tasic, H., Nnachebe Onah, M., Wigle, J., R, R., D, S.-H., J, A., Re, B., Bl, H., N, N., R, S., K, W., N, J., M, S., & J, H. (2022). Economic costs of childhood stunting to the private sector in low- and middle-income countries. *EClinicalMedicine*, 45. <https://doi.org/10.1016/j.eclinm.2022.101320>
- Al Taleb, A. R., Hoque, M., Hasanat, A., & Khan, M. B. (2017). Application of data mining techniques to predict length of stay of stroke patients. *2017 International Conference on Informatics, Health & Technology (ICIHT)*, 1–5. <https://doi.org/10.1109/ICIHT.2017.7899004>
- Alam, Md. Z., Rahman, M. Saifur, & Rahman, M. Sohel. (2019). A Random Forest based predictor for medical data classification using feature ranking. *Informatics in Medicine Unlocked*, 15, 100180. <https://doi.org/10.1016/j.imu.2019.100180>
- Alam, N. H., & Ashraf, H. (2003). Treatment of Infectious Diarrhea in Children. *Pediatric Drugs*, 5(3), 151–165. <https://doi.org/10.2165/00128072-200305030-00002>
- Alam, S., Ahmad, S. H., & Lamba, I. M. S. (1997). Clinical Profile of Persistent Diarrhea in a DTTU. *INDIAN PEDIATRICS*.
- Alghatani, K., Ammar, N., Rezgui, A., & Shaban-Nejad, A. (2021). Predicting Intensive Care Unit Length of Stay and Mortality Using Patient Vital Signs: Machine Learning Model Development and Validation. *JMIR Medical Informatics*, 9(5), e21347. <https://doi.org/10.2196/21347>
- Ali Abdu, N. A., & Basulaim, K. O. (2023). Machine learning in concept drift detection using statistical measures. *International Journal of Computers and Applications*, 0(0), 1–11. <https://doi.org/10.1080/1206212X.2023.2289706>
- Alnsour, Y., Hadidi, R., & Singh, N. (2019). Using Data Analytics to Predict Hospital Mortality in Sepsis Patients. *International Journal of Healthcare Information Systems and Informatics (IJHISI)*, 14(3), 40–57. <https://doi.org/10.4018/IJHISI.2019070104>

- Alsinglawi, B., Alshari, O., Alorjani, M., Mubin, O., Alnajjar, F., Novoa, M., & Darwish, O. (2022). An explainable machine learning framework for lung cancer hospital length of stay prediction. *Scientific Reports*, 12(1), 607. <https://doi.org/10.1038/s41598-021-04608-7>
- Aluisio, A. R., Maroof, Z., Chandramohan, D., Bruce, J., Masher, M. I., Manaseki-Holland, S., & Ensink, J. H. J. (2015). Risk Factors Associated with Recurrent Diarrheal Illnesses among Children in Kabul, Afghanistan: A Prospective Cohort Study. *PLOS ONE*, 10(2), e0116342. <https://doi.org/10.1371/journal.pone.0116342>
- Alwanin, R., Bchir, O., & Ben Ismail, M. M. (2025). Gradient Boosting Based Simultaneous Classification and Regression Approach. *IEEE Access*, PP, 1–1. <https://doi.org/10.1109/ACCESS.2025.3576086>
- Amamra, S.-A. (2025). Random Forest-Based Machine Learning Model Design for 21,700/5 Ah Lithium Cell Health Prediction Using Experimental Data. *Physchem*, 5(1). <https://doi.org/10.3390/physchem5010012>
- An, Q., Rahman, S., Zhou, J., & Kang, J. J. (2023). A Comprehensive Review on Machine Learning in Healthcare Industry: Classification, Restrictions, Opportunities and Challenges. *Sensors (Basel, Switzerland)*, 23(9), 4178. <https://doi.org/10.3390/s23094178>
- Arslan, R. C., Walther, M. P., & Tata, C. S. (2020). formr: A study framework allowing for automated feedback generation and complex longitudinal experience-sampling studies using R. *Behavior Research Methods*, 52(1), 376–387. <https://doi.org/10.3758/s13428-019-01236-y>
- Atkins, A., Allen, T., Wickham, H., McPherson, J., Allaire, J. J., Software, P., & PBC. (2023). *rsconnect: Deploy Docs, Apps, and APIs to “Posit Connect”, “shinyapps.io”, and “RPods”* (Version 1.0.1) [Computer software]. <https://cran.r-project.org/web/packages/rsconnect/index.html>
- Atlas, H. E., Conteh, B., Islam, M. T., Jere, K. C., Omoro, R., Sanogo, D., Schiaffino, F., Yousafzai, M. T., Ahmed, N., Awuor, A. O., Badji, H., Cornick, J., Feutz, E., Galagan, S. R., Haidara, F. C., Horne, B., Hossen, M. I., Hotwani, A., Houpt, E. R., ... Pavlinac, P. B. (2024). Diarrhea Case Surveillance in

- the Enterics for Global Health Shigella Surveillance Study: Epidemiologic Methods. *Open Forum Infectious Diseases*, 11(Suppl 1), S6–S16. <https://doi.org/10.1093/ofid/ofad664>
- Auret, L., & Aldrich, C. (2012). Interpretation of nonlinear relationships between process variables by use of random forests. *Minerals Engineering*, 35, 27–42.
<https://doi.org/10.1016/j.mineng.2012.05.008>
- Auria, L., & Moro, R. A. (2008). Support Vector Machines (SVM) as a Technique for Solvency Analysis. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.1424949>
- Austin, P. C., van Klaveren, D., Vergouwe, Y., Nieboer, D., Lee, D. S., & Steyerberg, E. W. (2017). Validation of prediction models: Examining temporal and geographic stability of baseline risk and estimated covariate effects. *Diagnostic and Prognostic Research*, 1(1), 12.
<https://doi.org/10.1186/s41512-017-0012-3>
- Ayelign, A., & Zerfu, T. (2021). Household, dietary and healthcare factors predicting childhood stunting in Ethiopia. *Heliyon*, 7(4), e06733. <https://doi.org/10.1016/j.heliyon.2021.e06733>
- Ayers, T. L. (2016). *Machine learning approaches for assessing moderate-to-severe diarrhea in children < 5 years of age, rural western Kenya 2008-2012*. 67.
- Ayoubi, C., & Foray, D. (2025). Machine learning in healthcare: A new pattern of diffusion for general purpose technologies. *Economics of Innovation and New Technology*, 0(0), 1–32.
<https://doi.org/10.1080/10438599.2025.2584210>
- Azur, M. J., Stuart, E. A., Frangakis, C., & Leaf, P. J. (2011). Multiple imputation by chained equations: What is it and how does it work? *International Journal of Methods in Psychiatric Research*, 20(1), 40–49. <https://doi.org/10.1002/mpr.329>
- Bado, A. R., Susuman, A. S., & Nebie, E. I. (2016). Trends and risk factors for childhood diarrhea in sub-Saharan countries (1990–2013): Assessing the neighborhood inequalities. *Global Health Action*, 9(1), 30166. <https://doi.org/10.3402/gha.v9.30166>

- Bahati, F., Mcknight, J., Swaleh, F., Malaba, R., Karimi, L., Ramadhan, M., Kiptim, P. K., Okiro, E. A., & English, M. (2022). Reporting of diagnostic and laboratory tests by general hospitals as an indication of access to diagnostic laboratory services in Kenya. *PLoS ONE*, 17(4), e0266667. <https://doi.org/10.1371/journal.pone.0266667>
- Balaha, H. M., & Hassan, A. E.-S. (2023). Skin cancer diagnosis based on deep transfer learning and sparrow search algorithm. *Neural Computing and Applications*, 35(1), 815–853. <https://doi.org/10.1007/s00521-022-07762-9>
- Bandi, M., Kumar, A., Vemula, R., & Vallu, S. (2024). Predictive Analytics in Healthcare: Enhancing Patient Outcomes through Data-Driven Forecasting and Decision-Making. *International Journal of Machine Learning*, 8, 20.
- Banerjee, M., Reynolds, E., Andersson, H. B., & Nallamotheu, B. (2019). Tree-Based Analysis: A Practical Approach to Create Clinical Decision Making Tools. *Circulation. Cardiovascular Quality and Outcomes*, 12(5), e004879. <https://doi.org/10.1161/CIRCOUTCOMES.118.004879>
- Bangura, J. B., Xiao, S., Qiu, D., Ouyang, F., & Chen, L. (2020). Barriers to childhood immunization in sub-Saharan Africa: A systematic review. *BMC Public Health*, 20(1), 1108. <https://doi.org/10.1186/s12889-020-09169-4>
- Baral, R., Nonvignon, J., Debellut, F., Agyemang, S. A., Clark, A., & Pecenka, C. (2020). Cost of illness for childhood diarrhea in low- and middle-income countries: A systematic review of evidence and modelled estimates. *BMC Public Health*, 20(1), 619. <https://doi.org/10.1186/s12889-020-08595-8>
- Bar-Zeev, N., King, C., Phiri, T., Beard, J., Mvula, H., Crampin, A. C., Heinsbroek, E., Lewycka, S., Tate, J. E., Parashar, U. D., Costello, A., Mwansambo, C., Heyderman, R. S., French, N., Cunliffe, N. A., Nakagomi, O., Verani, J. R., & Whitney, C. G. (2018). Impact of monovalent rotavirus vaccine on diarrhoea-associated post-neonatal infant mortality in rural communities in Malawi: A

- population-based birth cohort study. *The Lancet Global Health*, 6(9), e1036–e1044.
[https://doi.org/10.1016/S2214-109X\(18\)30314-0](https://doi.org/10.1016/S2214-109X(18)30314-0)
- Basaleem, H. O., Al-Sakkaf, K. A., & Bawazir, A. A. (2021). Integrated Management of Childhood Health in the Eastern Mediterranean Region. In I. Laher (Ed.), *Handbook of Healthcare in the Arab World* (pp. 2589–2618). Springer International Publishing. https://doi.org/10.1007/978-3-030-36811-1_113
- Bauer, J. M., & Herder, P. M. (2009). Designing Socio-Technical Systems. In A. Meijers (Ed.), *Philosophy of Technology and Engineering Sciences* (pp. 601–630). North-Holland.
<https://doi.org/10.1016/B978-0-444-51667-1.50026-4>
- Bayram, F., Ahmed, B. S., & Kassler, A. (2022). From concept drift to model degradation: An overview on performance-aware drift detectors. *Knowledge-Based Systems*, 245, 108632.
<https://doi.org/10.1016/j.knosys.2022.108632>
- Beal, T., Le, D. T., Trinh, T. H., Burra, D. D., Huynh, T., Duong, T. T., Truong, T. M., Nguyen, D. S., Nguyen, K. T., de Haan, S., & Jones, A. D. (2019). Child stunting is associated with child, maternal, and environmental factors in Vietnam. *Maternal & Child Nutrition*, 15(4), e12826.
<https://doi.org/10.1111/mcn.12826>
- Beal, T., Tumilowicz, A., Sutrisna, A., Izwardy, D., & Neufeld, L. M. (2018). A review of child stunting determinants in Indonesia. *Maternal & Child Nutrition*, 14(4), e12617.
<https://doi.org/10.1111/mcn.12617>
- Beckman, M., Guerrier, S., Lee, J., Molinari, R., Orso, S., & Rudnytskyi, I. (2020). *An Introduction to Statistical Programming Methods with R*. <https://smac-group.github.io/ds/>
- Belle, V., & Papantonis, I. (2021). Principles and Practice of Explainable Machine Learning. *Frontiers in Big Data*, 4. <https://www.frontiersin.org/articles/10.3389/fdata.2021.688969>

- Bello-Chavolla, O. Y., Bahena-López, J. P., Antonio-Villa, N. E., Vargas-Vázquez, A., González-Díaz, A., Márquez-Salinas, A., Fermín-Martínez, C. A., Naveja, J. J., & Aguilar-Salinas, C. A. (2020). Predicting Mortality Due to SARS-CoV-2: A Mechanistic Score Relating Obesity and Diabetes to COVID-19 Outcomes in Mexico. *The Journal of Clinical Endocrinology & Metabolism*, 105(8), 2752–2761. <https://doi.org/10.1210/clinem/dgaa346>
- Benjamin-Chung, J., Mertens, A., Colford, J. M., Hubbard, A. E., van der Laan, M. J., Coyle, J., Sofrygin, O., Cai, W., Nguyen, A., Pokpongkiat, N. N., Djajadi, S., Seth, A., Jilek, W., Jung, E., Chung, E. O., Rosete, S., Hejazi, N., Malenica, I., Li, H., ... Arnold, B. F. (2023). Early-childhood linear growth faltering in low- and middle-income countries. *Nature*, 621(7979), Article 7979. <https://doi.org/10.1038/s41586-023-06418-5>
- Bennett, A., Pollock, L., Jere, K. C., Pitzer, V. E., Parashar, U., Tate, J. E., Heyderman, R. S., Mwansambo, C., French, N., Nakagomi, O., Iturriza-Gomara, M., Everett, D., Cunliffe, N. A., Bar-Zeev, N., & VacSurv Consortium. (2018). Direct and possible indirect effects of vaccination on rotavirus hospitalisations among children in Malawi four years after programmatic introduction. *Vaccine*, 36(47), 7142–7148. <https://doi.org/10.1016/j.vaccine.2018.04.030>
- Bennish, M. L., Harris, J. R., Wojtyniak, B. J., & Struelens, M. (1990). Death in Shigellosis: Incidence and Risk Factors in Hospitalized Patients. *The Journal of Infectious Diseases*, 161(3), 500–506. <https://doi.org/10.1093/infdis/161.3.500>
- Bentéjac, C., Csörgő, A., & Martínez-Muñoz, G. (2021). A comparative analysis of gradient boosting algorithms. *Artificial Intelligence Review*, 54(3), 1937–1967. <https://doi.org/10.1007/s10462-020-09896-5>
- Berberi, L., Kozlov, V., Nguyen, G., Sáinz-Pardo Díaz, J., Calatrava, A., Moltó, G., Tran, V., & Lopez Garcia, A. (2025). Machine learning operations landscape: Platforms and tools. *Artificial Intelligence Review*, 58. <https://doi.org/10.1007/s10462-025-11164-3>

- Bernaola Aponte, G., Bada Mancilla, C. A., Carreazo, N. Y., & Rojas Galarza, R. A. (2013). Probiotics for treating persistent diarrhoea in children. *The Cochrane Database of Systematic Reviews*, 2013(8), CD007401. <https://doi.org/10.1002/14651858.CD007401.pub3>
- Berrar, D. (2019). Bayes' Theorem and Naive Bayes Classifier. In *Encyclopedia of Bioinformatics and Computational Biology* (pp. 403–412). Elsevier. <https://doi.org/10.1016/B978-0-12-809633-8.20473-1>
- Bertoncelli, C. M., Altamura, P., Vieira, E. R., Iyengar, S. S., Solla, F., & Bertoncelli, D. (2020). PredictMed: A logistic regression–based model to predict health conditions in cerebral palsy. *Health Informatics Journal*, 26(3), 2105–2118. <https://doi.org/10.1177/1460458219898568>
- Bhatikar, S. R., DeGroff, C., & Mahajan, R. L. (2005). A classifier based on the artificial neural network approach for cardiologic auscultation in pediatrics. *Artificial Intelligence in Medicine*, 33(3), 251–260. <https://doi.org/10.1016/j.artmed.2004.07.008>
- Bhutta, Z. A., Ghishan, F., Lindley, K., Memon, I. A., Mittal, S., & Rhoads, J. M. (2004). Persistent and Chronic Diarrhea and Malabsorption: Working Group Report of the Second World Congress of Pediatric Gastroenterology, Hepatology, and Nutrition. *Journal of Pediatric Gastroenterology and Nutrition*, 39, S711.
- Bhutta, Z. A., Nelson, E. A., Lee, W. S., Tarr, P. I., Zablah, R., Phua, K. B., Lindley, K., Bass, D., Phillips, A., & Group, P. D. W. (2008). Recent Advances and Evidence Gaps in Persistent Diarrhea. *Journal of Pediatric Gastroenterology and Nutrition*, 47(2), 260. <https://doi.org/10.1097/MPG.0b013e318181b334>
- Bhutta, Z. A., Nelson, E. A., Lee, W. S., Tarr, P. I., Zablah, R., Phua, K. B., Lindley, K., Bass, D., Phillips, A., & Persistent Diarrhea Working Group. (2008). Recent advances and evidence gaps in persistent diarrhea. *Journal of Pediatric Gastroenterology and Nutrition*, 47(2), 260–265. <https://doi.org/10.1097/MPG.0b013e318181b334>

- Bhutta, Z., Punjwani, N., & Lindblad, B. (1996). Concomitant bacteraemia as a risk factor for diarrhoeal disease mortality in Karachi: A case-control study of hospitalized children. *Acta Paediatrica*, 85(7), 809–813. <https://doi.org/10.1111/j.1651-2227.1996.tb14156.x>
- Bhutta, Z., & Syed, S. (2016). Diarrheal Diseases. In B. Caballero, P. M. Finglas, & F. Toldrá (Eds.), *Encyclopedia of Food and Health* (pp. 361–372). Academic Press. <https://doi.org/10.1016/B978-0-12-384947-2.00223-3>
- Biau, G., & Scornet, E. (2016). A random forest guided tour. *TEST*, 25(2), 197–227. <https://doi.org/10.1007/s11749-016-0481-7>
- Biecek, P., & Burzykowski, T. (2020). *Explanatory Model Analysis*. <https://ema.drwhy.ai/InstanceLevelExploration.html>
- Biecek, P., Maksymiuk, S., & Baniecki, H. (2023). *DALEX: moDel Agnostic Language for Exploration and eXplanation* (Version 2.4.3) [Computer software]. <https://CRAN.R-project.org/package=DALEX>
- Bilimoria, K. Y., Liu, Y., Paruch, J. L., Zhou, L., Kmiecik, T. E., Ko, C. Y., & Cohen, M. E. (2013). Development and Evaluation of the Universal ACS NSQIP Surgical Risk Calculator: A Decision Aide and Informed Consent Tool for Patients and Surgeons. *Journal of the American College of Surgeons*, 217(5), 833-842.e3. <https://doi.org/10.1016/j.jamcollsurg.2013.07.385>
- Binagwaho, A., Rukundo, A., Powers, S., Donahoe, K. B., Agbonyitor, M., Ngabo, F., Karema, C., Scott, K. W., & Smith Fawzi, M. C. (2020). Trends in burden and risk factors associated with childhood stunting in Rwanda from 2000 to 2015: Policy and program implications. *BMC Public Health*, 20(1), 83. <https://doi.org/10.1186/s12889-020-8164-4>
- Biondi, D., Kipp, W., Jhangri, G. S., Alibhai, A., Rubaale, T., & Saunders, L. D. (2011). Risk Factors and Trends in Childhood Stunting in a District in Western Uganda. *Journal of Tropical Pediatrics*, 57(1), 24–33. <https://doi.org/10.1093/tropej/fmq043>

- Bitew, F. H., Sparks, C. S., & Nyarko, S. H. (2021). Machine learning algorithms for predicting undernutrition among under-five children in Ethiopia. *Public Health Nutrition*, 25(2), 269–280. <https://doi.org/10.1017/S1368980021004262>
- Black, R. E., Morris, S. S., & Bryce, J. (2003). Where and why are 10 million children dying every year? *The Lancet*, 361(9376), 2226–2234. [https://doi.org/10.1016/S0140-6736\(03\)13779-8](https://doi.org/10.1016/S0140-6736(03)13779-8)
- Black, R., Fontaine, O., Lamberti, L., Bhan, M., Huicho, L., El Arifeen, S., Masanja, H., Walker, C. F., Mengestu, T. K., Pearson, L., Young, M., Orobato, N., Chu, Y., Jackson, B., Bateman, M., Walker, N., & Merson, M. (2019). Drivers of the reduction in childhood diarrhea mortality 1980-2015 and interventions to eliminate preventable diarrhea deaths by 2030. *Journal of Global Health*, 9(2), 020801. <https://doi.org/10.7189/jogh.09.020801>
- Blankenship, J. L., Cashin, J., Nguyen, T. T., & Ip, H. (2020). Childhood stunting and wasting in Myanmar: Key drivers and implications for policies and programmes. *Maternal & Child Nutrition*, 16(S2), e12710. <https://doi.org/10.1111/mcn.12710>
- Blockeel, H., Devos, L., Frénay, B., Nanack, G., & Nijssen, S. (2023). Decision trees: From efficient prediction to responsible AI. *Frontiers in Artificial Intelligence*, 6. <https://doi.org/10.3389/frai.2023.1124553>
- Boateng, E. Y., & Abaye, D. A. (2019). A Review of the Logistic Regression Model with Emphasis on Medical Research. *Journal of Data Analysis and Information Processing*, 7(4), Article 4. <https://doi.org/10.4236/jdaip.2019.74012>
- Bojer, C. S. (2022). Understanding machine learning-based forecasting methods: A decomposition framework and research opportunities. *International Journal of Forecasting, Special Issue: M5 Competition*, 38(4), 1555–1561. <https://doi.org/10.1016/j.ijforecast.2021.11.003>

- Bontempi, G., Ben Taieb, S., & Le Borgne, Y.-A. (2013). Machine Learning Strategies for Time Series Forecasting. In *Lecture Notes in Business Information Processing* (Vol. 138).
https://doi.org/10.1007/978-3-642-36318-4_3
- Brander, R. L., Pavlinac, P. B., Walson, J. L., John-Stewart, G. C., Weaver, M. R., Faruque, A. S. G., Zaidi, A. K. M., Sur, D., Sow, S. O., Hossain, M. J., Alonso, P. L., Breiman, R. F., Nasrin, D., Nataro, J. P., Levine, M. M., & Kotloff, K. L. (2019). Determinants of linear growth faltering among children with moderate-to-severe diarrhea in the Global Enteric Multicenter Study. *BMC Medicine*, 17(1), 214. <https://doi.org/10.1186/s12916-019-1441-3>
- Brocke, J. vom, Hevner, A., & Maedche, A. (2020). *Introduction to Design Science Research* (pp. 1–13).
https://doi.org/10.1007/978-3-030-46781-4_1
- Bui, E. N. (2021). Chapter 4—Machine learning in the Australian critical zone. In J. Dunn & P. Balaprakash (Eds.), *Data Science Applied to Sustainability Analysis* (pp. 43–78). Elsevier.
<https://doi.org/10.1016/B978-0-12-817976-5.00004-8>
- Burkhardt, M. C., & Shan, K. (2020). Deep Low-Density Separation for Semi-supervised Classification. *Computational Science – ICCS 2020*, 12139, 297–311. https://doi.org/10.1007/978-3-030-50420-5_22
- Caballero, S., Guix, S., El-Senousy, W. M., Calicó, I., Pintó, R. M., & Bosch, A. (2003). Persistent gastroenteritis in children infected with astrovirus: Association with serotype-3 strains. *Journal of Medical Virology*, 71(2), 245–250. <https://doi.org/10.1002/jmv.10476>
- Cai, X., Perez-Concha, O., Coiera, E., Martin-Sanchez, F., Day, R., Roffe, D., & Gallego, B. (2016). Real-time prediction of mortality, readmission, and length of stay using electronic health record data. *Journal of the American Medical Informatics Association*, 23(3), 553–561.
<https://doi.org/10.1093/jamia/ocv110>

- Carbonell, J. G., Michalski, R. S., & Mitchell, T. M. (1983). 1—AN OVERVIEW OF MACHINE LEARNING. In R. S. Michalski, J. G. Carbonell, & T. M. Mitchell (Eds.), *Machine Learning* (pp. 3–23). Morgan Kaufmann. <https://doi.org/10.1016/B978-0-08-051054-5.50005-4>
- CDC. (2018, November 9). *Global Diarrhea Burden | Global Water, Sanitation and Hygiene | Healthy Water | CDC*. <https://www.cdc.gov/healthywater/global/diarrhea-burden.html>
- CDC. (2021, December 20). *Principles of Epidemiology | Lesson 1—Section 8*. <https://www.cdc.gov/csels/dsepd/ss1978/lesson1/section8.html>
- CDC. (2022). *Diarrhea: Common Illness, Global Killer*. <https://www.cdc.gov/healthywater/pdf/global/programs/globaldiarrhea508c.pdf>
- Chakraborty, R. K., & Burns, B. (2023). Systemic Inflammatory Response Syndrome. In *StatPearls*. StatPearls Publishing. <http://www.ncbi.nlm.nih.gov/books/NBK547669/>
- Chander, S., & Vijaya, P. (2021). 3—Unsupervised learning methods for data clustering. In D. Binu & B. R. Rajakumar (Eds.), *Artificial Intelligence in Data Mining* (pp. 41–64). Academic Press. <https://doi.org/10.1016/B978-0-12-820601-0.00002-1>
- Chandra, S., Pandey, J., Tamrakar, A. K., & Siddiqi, M. I. (2017). SVMDF: A novel R-based Web application for prediction of dipeptidyl peptidase 4 inhibitors. *Chemical Biology & Drug Design*, 90(6), 1173–1183. <https://doi.org/10.1111/cbdd.13037>
- Chandrashekar, G., & Sahin, F. (2014). A survey on feature selection methods. *Computers & Electrical Engineering*, 40th-Year Commemorative Issue, 40(1), 16–28. <https://doi.org/10.1016/j.compeleceng.2013.11.024>
- Chang, K.-C., Tseng, M.-C., Weng, H.-H., Lin, Y.-H., Liou, C.-W., & Tan, T.-Y. (2002). Prediction of Length of Stay of First-Ever Ischemic Stroke. *Stroke*, 33(11), 2670–2674. <https://doi.org/10.1161/01.STR.0000034396.68980.39>

- Chang, W., Cheng, J., Allaire, J. J., Sievert, C., Schloerke, B., Xie, Y., Allen, J., McPherson, J., Dipert, A., Borges, B., RStudio, library), jQuery F. (jQuery library and jQuery U., inst/www/shared/jquery-AUTHORS.txt), jQuery contributors (jQuery library; authors listed in, inst/www/shared/jqueryui/AUTHORS.txt), jQuery U. contributors (jQuery U. library; authors listed in, library), M. O. (Bootstrap, library), J. T. (Bootstrap, library), B. contributors (Bootstrap, Twitter, library), I. (Bootstrap, ... R), R. C. T. (tar implementation from. (2023). *shiny: Web Application Framework for R* (Version 1.7.4.1) [Computer software]. <https://cran.r-project.org/web/packages/shiny/index.html>
- Chang, Y.-J., Lin, Y.-L., & Pai, P.-F. (2025). Support Vector Machines with Hyperparameter Optimization Frameworks for Classifying Mobile Phone Prices in Multi-Class. *Electronics*, 14(11). <https://doi.org/10.3390/electronics14112173>
- Charles A Janeway, J., Travers, P., Walport, M., & Shlomchik, M. J. (2001). Infectious agents and how they cause disease. *Immunobiology: The Immune System in Health and Disease. 5th Edition*. <https://www.ncbi.nlm.nih.gov/books/NBK27114/>
- Checkley, W., Buckley, G., Gilman, R. H., Assis, A. M., Guerrant, R. L., Morris, S. S., Mølbak, K., Valentiner-Branth, P., Lanata, C. F., Black, R. E., & Network, and T. C. M. and I. (2008). Multi-country analysis of the effects of diarrhoea on childhood stunting. *International Journal of Epidemiology*, 37(4), 816. <https://doi.org/10.1093/ije/dyn099>
- Chen, J. H., & Asch, S. M. (2017). Machine Learning and Prediction in Medicine—Beyond the Peak of Inflated Expectations. *New England Journal of Medicine*, 376(26), 2507–2509. <https://doi.org/10.1056/NEJMp1702071>
- Chilyabanyama, O. N., Chilengi, R., Simuyandi, M., Chisenga, C. C., Chirwa, M., Hamusonde, K., Saroj, R. K., Iqbal, N. T., Ngaruye, I., & Bosomprah, S. (2022). Performance of Machine Learning Classifiers

- in Classifying Stunting among Under-Five Children in Zambia. *Children*, 9(7), Article 7.
<https://doi.org/10.3390/children9071082>
- Chiu, C.-C., Wu, C.-M., Chien, T.-N., Kao, L.-J., & Qiu, J. T. (2022). Predicting the Mortality of ICU Patients by Topic Model with Machine-Learning Techniques. *Healthcare*, 10(6), Article 6.
<https://doi.org/10.3390/healthcare10061087>
- Chowdhury, T. R., Chakrabarty, S., Rakib, M., Winn, S., & Bennie, J. (2022). Risk factors for child stunting in Bangladesh: An analysis using MICS 2019 data. *Archives of Public Health*, 80(1), 126.
<https://doi.org/10.1186/s13690-022-00870-x>
- Cihula, K.-T. (2023, April 5). *Current eHealth trends in Central and Eastern Europe and Turkey*. Lexology.
<https://www.lexology.com/library/detail.aspx?g=6e060fd8-bff2-42a3-a154-4672de5d070f>
- Clark, A., Mahmud, S., Debellut, F., Pecenka, C., Jit, M., Perin, J., Tate, J., Soeters, H. M., Black, R. E., Santosham, M., & Sanderson, C. (2023). Estimating the global impact of rotavirus vaccines on child mortality. *International Journal of Infectious Diseases*, 137, 90–97.
<https://doi.org/10.1016/j.ijid.2023.10.005>
- Climate-Data. (2023). *Siaya climate: Weather Siaya & temperature by month*. <https://en.climate-data.org/africa/kenya/siaya/siaya-11165/>
- Cohen, J. P., Cao, T., Viviano, J. D., Huang, C.-W., Fralick, M., Ghassemi, M., Mamdani, M., Greiner, R., & Bengio, Y. (2021). Problems in the deployment of machine-learned models in health care. *CMAJ : Canadian Medical Association Journal*, 193(35), E1391–E1394.
<https://doi.org/10.1503/cmaj.202066>
- Collins, G. S., Moons, K. G. M., Dhiman, P., Riley, R. D., Beam, A. L., Calster, B. V., Ghassemi, M., Liu, X., Reitsma, J. B., Smeden, M. van, Boulesteix, A.-L., Camaradou, J. C., Celi, L. A., Denaxas, S., Denniston, A. K., Glocker, B., Golub, R. M., Harvey, H., Heinze, G., ... Logullo, P. (2024). TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use

- regression or machine learning methods. *BMJ*, 385, e078378. <https://doi.org/10.1136/bmj-2023-078378>
- Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. (2015). Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD Statement. *BMC Medicine*, 13(1), 1. <https://doi.org/10.1186/s12916-014-0241-z>
- Cooke, R., Goulet, O., Huysentruyt, K., Joosten, K., Khadilkar, A. V., Mao, M., Meyer, R., Prentice, A. M., & Singhal, A. (2023). Catch-Up Growth in Infants and Young Children With Faltering Growth: Expert Opinion to Guide General Clinicians. *Journal of Pediatric Gastroenterology and Nutrition*, 77(1), 7–15. <https://doi.org/10.1097/MPG.0000000000003784>
- Corey, K. M., Kashyap, S., Lorenzi, E., Lagoo-Deenadayalan, S. A., Heller, K., Whalen, K., Balu, S., Heflin, M. T., McDonald, S. R., Swaminathan, M., & Sendak, M. (2018). Development and validation of machine learning models to identify high-risk surgical patients using automatically curated electronic health record data (Pythia): A retrospective, single-site study. *PLOS Medicine*, 15(11), e1002701. <https://doi.org/10.1371/journal.pmed.1002701>
- County Government of Siaya. (2017). *Siaya County Annual Development Plan 2017-2018*. <https://siaya.go.ke/wp-content/uploads/2017/09/SIAYA-COUNTY-2017-18-ADP-FINAL.pdf>
- Cowley, L. E., Farewell, D. M., Maguire, S., & Kemp, A. M. (2019). Methodological standards for the development and evaluation of clinical prediction rules: A review of the literature. *Diagnostic and Prognostic Research*, 3(1), 16. <https://doi.org/10.1186/s41512-019-0060-y>
- Crawford, M. (2014). *Strengths and Limitations of Correlational Design*. <https://doi.org/10.13140/RG.2.1.4195.7920>
- Cruz, J., Bartlett, A., Mendez, H., & Sibrian, R. (1992). Epidemiology of persistent diarrhea among Guatemalan rural children. *Acta Paediatrica*, 81(s383), 22–26. <https://doi.org/10.1111/j.1651-2227.1992.tb12367.x>

- Cunningham, P., Cord, M., & Delany, S. J. (2008). Supervised Learning. In M. Cord & P. Cunningham (Eds.), *Machine Learning Techniques for Multimedia: Case Studies on Organization and Retrieval* (pp. 21–49). Springer. https://doi.org/10.1007/978-3-540-75171-7_2
- Daghistani, T. A., Elshaw, R., Sakr, S., Ahmed, A. M., Al-Thwayee, A., & Al-Mallah, M. H. (2019). Predictors of in-hospital length of stay among cardiac patients: A machine learning approach. *International Journal of Cardiology*, 288, 140–147. <https://doi.org/10.1016/j.ijcard.2019.01.046>
- Dai, B., Chen, R.-C., Zhu, S.-Z., & Zhang, W.-W. (2018). Using Random Forest Algorithm for Breast Cancer Diagnosis. *2018 International Symposium on Computer, Consumer and Control (IS3C)*, 449–452. <https://doi.org/10.1109/IS3C.2018.00119>
- Danaei, G., Andrews, K. G., Sudfeld, C. R., Fink, G., McCoy, D. C., Peet, E., Sania, A., Smith Fawzi, M. C., Ezzati, M., & Fawzi, W. W. (2016). Risk Factors for Childhood Stunting in 137 Developing Countries: A Comparative Risk Assessment Analysis at Global, Regional, and Country Levels. *PLoS Medicine*, 13(11), e1002164. <https://doi.org/10.1371/journal.pmed.1002164>
- D'Anci, K. E., Constant, F., & Rosenberg, I. H. (2006). Hydration and cognitive function in children. *Nutrition Reviews*, 64(10 Pt 1), 457–464. <https://doi.org/10.1301/nr.2006.oct.457-464>
- Das, J. K., Tripathi, A., Ali, A., Hassan, A., Dojoseandy, C., & Bhutta, Z. A. (2013). Vaccines for the prevention of diarrhea due to cholera, shigella, ETEC and rotavirus. *BMC Public Health*, 13(3), S11. <https://doi.org/10.1186/1471-2458-13-S3-S11>
- Davis, S. E., Lasko, T. A., Chen, G., Siew, E. D., & Matheny, M. E. (2017). Calibration drift in regression and machine learning models for acute kidney injury. *Journal of the American Medical Informatics Association : JAMIA*, 24(6), 1052–1061. <https://doi.org/10.1093/jamia/ocx030>
- de Onis, M., & Branca, F. (2016). Childhood stunting: A global perspective. *Maternal & Child Nutrition*, 12(S1), 12–26. <https://doi.org/10.1111/mcn.12231>

- Debsarma, D., Saha, J., & Ghosh, S. (2022). Factors associated with delay in treatment-seeking behaviour for fever cases among caregivers of under-five children in India: Evidence from the National Family Health Survey-4, 2015–16. *PLoS ONE*, 17(6), e0269844.
<https://doi.org/10.1371/journal.pone.0269844>
- Delahanty, R. J., Alvarez, J., Flynn, L. M., Sherwin, R. L., & Jones, S. S. (2019). Development and Evaluation of a Machine Learning Model for the Early Identification of Patients at Risk for Sepsis. *Annals of Emergency Medicine*, 73(4), 334–344.
<https://doi.org/10.1016/j.annemergmed.2018.11.036>
- Demšar, J. (2006). Statistical Comparisons of Classifiers over Multiple Data Sets. *The Journal of Machine Learning Research*, 7, 1–30.
- Deppen, S. A., Blume, J. D., Aldrich, M. C., Fletcher, S. A., Massion, P. P., Walker, R. C., Chen, H. C., Speroff, T., Degesys, C. A., Pinkerman, R., Lambright, E. S., Nesbitt, J. C., Putnam, J. B., & Grogan, E. L. (2014). Predicting lung cancer prior to surgical resection in patients with lung nodules. *Journal of Thoracic Oncology: Official Publication of the International Association for the Study of Lung Cancer*, 9(10), 1477–1484. <https://doi.org/10.1097/JTO.0000000000000287>
- Deribew, A., Tessema, G. A., Deribe, K., Melaku, Y. A., Lakew, Y., Amare, A. T., Abera, S. F., Mohammed, M., Hiruye, A., Teklay, E., Misganaw, A., & Kassebaum, N. (2016). Trends, causes, and risk factors of mortality among children under 5 in Ethiopia, 1990–2013: Findings from the Global Burden of Disease Study 2013. *Population Health Metrics*, 14(1), 42. <https://doi.org/10.1186/s12963-016-0112-2>
- DeSalvo, K. B., Fan, V. S., McDonell, M. B., & Fihn, S. D. (2005). Predicting Mortality and Healthcare Utilization with a Single Question. *Health Services Research*, 40(4), 1234–1246.
<https://doi.org/10.1111/j.1475-6773.2005.00404.x>

- Devi, B., Lepcha, M., & Basnet, S. (2023). Application Of Correlational Research Design In Nursing And Medical Research. *Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/Journal of Xi'an Shiyou University*, 65, 60–69. <https://doi.org/10.17605/OSF.IO/YRZ68>
- Dhal, P., & Azad, C. (2022). A comprehensive survey on feature selection in the various fields of machine learning. *Applied Intelligence*, 52(4), 4543–4581. <https://doi.org/10.1007/s10489-021-02550-9>
- Di, M.-Y., Liu, H., Yang, Z.-Y., Bonis, P. A. L., Tang, J.-L., & Lau, J. (2016). Prediction Models of Mortality in Acute Pancreatitis in Adults. *Annals of Internal Medicine*, 165(7), 482–490. <https://doi.org/10.7326/M16-0650>
- Dimitrova, A., Carrasco-Escobar, G., Richardson, R., & Benmarhnia, T. (2023). Essential childhood immunization in 43 low- and middle-income countries: Analysis of spatial trends and socioeconomic inequalities in vaccine coverage. *PLOS Medicine*, 20(1), e1004166. <https://doi.org/10.1371/journal.pmed.1004166>
- Ding, Z., Huang, Y., Yuan, H., & Dong, H. (2020). Introduction to Reinforcement Learning. In H. Dong, Z. Ding, & S. Zhang (Eds.), *Deep Reinforcement Learning: Fundamentals, Research and Applications* (pp. 47–123). Springer. https://doi.org/10.1007/978-981-15-4095-0_2
- Drysdale, E., Dolatabadi, E., Chivers, C., Liu, V., Saria, S., Sendak, M., Wiens, J., Brudno, M., Hoyt, A., Mazwi, M., Mamdani, M., Allen, V., McGregor, C., Ross, H., Szeto, A., Anand, A., Wang, B., Paprica, P. A., & Goldenberg, A. (2020). *Implementing AI in healthcare*.
- Durairaj, P., Raju, S., & Thirumalaikumarasamy, S. (2017). Clinical profile and risk factors for persistent diarrhoea in children under five years of age in an urban referral centre. *International Journal of Contemporary Pediatrics*, 4(6), 1986–1994. <https://doi.org/10.18203/2349-3291.ijcp20174156>
- Ebinger, J., Wells, M., Ouyang, D., Davis, T., Kaufman, N., Cheng, S., & Chugh, S. (2021). A Machine Learning Algorithm Predicts Duration of hospitalization in COVID-19 patients. *Intelligence-Based Medicine*, 5, 100035. <https://doi.org/10.1016/j.ibmed.2021.100035>

- Eddington, H. S., Trickey, A. W., Shah, V., & Harris, A. H. S. (2022). Tutorial: Implementing and visualizing machine learning (ML) clinical prediction models into web-accessible calculators using Shiny R. *Annals of Translational Medicine*, 10(24), 1414. <https://doi.org/10.21037/atm-22-847>
- Ebrahim, S. M. A., & Abraham, A. (2013). *A Review of Class Imbalance Problem*.
- Emukule, G. O., McMorrow, M., Ulloa, C., Khagayi, S., Njuguna, H. N., Burton, D., Montgomery, J. M., Muthoka, P., Katz, M. A., Breiman, R. F., & Mott, J. A. (2014). Predicting Mortality among Hospitalized Children with Respiratory Illness in Western Kenya, 2009–2012. *PLOS ONE*, 9(3), e92968. <https://doi.org/10.1371/journal.pone.0092968>
- Engel, T., Könings, M., von Heydebrand, N., Goswami, S., & Krcmar, H. (2013). *A NFC-based Concept for Medication Related Patient Services*.
- Espadoto, M., Martins, R. M., Kerren, A., Hirata, N. S. T., & Telea, A. C. (2021). Toward a Quantitative Survey of Dimension Reduction Techniques. *IEEE Transactions on Visualization and Computer Graphics*, 27(3), 2153–2173. *IEEE Transactions on Visualization and Computer Graphics*. <https://doi.org/10.1109/TVCG.2019.2944182>
- Ettema, R. G. A., Peelen, L. M., Schuurmans, M. J., Nierich, A. P., Kalkman, C. J., & Moons, K. G. M. (2010). Prediction Models for Prolonged Intensive Care Unit Stay After Cardiac Surgery. *Circulation*, 122(7), 682–689. <https://doi.org/10.1161/CIRCULATIONAHA.109.926808>
- Fawagreh, K., Gaber, M. M., & Elyan, E. (2014). Random forests: From early developments to recent advancements. *Systems Science & Control Engineering*, 2(1), 602–609. <https://doi.org/10.1080/21642583.2014.956265>
- Fawzi, M. C. S., Andrews, K. G., Fink, G., Danaei, G., McCoy, D. C., Sudfeld, C. R., Peet, E. D., Cho, J., Liu, Y., Finlay, J. E., Ezzati, M., Kaaya, S. F., & Fawzi, W. W. (2019). Lifetime economic impact of the burden of childhood stunting attributable to maternal psychosocial risk factors in 137

- low/middle-income countries. *BMJ Global Health*, 4(1), e001144.
<https://doi.org/10.1136/bmjgh-2018-001144>
- Fenske, N., Burns, J., Hothorn, T., & Rehfuess, E. A. (2013). Understanding Child Stunting in India: A Comprehensive Analysis of Socio-Economic, Nutritional and Environmental Determinants Using Additive Quantile Regression. *PLOS ONE*, 8(11), e78692.
<https://doi.org/10.1371/journal.pone.0078692>
- Fenta, H. M., Zewotir, T., & Muluneh, E. K. (2021). A machine learning classifier approach for identifying the determinants of under-five child undernutrition in Ethiopian administrative zones. *BMC Medical Informatics and Decision Making*, 21(1), 291. <https://doi.org/10.1186/s12911-021-01652-1>
- Ferdous, F., Das, S. K., Ahmed, S., Farzana, F. D., Latham, J. R., Chisti, M. J., Ud-Din, A. I. M. S., Azmi, I. J., Talukder, K. A., & Faruque, A. S. G. (2013). Severity of Diarrhea and Malnutrition among Under Five-Year-Old Children in Rural Bangladesh. *The American Journal of Tropical Medicine and Hygiene*, 89(2), 223–228. <https://doi.org/10.4269/ajtmh.12-0743>
- Fernandez-Felix, B. M., Barca, L. V., Garcia-Esquinas, E., Correa-Pérez, A., Fernández-Hidalgo, N., Muriel, A., Lopez-Alcalde, J., Álvarez-Díaz, N., Pijoan, J. I., Ribera, A., Elorza, E. N., Muñoz, P., Fariñas, M. del C., Goenaga, M. Á., & Zamora, J. (2021). Prognostic models for mortality after cardiac surgery in patients with infective endocarditis: A systematic review and aggregation of prediction models. *Clinical Microbiology and Infection*, 27(10), 1422–1430.
<https://doi.org/10.1016/j.cmi.2021.05.051>
- Feutz, E., Biswas, P. K., Ndeketa, L., Ogwel, B., Onwuchekwa, U., Sarwar, G., Sultana, S., Peñataro Yori, P., Acebedo, A., Ahmed, N., Ahmed, I., Atlas, H. E., Awuor, A. O., Bhuiyan, M. A. I., Conteh, B., Diawara, O., Elwood, S., Fane, M., Hossen, M. I., ... Galagan, S. R. (2024). Data Management in Multicountry Consortium Studies: The Enterics For Global Health (EFGH) Shigella Surveillance

- Study Example. *Open Forum Infectious Diseases*, 11(Suppl 1), S48–S57.
<https://doi.org/10.1093/ofid/ofad573>
- Foukalas, F. (2024). A Survey of Artificial Neural Network Computing Systems. *Cognitive Computation*, 17(1), 4. <https://doi.org/10.1007/s12559-024-10383-0>
- Fuchs, S. C., & Victora, C. G. (2002). Risk and prognostic factors for diarrheal disease in Brazilian infants: A special case-control design application. *Cadernos de Saúde Pública*, 18(3), 773–782.
<https://doi.org/10.1590/S0102-311X2002000300021>
- Gabriel, R. A., Sharma, B. S., Doan, C. N., Jiang, X., Schmidt, U. H., & Vaida, F. (2019). A Predictive Model for Determining Patients Not Requiring Prolonged Hospital Length of Stay After Elective Primary Total Hip Arthroplasty. *Anesthesia & Analgesia*, 129(1), 43.
<https://doi.org/10.1213/ANE.0000000000003798>
- Gabriel, R. A., Waterman, R. S., Kim, J., & Ohno-Machado, L. (2017). A Predictive Model for Extended Postanesthesia Care Unit Length of Stay in Outpatient Surgeries. *Anesthesia & Analgesia*, 124(5), 1529–1536. <https://doi.org/10.1213/ANE.0000000000001827>
- Ganguly, E., Sharma, P. K., & Bunker, C. H. (2015). Prevalence and risk factors of diarrhea morbidity among under-five children in India: A systematic review and meta-analysis. *Indian Journal of Child Health*, 2(4), 152–160.
- Gardner, J., & Brooks, C. (2017). A Statistical Framework for Predictive Model Evaluation in MOOCs. *Proceedings of the Fourth (2017) ACM Conference on Learning @ Scale*, 269–272.
<https://doi.org/10.1145/3051457.3054002>
- Garenne, M., Myatt, M., Khara, T., Dolan, C., & Briend, A. (2018). Concurrent wasting and stunting among under-five children in Niakhar, Senegal. *Maternal & Child Nutrition*, 15(2), e12736.
<https://doi.org/10.1111/mcn.12736>

- Geberselassie, S. B., Abebe, S. M., Melsew, Y. A., Mutuku, S. M., & Wassie, M. M. (2018). Prevalence of stunting and its associated factors among children 6-59 months of age in Libo-Kemekem district, Northwest Ethiopia; A community based cross sectional study. *PLOS ONE*, 13(5), e0195361. <https://doi.org/10.1371/journal.pone.0195361>
- Gebreayohanes, M., & Dessie, A. (2022). Prevalence of stunting and its associated factors among children 6–59 months of age in pastoralist community, Northeast Ethiopia: A community-based cross-sectional study. *PLOS ONE*, 17(2), e0256722. <https://doi.org/10.1371/journal.pone.0256722>
- Geldsetzer, P., Williams, T. C., Kirolos, A., Mitchell, S., Ratcliffe, L. A., Kohli-Lynch, M. K., Bischoff, E. J. L., Cameron, S., & Campbell, H. (2014). The recognition of and care seeking behaviour for childhood illness in developing countries: A systematic review. *PloS One*, 9(4), e93427. <https://doi.org/10.1371/journal.pone.0093427>
- Gennatas, E. D., & Chen, J. H. (2021). Chapter 1 - Artificial intelligence in medicine: Past, present, and future. In L. Xing, M. L. Giger, & J. K. Min (Eds.), *Artificial Intelligence in Medicine* (pp. 3–18). Academic Press. <https://doi.org/10.1016/B978-0-12-821259-2.00001-6>
- Gerber, A., & Kotze, P. (2015). *Design Science Research as Research Approach in Doctoral Studies*. https://www.cair.org.za/sites/default/files/2019-08/Paper_2_0.pdf
- Ghiasi, M. M., & Zendehboudi, S. (2021). Application of decision tree-based ensemble learning in the classification of breast cancer. *Computers in Biology and Medicine*, 128, 104089. <https://doi.org/10.1016/j.combiomed.2020.104089>
- Gholipour, C., Rahim, F., Fakhree, A., & Ziapour, B. (2015). Using an Artificial Neural Networks (ANNs) Model for Prediction of Intensive Care Unit (ICU) Outcome and Length of Stay at Hospital in Traumatic Patients. *Journal of Clinical and Diagnostic Research : JCDR*, 9(4), OC19–OC23. <https://doi.org/10.7860/JCDR/2015/9467.5828>

- Giannattasio, A., Guarino, A., & Lo Vecchio, A. (2016). Management of children with prolonged diarrhea. *F1000Research*, 5. <https://doi.org/10.12688/f1000research.7469.1>
- Gilson, L., Alliance for Health Policy and Systems Research, & World Health Organization. (2012). *Health Policy and Systems Research: A Methodology Reader*. 18.
- Giuffrè, M., & Shung, D. L. (2023). Harnessing the power of synthetic data in healthcare: Innovation, application, and privacy. *Npj Digital Medicine*, 6(1), Article 1. <https://doi.org/10.1038/s41746-023-00927-3>
- Global Burden of Disease 2017 collaborators. (2017). *Quantifying risks and interventions that have reduced mortality due to diarrhea among children younger than 5 years old: An analysis of the Global Burden of Disease study 2017*. https://www.open-access.bcu.ac.uk/8324/1/diarrhea_manuscript_accepted.pdf
- Gonçalves, C. P., & Rouco, J. (2020). *Comparing Decision Tree-Based Ensemble Machine Learning Models for COVID-19 Death Probability Profiling* (p. 2020.12.06.20244756). medRxiv. <https://doi.org/10.1101/2020.12.06.20244756>
- Gonzales, A., Guruswamy, G., & Smith, S. R. (2023). Synthetic data in health care: A narrative review. *PLOS Digital Health*, 2(1), e0000082. <https://doi.org/10.1371/journal.pdig.0000082>
- González-Gonzalo, C., Thee, E. F., Klaver, C. C. W., Lee, A. Y., Schlingemann, R. O., Tufail, A., Verbraak, F., & Sánchez, C. I. (2022). Trustworthy AI: Closing the gap between development and integration of AI systems in ophthalmic practice. *Progress in Retinal and Eye Research*, 90, 101034. <https://doi.org/10.1016/j.preteyeres.2021.101034>
- Google for Developers. (2022, July 18). *Descending into ML: Training and Loss | Machine Learning*. Google for Developers. <https://developers.google.com/machine-learning/crash-course/descending-into-ml/training-and-loss>

- Goweke, V. F., Kinabo, J., Jumbe, T., Rybak, C., & Stuetz, W. (2021). High Prevalence of Stunting and Anaemia Is Associated with Multiple Micronutrient Deficiencies in School Children of Small-Scale Farmers from Chamwino and Kilosa Districts, Tanzania. *Nutrients*, 13(5), Article 5. <https://doi.org/10.3390/nu13051576>
- Grandini, M., Bagli, E., & Visani, G. (2020). *Metrics for Multi-Class Classification: An Overview* (arXiv:2008.05756). arXiv. <https://doi.org/10.48550/arXiv.2008.05756>
- Griffin, P. M., Ryan, C. A., Nyaphisi, M., Hargrett-Bean, N., Waldman, R. J., & Blake, P. A. (1988). Risk factors for fatal diarrhea: A case-control study of African children. *American Journal of Epidemiology*, 128(6), 1322–1329. <https://doi.org/10.1093/oxfordjournals.aje.a115085>
- Guandalini, S. (2020). *Diarrhea: Practice Essentials, Background, Pathophysiology*. <https://emedicine.medscape.com/article/928598-overview?form=fpf>
- Guarino, A., & Giannattasio, A. (2011). New molecular approaches in the diagnosis of acute diarrhea: Advantages for clinicians and researchers. *Current Opinion in Gastroenterology*, 27(1), 24–29. <https://doi.org/10.1097/MOG.0b013e3283413750>
- Guerrant, R. L., DeBoer, M. D., Moore, S. R., Scharf, R. J., & Lima, A. A. M. (2013). The impoverished gut—A triple burden of diarrhoea, stunting and chronic disease. *Nature Reviews. Gastroenterology & Hepatology*, 10(4), 220–229. <https://doi.org/10.1038/nrgastro.2012.239>
- Gupta, A., Kumar, L., Jain, R., & Nagrath, P. (2020). Heart Disease Prediction Using Classification (Naive Bayes). In P. K. Singh, W. Pawłowski, S. Tanwar, N. Kumar, J. J. P. C. Rodrigues, & M. S. Obaidat (Eds.), *Proceedings of First International Conference on Computing, Communications, and Cyber-Security (IC4S 2019)* (pp. 561–573). Springer. https://doi.org/10.1007/978-981-15-3369-3_42
- Gutiérrez-Gutiérrez, B., Salamanca, E., de Cueto, M., Hsueh, P.-R., Viale, P., Paño-Pardo, J. R., Venditti, M., Tumbarello, M., Daikos, G., Pintado, V., Doi, Y., Tuon, F. F., Karaïskos, I., Machuca, I., Schwaber, M. J., Azap, Özlem Kurt, Souli, M., Roilides, E., Pournaras, S., ... Jové, E. (2016). A

- Predictive Model of Mortality in Patients With Bloodstream Infections due to Carbapenemase-Producing Enterobacteriaceae. *Mayo Clinic Proceedings*, 91(10), 1362–1371.
<https://doi.org/10.1016/j.mayocp.2016.06.024>
- Habebh, H., & Gohel, S. (2021). Machine Learning in Healthcare. *Current Genomics*, 22(4), 291–300.
<https://doi.org/10.2174/1389202922666210705124359>
- Hakami, A. (2024). Strategies for overcoming data scarcity, imbalance, and feature selection challenges in machine learning models for predictive maintenance. *Scientific Reports*, 14(1), 9645.
<https://doi.org/10.1038/s41598-024-59958-9>
- Halder, R., Uddin, M., Uddin, M. A., Aryal, S., & Khraisat, A. (2024). Enhancing K-nearest neighbor algorithm: A comprehensive review and performance analysis of modifications. *Journal of Big Data*, 11. <https://doi.org/10.1186/s40537-024-00973-y>
- Hanchuk, D., & Semerikov, S. (2025). *Implementing MLOps practices for effective machine learning model deployment: A meta synthesis*.
- Hanieh, S., Braat, S., Simpson, J. A., Ha, T. T. T., Tran, T. D., Tuan, T., Fisher, J., & Biggs, B.-A. (2019). The Stunting Tool for Early Prevention: Development and external validation of a novel tool to predict risk of stunting in children at 3 years of age. *BMJ Global Health*, 4(6), e001801.
<https://doi.org/10.1136/bmjgh-2019-001801>
- Hao, X., Zhang, G., & Ma, S. (2016). Deep Learning. *International Journal of Semantic Computing*, 10(03), 417–439. <https://doi.org/10.1142/S1793351X16500045>
- Haque, M. A., Choudhury, N., Wahid, B. Z., Ahmed, S. T., Farzana, F. D., Ali, M., Naz, F., Siddiqua, T. J., Rahman, S. S., Faruque, A. S. G., & Ahmed, T. (2023). A predictive modelling approach to illustrate factors correlating with stunting among children aged 12–23 months: A cluster randomised pre-post study. *BMJ Open*, 13(4), e067961. <https://doi.org/10.1136/bmjopen-2022-067961>

- Harish, K. B., Price, W. N., & Aphinyanaphongs, Y. (2022). Open-Source Clinical Machine Learning Models: Critical Appraisal of Feasibility, Advantages, and Challenges. *JMIR Formative Research*, 6(4), e33970. <https://doi.org/10.2196/33970>
- Harrell Jr, F. E. (2023). *rms: Regression Modeling Strategies* (Version 6.7-1) [Computer software]. <https://cran.r-project.org/web/packages/rms/index.html>
- Hartman, R. M., Cohen, A. L., Antoni, S., Mwenda, J., Weldegebriel, G., Biey, J., Shaba, K., de Oliveira, L., Rey, G., Ortiz, C., Tereza, M., Fahmy, K., Ghoniem, A., Ashmony, H., Videbaek, D., Singh, S., Tondo, E., Sharifuzzaman, M., Liyanage, J., ... Nakamura, T. (2022). Risk Factors for Mortality Among Children Younger Than Age 5 Years With Severe Diarrhea in Low- and Middle-income Countries: Findings From the World Health Organization-coordinated Global Rotavirus and Pediatric Diarrhea Surveillance Networks. *Clinical Infectious Diseases*, ciac561. <https://doi.org/10.1093/cid/ciac561>
- Hartzband, P., & Groopman, J. (2020). Physician Burnout, Interrupted. *New England Journal of Medicine*, 382(26), 2485–2487. <https://doi.org/10.1056/NEJMp2003149>
- Hasan, Md. Z., Mehdi, G. G., De Broucker, G., Ahmed, S., Ali, Md. W., Martin Del Campo, J., Constenla, D., Patenaude, B., & Uddin, Md. J. (2021). The economic burden of diarrhea in children under 5 years in Bangladesh. *International Journal of Infectious Diseases*, 107, 37–46. <https://doi.org/10.1016/j.ijid.2021.04.038>
- Hassanipour, S., Ghaem, H., Arab-Zozani, M., Seif, M., Fararouei, M., Abdzadeh, E., Sabetian, G., & Paydar, S. (2019). Comparison of artificial neural network and logistic regression models for prediction of outcomes in trauma patients: A systematic review and meta-analysis. *Injury*, 50(2), 244–250. <https://doi.org/10.1016/j.injury.2019.01.007>
- Heber, S., Pereyra, D., Schrottmaier, W. C., Kammerer, K., Santol, J., Rumpf, B., Pawelka, E., Hanna, M., Scholz, A., Liu, M., Hell, A., Heiplik, K., Lickefett, B., Havervall, S., Traugott, M. T., Neuböck, M. J.,

- Schörgenhofer, C., Seitz, T., Firbas, C., ... Assinger, A. (2022). A Model Predicting Mortality of Hospitalized Covid-19 Patients Four Days After Admission: Development, Internal and Temporal-External Validation. *Frontiers in Cellular and Infection Microbiology*, 11.
<https://www.frontiersin.org/articles/10.3389/fcimb.2021.795026>
- Hemingway, H., Croft, P., Perel, P., Hayden, J. A., Abrams, K., Timmis, A., Briggs, A., Udumyan, R., Moons, K. G. M., Steyerberg, E. W., Roberts, I., Schroter, S., Altman, D. G., & Riley, R. D. (2013). Prognosis research strategy (PROGRESS) 1: A framework for researching clinical outcomes. *BMJ*, 346, e5595. <https://doi.org/10.1136/bmj.e5595>
- Herrera, F., Charte, F., Rivera, A. J., & del Jesus, M. J. (2016). Multilabel Classification. In F. Herrera, F. Charte, A. J. Rivera, & M. J. del Jesus (Eds.), *Multilabel Classification: Problem Analysis, Metrics and Techniques* (pp. 17–31). Springer International Publishing. https://doi.org/10.1007/978-3-319-41111-8_2
- Hessel, F. (2008). Burden of DiseaseBurdenof disease(s). In W. Kirch (Ed.), *Encyclopedia of Public Health* (pp. 94–96). Springer Netherlands. https://doi.org/10.1007/978-1-4020-5614-7_297
- Hevner, A., & Wickramasinghe, N. (2018). *Design Science Research Opportunities in Health Care* (pp. 3–18). https://doi.org/10.1007/978-3-319-72287-0_1
- Heymann, H., Kies, A. D., Frye, M., Schmitt, R. H., & Boza, A. (2022). Guideline for Deployment of Machine Learning Models for Predictive Quality in Production. *Procedia CIRP*, 107, 815–820. <https://doi.org/10.1016/j.procir.2022.05.068>
- Hinostroza Fuentes, V. G., Karim, H. A., Tan, M. J. T., & AlDahoul, N. (2025). AI with agency: A vision for adaptive, efficient, and ethical healthcare. *Frontiers in Digital Health*, 7, 1600216. <https://doi.org/10.3389/fdgth.2025.1600216>

- Hoddinott, J., Alderman, H., Behrman, J. R., Haddad, L., & Horton, S. (2013). The economic rationale for investing in stunting reduction. *Maternal & Child Nutrition*, 9(S2), 69–82.
<https://doi.org/10.1111/mcn.12080>
- Holm, J., Frumento, P., Almondo, G., Gémes, K., Bottai, M., Alexanderson, K., Friberg, E., & Farrants, K. (2021). Predicting the duration of sickness absence due to knee osteoarthritis: A prognostic model developed in a population-based cohort in Sweden. *BMC Musculoskeletal Disorders*, 22(1), 603. <https://doi.org/10.1186/s12891-021-04400-8>
- Holmström, J., Ketokivi, M., & Hameri, A.-P. (2009). Bridging Practice and Theory: A Design Science Approach. *Decision Sciences*, 40(1), 65–87. <https://doi.org/10.1111/j.1540-5915.2008.00221.x>
- Horne, B., Badji, H., Bhuiyan, M. T. R., Romaina Cachique, L., Cornick, J., Hotwani, A., Juma, J., Ochieng, J. B., Abdou, M., Apondi, E., Atlas, H. E., Awuor, A. O., Baker, K. S., Ceesay, B. E., Charles, M., Cunliffe, N. A., Feutz, E., Galagan, S. R., Guindo, I., ... Tennant, S. M. (2024). Microbiological Methods Used in the Enterics for Global Health Shigella Surveillance Study. *Open Forum Infectious Diseases*, 11(Supplement_1), S25–S33. <https://doi.org/10.1093/ofid/ofad576>
- Hossain, M. I., Faruque, A. S. G., Sarmin, M., Chisti, M. J., & Ahmed, T. (2022). Prolonged diarrhea among under-five children in Bangladesh: Burden and risk factors. *PLOS ONE*, 17(10), e0273148. <https://doi.org/10.1371/journal.pone.0273148>
- Houthooft, R., Ruyssinck, J., van der Hertten, J., Stijven, S., Couckuyt, I., Gadeyne, B., Ongenaes, F., Colpaert, K., Decruyenaere, J., Dhaene, T., & De Turck, F. (2015). Predictive modelling of survival and length of stay in critically ill patients using sequential organ failure scores. *Artificial Intelligence in Medicine*, 63(3), 191–207. <https://doi.org/10.1016/j.artmed.2014.12.009>
- Hu, Y., Zhang, X., Slavin, V., Belsti, Y., Tiruneh, S. A., Callander, E., & Enticott, J. (2025). Beyond Comparing Machine Learning and Logistic Regression in Clinical Prediction Modelling: Shifting

- from Model Debate to Data Quality. *Journal of Medical Internet Research*, 27, e77721.
<https://doi.org/10.2196/77721>
- Huang, M.-W., Chiu, C.-H., Tsai, C.-F., & Lin, W.-C. (2021). On Combining Feature Selection and Over-Sampling Techniques for Breast Cancer Prediction. *Applied Sciences*, 11(14), Article 14.
<https://doi.org/10.3390/app11146574>
- Huang, Y., Li, W., Macheret, F., Gabriel, R. A., & Ohno-Machado, L. (2020). A tutorial on calibration measurements and calibration models for clinical prediction models. *Journal of the American Medical Informatics Association*, 27(4), 621–633. <https://doi.org/10.1093/jamia/ocz228>
- IBM. (2020). *What is Machine Learning?* | IBM. <https://www.ibm.com/cloud/learn/machine-learning>
- IBM. (2021). *Evaluate and select a machine learning algorithm—IBM Garage Practices*.
<https://www.ibm.com/garage/method/practices/reason/evaluate-and-select-machine-learning-algorithm/>
- IBM. (2023). *What is the k-nearest neighbors algorithm?* | IBM. <https://www.ibm.com/topics/knn>
- IBM Institute for Business Value. (2018). *Consumers at the heart of healthcare*.
<https://www.ibm.com/downloads/cas/5MWZ9BWQ>
- Ikeda, N., Irie, Y., & Shibuya, K. (2013). Determinants of reduced child stunting in Cambodia: Analysis of pooled data from three Demographic and Health Surveys. *Bulletin of the World Health Organization*, 91, 341–349. <https://doi.org/10.2471/BLT.12.113381>
- Institute for Health Metrics and Evaluation. (2020, November 25). *Kenya—Siaya*. Institute for Health Metrics and Evaluation. <https://www.healthdata.org/kenya-siaya>
- Ippoliti, R., Falavigna, G., Zanelli, C., Bellini, R., & Numico, G. (2021). Neural networks and hospital length of stay: An application to support healthcare management with national benchmarks and thresholds. *Cost Effectiveness and Resource Allocation*, 19(1), 67.
<https://doi.org/10.1186/s12962-021-00322-3>

- Irizarry, R. A. (2019). *Chapter 29 Cross validation / Introduction to Data Science*.
<https://rafalab.github.io/dsbook/cross-validation.html>
- Iwendi, C., Bashir, A. K., Peshkar, A., Sujatha, R., Chatterjee, J. M., Pasupuleti, S., Mishra, R., Pillai, S., & Jo, O. (2020). COVID-19 Patient Health Prediction Using Boosted Random Forest Algorithm. *Frontiers in Public Health*, 8. <https://www.frontiersin.org/articles/10.3389/fpubh.2020.00357>
- Jabbar, M. A., & Samreen, S. (2016). Heart disease prediction system based on hidden naïve bayes classifier. *2016 International Conference on Circuits, Controls, Communications and Computing (I4C)*, 1–5. <https://doi.org/10.1109/CIMCA.2016.8053261>
- Jackson, S. L. (2008). *Research methods and statistics: A critical thinking approach* (3rd ed). Heinle Cengage Learning.
- Jackson, S., Yaqub, M., & Li, C.-X. (2019). The Agile Deployment of Machine Learning Models in Healthcare. *Frontiers in Big Data*, 1.
<https://www.frontiersin.org/articles/10.3389/fdata.2018.00007>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2022). An introduction to statistical learning with applications in R. *Statistical Theory and Related Fields*, 6(1), 87–87.
<https://doi.org/10.1080/24754269.2021.1980261>
- James, G., Witten, D., Hastie, T., Tibshirani, R., & Taylor, J. (2023). Unsupervised Learning. In G. James, D. Witten, T. Hastie, R. Tibshirani, & J. Taylor (Eds.), *An Introduction to Statistical Learning: With Applications in Python* (pp. 503–556). Springer International Publishing.
https://doi.org/10.1007/978-3-031-38747-0_12
- Janardhanan, P., L., H., & Sabika, F. (2015). Effectiveness of Support Vector Machines in Medical Data mining. *Journal of Communications Software and Systems*, 11(1), 25.
<https://doi.org/10.24138/jcomss.v11i1.114>

- Jannah, D., & S, A. (2025). KNN Vs Naive Bayes: An Innovative Comparison in Predictive AI Learning With Association Data Support. *Journal of Digital Technology and Computer Science*, 3, 66–79.
<https://doi.org/10.61220/digitech.v3i1.20251>
- Javaid, M., Haleem, A., Pratap Singh, R., Suman, R., & Rab, S. (2022). Significance of machine learning in healthcare: Features, pillars and applications. *International Journal of Intelligent Networks*, 3, 58–73. <https://doi.org/10.1016/j.ijin.2022.05.002>
- Jayatilake, S. M. D. A. C., & Ganegoda, G. U. (2021). Involvement of Machine Learning Tools in Healthcare Decision Making. *Journal of Healthcare Engineering*, 2021, e6679512.
<https://doi.org/10.1155/2021/6679512>
- Ji, X., & Kattan, M. W. (2018). Tutorial: Development of an online risk calculator platform. *Annals of Translational Medicine*, 6(3), 46. <https://doi.org/10.21037/atm.2017.11.37>
- Jiang, L., Cai, Z., Zhang, H., & Wang, D. (2013). Naive Bayes text classifiers: A locally weighted learning approach. *Journal of Experimental & Theoretical Artificial Intelligence*, 25.
<https://doi.org/10.1080/0952813X.2012.721010>
- Jiang, Y., Su, X., Wang, C., Zhang, L., Zhang, X., Wang, L., & Cui, Y. (2015). Prevalence and risk factors for stunting and severe stunting among children under three years old in mid-western rural areas of China. *Child: Care, Health and Development*, 41(1), 45–51. <https://doi.org/10.1111/cch.12148>
- Johnson, Y., & Kaneene, J. (2018). *Epidemiology: From Recognition to Results* (pp. 1–30).
<https://doi.org/10.1002/9781119194521.ch1>
- Jones, L. D., Golan, D., Hanna, S. A., & Ramachandran, M. (2018). Artificial intelligence, machine learning and the evolution of healthcare. *Bone & Joint Research*, 7(3), 223–225.
<https://doi.org/10.1302/2046-3758.73.BJR-2017-0147.R1>
- Jun, Z. (2021). The Development and Application of Support Vector Machine. *Journal of Physics: Conference Series*, 1748(5), 052006. <https://doi.org/10.1088/1742-6596/1748/5/052006>

- Kaiser, L., Babaeizadeh, M., Milos, P., Osinski, B., Campbell, R. H., Czechowski, K., Erhan, D., Finn, C., Kozakowski, P., Levine, S., Mohiuddin, A., Sepassi, R., Tucker, G., & Michalewski, H. (2020). *Model-Based Reinforcement Learning for Atari* (arXiv:1903.00374). arXiv. <https://doi.org/10.48550/arXiv.1903.00374>
- Kalagara, S., Eltorai, A. E. M., Durand, W. M., DePasse, J. M., & Daniels, A. H. (2018). Machine learning modeling for predicting hospital readmission following lumbar laminectomy. *Journal of Neurosurgery. Spine*, 30(3), 344–352. <https://doi.org/10.3171/2018.8.SPINE1869>
- Kang, Y., Aguayo, V. M., Campbell, R. K., Dzed, L., Joshi, V., Waid, J. L., Gupta, S. D., Haselow, N. J., & West Jr., K. P. (2018). Nutritional status and risk factors for stunting in preschool children in Bhutan. *Maternal & Child Nutrition*, 14(S4), e12653. <https://doi.org/10.1111/mcn.12653>
- Karambizi, N. U., McMahan, C. S., Blue, C. N., & Temesvari, L. A. (2021). Global estimated Disability-Adjusted Life-Years (DALYs) of diarrheal diseases: A systematic analysis of data from 28 years of the global burden of disease study. *PLoS ONE*, 16(10), e0259077. <https://doi.org/10.1371/journal.pone.0259077>
- Karamitsos, I., Albarhami, S., & Apostolopoulos, C. (2020). Applying DevOps Practices of Continuous Automation for Machine Learning. *Information*, 11(7), Article 7. <https://doi.org/10.3390/info11070363>
- Karamizadeh, S., Abdullah, S., Asl, M., Shayan, J., & Rajabi, M. (2014). Advantage and Drawback of Support Vector Machine Functionality. In *I4CT 2014—1st International Conference on Computer, Communications, and Control Technology, Proceedings*. <https://doi.org/10.1109/I4CT.2014.6914146>
- Kariri, E., Louati, H., Louati, A., & Masmoudi, F. (2023). Exploring the Advancements and Future Research Directions of Artificial Neural Networks: A Text Mining Approach. *Applied Sciences*, 13, 3186. <https://doi.org/10.3390/app13053186>

- Karlsson, O., Kim, R., Moloney, G. M., Hasman, A., & Subramanian, S. V. (2023). Patterns in child stunting by age: A cross-sectional study of 94 low- and middle-income countries. *Maternal & Child Nutrition*, 19(4), e13537. <https://doi.org/10.1111/mcn.13537>
- Kasprzak, P., Mitchell, L., Kravchuk, O., & Timmins, A. (2020). Six Years of Shiny in Research—Collaborative Development of Web Tools in R. *The R Journal*, 12(2), 155. <https://doi.org/10.32614/RJ-2021-004>
- Kassambara. (2018). *Cross-Validation Essentials in R - Articles—STHDA*. <http://www.sthda.com/english/articles/38-regression-model-validation/157-cross-validation-essentials-in-r/>
- KEMRI-CGHR. (2020). *Health and Demographic Surveillance System (HDSS) Fact Sheet 2020*.
- Kent, P., Cancelliere, C., Boyle, E., Cassidy, J. D., & Kongsted, A. (2020). A conceptual framework for prognostic research. *BMC Medical Research Methodology*, 20(1), 172. <https://doi.org/10.1186/s12874-020-01050-7>
- Kenya National Bureau of Statistics. (2015). *Kenya Demographic and Health Survey, 2014*. <https://dhsprogram.com/pubs/pdf/fr308/fr308.pdf>
- Kenya National Bureau of Statistics, & Ministry of Health. (2023, January). 2022 Kenya Demographic and Health Survey Key Indicators Report. *Kenya National Bureau of Statistics*. <https://www.knbs.or.ke/download/2022-kdhs-key-indicators-report/>
- Khagayi, S., Burton, D. C., Onkoba, R., Ochieng, B., Ismail, A., Mutonga, D., Muthoni, J., Feikin, D. R., Breiman, R. F., Mwenda, J. M., Odhiambo, F., & Laserson, K. F. (2014). High Burden of Rotavirus Gastroenteritis in Young Children in Rural Western Kenya, 2010–2011. *The Pediatric Infectious Disease Journal*, 33, S34. <https://doi.org/10.1097/INF.0000000000000049>
- Khagayi, S., Omere, R., Otieno, G. P., Ogwel, B., Ochieng, J. B., Juma, J., Apondi, E., Bigogo, G., Onyango, C., Ngama, M., Njeru, R., Owor, B. E., Mwanga, M. J., Addo, Y., Tabu, C., Amwayi, A., Mwenda, J.

- M., Tate, J. E., Parashar, U. D., ... Verani, J. R. (2020). Effectiveness of Monovalent Rotavirus Vaccine Against Hospitalization With Acute Rotavirus Gastroenteritis in Kenyan Children. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, 70(11), 2298–2305. <https://doi.org/10.1093/cid/ciz664>
- Khalaf, M., Hussain, A., Alfandi, O., Al-Jumeily Obe, D., Alloghani, M., Alsaadi, M., Dawood, O. A., & Abd, D. (2019). *An Application of Using Support Vector Machine Based on Classification Technique for Predicting Medical Data Sets* (pp. 580–591). https://doi.org/10.1007/978-3-030-26969-2_55
- Khalid, S., Khalil, T., & Nasreen, S. (2014). A survey of feature selection and feature extraction techniques in machine learning. *2014 Science and Information Conference*, 372–378. <https://doi.org/10.1109/SAI.2014.6918213>
- Khalilia, M., Chakraborty, S., & Popescu, M. (2011). Predicting disease risks from highly imbalanced data using random forest. *BMC Medical Informatics and Decision Making*, 11(1), 51. <https://doi.org/10.1186/1472-6947-11-51>
- Khalilia, M., Choi, M., Henderson, A., Iyengar, S., Braunstein, M., & Sun, J. (2015). Clinical Predictive Modeling Development and Deployment through FHIR Web Services. *AMIA Annual Symposium Proceedings*, 2015, 717–726.
- Khan, N., Raza, M. A., Mirjat, N. H., Balouch, N., Abbas, G., Yousef, A., & Touti, E. (2024). Unveiling the predictive power: A comprehensive study of machine learning model for anticipating chronic kidney disease. *Frontiers in Artificial Intelligence*, 6. <https://doi.org/10.3389/frai.2023.1339988>
- Kharya, S., Agrawal, S., & Soni, S. (2014). Naive Bayes Classifiers: A Probabilistic Detection Model for Breast Cancer. *International Journal of Computer Applications*, 92(10), 26–31. <https://doi.org/10.5120/16045-5206>
- Kimani-Murage, E. W., Kahn, K., Pettifor, J. M., Tollman, S. M., Dunger, D. B., Gómez-Olivé, X. F., & Norris, S. A. (2010). The prevalence of stunting, overweight and obesity, and metabolic disease

- risk in rural South African children. *BMC Public Health*, 10(1), 158.
<https://doi.org/10.1186/1471-2458-10-158>
- Kismul, H., Acharya, P., Mapatano, M. A., & Hatløy, A. (2017). Determinants of childhood stunting in the Democratic Republic of Congo: Further analysis of Demographic and Health Survey 2013–14. *BMC Public Health*, 18(1), 74. <https://doi.org/10.1186/s12889-017-4621-0>
- Kleinke, K., Reinecke, J., Salfrán, D., & Spiess, M. (2020). *Applied Multiple Imputation: Advantages, Pitfalls, New Developments and Applications in R*. Springer International Publishing.
<https://doi.org/10.1007/978-3-030-38164-6>
- Klug, M., Barash, Y., Bechler, S., Resheff, Y. S., Tron, T., Ironi, A., Soffer, S., Zimlichman, E., & Klang, E. (2020). A Gradient Boosting Machine Learning Model for Predicting Early Mortality in the Emergency Department Triage: Devising a Nine-Point Triage Score. *Journal of General Internal Medicine*, 35(1), 220–227. <https://doi.org/10.1007/s11606-019-05512-7>
- Kokonya, D. (2014). Burnout Syndrome among Medical Workers at Kenyatta National Hospital (KNH), Nairobi, Kenya. *African Journal of Psychiatry*, 17. <https://doi.org/10.4172/1994-8220.1000142>
- Kong, J., Rios, T., Kowalczyk, W., Menzel, S., & Bäck, T. (2020). On the Performance of Oversampling Techniques for Class Imbalance Problems. *Advances in Knowledge Discovery and Data Mining*, 12085, 84–96. https://doi.org/10.1007/978-3-030-47436-2_7
- Kotloff, K. L., Blackwelder, W. C., Nasrin, D., Nataro, J. P., Farag, T. H., van Eijk, A., Adegbola, R. A., Alonso, P. L., Breiman, R. F., Golam Faruque, A. S., Saha, D., Sow, S. O., Sur, D., Zaidi, A. K. M., Biswas, K., Panchalingam, S., Clemens, J. D., Cohen, D., Glass, R. I., ... Levine, M. M. (2012). The Global Enteric Multicenter Study (GEMS) of Diarrheal Disease in Infants and Young Children in Developing Countries: Epidemiologic and Clinical Methods of the Case/Control Study. *Clinical Infectious Diseases*, 55(suppl_4), S232–S245. <https://doi.org/10.1093/cid/cis753>

- Kotloff, K. L., Nataro, J. P., Blackwelder, W. C., Nasrin, D., Farag, T. H., Panchalingam, S., Wu, Y., Sow, S. O., Sur, D., Breiman, R. F., Faruque, A. S., Zaidi, A. K., Saha, D., Alonso, P. L., Tamboura, B., Sanogo, D., Onwuchekwa, U., Manna, B., Ramamurthy, T., ... Levine, M. M. (2013). Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): A prospective, case-control study. *Lancet (London, England)*, 382(9888), 209–222. [https://doi.org/10.1016/S0140-6736\(13\)60844-2](https://doi.org/10.1016/S0140-6736(13)60844-2)
- Kotu, V., & Deshpande, B. (2015). Chapter 2—Data Mining Process. In V. Kotu & B. Deshpande (Eds.), *Predictive Analytics and Data Mining* (pp. 17–36). Morgan Kaufmann. <https://doi.org/10.1016/B978-0-12-801460-8.00002-1>
- Kramer, A. A., & Zimmerman, J. E. (2010). A predictive model for the early identification of patients at risk for a prolonged intensive care unit length of stay. *BMC Medical Informatics and Decision Making*, 10(1), 27. <https://doi.org/10.1186/1472-6947-10-27>
- Kuhn, M. (2019). *11 Subsampling For Class Imbalances | The caret Package*. <https://topepo.github.io/caret/subsampling-for-class-imbalances.html>
- Kuhn, M., cre, Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Kenkel, B., R Core Team, Benesty, M., Lescarbeau, R., Ziem, A., Scrucca, L., Tang, Y., Candan, C., & Hunt, T. (2022). *caret: Classification and Regression Training* (Version 6.0-93) [Computer software]. <https://CRAN.R-project.org/package=caret>
- Kumari, R., & Kr., S. (2017). Machine Learning: A Review on Binary Classification. *International Journal of Computer Applications*, 160(7), 11–15. <https://doi.org/10.5120/ijca2017913083>
- Kursa, M. B., & Rudnicki, W. R. (2020, May 21). *Package ‘Boruta.’* <https://cran.r-project.org/web/packages/Boruta/Boruta.pdf>

- Kyrychek, M. (2025). Gradient boosting as a tool for solving classification problems in data-constrained environments. *Technologies and Engineering*, 26, 37–47. <https://doi.org/10.30857/2786-5371.2025.2.3>
- Lamberti, L. M., Fischer Walker, C. L., & Black, R. E. (2012). Systematic review of diarrhea duration and severity in children and adults in low- and middle-income countries. *BMC Public Health*, 12, 276. <https://doi.org/10.1186/1471-2458-12-276>
- Lanata, C. F., Black, R. E., Gilman, R. H., Lazo, F., & Aguila, R. D. (1991). Epidemiologic, Clinical, and Laboratory Characteristics of Acute vs. Persistent Diarrhea in Periurban Lima, Peru. *Journal of Pediatric Gastroenterology and Nutrition*, 12(1), 82.
- Lanata, C. F., Fischer-Walker, C. L., Olascoaga, A. C., Torres, C. X., Aryee, M. J., Black, R. E., & Child Health Epidemiology Reference Group of the World Health Organization and UNICEF. (2013). Global causes of diarrheal disease mortality in children <5 years of age: A systematic review. *PloS One*, 8(9), e72788. <https://doi.org/10.1371/journal.pone.0072788>
- Lanata, C. F., Fischer-Walker, C. L., Olascoaga, A. C., Torres, C. X., Aryee, M. J., Black, R. E., & Unicef, for the C. H. E. R. G. of the W. H. O. and. (2013). Global Causes of Diarrheal Disease Mortality in Children <5 Years of Age: A Systematic Review. *PLOS ONE*, 8(9), e72788. <https://doi.org/10.1371/journal.pone.0072788>
- Langarizadeh, M., & Moghbeli, F. (2016). Applying Naive Bayesian Networks to Disease Prediction: A Systematic Review. *Acta Informatica Medica*, 24(5), 364–369. <https://doi.org/10.5455/aim.2016.24.364-369>
- Lau, F. (2017). Chapter 12 Methods for Correlational Studies. In *Handbook of eHealth Evaluation: An Evidence-based Approach [Internet]*. University of Victoria. <https://www.ncbi.nlm.nih.gov/books/NBK481614/>

- Lea-Pereira, M. C., Amaya-Pascasio, L., Martínez-Sánchez, P., Rodríguez Salvador, M. del M., Galván-Espinosa, J., Téllez-Ramírez, L., Reche-Lorite, F., Sánchez, M.-J., & García-Torrecillas, J. M. (2022). Predictive Model and Mortality Risk Score during Admission for Ischaemic Stroke with Conservative Treatment. *International Journal of Environmental Research and Public Health*, 19(6), Article 6. <https://doi.org/10.3390/ijerph19063182>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), Article 7553. <https://doi.org/10.1038/nature14539>
- Lee, G., Yori, P. P., Olortegui, M. P., Caulfield, L. E., Sack, D. A., Fischer-Walker, C., Black, R. E., & Kosek, M. (2014). An instrument for the assessment of diarrhoeal severity based on a longitudinal community-based study. *BMJ Open*, 4(6), e004816. <https://doi.org/10.1136/bmjopen-2014-004816>
- Lee, M., Gatton, T. M., & Lee, K.-K. (2010). A Monitoring and Advisory System for Diabetes Patient Management Using a Rule-Based Method and KNN. *Sensors*, 10(4), Article 4. <https://doi.org/10.3390/s100403934>
- Leisman, D. E., Harhay, M. O., Lederer, D. J., Abramson, M., Adjei, A. A., Bakker, J., Ballas, Z. K., Barreiro, E., Bell, S. C., Bellomo, R., Bernstein, J. A., Branson, R. D., Brusasco, V., Chalmers, J. D., Chokroverty, S., Citerio, G., Collop, N. A., Cooke, C. R., Crapo, J. D., ... Maslove, D. M. (2020). Development and Reporting of Prediction Models: Guidance for Authors From Editors of Respiratory, Sleep, and Critical Care Journals. *Critical Care Medicine*, 48(5), 623–633. <https://doi.org/10.1097/CCM.0000000000004246>
- Lenters, L., Wazny, K., & Bhutta, Z. A. (2016). Management of Severe and Moderate Acute Malnutrition in Children. In R. E. Black, R. Laxminarayan, M. Temmerman, & N. Walker (Eds.), *Reproductive, Maternal, Newborn, and Child Health: Disease Control Priorities, Third Edition (Volume 2)*. The

International Bank for Reconstruction and Development / The World Bank.

<http://www.ncbi.nlm.nih.gov/books/NBK361900/>

Lequertier, V., Wang, T., Fondrevelle, J., Augusto, V., & Duclos, A. (2021). Hospital Length of Stay

Prediction Methods: A Systematic Review. *Medical Care*, 59(10), 929–938.

<https://doi.org/10.1097/MLR.0000000000001596>

Levin, S., Barnes, S., Toerper, M., Debraine, A., DeAngelo, A., Hamrock, E., Hinson, J., Hoyer, E.,

Dungarani, T., & Howell, E. (2021). Machine-learning-based hospital discharge predictions can support multidisciplinary rounds and decrease hospital length-of-stay. *BMJ Innovations*, 7(2).

<https://doi.org/10.1136/bmjinnov-2020-000420>

Levine, M. M., Nasrin, D., Acácio, S., Bassat, Q., Powell, H., Tennant, S. M., Sow, S. O., Sur, D., Zaidi, A. K.

M., Faruque, A. S. G., Hossain, M. J., Alonso, P. L., Breiman, R. F., O'Reilly, C. E., Mintz, E. D.,

Omoro, R., Ochieng, J. B., Oundo, J. O., Tamboura, B., ... Kotloff, K. L. (2020). Diarrhoeal disease and subsequent risk of death in infants and children residing in low-income and middle-income

countries: Analysis of the GEMS case-control study and 12-month GEMS-1A follow-on study. *The Lancet Global Health*, 8(2), e204–e214. [https://doi.org/10.1016/S2214-109X\(19\)30541-8](https://doi.org/10.1016/S2214-109X(19)30541-8)

Lewis, T., Synowiec, C., Lagomarsino, G., & Schweitzer, J. (2012). E-health in low- and middle-income

countries: Findings from the Center for Health Market Innovations. *Bulletin of the World Health Organization*, 90(5), 332–340. <https://doi.org/10.2471/BLT.11.099820>

Lezo, A., Baldini, L., & Asteggiano, M. (2020). Failure to Thrive in the Outpatient Clinic: A New Insight.

Nutrients, 12(8), 2202. <https://doi.org/10.3390/nu12082202>

Li, C., Zhang, Shuheng, Zhang, H., Pang, L., Lam, K., Hui, C., & Zhang, Su. (2012). Using the K-Nearest

Neighbor Algorithm for the Classification of Lymph Node Metastasis in Gastric Cancer.

Computational and Mathematical Methods in Medicine, 2012.

<https://doi.org/10.1155/2012/876545>

- Li, H. L. and M. (2022). 7.2 *Data Splitting and Resampling | Introduction to Data Science*.
<http://scientistcafe.com/IDS/>
- Li, W., & Guo, Q. (2021). Plotting receiver operating characteristic and precision–recall curves from presence and background data. *Ecology and Evolution*, 11(15), 10192–10206.
<https://doi.org/10.1002/ece3.7826>
- Li, Y., Xia, S., Jiang, X., Feng, C., Gong, S., Ma, J., Fang, Z., Yin, J., & Yin, Y. (2021). Gut Microbiota and Diarrhea: An Updated Review. *Frontiers in Cellular and Infection Microbiology*, 11.
<https://doi.org/10.3389/fcimb.2021.625210>
- Li, Y., Yang, Y., Che, J., & Zhang, L. (2019). *Predicting the Number of Nearest Neighbor for kNN Classifier*.
- Li, Z., Kim, R., Vollmer, S., & Subramanian, S. V. (2020). Factors Associated With Child Stunting, Wasting, and Underweight in 35 Low- and Middle-Income Countries. *JAMA Network Open*, 3(4), e203386.
<https://doi.org/10.1001/jamanetworkopen.2020.3386>
- Lim, Y. (2022, April 5). *Stacked ensembles—Improving model performance on a higher level*. Medium.
<https://towardsdatascience.com/stacked-ensembles-improving-model-performance-on-a-higher-level-99ffc4ea5523>
- Lima, A. A., Fang, G., Schorling, J. B., de Albuquerque, L., McAuliffe, J. F., Mota, S., Leite, R., & Guerrant, R. L. (1992). Persistent diarrhea in Northeast Brazil: Etiologies and interactions with malnutrition. *Acta Paediatrica*, 81(s383), 39–44. <https://doi.org/10.1111/j.1651-2227.1992.tb12370.x>
- Lima, A. A. M., & Guerrant, R. L. (1992a). Persistent Diarrhea in Children: Epidemiology, Risk Factors, Pathophysiology, Nutritional Impact, and Management. *Epidemiologic Reviews*, 14(1), 222–242.
<https://doi.org/10.1093/oxfordjournals.epirev.a036088>

- Lima, A. A. M., & Guerrant, R. L. (1992b). Persistent Diarrhea in Children: Epidemiology, Risk Factors, Pathophysiology, Nutritional Impact, and Management. *Epidemiologic Reviews*, 14(1), 222–242. <https://doi.org/10.1093/oxfordjournals.epirev.a036088>
- Lima, A. A., Moore, S. R., Barboza, M. S., Soares, A. M., Schleupner, M. A., Newman, R. D., Sears, C. L., Nataro, J. P., Fedorko, D. P., Wuhib, T., Schorling, J. B., & Guerrant, R. L. (2000). Persistent diarrhea signals a critical period of increased diarrhea burdens and nutritional shortfalls: A prospective cohort study among children in northeastern Brazil. *The Journal of Infectious Diseases*, 181(5), 1643–1651. <https://doi.org/10.1086/315423>
- Linardatos, P., Papastefanopoulos, V., & Kotsiantis, S. (2021). Explainable AI: A Review of Machine Learning Interpretability Methods. *Entropy*, 23(1), Article 1. <https://doi.org/10.3390/e23010018>
- Ling, Y., Chen, Y., Chirikov, V., Xie, J., Qiu, H., Otgonsuren, Z., Dong, P., & Gao, X. (2018). A Prediction Model for Length of Stay in the Icu Among Septic Patients: A Machine Learning Approach. *Value in Health*, 21, S5. <https://doi.org/10.1016/j.jval.2018.04.016>
- Lins, M. das G. M., Motta, M. E. F. A., & Silva, G. A. P. da. (2003). Risk factors to persistent diarrhea in infants. *Arquivos de Gastroenterologia*, 40, 239–246. <https://doi.org/10.1590/S0004-28032003000400008>
- Lisk, R., Uddin, M., Parbhoo, A., Yeong, K., Fluck, D., Sharma, P., Lean, M. E. J., & Han, T. S. (2019). Predictive model of length of stay in hospital among older patients. *Aging Clinical and Experimental Research*, 31(7), 993–999. <https://doi.org/10.1007/s40520-018-1033-7>
- Liu, H. (2010). Feature Selection. In C. Sammut & G. I. Webb (Eds.), *Encyclopedia of Machine Learning* (pp. 402–406). Springer US. https://doi.org/10.1007/978-0-387-30164-8_306
- Liu, J., Garcia Bardales, P. F., Islam, K., Jarju, S., Juma, J., Mhango, C., Naumanga, Q., Qureshi, S., Sonye, C., Ahmed, N., Aziz, F., Bhuiyan, M. T. R., Charles, M., Cunliffe, N. A., Abdou, M., Galagan, S. R., Gitte, E., Guindo, I., Jahangir Hossain, M., ... Houpt, E. R. (2024). Shigella Detection and

- Molecular Serotyping With a Customized TaqMan Array Card in the Enterics for Global Health (EFGH): Shigella Surveillance Study. *Open Forum Infectious Diseases*, 11(Supplement_1), S34–S40. <https://doi.org/10.1093/ofid/ofad574>
- Liu, L., Oza, S., Hogan, D., Chu, Y., Perin, J., Zhu, J., Lawn, J. E., Cousens, S., Mathers, C., & Black, R. E. (2016). Global, regional, and national causes of under-5 mortality in 2000–15: An updated systematic analysis with implications for the Sustainable Development Goals. *The Lancet*, 388(10063), 3027–3035. [https://doi.org/10.1016/S0140-6736\(16\)31593-8](https://doi.org/10.1016/S0140-6736(16)31593-8)
- Liu, L., Oza, S., Hogan, D., Perin, J., Rudan, I., Lawn, J. E., Cousens, S., Mathers, C., & Black, R. E. (2015). Global, regional, and national causes of child mortality in 2000–13, with projections to inform post-2015 priorities: An updated systematic analysis. *Lancet (London, England)*, 385(9966), 430–440. [https://doi.org/10.1016/S0140-6736\(14\)61698-6](https://doi.org/10.1016/S0140-6736(14)61698-6)
- Liu, M., Li, S., Yuan, H., Ong, M. E. H., Ning, Y., Xie, F., Saffari, S. E., Shang, Y., Volovici, V., Chakraborty, B., & Liu, N. (2023). Handling missing values in healthcare data: A systematic review of deep learning-based imputation techniques. *Artificial Intelligence in Medicine*, 142, 102587. <https://doi.org/10.1016/j.artmed.2023.102587>
- Lo Vecchio, A., Conelli, M. L., & Guarino, A. (2021). Infections and Chronic Diarrhea in Children. *The Pediatric Infectious Disease Journal*, 40(7), e255. <https://doi.org/10.1097/INF.0000000000003182>
- Lockhart, A., Mucida, D., & Parsa, R. (2022). Immunity to enteric viruses. *Immunity*, 55(5), 800–818. <https://doi.org/10.1016/j.immuni.2022.04.007>
- Loli, S., & Carcamo, C. P. (2020). Rotavirus vaccination and stunting: Secondary Data Analysis from the Peruvian Demographic and Health Survey. *Vaccine*, 38(50), 8010–8015. <https://doi.org/10.1016/j.vaccine.2020.10.044>

- Lucy Lawrence, S. (2021). *Predicting stunting status among children under five years: The case study of Tanzania* [Thesis, University of Rwanda]. <http://dr.ur.ac.rw/handle/123456789/1389>
- Luijken, K., Wynants, L., van Smeden, M., Van Calster, B., Steyerberg, E. W., Groenwold, R. H. H., & Collaborators. (2020). Changing predictor measurement procedures affected the performance of prediction models in clinical examples. *Journal of Clinical Epidemiology*, 119, 7–18. <https://doi.org/10.1016/j.jclinepi.2019.11.001>
- Lumumba, V., Sang, D., Mpaine, M., Njoka, G., & Kavita, M. (2025). Comparative Analysis of Cross-Validation Techniques: LOOCV, K-folds Cross-Validation, and Repeated K-folds Cross-Validation in Machine Learning Models. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.5266507>
- Maggini, S., Pierre, A., & Calder, P. C. (2018). Immune Function and Micronutrient Requirements Change over the Life Course. *Nutrients*, 10(10), 1531. <https://doi.org/10.3390/nu10101531>
- Maggioni, F., & Spinelli, A. (2025). A novel robust optimization model for nonlinear Support Vector Machine. *European Journal of Operational Research*, 322(1), 237–253. <https://doi.org/10.1016/j.ejor.2024.12.014>
- Mahadevan, A., & Mathioudakis, M. (2023). *Cost-Effective Retraining of Machine Learning Models* (arXiv:2310.04216). arXiv. <http://arxiv.org/abs/2310.04216>
- Majhi, B. (2023). A Modified Artificial Neural Network (ANN)-Based Time Series Prediction of COVID-19 Cases from Multi-Country Data. *Journal of The Institution of Engineers (India): Series B*, 104(2), 335–350. <https://doi.org/10.1007/s40031-022-00849-w>
- Mank, I., Vandormael, A., Traoré, I., Ouédraogo, W. A., Sauerborn, R., & Danquah, I. (2020). Dietary habits associated with growth development of children aged < 5 years in the Nouna Health and Demographic Surveillance System, Burkina Faso. *Nutrition Journal*, 19(1), 81. <https://doi.org/10.1186/s12937-020-00591-3>

- Manokhin, V., & Grønhaug, D. (2026). *Classifier Calibration at Scale: An Empirical Study of Model-Agnostic Post-Hoc Methods*. <https://doi.org/10.5281/zenodo.18299333>
- Markus, A. F., Kors, J. A., & Rijnbeek, P. R. (2021). The role of explainability in creating trustworthy artificial intelligence for health care: A comprehensive survey of the terminology, design choices, and evaluation strategies. *Journal of Biomedical Informatics*, 113, 103655. <https://doi.org/10.1016/j.jbi.2020.103655>
- Marwala, T., Fournier-Tombs, E., & Stinckwich, S. (2023). *The Use of Synthetic Data to Train AI Models: Opportunities and Risks for Sustainable Development*. (1).
- Masanja, H., Mongi, P., Baraka, J., Jackson, B., Kisisiwe, Y., Manji, K., Iriya, N., John, T., Kimatta, S., Walker, N., & Black, R. E. (2019). Factors associated with the decline in under five diarrhea mortality in Tanzania from 1980-2015. *Journal of Global Health*, 9(2), 020806. <https://doi.org/10.7189/jogh.09.020806>
- Matthews, S., Ward, M., Nolan, A., Normand, C., Kenny, R. A., & May, P. (2022). Predicting mortality in The Irish Longitudinal Study on Ageing (TILDA): Development of a four-year index and comparison with international measures. *BMC Geriatrics*, 22(1), 510. <https://doi.org/10.1186/s12877-022-03196-z>
- Mbunge, E., & Batani, J. (2023). Application of deep learning and machine learning models to improve healthcare in sub-Saharan Africa: Emerging opportunities, trends and implications. *Telematics and Informatics Reports*, 11, 100097. <https://doi.org/10.1016/j.teler.2023.100097>
- McCombes, S., & Bhandari, P. (2021, June 7). *What Is a Research Design | Types, Guide & Examples*. Scribbr. <https://www.scribbr.com/methodology/research-design/>
- McGovern, M. E., Krishna, A., Aguayo, V. M., & Subramanian, S. (2017). A review of the evidence linking child stunting to economic outcomes. *International Journal of Epidemiology*, 46(4), 1171–1191. <https://doi.org/10.1093/ije/dyx017>

- McNerney, R. (2015). Diagnostics for Developing Countries. *Diagnostics*, 5(2), 200–209.
<https://doi.org/10.3390/diagnostics5020200>
- Md, N., Sr, M., Pd, P., Ll, D., B, L., Aa, L., & Ri, G. (2002, May). *Early childhood diarrhea is associated with diminished cognitive function 4 to 7 years later in children in a northeast Brazilian shantytown*. *Am J Trop Med Hyg*. The American Journal of Tropical Medicine and Hygiene.
<https://doi.org/10.4269/ajtmh.2002.66.590>
- Medlock, S., Ravelli, A. C. J., Tamminga, P., Mol, B. W. M., & Abu-Hanna, A. (2011). Prediction of Mortality in Very Premature Infants: A Systematic Review of Prediction Models. *PLoS ONE*, 6(9), e23441. <https://doi.org/10.1371/journal.pone.0023441>
- Mehal, J. M., Esposito, D. H., Holman, R. C., Tate, J. E., Callinan, L. S., & Parashar, U. D. (2012). Risk Factors for Diarrhea-associated Infant Mortality in the United States, 2005–2007. *The Pediatric Infectious Disease Journal*, 31(7), 717. <https://doi.org/10.1097/INF.0b013e318253a78b>
- Mehta, N. M., Corkins, M. R., Lyman, B., Malone, A., Goday, P. S., Carney, L. N., Monczka, J. L., Plogsted, S. W., Schwenk, W. F., & American Society for Parenteral and Enteral Nutrition Board of Directors. (2013). Defining pediatric malnutrition: A paradigm shift toward etiology-related definitions. *JPEN. Journal of Parenteral and Enteral Nutrition*, 37(4), 460–481.
<https://doi.org/10.1177/0148607113479972>
- Mekonnen, W. (2020). *REVIEW ON CORRELATION RESEARCH*.
- Melfi, V., Finley, A. O., & Doser, J. W. (2020). *R Programming for Data Sciences*.
<https://www.stt.msu.edu/~melfi/STT180Text/data.html>
- Melissa, C. S. (2021, March 29). *Medical Definition of Risk factor*. RxList.
https://www.rxlist.com/risk_factor/definition.htm

- Mengesha, A., Hailu, S., Birhane, M., & Belay, M. M. (2021). The Prevalence of Stunting and Associated Factors among Children Under Five years of age in Southern Ethiopia: Community Based Cross-Sectional Study. *Annals of Global Health*, 87(1), 111. <https://doi.org/10.5334/aogh.3432>
- Mijwil, M. (2021). Artificial Neural Networks Advantages and Disadvantages. *Mesopotamian Journal of Big Data*, 2021, 29–31. <https://doi.org/10.58496/MJBD/2021/006>
- Militello, L., Arbuckle, N., Saleem, J., Patterson, E., Flanagan, M., Haggstrom, D., & Doebbeling, B. (2013). Sources of variation in primary care clinical workflow: Implications for the design of cognitive support. *Health Informatics Journal*. <https://doi.org/10.1177/1460458213476968>
- Miller, J. L., Tada, M., Goto, M., Chen, H., Dang, E., Mohr, N. M., & Lee, S. (2022). Prediction models for severe manifestations and mortality due to COVID-19: A systematic review. *Academic Emergency Medicine*, 29(2), 206–216. <https://doi.org/10.1111/acem.14447>
- Mills, A. (2014, February 5). *Health Care Systems in Low- and Middle-Income Countries* [Review-article]. Massachusetts Medical Society. (world). [Http://Dx.Doi.Org/10.1056/NEJMra1110897](http://Dx.Doi.Org/10.1056/NEJMra1110897). <https://doi.org/10.1056/NEJMra1110897>
- Ministry of Health. (2014). *Kenya Health policy 2014 to 2030*. <http://publications.universalhealth2030.org/ref/d6e32af10e5c515876d34f801774aa9a>
- Ministry of Health, Kenya National Bureau of Statistics, & The DHS Program. (2021, September). *The Kenya Malaria Indicator Survey 2020*. <https://dhsprogram.com/pubs/pdf/MIS36/MIS36.pdf>
- Mirza, N. M., Caulfield, L. E., Black, R. E., & Macharia, W. M. (1997). Risk Factors for Diarrheal Duration. *American Journal of Epidemiology*, 146(9), 776–785. <https://doi.org/10.1093/oxfordjournals.aje.a009354>
- Mitra, A. K., Rahman, M. M., & Fuchs, G. J. (2000). Risk Factors and Gender Differentials for Death among Children Hospitalized with Diarrhoea in Bangladesh. *Journal of Health, Population and Nutrition*, 18(3), 151–156.

- Moench-Pfanner, R., Silo, S., Laillou, A., Wieringa, F., Hong, Rathamony, Hong, Rathavuth, Poirot, E., & Bagriansky, J. (2016). The Economic Burden of Malnutrition in Pregnant Women and Children under 5 Years of Age in Cambodia. *Nutrients*, 8(5), Article 5. <https://doi.org/10.3390/nu8050292>
- Moerland, T. M., Broekens, J., Plaat, A., & Jonker, C. M. (2023). Model-based Reinforcement Learning: A Survey. *Foundations and Trends® in Machine Learning*, 16(1), 1–118. <https://doi.org/10.1561/22000000086>
- Mohammed, N. I., Jarde, A., Mackenzie, G., D'Alessandro, U., & Jeffries, D. (2022). Deploying Machine Learning Models Using Progressive Web Applications: Implementation Using a Neural Network Prediction Model for Pneumonia Related Child Mortality in The Gambia. *Frontiers in Public Health*, 9, 772620. <https://doi.org/10.3389/fpubh.2021.772620>
- Mølbak, K., Jensen, H., Ingholt, L., & Aaby, P. (1997). Risk Factors for Diarrheal Disease Incidence in Early Childhood: A Community Cohort Study from Guinea-Bissau. *American Journal of Epidemiology*, 146(3), 273–282. <https://doi.org/10.1093/oxfordjournals.aje.a009263>
- Monteiro, C. A., Benicio, M. H. D., Conde, W. L., Konno, S., Lovadino, A. L., Barros, A. J., & Victora, C. G. (2010). Narrowing socioeconomic inequality in child stunting: The Brazilian experience, 1974–2007. *Bulletin of the World Health Organization*, 88(4), 305–311. <https://doi.org/10.2471/BLT.09.069195>
- Montenegro, C. R., Gomez, G., Hincapie, O., Dvoretzkiy, S., DeWitt, T., Gracia, D., & Misas, J. D. (2022). The pediatric global burden of stunting: Focus on Latin America. *Lifestyle Medicine*, 3(3), e67. <https://doi.org/10.1002/lim2.67>
- Montomoli, J., Romeo, L., Moccia, S., Bernardini, M., Migliorelli, L., Berardini, D., Donati, A., Carsetti, A., Bocci, M. G., Wendel Garcia, P. D., Fumeaux, T., Guerci, P., Schüpbach, R. A., Ince, C., Frontoni, E., Hilty, M. P., Alfaro-Farias, M., Vizmanos-Lamotte, G., Tschoellitsch, T., ... Colak, E. (2021). Machine learning using the extreme gradient boosting (XGBoost) algorithm predicts 5-day delta

- of SOFA score at ICU admission in COVID-19 patients. *Journal of Intensive Medicine*, 1(2), 110–116. <https://doi.org/10.1016/j.jointm.2021.09.002>
- Moons, K. G. M., Damen, J. A. A., Kaul, T., Hooft, L., Navarro, C. A., Dhiman, P., Beam, A. L., Calster, B. V., Celi, L. A., Denaxas, S., Denniston, A. K., Ghassemi, M., Heinze, G., Kengne, A. P., Maier-Hein, L., Liu, X., Logullo, P., McCradden, M. D., Liu, N., ... Smeden, M. van. (2025). PROBAST+AI: An updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. *BMJ*, 388, e082505. <https://doi.org/10.1136/bmj-2024-082505>
- Moore, S. R., Lima, N. L., Soares, A. M., Oriá, R. B., Pinkerton, R. C., Barrett, L. J., Guerrant, R. L., & Lima, A. A. M. (2010). Prolonged episodes of acute diarrhea reduce growth and increase risk of persistent diarrhea in children. *Gastroenterology*, 139(4), 1156–1164. <https://doi.org/10.1053/j.gastro.2010.05.076>
- Morales-Morales, J., Ogueda-Oliva, A., Caiseda, C., & Seshaiyer, P. (2025). Enhancing pedagogical practices with Artificial Neural Networks in the age of AI to engage the next generation in Biomathematics. *Bulletin of Mathematical Biology*, 87(10), 139. <https://doi.org/10.1007/s11538-025-01511-4>
- Morozoff, C., Ahmed, N., Chinkhumba, J., Islam, M. T., Jallow, A. F., Ogwel, B., Zegarra Paredes, L. F., Sanogo, D., Atlas, H. E., Badji, H., Bar-Zeev, N., Conteh, B., Güimack Fajardo, M., Feutz, E., Haidara, F. C., Karim, M., Mamby Keita, A., Keita, Y., Khanam, F., ... Means, A. R. (2024). Quantifying the Cost of Shigella Diarrhea in the Enterics for Global Health (EFGH) Shigella Surveillance Study. *Open Forum Infectious Diseases*, 11(Supplement_1), S41–S47. <https://doi.org/10.1093/ofid/ofad575>
- Morris, S. S., Cousens, S. N., Lanata, C. F., & Kirkwood, B. R. (1994). Diarrhoea—Defining the Episode. *International Journal of Epidemiology*, 23(3), 617–623. <https://doi.org/10.1093/ije/23.3.617>

- Motluk, A. (2018). Do doctors experiencing burnout make more errors? *CMAJ : Canadian Medical Association Journal*, 190(40), E1216–E1217. <https://doi.org/10.1503/cmaj.109-5663>
- Moubayed, A., Injadat, M., Nassif, A., Lutfiyya, H., & Shami, A. (2018). E-Learning: Challenges and Research Opportunities Using Machine Learning & Data Analytics. *IEEE Access*, PP, 1–1. <https://doi.org/10.1109/ACCESS.2018.2851790>
- Mtambo, O. P. L., Katoma, V., & Kazembe, L. N. M. (2016). *Analysis of severe childhood stunting in Namibia*.
- Muga, R., Kizito, M., Mbayah, T., & Gakuruh. (2005). *Chapter 2 Overview of the Health System in Kenya*.
- Muinga, N., Magare, S., Monda, J., English, M., Fraser, H., Powell, J., & Paton, C. (2020). Digital health Systems in Kenyan Public Hospitals: A mixed-methods survey. *BMC Medical Informatics and Decision Making*, 20(1), 2. <https://doi.org/10.1186/s12911-019-1005-7>
- Mulatya, D. M., & Mutuku, F. W. (2020). Assessing Comorbidity of Diarrhea and Acute Respiratory Infections in Children Under 5 Years: Evidence From Kenya’s Demographic Health Survey 2014. *Journal of Primary Care & Community Health*, 11, 2150132720925190. <https://doi.org/10.1177/2150132720925190>
- Mumford, E. (2000). A Socio-Technical Approach to Systems Design. *Requir. Eng.*, 5, 125–133. <https://doi.org/10.1007/PL00010345>
- Murphree, D. H., Quest, D. J., Allen, R. M., Ngufor, C., & Storlie, C. B. (2018). Deploying Predictive Models In A Healthcare Environment—An Open Source Approach. *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 6112–6116. <https://doi.org/10.1109/EMBC.2018.8513689>
- Mwenda, J. M., Tate, J. E., Steele, A. D., & Parashar, U. D. (2014). Preparing for the Scale-up of Rotavirus Vaccine Introduction in Africa: Establishing Surveillance Platforms to Monitor Disease Burden

- and Vaccine Impact. *The Pediatric Infectious Disease Journal*, 33, S1.
<https://doi.org/10.1097/INF.0000000000000132>
- Naemi, A., Schmidt, T., Mansourvar, M., Naghavi-Behzad, M., Ebrahimi, A., & Wiil, U. K. (2021). Machine learning techniques for mortality prediction in emergency departments: A systematic review. *BMJ Open*, 11(11), e052663. <https://doi.org/10.1136/bmjopen-2021-052663>
- Nagelkerke, J. (2018, September 20). *Introduction to modelplotr | R-bloggers*. <https://www.r-bloggers.com/2018/09/introduction-to-modelplotr/>
- Nagelkerke, J. (2020). *modelplotr: Plots to evaluate the business value of predictive models* [Computer software]. <https://cran.r-project.org/web/packages/modelplotr/vignettes/modelplotr.html>
- Nanga, S., Bawah, A. T., Acquaye, B. A., Billa, M.-I., Baeta, F. D., Odai, N. A., Obeng, S. K., & Nsiah, A. D. (2021). Review of Dimension Reduction Methods. *Journal of Data Analysis and Information Processing*, 9(3), Article 3. <https://doi.org/10.4236/jdaip.2021.93013>
- Nasrin, D., Liang, Y., Powell, H., Casanova, I. G., Sow, S. O., Hossain, M. J., Omoro, R., Sanogo, D., Tamboura, B., Zaman, S. M. A., Antonio, M., Jones, J. C. M., Awuor, A. O., Kasumba, I. N., Ochieng, J. B., Badji, H., Verani, J. R., Widdowson, M.-A., Roose, A., ... Kotloff, K. L. (2023). Moderate-to-Severe Diarrhea and Stunting Among Children Younger Than 5 Years: Findings From the Vaccine Impact on Diarrhea in Africa (VIDA) Study. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, 76(Suppl 1), S41–S48.
<https://doi.org/10.1093/cid/ciac945>
- Nasrin, D., Wu, Y., Blackwelder, W. C., Farag, T. H., Saha, D., Sow, S. O., Alonso, P. L., Breiman, R. F., Sur, D., Faruque, A. S. G., Zaidi, A. K. M., Biswas, K., Van Eijk, A. M., Walker, D. G., Levine, M. M., & Kotloff, K. L. (2013). Health care seeking for Childhood Diarrhea in Developing Countries: Evidence from Seven Sites in Africa and Asia. *The American Journal of Tropical Medicine and Hygiene*, 89(1 Suppl), 3–12. <https://doi.org/10.4269/ajtmh.12-0749>

- Natekin, A., & Knoll, A. (2013). Gradient boosting machines, a tutorial. *Frontiers in Neurorobotics*, 7.
<https://www.frontiersin.org/articles/10.3389/fnbot.2013.00021>
- National AIDS and STI Control Programme (NASCOP). (2020). *KENPHIA 2018 Preliminary Report Vol 1*.
<https://www.health.go.ke/wp-content/uploads/2020/02/KENPHIA-2018-PREL-REP-2020-HR3-final.pdf>
- Ndagijimana, S., Kabano, I. H., Masabo, E., & Ntaganda, J. M. (2023). Prediction of Stunting Among Under-5 Children in Rwanda Using Machine Learning Techniques. *Journal of Preventive Medicine and Public Health*, 56(1), 41–49. <https://doi.org/10.3961/jpmph.22.388>
- Nejatian, S., Parvin, H., & Faraji, E. (2018). Using sub-sampling and ensemble clustering techniques to improve performance of imbalanced classification. *Neurocomputing, Machine Learning and Data Mining Techniques for Medical Complex Data Analysis*, 276, 55–66.
<https://doi.org/10.1016/j.neucom.2017.06.082>
- Ngabo, F., Mvundura, M., Gazley, L., Gatera, M., Rugambwa, C., Kayonga, E., Tuyishime, Y., Niyibaho, J., Mwenda, J. M., Donnen, P., Lepage, P., Binagwaho, A., & Atherly, D. (2016). The Economic Burden Attributable to a Child's Inpatient Admission for Diarrheal Disease in Rwanda. *PLOS ONE*, 11(2), e0149805. <https://doi.org/10.1371/journal.pone.0149805>
- Ngan, P. K., Khanh, N. G., van Tuong, C., Quy, P. P., Anh, D. N., & Thuy, H. T. (1992). Persistent diarrhea in Vietnamese children: A preliminary report. *Acta Paediatrica*, 81(s383), 124–126.
<https://doi.org/10.1111/j.1651-2227.1992.tb12385.x>
- Ngari, M. M., Iversen, P. O., Thitiri, J., Mwalekwa, L., Timbwa, M., Fegan, G. W., & Berkley, J. A. (2019). Linear growth following complicated severe malnutrition: 1-year follow-up cohort of Kenyan children. *Archives of Disease in Childhood*, 104(3), 229–235.
<https://doi.org/10.1136/archdischild-2018-315641>

Nguyen, Q. H., Ly, H.-B., Ho, L. S., Al-Ansari, N., Le, H. V., Tran, V. Q., Prakash, I., & Pham, B. T. (2021).

Influence of Data Splitting on Performance of Machine Learning Models in Prediction of Shear Strength of Soil. *Mathematical Problems in Engineering*, 2021, e4832864.

<https://doi.org/10.1155/2021/4832864>

Niazkar, H. R., & Niazkar, M. (2020). Application of artificial neural networks to predict the COVID-19

outbreak. *Global Health Research and Policy*, 5(1), 50. <https://doi.org/10.1186/s41256-020-00175-y>

Niehaus, M. D., Moore, S. R., Patrick, P. D., Derr, L. L., Lorntz, B., Lima, A. A., & Guerrant, R. L. (2002).

Early childhood diarrhea is associated with diminished cognitive function 4 to 7 years later in children in a northeast Brazilian shantytown. *The American Journal of Tropical Medicine and Hygiene*, 66(5), 590–593. <https://doi.org/10.4269/ajtmh.2002.66.590>

Nieuwenhuijsen, K., Verbeek, J. H. A. M., de Boer, A. G. E. M., Blonk, R. W. B., & van Dijk, F. J. H. (2006).

Predicting the duration of sickness absence for patients with common mental disorders in occupational health care. *Scandinavian Journal of Work, Environment & Health*, 32(1), 67–74.

<https://doi.org/10.5271/sjweh.978>

Nikolić, M., Nikolić, D., Stefanović, M., Koprivica, S., & Stefanović, D. (2025). Mitigating Algorithmic Bias

Through Probability Calibration: A Case Study on Lead Generation Data. *Mathematics*, 13(13).

<https://doi.org/10.3390/math13132183>

Niyibitegeka, F., Riewpaiboon, A., Youngkong, S., & Thavorncharoensap, M. (2021). Economic burden of

childhood diarrhea in Burundi. *Global Health Research and Policy*, 6(1), 13.

<https://doi.org/10.1186/s41256-021-00194-3>

Njuguna, J., & Muruka, C. (2011). Diarrhoea and malnutrition among children in a Kenyan District: A

correlational study. *Undefined*. <https://www.semanticscholar.org/paper/ORIGINAL-RESEARCH->

DIARRHOEA-AND-MALNUTRITION-AMONG-Njuguna-

Muruka/01499af9c6b6cf352f15c7a0cc3a83319f06ed37

Nnamoko, N., Arshad, F., England, D., Vora, J., & Norman, J. (2014). *Evaluation of Filter and Wrapper Methods for Feature Selection in Supervised Machine Learning*.

Nopour, R., Erfannia, L., Mehrabi, N., Mashoufi, M., Mahdavi, A., & Shanbehzadeh, M. (2022).

Comparison of Two Statistical Models for Predicting Mortality in COVID-19 Patients in Iran.

Shiraz E-Medical Journal, 23(6), Article 6. <https://doi.org/10.5812/semj.119172>

Nuzhat, S., Shahunja, K. M., Shahid, A. S. M. S. B., Khan, S. H., Islam, S. B., Islam, Md. R., Ahmed, T.,

Chisti, M. J., Hossain, M. I., & Faruque, A. S. G. (2020). Diarrhoeal children with concurrent severe wasting and stunting compared to severe wasting or severe stunting. *Tropical Medicine & International Health*, 25(8), 928–935. <https://doi.org/10.1111/tmi.13446>

Nyawanda, B. O., Mott, J. A., Njuguna, H. N., Mayieka, L., Khagayi, S., Onkoba, R., Makokha, C., Otieno,

N. A., Bigogo, G. M., Katz, M. A., Feikin, D. R., & Verani, J. R. (2016). Evaluation of case definitions to detect respiratory syncytial virus infection in hospitalized children below 5 years in Rural Western Kenya, 2009-2013. *BMC Infectious Diseases*, 16, 218.

<https://doi.org/10.1186/s12879-016-1532-0>

Ochoa, T. J., Salazar-Lindo, E., & Cleary, T. G. (2004). Management of children with infection-associated persistent diarrhea. *Seminars in Pediatric Infectious Diseases, Gastrointestinal Infections in the Pediatric Population*, 15(4), 229–236. <https://doi.org/10.1053/j.spid.2004.07.003>

Odhambo, F. O., Laserson, K. F., Sewe, M., Hamel, M. J., Feikin, D. R., Adazu, K., Ogwang, S., Obor, D., Amek, N., Bayoh, N., Ombok, M., Lindblade, K., Desai, M., ter Kuile, F., Phillips-Howard, P., van Eijk, A. M., Rosen, D., Hightower, A., Ofware, P., ... Vulule, J. M. (2012). Profile: The KEMRI/CDC Health and Demographic Surveillance System—Western Kenya. *International Journal of Epidemiology*, 41(4), 977–987. <https://doi.org/10.1093/ije/dys108>

- Ogwel, B., Mzazi, V. H., Awuor, A. O., Okonji, C., Anyango, R. O., Oreso, C., Ochieng, J. B., Munga, S., Nasrin, D., Tickell, K. D., Pavlinac, P. B., Kotloff, K. L., & Omore, R. (2024). Predictive modelling of linear growth faltering among pediatric patients with Diarrhea in Rural Western Kenya: An explainable machine learning approach. *BMC Medical Informatics and Decision Making*, 24(1), 368. <https://doi.org/10.1186/s12911-024-02779-7>
- Ogwel, B., Mzazi, V. H., Awuor, A. O., Okonji, C., Anyango, R. O., Oreso, C., Ochieng, J. B., Munga, S., Nasrin, D., Tickell, K. D., Pavlinac, P. B., Kotloff, K. L., & Omore, R. (2025). Derivation and validation of a clinical predictive model for longer duration diarrhea among pediatric patients in Kenya using machine learning algorithms. *BMC Medical Informatics and Decision Making*, 25(1), 28. <https://doi.org/10.1186/s12911-025-02855-6>
- Ogwel, B., Mzazi, V. H., Nyawanda, B. O., Otieno, G., Tickell, K. D., & Omore, R. (2024). A machine learning approach to predicting inpatient mortality among pediatric acute gastroenteritis patients in Kenya. *Learning Health Systems*, e10478. <https://doi.org/10.1002/lrh2.10478>
- Ogwel, B., Mzazi, V., Nyawanda, B. O., Otieno, G., & Omore, R. (2023). Predictive modeling for infectious diarrheal disease in pediatric populations: A systematic review. *Learning Health Systems*, n/a(n/a), e10382. <https://doi.org/10.1002/lrh2.10382>
- Ojeda, F. M., Jansen, M. L., Thiéry, A., Blankenberg, S., Weimar, C., Schmid, M., & Ziegler, A. (2023). Calibrating machine learning approaches for probability estimation: A comprehensive comparison. *Statistics in Medicine*, 42(29), 5451–5478. <https://doi.org/10.1002/sim.9921>
- Oluoko, E. A. (2021). *Factors Associated With Stunting Among Under-five Year Old Children in Kenya: A Multilevel Logistic Regression Approach*.
- Omar, S., Ngadi, A., & H. Jebur, H. (2013). Machine Learning Techniques for Anomaly Detection: An Overview. *International Journal of Computer Applications*, 79(2), 33–41. <https://doi.org/10.5120/13715-1478>

- Omoro, R., Khagayi, S., Ogwel, B., Onkoba, R., Ochieng, J. B., Juma, J., Munga, S., Tabu, C., Kibet, S., Nuorti, J. P., Odhiambo, F., Mwenda, J. M., Breiman, R. F., Parashar, U. D., & Tate, J. E. (2019). Rates of hospitalization and death for all-cause and rotavirus acute gastroenteritis before rotavirus vaccine introduction in Kenya, 2010–2013. *BMC Infectious Diseases*, 19(1), 47. <https://doi.org/10.1186/s12879-018-3615-6>
- Omoro, R., O'Reilly, C. E., Williamson, J., Moke, F., Were, V., Farag, T. H., van Eijk, A. M., Kotloff, K. L., Levine, M. M., Obor, D., Odhiambo, F., Vulule, J., Laserson, K. F., Mintz, E. D., & Breiman, R. F. (2013). Health Care-Seeking Behavior During Childhood Diarrheal Illness: Results of Health Care Utilization and Attitudes Surveys of Caretakers in Western Kenya, 2007–2010. *The American Journal of Tropical Medicine and Hygiene*, 89(1 Suppl), 29–40. <https://doi.org/10.4269/ajtmh.12-0755>
- Ooka, T., Johno, H., Nakamoto, K., Yoda, Y., Yokomichi, H., & Yamagata, Z. (2021). Random forest approach for determining risk prediction and predictive factors of type 2 diabetes: Large-scale health check-up data in Japan. *BMJ Nutrition, Prevention & Health*, bmjnph. <https://doi.org/10.1136/bmjnph-2020-000200>
- Ooms, J. (2014). *The OpenCPU System: Towards a Universal Interface for Scientific Computing through Separation of Concerns*.
- Opara, D. C., Ikpeme, E. E., & Ekanem, U. S. (2010). Prevalence of stunting, underweight and obesity in school aged children in Uyo, Nigeria. *Pakistan Journal of Nutrition*, 9(5), 459–466.
- O'Reilly, C. E., Jaron, P., Ochieng, B., Nyaguara, A., Tate, J. E., Parsons, M. B., Bopp, C. A., Williams, K. A., Vinjé, J., Blanton, E., Wannemuehler, K. A., Vulule, J., Laserson, K. F., Breiman, R. F., Feikin, D. R., Widdowson, M.-A., & Mintz, E. (2012). Risk Factors for Death among Children Less than 5 Years Old Hospitalized with Diarrhea in Rural Western Kenya, 2005–2007: A Cohort Study. *PLOS Medicine*, 9(7), e1001256. <https://doi.org/10.1371/journal.pmed.1001256>

- Otieno, G. P., Bottomley, C., Khagayi, S., Adetifa, I., Ngama, M., Omere, R., Ogwel, B., Owor, B. E., Bigogo, G., Ochieng, J. B., Onyango, C., Juma, J., Mwenda, J., Tabu, C., Tate, J. E., Addo, Y., Britton, T., Parashar, U. D., Breiman, R. F., ... Nokes, D. J. (2020). Impact of the Introduction of Rotavirus Vaccine on Hospital Admissions for Diarrhea Among Children in Kenya: A Controlled Interrupted Time-Series Analysis. *Clinical Infectious Diseases*, 70(11), 2306–2313. <https://doi.org/10.1093/cid/ciz912>
- Ottenhoff, M. C., Ramos, L. A., Potters, W., Janssen, M. L. F., Hubers, D., Hu, S., Fridgeirsson, E. A., Piña-Fuentes, D., Thomas, R., Horst, I. C. C. van der, Herff, C., Kubben, P., Elbers, P. W. G., Marquering, H. A., Welling, M., Simsek, S., Kruif, M. D. de, Dormans, T., Fleuren, L. M., ... Beudel, M. (2021). Predicting mortality of individual patients with COVID-19: A multicentre Dutch cohort. *BMJ Open*, 11(7), e047347. <https://doi.org/10.1136/bmjopen-2020-047347>
- Oymak, S., & Gulcu, T. C. (2020). *Statistical and Algorithmic Insights for Semi-supervised Learning with Self-training* (arXiv:2006.11006). arXiv. <https://doi.org/10.48550/arXiv.2006.11006>
- Page, A.-L., Ciglenecki, I., Jasmin, E. R., Desvignes, L., Grandesso, F., Polonsky, J., Nicholas, S., Alberti, K. P., Porten, K., & Luquero, F. J. (2015). Geographic Distribution and Mortality Risk Factors during the Cholera Outbreak in a Rural Region of Haiti, 2010-2011. *PLOS Neglected Tropical Diseases*, 9(3), e0003605. <https://doi.org/10.1371/journal.pntd.0003605>
- Pajila, B., Sheena, B. G., Gayathri, A., Aswini, J., Manogaran, N., & R, S. (2023). *A Comprehensive Survey on Naive Bayes Algorithm: Advantages, Limitations and Applications* (p. 1234). <https://doi.org/10.1109/ICOSEC58147.2023.10276274>
- Panchalingam, S., Antonio, M., Hossain, A., Mandomando, I., Ochieng, B., Oundo, J., Ramamurthy, T., Tamboura, B., Zaidi, A. K. M., Petri, W., Houpt, E., Murray, P., Prado, V., Vidal, R., Steele, D., Strockbine, N., Sansonetti, P., Glass, R. I., Robins-Browne, R. M., ... Nataro, J. P. (2012). Diagnostic microbiologic methods in the GEMS-1 case/control study. *Clinical Infectious Diseases*:

- An Official Publication of the Infectious Diseases Society of America, 55 Suppl 4(Suppl 4), S294-302. <https://doi.org/10.1093/cid/cis754>*
- Pantic, I. V., Paunovic Pantic, J., Valjarevic, S., Corridon, P. R., & Topalovic, N. (2025). Artificial intelligence – based approaches based on random forest algorithm for signal analysis: Potential applications in detection of chemico - biological interactions. *Chemico-Biological Interactions, 418*, 111624. <https://doi.org/10.1016/j.cbi.2025.111624>
- Paquette-Rufiange, A., Laforest, M., & Prudhomme, S. (2026). Goal-oriented calibration of a model and its modeling errors. *Computer Methods in Applied Mechanics and Engineering, 448*, 118486. <https://doi.org/10.1016/j.cma.2025.118486>
- Parashar, U. D., Kilgore, P. E., Holman, R. C., Clarke, M. J., Bresee, J. S., & Glass, R. I. (1998). Diarrheal Mortality in US Infants: Influence of Birth Weight on Risk Factors for Death. *Archives of Pediatrics & Adolescent Medicine, 152*(1), 47–51. <https://doi.org/10.1001/archpedi.152.1.47>
- Park, C., Hsiung, J.-T., Soohoo, M., & Streja, E. (2019). *Choosing Wisely: Using the Appropriate Statistical Test for Trend in SAS.*
- Patel, A. B., Ovung, R., Badhoniya, N. B., & Dibley, M. J. (2012). Risk Factors for Predicting Diarrheal Duration and Morbidity in Children with Acute Diarrhea. *The Indian Journal of Pediatrics, 79*(4), 472–477. <https://doi.org/10.1007/s12098-011-0561-3>
- Patel, H. J., & Saini, J. R. (2021). A Patient Care Predictive Model using Logistic Regression. *International Journal of Advanced Computer Science and Applications, 12*(12). <https://doi.org/10.14569/IJACSA.2021.0121278>
- Patel, K., & Rana, R. (2006). Pedimune in recurrent respiratory infection and diarrhoea—The Indian experience—The pride study. *Indian Journal of Pediatrics, 73*(7), 585–591. <https://doi.org/10.1007/BF02759923>

- Patel, M., Pedreira, C., De Oliveira, L. H., Tate, J., Orozco, M., Mercado, J., Gonzalez, A., Malespin, O., Amador, J. J., Umaña, J., Balmaseda, A., Perez, M. C., Gentsch, J., Kerin, T., Hull, J., Mijatovic, S., Andrus, J., & Parashar, U. (2009). Association Between Pentavalent Rotavirus Vaccine and Severe Rotavirus Diarrhea Among Children in Nicaragua. *JAMA*, *301*(21), 2243–2251. <https://doi.org/10.1001/jama.2009.756>
- Patnayakuni, R., & Ruppel, C. P. (2010). A socio-technical approach to improving the systems development process. *Information Systems Frontiers*, *12*(2), 219–234. <https://doi.org/10.1007/s10796-008-9093-4>
- Paudel, R., Pradhan, B., Wagle, R. R., Pahari, D. P., & Onta, S. R. (2012). Risk Factors for Stunting Among Children: A Community Based Case Control Study in Nepal. *Kathmandu University Medical Journal*, *10*(3), Article 3. <https://doi.org/10.3126/kumj.v10i3.8012>
- Pavlinac, P. B., Brander, R. L., Atlas, H. E., John-Stewart, G. C., Denno, D. M., & Walson, J. L. (2018). Interventions to reduce post-acute consequences of diarrheal disease in children: A systematic review. *BMC Public Health*, *18*(1), 208. <https://doi.org/10.1186/s12889-018-5092-7>
- Pavlovic, M. (2025). *Understanding Model Calibration—A gentle introduction and visual exploration of calibration and the expected calibration error (ECE)*. <https://doi.org/10.48550/arXiv.2501.19047>
- Pawlowski, S. W., Warren, C. A., & Guerrant, R. (2009). Diagnosis and Treatment of Acute or Persistent Diarrhea. *Gastroenterology*, *136*(6), 1874–1886. <https://doi.org/10.1053/j.gastro.2009.02.072>
- Peduzzi, P., Concato, J., Kemper, E., Holford, T. R., & Feinstein, A. R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology*, *49*(12), 1373–1379. [https://doi.org/10.1016/s0895-4356\(96\)00236-3](https://doi.org/10.1016/s0895-4356(96)00236-3)
- Pensa, R. G., Crombach, A., Peignier, S., & Rigotti, C. (2025). Explaining Random Forest and XGBoost with Shallow Decision Trees by Co-clustering Feature Importance. *Machine Learning*, *114*(12), 287. <https://doi.org/10.1007/s10994-025-06932-9>

- Petersen, P., & Zech, J. (2026). *Mathematical theory of deep learning* (arXiv:2407.18384). arXiv.
<https://doi.org/10.48550/arXiv.2407.18384>
- Phatcharathada, B., & Srisuradetchai, P. (2025). Randomized Feature and Bootstrapped Naive Bayes Classification. *Applied System Innovation*, 8(4). <https://doi.org/10.3390/asi8040094>
- Piccolo, S. R., Lee, T. J., Suh, E., & Hill, K. (2020). ShinyLearner: A containerized benchmarking tool for machine-learning classification of tabular data. *GigaScience*, 9(4), giaa026.
<https://doi.org/10.1093/gigascience/giaa026>
- Pienaar, M. A., Sempa, J. B., Luwes, N., & Solomon, L. J. (2022). An Artificial Neural Network Model for Pediatric Mortality Prediction in Two Tertiary Pediatric Intensive Care Units in South Africa. A Development Study. *Frontiers in Pediatrics*, 10.
<https://www.frontiersin.org/articles/10.3389/fped.2022.797080>
- Pinto, E. A., Barros-Filho, A. de A., & Barros, M. B. (1998). Risk factors for persistent diarrhea in hospitalized children. *Arquivos de gastroenterologia*, 35(2), 126–131.
- Platts-Mills, J. A., Liu, J., Rogawski, E. T., Kabir, F., Lertsethtakarn, P., Sigua, M., Khan, S. S., Praharaj, I., Murei, A., Nshama, R., Mujaga, B., Havt, A., Maciel, I. A., McMurtry, T. L., Operario, D. J., Taniuchi, M., Gratz, J., Stroup, S. E., Roberts, J. H., ... Nyathi, E. (2018). Use of quantitative molecular diagnostic methods to assess the aetiology, burden, and clinical characteristics of diarrhoea in children in low-resource settings: A reanalysis of the MAL-ED cohort study. *The Lancet Global Health*, 6(12), e1309–e1318. [https://doi.org/10.1016/S2214-109X\(18\)30349-8](https://doi.org/10.1016/S2214-109X(18)30349-8)
- Pourhomayoun, M., & Shakibi, M. (2021). Predicting mortality risk in patients with COVID-19 using machine learning to help medical decision-making. *Smart Health*, 20, 100178.
<https://doi.org/10.1016/j.smhl.2020.100178>
- Powell, H., Liang, Y., Neuzil, K. M., Jamka, L. P., Nasrin, D., Sow, S. O., Hossain, M. J., Omore, R., & Kotloff, K. L. (2023). A Description of the Statistical Methods for the Vaccine Impact on Diarrhea in Africa

- (VIDA) Study. *Clinical Infectious Diseases*, 76(Supplement_1), S5–S11.
<https://doi.org/10.1093/cid/ciac968>
- PrabhuDas, M., Adkins, B., Gans, H., King, C., Levy, O., Ramilo, O., & Siegrist, C.-A. (2011). Challenges in infant immunity: Implications for responses to infection and vaccines. *Nature Immunology*, 12(3), 189–194. <https://doi.org/10.1038/ni0311-189>
- Prendergast, A. J., & Humphrey, J. H. (2014). The stunting syndrome in developing countries. *Paediatrics and International Child Health*, 34(4), 250–265.
<https://doi.org/10.1179/2046905514Y.00000000158>
- Pudjihartono, N., Fadason, T., Kempa-Liehr, A. W., & O’Sullivan, J. M. (2022). A Review of Feature Selection Methods for Machine Learning-Based Disease Risk Prediction. *Frontiers in Bioinformatics*, 2. <https://doi.org/10.3389/fbinf.2022.927312>
- Qu, W., Balki, I., Mendez, M., Valen, J., Levman, J., & Tyrrell, P. N. (2020). Assessing and mitigating the effects of class imbalance in machine learning with application to X-ray imaging. *International Journal of Computer Assisted Radiology and Surgery*, 15(12), 2041–2048.
<https://doi.org/10.1007/s11548-020-02260-6>
- Quamme, S. H., & Iversen, P. O. (2022). Prevalence of child stunting in Sub-Saharan Africa and its risk factors. *Clinical Nutrition Open Science*, 42, 49–61. <https://doi.org/10.1016/j.nutos.2022.01.009>
- R Core Team. (2021). *R: The R Project for Statistical Computing*. <https://www.r-project.org/>
- Rachmi, C. N., Agho, K. E., Li, M., & Baur, L. A. (2016). Stunting, Underweight and Overweight in Children Aged 2.0–4.9 Years in Indonesia: Prevalence Trends and Associated Risk Factors. *PLOS ONE*, 11(5), e0154756. <https://doi.org/10.1371/journal.pone.0154756>
- Rahaman, Md. J. (2024). A Comprehensive Review to Understand the Definitions, Advantages, Disadvantages and Applications of Machine Learning Algorithms. *International Journal of Computer Applications*, 186, 43–47. <https://doi.org/10.5120/ijca2024923868>

- Rahman, S. M. J., Ahmed, N. A. M. F., Abedin, M. M., Ahammed, B., Ali, M., Rahman, M. J., & Maniruzzaman, M. (2021). Investigate the risk factors of stunting, wasting, and underweight among under-five Bangladeshi children and its prediction based on machine learning approach. *PLOS ONE*, 16(6), e0253172. <https://doi.org/10.1371/journal.pone.0253172>
- Rahmani, K., Thapa, R., Tsou, P., Casie Chetty, S., Barnes, G., Lam, C., & Foon Tso, C. (2023). Assessing the effects of data drift on the performance of machine learning models used in clinical sepsis prediction. *International Journal of Medical Informatics*, 173, 104930. <https://doi.org/10.1016/j.ijmedinf.2022.104930>
- Rahutomo, R., Elwirehardja, G. N., Isnani, M., Asadi, F., & Pardamean, B. (2023). Machine Learning Implementations in Childhood Stunting Research: A Systematic Literature Review. 2023 *International Conference on Information Management and Technology (ICIMTech)*, 229–234. <https://doi.org/10.1109/ICIMTech59029.2023.10277881>
- Rajić, V. (2026). Statistical Hypothesis Testing: A Comprehensive Review of Theory, Methods, and Applications. *Mathematics*, 14(2). <https://doi.org/10.3390/math14020300>
- Ramspek, C. L., Jager, K. J., Dekker, F. W., Zoccali, C., & van Diepen, M. (2021). External validation of prognostic models: What, why, how, when and where? *Clinical Kidney Journal*, 14(1), 49–58. <https://doi.org/10.1093/ckj/sfaa188>
- Rana, R., McGrath, M., Gupta, P., Thakur, E., & Kerac, M. (2020). Feeding Interventions for Infants with Growth Failure in the First Six Months of Life: A Systematic Review. *Nutrients*, 12(7), 2044. <https://doi.org/10.3390/nu12072044>
- Rannan-Eliya, R., Hossain, S., Anuranga, C., Wickramasinghe, R., Jayatissa, R., & Abeykoon, A. (2013). Trends and determinants of childhood stunting and underweight in Sri Lanka. *Ceylon Medical Journal*, 58(1), 10. <https://doi.org/10.4038/cmj.v58i1.5357>

- Raptis, D. A., Mettler, T., Tzanas, K., & Graf, R. (2012). A novel open-source web-based platform promoting collaboration of healthcare professionals and biostatisticians: A design science approach. *Informatics for Health and Social Care*, 37(1), 22–36.
<https://doi.org/10.3109/17538157.2011.590257>
- Reis, I., Baron, D., & Shahaf, S. (2018). Probabilistic Random Forest: A Machine Learning Algorithm for Noisy Data Sets. *The Astronomical Journal*, 157(1), 16. <https://doi.org/10.3847/1538-3881/aaf101>
- Rigatti, S. J. (2017). Random Forest. *Journal of Insurance Medicine*, 47(1), 31–39.
<https://doi.org/10.17849/insm-47-01-31-39.1>
- Riley, R. D., Ensor, J., Snell, K. I. E., Debray, T. P. A., Altman, D. G., Moons, K. G. M., & Collins, G. S. (2016). External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: Opportunities and challenges. *BMJ (Clinical Research Ed.)*, 353, i3140.
<https://doi.org/10.1136/bmj.i3140>
- Riley, R. D., Ensor, J., Snell, K. I. E., Harrell, F. E., Martin, G. P., Reitsma, J. B., Moons, K. G. M., Collins, G., & Smeden, M. van. (2020). Calculating the sample size required for developing a clinical prediction model. *BMJ*, 368, m441. <https://doi.org/10.1136/bmj.m441>
- Rodríguez, L., Cervantes, E., & Ortiz, R. (2011). Malnutrition and Gastrointestinal and Respiratory Infections in Children: A Public Health Problem. *International Journal of Environmental Research and Public Health*, 8(4), 1174–1205. <https://doi.org/10.3390/ijerph8041174>
- Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M., & Rigol-Sanchez, J. P. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 67, 93–104.
<https://doi.org/10.1016/j.isprsjprs.2011.11.002>

- ROTA Council. (2019). *The broader impact of early childhood diarrhea*. <https://preventrotavirus.org/wp-content/uploads/2019/05/ROTA-Brief6-LastingImpact-SP-1-3.pdf>
- Roy, S. K., Buis, M., Weersma, R., Khatun, W., Chowdhury, S., Begum, A., Sarker, D., Thakur, S. K., & Khanam, M. (2011). Risk Factors of Mortality in Severely-malnourished Children Hospitalized with Diarrhoea. *Journal of Health, Population, and Nutrition*, 29(3), 229–235.
- Roy, S., Karim, A. B., Das, S. K., Maksud, S. I., Sultana, S. N., & Sultana, A. (2019). Risk Factors of Persistent Diarrhoea in Children below Five Years of Age. *Journal of Pediatric Health and Nutrition*, 1(1), 13–18. <https://doi.org/10.14302/issn.2691-5014.jphn-19-3038>
- Ryo, M., & Rillig, M. C. (2017). Statistically reinforced machine learning for nonlinear patterns and variable interactions. *Ecosphere*, 8(11), e01976. <https://doi.org/10.1002/ecs2.1976>
- Rytter, M. J. H., Kolte, L., Briend, A., Friis, H., & Christensen, V. B. (2014). The immune system in children with malnutrition—A systematic review. *PloS One*, 9(8), e105017. <https://doi.org/10.1371/journal.pone.0105017>
- Saad, K., Abo-Elela, M. G. M., El-Baseer, K. A. A., Ahmed, A. E., Ahmad, F.-A., Tawfeek, M. S. K., El-Houfey, A. A., Aboul_Khair, M. D., Abdel-Salam, A. M., Abo-elgheit, A., Qubaisy, H., Ali, A. M., & Abdel-Mawgoud, E. (2016). Effects of bovine colostrum on recurrent respiratory tract infections and diarrhea in children. *Medicine*, 95(37), e4560. <https://doi.org/10.1097/MD.0000000000004560>
- Sachdev, H. P. S., Kumar, S., Singh, K. K., Satyanarayana, L., & Puri, R. K. (1991). Risk Factors for Fatal Diarrhea in Hospitalized Children in India. *Journal of Pediatric Gastroenterology and Nutrition*, 12(1), 76.
- Saeed, N. K., Al-Beltagi, M., Bediwy, A. S., El-Sawaf, Y., & Toema, O. (2022). Gut microbiota in various childhood disorders: Implication and indications. *World Journal of Gastroenterology*, 28(18), 1875–1901. <https://doi.org/10.3748/wjg.v28.i18.1875>

- Sagaro, E. (2016). Risk factors for persistent diarrhea. *International Journal of Gastroenterology, Hepatology, Transplant and Nutrition*, 1(1), 20.
- Sahare, M., & Gupta, H. (2012). A Review of Multi-Class Classification for Imbalanced Data. *International Journal of Advanced Computer Research*, 2(3).
- Sahiner, B., Chen, W., Samala, R. K., & Petrick, N. (2023). Data drift in medical machine learning: Implications and potential remedies. *The British Journal of Radiology*, 20220878.
<https://doi.org/10.1259/bjr.20220878>
- Saito, T., & Rehmsmeier, M. (2023). *precrec: Calculate Accurate Precision-Recall and ROC (Receiver Operator Characteristics) Curves* (Version 0.14.1) [Computer software]. <https://CRAN.R-project.org/package=precrec>
- Salah, E. T., Algasim, S. H., Mhamoud, A. S., & Husian, N. E. O. S. A. (2015). Prevalence of hypoxemia in under-five children with pneumonia in an emergency pediatrics hospital in Sudan. *Indian Journal of Critical Care Medicine : Peer-Reviewed, Official Publication of Indian Society of Critical Care Medicine*, 19(4), 203–207. <https://doi.org/10.4103/0972-5229.154549>
- Saleh, H., & Layous, J. (2022). *Machine Learning -Regression*.
<https://doi.org/10.13140/RG.2.2.35768.67842>
- Salman, H., Kalakech, A., & Steiti, A. (2024). Random Forest Algorithm Overview. *Babylonian Journal of Machine Learning*, 2024, 69–79. <https://doi.org/10.58496/BJML/2024/007>
- Sarica, A., Cerasa, A., & Quattrone, A. (2017). Random Forest Algorithm for the Classification of Neuroimaging Data in Alzheimer’s Disease: A Systematic Review. *Frontiers in Aging Neuroscience*, 9, 329.
- Sarker, I. H. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*, 2(3), 160. <https://doi.org/10.1007/s42979-021-00592-x>

- Sarker, S. A., Ahmed, T., & Brüssow, H. (2017). Persistent diarrhea: A persistent infection with enteropathogens or a gut commensal dysbiosis? *Environmental Microbiology*, 19(10), 3789–3801. <https://doi.org/10.1111/1462-2920.13873>
- Scacchi, W. (2003). *Socio-Technical Design*.
- Schein, A. I., & Ungar, L. H. (2007). Active learning for logistic regression: An evaluation. *Machine Learning*, 68(3), 235–265. <https://doi.org/10.1007/s10994-007-5019-5>
- Schilling, K. A., Omore, R., Derado, G., Ayers, T., Ochieng, J. B., Farag, T. H., Nasrin, D., Panchalingam, S., Nataro, J. P., Kotloff, K. L., Levine, M. M., Oundo, J., Parsons, M. B., Bopp, C., Laserson, K., Stauber, C. E., Rothenberg, R., Breiman, R. F., O'Reilly, C. E., & Mintz, E. D. (2017). Factors Associated with the Duration of Moderate-to-Severe Diarrhea among Children in Rural Western Kenya Enrolled in the Global Enteric Multicenter Study, 2008–2012. *The American Journal of Tropical Medicine and Hygiene*, 97(1), 248–258. <https://doi.org/10.4269/ajtmh.16-0898>
- Schloerke, B., Allen, J., Tremblay, B., Dunné, F. van, Vandewoude, S., & RStudio. (2022). *plumber: An API Generator for R* (Version 1.2.1) [Computer software]. <https://cran.r-project.org/web/packages/plumber/index.html>
- Schonlau, M., & Zou, R. Y. (2020). The random forest algorithm for statistical learning. *The Stata Journal*, 20(1), 3–29. <https://doi.org/10.1177/1536867X20909688>
- Schreier, G., Pfeifer, B., & Baumgartner, M. (2022). *Development of an Architecture to Implement Machine Learning Based Risk Prediction in Clinical Routine: A Service-Oriented Approach*. IOS Press.
- Seeram, E. (2019). An Overview of Correlational Research. *Radiologic Technology*, 91(2), 176–179.
- Selbuz, S., Kırşacıoğlu, C. T., Kuloğlu, Z., Yılmaz, M., Penezoğlu, N., Sayıcı, U., Altuntaş, C., & Kansu, A. (2019). Diagnostic Workup and Micronutrient Deficiencies in Children With Failure to Thrive

- Without Underlying Diseases. *Nutrition in Clinical Practice: Official Publication of the American Society for Parenteral and Enteral Nutrition*, 34(4), 581–588. <https://doi.org/10.1002/ncp.10229>
- Senarath, U., & Sujendran, S. (2015). Prevalence of Stunting among Children Aged 6 to 36 Months, in the Eastern Province of Sri Lanka. *Journal of Nutritional Disorders & Therapy*, 05(01).
<https://doi.org/10.4172/2161-0509.1000154>
- Sendak, M. P., D’Arcy, J., Kashyap, S., Gao, M., Nichols, M., Corey, K., Ratliff, W., & Balu, S. (2020). A Path for Translation of Machine Learning Products into Healthcare Delivery. *EMJ Innovations*.
<https://doi.org/10.33590/emjinnov/19-00172>
- Serra, A., & Tagliaferri, R. (2019). Unsupervised Learning: Clustering. In *Encyclopedia of Bioinformatics and Computational Biology* (pp. 350–357). Elsevier. <https://doi.org/10.1016/B978-0-12-809633-8.20487-1>
- Sey, I. C. M., Ehimiyein, A. M., Bottomley, C., Riley, E. M., & Mooney, J. P. (2020). Does Malaria Cause Diarrhoea? A Systematic Review. *Frontiers in Medicine*, 7, 589379.
<https://doi.org/10.3389/fmed.2020.589379>
- Shah, G., Das, B., Kumar, S., Singh, M., & Bhandari, G. (2007). Acid base and electrolyte disturbance in diarrhoea. *Kathmandu University Medical Journal (KUMJ)*, 5(1).
<https://pubmed.ncbi.nlm.nih.gov/18603987/>
- Shah, M. (2016, September 14). *RServe – Introduction and Installation | R-bloggers*. <https://www.r-bloggers.com/2016/09/rserve-introduction-and-installation/>
- Shahid, N., Rappon, T., & Berta, W. (2019). Applications of artificial neural networks in health care organizational decision-making: A scoping review. *PLOS ONE*, 14(2), e0212356.
<https://doi.org/10.1371/journal.pone.0212356>

- Shahid, N. S., Sack, D. A., Rahman, M., Alam, A. N., & Rahman, N. (1988). Risk factors for persistent diarrhoea. *BMJ (Clinical Research Ed.)*, 297(6655), 1036–1038.
<https://doi.org/10.1136/bmj.297.6655.1036>
- Shanthakumari, R., Nalini, C., Vinothkumar, S., Roopadevi, E. M., & Govindaraj, B. (2022). Multi Disease Prediction System using Random Forest Algorithm in Healthcare System. *2022 International Mobile and Embedded Technology Conference (MECON)*, 242–247.
<https://doi.org/10.1109/MECON53876.2022.9752432>
- Sharma, S., Gosain, A., & Jain, S. (2022). *A Review of the Oversampling Techniques in Class Imbalance Problem* (pp. 459–472). https://doi.org/10.1007/978-981-16-2594-7_38
- Shee, H., Cheruiyot, W., & Kimani, S. (2014). *Application of k-Nearest Neighbour Classification in Medical Data Mining*. 4.
- Shekar, M., Dayton Eberwein, J., & Kakietek, J. (2016). The costs of stunting in South Asia and the benefits of public investments in nutrition. *Maternal & Child Nutrition*, 12(S1), 186–195.
<https://doi.org/10.1111/mcn.12281>
- Shen, H., Zhao, H., & Jiang, Y. (2023). Machine Learning Algorithms for Predicting Stunting among Under-Five Children in Papua New Guinea. *Children*, 10(10), Article 10.
<https://doi.org/10.3390/children10101638>
- Shipe, M. E., Deppen, S. A., Farjah, F., & Grogan, E. L. (2019). Developing prediction models for clinical use using logistic regression: An overview. *Journal of Thoracic Disease*, 11(Suppl 4), S574–S584.
<https://doi.org/10.21037/jtd.2019.01.25>
- Shouman, M., Turner, T., & Stocker, R. (2012). Applying k-Nearest Neighbour in Diagnosing Heart Disease Patients. *International Journal of Information and Education Technology*, 2(3).

- Simon, A. K., Hollander, G. A., & McMichael, A. (2015). Evolution of the immune system in humans from infancy to old age. *Proceedings of the Royal Society B: Biological Sciences*, 282(1821), 20143085. <https://doi.org/10.1098/rspb.2014.3085>
- Simsekler, M. C. E., Alhashmi, N. H., Azar, E., King, N., Luqman, R. A. M. A., & Al Mulla, A. (2021). Exploring drivers of patient satisfaction using a random forest algorithm. *BMC Medical Informatics and Decision Making*, 21(1), 157. <https://doi.org/10.1186/s12911-021-01519-5>
- Smith, A. E., Shah, M., & Badireddy, M. (2023). Failure to Thrive. In *StatPearls*. StatPearls Publishing. <http://www.ncbi.nlm.nih.gov/books/NBK459287/>
- Södergård, C., Mildorf, T., Habyarimana, E., Berre, A. J., Fernandes, J. A., & Zinke-Wehlmann, C. (Eds.). (2021). *Big Data in Bioeconomy: Results from the European DataBio Project*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-71069-9>
- Son, Y.-J., Kim, H.-G., Kim, E.-H., Choi, S., & Lee, S.-K. (2010). Application of Support Vector Machine for Prediction of Medication Adherence in Heart Failure Patients. *Healthcare Informatics Research*, 16(4), 253–259. <https://doi.org/10.4258/hir.2010.16.4.253>
- Song, J., Gao, Y., Yin, P., Li, Yi, Li, Yang, Zhang, J., Su, Q., Fu, X., & Pi, H. (2021). The Random Forest Model Has the Best Accuracy Among the Four Pressure Ulcer Prediction Models Using Machine Learning Algorithms. *Risk Management and Healthcare Policy*, 14, 1175–1187. <https://doi.org/10.2147/RMHP.S297838>
- Sperrin, M., Riley, R. D., Collins, G. S., & Martin, G. P. (2022). Targeted validation: Validating clinical prediction models in their intended population and setting. *Diagnostic and Prognostic Research*, 6(1), 24. <https://doi.org/10.1186/s41512-022-00136-8>
- Sreeramareddy, C. T., Shankar, R. P., Sreekumaran, B. V., Subba, S. H., Joshi, H. S., & Ramachandran, U. (2006). Care seeking behaviour for childhood illness- a questionnaire survey in western Nepal. *BMC International Health and Human Rights*, 6(1), 7. <https://doi.org/10.1186/1472-698X-6-7>

- Steyerberg, E. (2009). *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating* (Vol. 19). <https://doi.org/10.1007/978-0-387-77244-8>
- Stone, K., Zwiggelaar, R., Jones, P., & Parthaláin, N. M. (2022). A systematic review of the prediction of hospital length of stay: Towards a unified framework. *PLOS Digital Health*, 1(4), e0000017. <https://doi.org/10.1371/journal.pdig.0000017>
- Strand, T. A., Sharma, P. R., Gjessing, H. K., Ulak, M., Chandyo, R. K., Adhikari, R. K., & Sommerfelt, H. (2012). Risk Factors for Extended Duration of Acute Diarrhea in Young Children. *PLoS ONE*, 7(5). <https://doi.org/10.1371/journal.pone.0036436>
- Sunguya, B. F., Zhu, S., Mpembeni, R., & Huang, J. (2019). Trends in prevalence and determinants of stunting in Tanzania: An analysis of Tanzania demographic health surveys (1991–2016). *Nutrition Journal*, 18(1), 85. <https://doi.org/10.1186/s12937-019-0505-8>
- Supriya, M., & Deepa, A. J. (2020). A novel approach for breast cancer prediction using optimized ANN classifier based on big data environment. *Health Care Management Science*, 23(3), 414–426. <https://doi.org/10.1007/s10729-019-09498-w>
- Suto, J., Oniga, S., & Sitar, P. P. (2016). Comparison of wrapper and filter feature selection algorithms on human activity recognition. *2016 6th International Conference on Computers Communications and Control (ICCCC)*, 124–129. <https://doi.org/10.1109/ICCCC.2016.7496749>
- Taguri, A. E., Betimal, I., Mahmud, S. M., Ahmed, A. M., Goulet, O., Galan, P., & Hercberg, S. (2009). Risk factors for stunting among under-fives in Libya. *Public Health Nutrition*, 12(8), 1141–1149. <https://doi.org/10.1017/S1368980008003716>
- Takele, K., Zewotir, T., & Ndanguza, D. (2019). Understanding correlates of child stunting in Ethiopia using generalized linear mixed models. *BMC Public Health*, 19(1), 626. <https://doi.org/10.1186/s12889-019-6984-x>

- Talbert, A., Ngari, M., Bauni, E., Mwangome, M., Mturi, N., Otiende, M., Maitland, K., Walson, J., & Berkley, J. A. (2019). Mortality after inpatient treatment for diarrhea in children: A cohort study. *BMC Medicine*, 17(1), 20. <https://doi.org/10.1186/s12916-019-1258-0>
- Talukder, A., & Ahammed, B. (2020). Machine learning algorithms for predicting malnutrition among under-five children in Bangladesh. *Nutrition*, 78, 110861. <https://doi.org/10.1016/j.nut.2020.110861>
- Tang, J., Alelyani, S., & Liu, H. (2014). *Feature Selection for Classification: A Review*.
- Tang, M. N., Adolphe, S., Rogers, S. R., & Frank, D. A. (2021). Failure to Thrive or Growth Faltering: Medical, Developmental/Behavioral, Nutritional, and Social Dimensions. *Pediatrics In Review*, 42(11), 590–603. <https://doi.org/10.1542/pir.2020-001883>
- Tarekegn, A. N., Giacobini, M., & Michalak, K. (2021). A review of methods for imbalanced multi-label classification. *Pattern Recognition*, 118, 107965. <https://doi.org/10.1016/j.patcog.2021.107965>
- Templ, M., Kowarik, A., Alfons, A., Cillia, G. de, Prantner, B., & Rannetbauer, W. (2022). *VIM: Visualization and Imputation of Missing Values* (Version 6.2.2) [Computer software]. <https://cran.r-project.org/web/packages/VIM/index.html>
- Thabtah, F., Abdelhamid, N., & Peebles, D. (2019). A machine learning autism classification based on logistic regression analysis. *Health Information Science and Systems*, 7(1), 12. <https://doi.org/10.1007/s13755-019-0073-5>
- Theerthagiri, P. (2022). Predictive analysis of cardiovascular disease using gradient boosting based learning and recursive feature elimination technique. *Intelligent Systems with Applications*, 16, 200121. <https://doi.org/10.1016/j.iswa.2022.200121>
- Thobari, J. A., Sutarman, Mulyadi, A. W. E., Watts, E., Carvalho, N., Debellut, F., Clark, A., Soenarto, Y., & Bines, J. E. (2022). Direct and indirect costs of acute diarrhea in children under five years of age

- in Indonesia: Health facilities and community survey. *The Lancet Regional Health – Western Pacific*, 19. <https://doi.org/10.1016/j.lanwpc.2021.100333>
- Tickell, K. D., Pavlinac, P. B., John-Stewart, G. C., Denno, D. M., Richardson, B. A., Naulikha, J. M., Kirera, R. K., Swierczewski, B. E., Singa, B. O., & Walson, J. L. (2017). Impact of Childhood Nutritional Status on Pathogen Prevalence and Severity of Acute Diarrhea. *The American Journal of Tropical Medicine and Hygiene*, 97(5), 1337–1344. <https://doi.org/10.4269/ajtmh.17-0139>
- Tickell, K. D., Sharmin, R., Deichsel, E. L., Lamberti, L. M., Walson, J. L., Faruque, A. S. G., Pavlinac, P. B., Kotloff, K. L., & Chisti, M. J. (2020). The effect of acute malnutrition on enteric pathogens, moderate-to-severe diarrhoea, and associated mortality in the Global Enteric Multicenter Study cohort: A post-hoc analysis. *The Lancet Global Health*, 8(2), e215–e224. [https://doi.org/10.1016/S2214-109X\(19\)30498-X](https://doi.org/10.1016/S2214-109X(19)30498-X)
- Toroitich, A. M., Dunford, L., Armitage, R., & Tanna, S. (2022). Patients Access to Medicines – A Critical Review of the Healthcare System in Kenya. *Risk Management and Healthcare Policy*, Volume 15, 361–374. <https://doi.org/10.2147/RMHP.S348816>
- Troeger, C., Blacker, B. F., Khalil, I. A., Rao, P. C., Cao, S., Zimsen, S. R., Albertson, S. B., Stanaway, J. D., Deshpande, A., Abebe, Z., Alvis-Guzman, N., Amare, A. T., Asgedom, S. W., Anteneh, Z. A., Antonio, C. A. T., Aremu, O., Asfaw, E. T., Atey, T. M., Atique, S., ... Reiner, R. C. (2018). Estimates of the global, regional, and national morbidity, mortality, and aetiologies of diarrhoea in 195 countries: A systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Infectious Diseases*, 18(11), 1211–1228. [https://doi.org/10.1016/S1473-3099\(18\)30362-1](https://doi.org/10.1016/S1473-3099(18)30362-1)
- Troeger, C., Redford, S., & Biehl, M. (2019, February 19). *Understanding the burden of diarrhea*. DefeatDD. <https://www.defeatdd.org/blog/understanding-burden-diarrhea>

- Tsui, A., Tudosiu, P.-D., Brudfors, M., Jha, A., Cardoso, J., Ourselin, S., Ashburner, J., Rees, G., Davis, D., & Nachev, P. (2023). Predicting mortality in acutely hospitalised older patients: The impact of model dimensionality. *BMC Medicine*, 21(1), 10. <https://doi.org/10.1186/s12916-022-02698-2>
- Tusting, L. S., Bradley, J., Bhatt, S., Gibson, H. S., Weiss, D. J., Shenton, F. C., & Lindsay, S. W. (2020). Environmental temperature and growth faltering in African children: A cross-sectional study. *The Lancet Planetary Health*, 4(3), e116–e123. [https://doi.org/10.1016/S2542-5196\(20\)30037-1](https://doi.org/10.1016/S2542-5196(20)30037-1)
- Uauy, R., Desjeux, J.-F., Ahmed, T., Hossain, M., Brewster, D., Forbes, D., Caton, H., & Kleinman, R. E. (2012). Global efforts to address severe acute malnutrition. *Journal of Pediatric Gastroenterology and Nutrition*, 55(5), 476–481. <https://doi.org/10.1097/MPG.0b013e318272af06>
- Uddin, S., Haque, I., Lu, H., Moni, M. A., & Gide, E. (2022). Comparative performance analysis of K-nearest neighbour (KNN) algorithm and its different variants for disease prediction. *Scientific Reports*, 12(1), Article 1. <https://doi.org/10.1038/s41598-022-10358-x>
- Uddin, S., Khan, A., Hossain, M. E., & Moni, M. A. (2019). Comparing different supervised machine learning algorithms for disease prediction. *BMC Medical Informatics and Decision Making*, 19(1), 281. <https://doi.org/10.1186/s12911-019-1004-8>
- Ukah, U. V., Payne, B., Karjalainen, H., Kortelainen, E., Seed, P. T., Conti-Ramsden, F. I., Cao, V., Laivuori, H., Hutcheon, J., Chappell, L., Mark Ansermino, J., Vatish, M., Redman, C., Lee, T., Li, L., Magee, L. A., & von Dadelszen, P. (2019). Temporal and external validation of the fullPIERS model for the prediction of adverse maternal outcomes in women with pre-eclampsia. *Pregnancy Hypertension*, 15, 42–50. <https://doi.org/10.1016/j.preghy.2018.01.004>
- Umamaheswari, B., Biswal, N., Adhisivam, B., Parija, S. C., & Srinivasan, S. (2010). Persistent diarrhea: Risk factors and outcome. *The Indian Journal of Pediatrics*, 77(8), 885–888. <https://doi.org/10.1007/s12098-010-0125-y>

- UNICEF. (2022, July). *Diarrhoea*. UNICEF DATA. <https://data.unicef.org/topic/child-health/diarrhoeal-disease/>
- Urbanek, S. (2022). *Rserve: Binary R server* (Version 1.8-11) [Computer software]. <https://cran.r-project.org/web/packages/Rserve/index.html>
- Uysal, G., Sökmen, A., & Vidinlisan, S. (2000). Clinical risk factors for fatal diarrhea in hospitalized children. *The Indian Journal of Pediatrics*, 67(5), 329–333. <https://doi.org/10.1007/BF02820679>
- Vaivada, T., Akseer, N., Akseer, S., Somaskandan, A., Stefopoulos, M., & Bhutta, Z. A. (2020). Stunting in childhood: An overview of global burden, trends, determinants, and drivers of decline. *The American Journal of Clinical Nutrition*, 112(Supplement_2), 777S-791S. <https://doi.org/10.1093/ajcn/nqaa159>
- Vake, D., Šinik, B., Vičič, J., & Tošić, A. (2025). Is Open Source the Future of AI? A Data-Driven Approach. *Applied Sciences*, 15(5). <https://doi.org/10.3390/app15052790>
- Vale, D., El-Sharif, A., & Ali, M. (2022). Explainable artificial intelligence (XAI) post-hoc explainability methods: Risks and limitations in non-discrimination law. *AI and Ethics*, 2, 1–12. <https://doi.org/10.1007/s43681-022-00142-y>
- van Buuren, S., Groothuis-Oudshoorn, K., Vink, G., Schouten, R., Robitzsch, A., Rockenschaub, P., Doove, L., & Jolan, S. (2021, January 27). *Package ‘mice.’* <https://cran.r-project.org/web/packages/mice/mice.pdf>
- Van Calster, B., McLernon, D. J., van Smeden, M., Wynants, L., Steyerberg, E. W., Bossuyt, P., Collins, G. S., Macaskill, P., McLernon, D. J., Moons, K. G. M., Steyerberg, E. W., Van Calster, B., van Smeden, M., Vickers, A. J., & On behalf of Topic Group ‘Evaluating diagnostic tests and prediction models’ of the STRATOS initiative. (2019). Calibration: The Achilles heel of predictive analytics. *BMC Medicine*, 17(1), 230. <https://doi.org/10.1186/s12916-019-1466-7>

- van Engelen, J. E., & Hoos, H. H. (2020). A survey on semi-supervised learning. *Machine Learning*, 109(2), 373–440. <https://doi.org/10.1007/s10994-019-05855-6>
- Varpio, L., Paradis, E., Uijtdehaage, S., & Young, M. (2020). The Distinctions Between Theory, Theoretical Framework, and Conceptual Framework. *Academic Medicine: Journal of the Association of American Medical Colleges*, 95(7), 989–994. <https://doi.org/10.1097/ACM.0000000000003075>
- Vega, R. M., & Avva, U. (2024). Pediatric Dehydration. In *StatPearls*. StatPearls Publishing.
<http://www.ncbi.nlm.nih.gov/books/NBK436022/>
- Verburg, I. W. M., Keizer, N. F. de, Jonge, E. de, & Peek, N. (2014). Comparison of Regression Methods for Modeling Intensive Care Length of Stay. *PLOS ONE*, 9(10), e109684.
<https://doi.org/10.1371/journal.pone.0109684>
- Victora, C. G., de Onis, M., Hallal, P. C., Blössner, M., & Shrimpton, R. (2010). Worldwide timing of growth faltering: Revisiting implications for interventions. *Pediatrics*, 125(3), e473-480.
<https://doi.org/10.1542/peds.2009-1519>
- Victora, C. G., Huttly, S. R., Fuchs, S. C., Nobre, L. C., & Barros, F. C. (1992). Deaths due to dysentery, acute and persistent diarrhoea among Brazilian infants. *Acta Paediatrica*, 81(s383), 7–11.
<https://doi.org/10.1111/j.1651-2227.1992.tb12364.x>
- VijayaKumar, K., Lavanya, B., Nirmala, I., & Caroline, S. S. (2019). Random Forest Algorithm for the Prediction of Diabetes. *2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN)*, 1–5. <https://doi.org/10.1109/ICSCAN.2019.8878802>
- Vilcins, D., Sly, P. D., & Jagals, P. (2018). Environmental Risk Factors Associated with Child Stunting: A Systematic Review of the Literature. *Annals of Global Health*, 84(4), 551–562.
<https://doi.org/10.29024/aogh.2361>

- Walker, C. L. F., Perin, J., Liu, J. L., Katz, J., Tielsch, J. M., & Black, R. (2013). Does comorbidity increase the risk of mortality among children under 3 years of age? *BMJ Open*, 3(8), e003457.
<https://doi.org/10.1136/bmjopen-2013-003457>
- Wambui, W. M., Kimani, S., & Odhiambo, E. (2018). Determinants of Health Seeking Behavior among Caregivers of Infants Admitted with Acute Childhood Illnesses at Kenyatta National Hospital, Nairobi, Kenya. *International Journal of Pediatrics*, 2018, e5190287.
<https://doi.org/10.1155/2018/5190287>
- Wang, H., Abbas, K. M., Abbasifard, M., Abbasi-Kangevari, M., Abbastabar, H., Abd-Allah, F., Abdelalim, A., Abolhassani, H., Abreu, L. G., Abrigo, M. R. M., Abushouk, A. I., Adabi, M., Adair, T., Adebayo, O. M., Adedeji, I. A., Adekanmbi, V., Adeoye, A. M., Adetokunboh, O. O., Advani, S. M., ... Murray, C. J. L. (2020). Global age-sex-specific fertility, mortality, healthy life expectancy (HALE), and population estimates in 204 countries and territories, 1950–2019: A comprehensive demographic analysis for the Global Burden of Disease Study 2019. *The Lancet*, 396(10258), 1160–1203. [https://doi.org/10.1016/S0140-6736\(20\)30977-6](https://doi.org/10.1016/S0140-6736(20)30977-6)
- Wang, J., Xu, S., Wang, Y., & EL Bouhissi, H. (2025). An analytics framework for healthcare expenditure forecasting with machine learning. *Healthcare Analytics*, 8, 100428.
<https://doi.org/10.1016/j.health.2025.100428>
- Wang, L., Wang, X., Chen, A., Jin, X., & Che, H. (2020). Prediction of Type 2 Diabetes Risk and Its Effect Evaluation Based on the XGBoost Model. *Healthcare*, 8(3), Article 3.
<https://doi.org/10.3390/healthcare8030247>
- Webb, G. I. (2016). Naïve Bayes. In C. Sammut & G. I. Webb (Eds.), *Encyclopedia of Machine Learning and Data Mining* (pp. 1–2). Springer US. https://doi.org/10.1007/978-1-4899-7502-7_581-1
- Were, V., Buff, A. M., Desai, M., Kariuki, S., Samuels, A. M., Phillips-Howard, P., Kuile, F. O. ter, Kachur, S. P., & Niessen, L. W. (2019). Trends in malaria prevalence and health related socioeconomic

- inequality in rural western Kenya: Results from repeated household malaria cross-sectional surveys from 2006 to 2013. *BMJ Open*, 9(9), e033883. <https://doi.org/10.1136/bmjopen-2019-033883>
- Werner de Vargas, V., Schneider Aranda, J. A., dos Santos Costa, R., da Silva Pereira, P. R., & Victória Barbosa, J. L. (2023). Imbalanced data preprocessing techniques for machine learning: A systematic mapping study. *Knowledge and Information Systems*, 65(1), 31–57. <https://doi.org/10.1007/s10115-022-01772-8>
- White, N. J. (2018). Anaemia and malaria. *Malaria Journal*, 17(1), 371. <https://doi.org/10.1186/s12936-018-2509-9>
- WHO. (2017). *Diarrhoeal disease*. <https://www.who.int/news-room/fact-sheets/detail/diarrhoeal-disease>
- WHO, World Bank, & UNICEF. (2020). *UNICEF/WHO/The World Bank Group joint child malnutrition estimates: Levels and trends in child malnutrition: key findings of the 2020 edition*. <https://www.who.int/publications-detail-redirect/jme-2020-edition>
- Wichmann, R. M., Fernandes, F. T., & Chiavegatto Filho, A. D. P. (2023). Improving the performance of machine learning algorithms for health outcomes predictions in multicentric cohorts. *Scientific Reports*, 13(1), Article 1. <https://doi.org/10.1038/s41598-022-26467-6>
- Wierzba, T. F., & Muhib, F. (2018). Exploring the broader consequences of diarrhoeal diseases on child health. *The Lancet Global Health*, 6(3), e230–e231. [https://doi.org/10.1016/S2214-109X\(18\)30047-0](https://doi.org/10.1016/S2214-109X(18)30047-0)
- Wijayanti, S., Utami, E., & Yaqin, A. (2022). *Comparison of Kernels on Support Vector Machine (SVM) Methods for Analysis of Cyberbullying* (p. 108). <https://doi.org/10.1109/ICITISEE57756.2022.10057761>

- Wilimitis, D., & Walsh, C. G. (2023). Practical Considerations and Applied Examples of Cross-Validation for Model Development and Evaluation in Health Care: Tutorial. *JMIR AI*, 2, e49023.
<https://doi.org/10.2196/49023>
- Woergoetter, F., & Porr, B. (2008). Reinforcement Learning. *Scholarpedia*, 3(3), 1448.
<https://doi.org/10.4249/scholarpedia.1448>
- Wolde, D., Tilahun, G. A., Kotiso, K. S., Medhin, G., & Eguale, T. (2022). The Burden of Diarrheal Diseases and Its Associated Factors among Under-Five Children in Welkite Town: A Community Based Cross-Sectional Study. *International Journal of Public Health*, 67, 1604960.
<https://doi.org/10.3389/ijph.2022.1604960>
- Wolff, R. F., Moons, K. G. M., Riley, R. D., Whiting, P. F., Westwood, M., Collins, G. S., Reitsma, J. B., Kleijnen, J., Mallett, S., & PROBAST Group†. (2019). PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. *Annals of Internal Medicine*, 170(1), 51–58.
<https://doi.org/10.7326/M18-1376>
- Wondimu, A., van der Schans, J., van Hulst, M., & Postma, M. J. (2020). Inequalities in Rotavirus Vaccine Uptake in Ethiopia: A Decomposition Analysis. *International Journal of Environmental Research and Public Health*, 17(8), 2696. <https://doi.org/10.3390/ijerph17082696>
- Wong, A., Young, A. T., Liang, A. S., Gonzales, R., Douglas, V. C., & Hadley, D. (2018). Development and Validation of an Electronic Health Record-Based Machine Learning Model to Estimate Delirium Risk in Newly Hospitalized Patients Without Known Cognitive Impairment. *JAMA Network Open*, 1(4), e181018. <https://doi.org/10.1001/jamanetworkopen.2018.1018>
- World Health Organization. (2016). *Global diffusion of eHealth: Making universal health coverage achievable*.
- World Health Organization. (2017). *Diarrhoeal disease*. <https://www.who.int/news-room/fact-sheets/detail/diarrhoeal-disease>

World Health Organization. (2020). *Global health estimates: Leading causes of DALYs*.

<https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/global-health-estimates-leading-causes-of-dalys>

World Health Organization. (2021a). *Ethics and Governance of Artificial Intelligence for Health: WHO Guidance* (1st ed). World Health Organization.

World Health Organization. (2021b, January 9). *Fact sheets—Malnutrition*. <https://www.who.int/news-room/fact-sheets/detail/malnutrition>

World Health Organization. (2024, March 7). *Diarrhoeal disease: Factsheet*. <https://www.who.int/news-room/fact-sheets/detail/diarrhoeal-disease>

World Health Organization, & UNICEF (Eds.). (2005). *Handbook IMCI: Integrated management of childhood illness*. World Health Organization.

Wyllie, R., Hyams, J. S., & Kay, M. (2015). *Pediatric Gastrointestinal and Liver Disease*. Elsevier Health Sciences.

Xie, J., Sun, L., & Zhao, Y. F. (2025). On the Data Quality and Imbalance in Machine Learning-based Design and Manufacturing—A Systematic Review. *Engineering*, 45, 105–131.
<https://doi.org/10.1016/j.eng.2024.04.024>

Xing, W., & Bei, Y. (2019). Medical Health Big Data Classification Based on KNN Classification Algorithm. *IEEE Access*, PP, 1–1. <https://doi.org/10.1109/ACCESS.2019.2955754>

Yang, C.-S., Wei, C.-P., Yuan, C.-C., & Schoung, J.-Y. (2010). Predicting the length of hospital stay of burn patients: Comparisons of prediction accuracy among different clinical stages. *Decision Support Systems*, 50(1), 325–335. <https://doi.org/10.1016/j.dss.2010.09.001>

Yang, S., Zhu, F., Ling, X., Liu, Q., & Zhao, P. (2021). Intelligent Health Care: Applications of Deep Learning in Computational Medicine. *Frontiers in Genetics*, 12.
<https://www.frontiersin.org/articles/10.3389/fgene.2021.607471>

- Yassin, K. (2000). Morbidity and risk factors of diarrheal diseases among under-five children in rural upper Egypt. *Journal of Tropical Pediatrics*, 46(5), 282–287.
<https://doi.org/10.1093/tropej/46.5.282>
- Ye-ting, Z., & Dao-ming, T. (2018). Systemic Inflammatory Response Syndrome (SIRS) and the Pattern and Risk of Sepsis Following Gastrointestinal Perforation. *Medical Science Monitor : International Medical Journal of Experimental and Clinical Research*, 24, 3888–3894.
<https://doi.org/10.12659/MSM.907922>
- Ygberg, S., & Nilsson, A. (2012). The developing immune system—From foetus to toddler. *Acta Paediatrica (Oslo, Norway: 1992)*, 101(2), 120–127. <https://doi.org/10.1111/j.1651-2227.2011.02494.x>
- Yoshikawa, H. (2022). Can naive Bayes classifier predict infection in a close contact of COVID-19? A comparative test for predictability of the predictive model and healthcare workers in Japan. *Journal of Infection and Chemotherapy*, 28(6), 774–779.
<https://doi.org/10.1016/j.jiac.2022.02.017>
- Yousefi, S., Haghighian, F., Naghdizadegan Jahromi, M., & Pourghasemi, H. R. (2022). Chapter 26—Pest-infected oak trees identify using remote sensing-based classification algorithms. In H. R. Pourghasemi (Ed.), *Computers in Earth and Environmental Sciences* (pp. 363–376). Elsevier.
<https://doi.org/10.1016/B978-0-323-89861-4.00020-8>
- Yu, W., Liu, T., Valdez, R., Gwinn, M., & Khoury, M. J. (2010). Application of support vector machine modeling for prediction of common diseases: The case of diabetes and pre-diabetes. *BMC Medical Informatics and Decision Making*, 10(1), 16. <https://doi.org/10.1186/1472-6947-10-16>
- Završnik, J., Kokol, P., Žlahtič, B., & Blažun Vošner, H. (2025). Machine Learning in Primary Health Care: The Research Landscape. *Healthcare*, 13(13), 1629.
<https://doi.org/10.3390/healthcare13131629>

- Zelege, A. J., Palumbo, P., Tubertini, P., Miglio, R., & Chiari, L. (2023). Machine learning-based prediction of hospital prolonged length of stay admission at emergency department: A Gradient Boosting algorithm analysis. *Frontiers in Artificial Intelligence*, 6.
<https://www.frontiersin.org/articles/10.3389/frai.2023.1179226>
- Zenebe, G. A., Gebretsadik, S., Muche, T., Sisay, D., Meno, A., Hareru, H. E., & Debela, B. G. (2022). Level of Mothers'/Caregivers' Healthcare-Seeking Behavior for Child's Diarrhea, Fever, and Respiratory Tract Infections and Associated Factors in Ethiopia: A Systematic Review and Meta-Analysis. *BioMed Research International*, 2022, e4053085.
<https://doi.org/10.1155/2022/4053085>
- Zhang, A., Xing, L., Zou, J., & Wu, J. C. (2022). Shifting machine learning for healthcare from development to deployment and from models to data. *Nature Biomedical Engineering*, 6(12), Article 12.
<https://doi.org/10.1038/s41551-022-00898-y>
- Zhang, X., Xiao, J., & Gu, F. (2019). Applying Support Vector Machine to Electronic Health Records for Cancer Classification. *2019 Spring Simulation Conference (SpringSim)*, 1–9.
<https://doi.org/10.23919/SpringSim.2019.8732906>
- Zhang, X., Zhivaki, D., & Lo-Man, R. (2017). Unique aspects of the perinatal immune system. *Nature Reviews. Immunology*, 17(8), 495–507. <https://doi.org/10.1038/nri.2017.54>
- Zhao, J., Feng, Q., Wu, P., Lupu, R. A., Wilke, R. A., Wells, Q. S., Denny, J. C., & Wei, W.-Q. (2019). Learning from Longitudinal Data in Electronic Health Record and Genetic Data to Improve Cardiovascular Event Prediction. *Scientific Reports*, 9(1), 717. <https://doi.org/10.1038/s41598-018-36745-x>
- Zheng, D., Liwinski, T., & Elinav, E. (2020). Interaction between microbiota and immunity in health and disease. *Cell Research*, 30(6), 492–506. <https://doi.org/10.1038/s41422-020-0332-7>

- Zhongheng Zhang, Y. Z. (2019). Predictive analytics with gradient boosting in clinical medicine. *Annals of Translational Medicine*, 7(7). <https://doi.org/10.21037/atm.2019.03.29>
- Zhou, X., Li, X., Zhang, Z., Han, Q., Deng, H., Jiang, Y., Tang, C., & Yang, L. (2022). Support vector machine deep mining of electronic medical records to predict the prognosis of severe acute myocardial infarction. *Frontiers in Physiology*, 13, 991990. <https://doi.org/10.3389/fphys.2022.991990>
- Zhu, M., Chen, W., Hirdes, J. P., & Stolee, P. (2007). The K-nearest neighbor algorithm predicted rehabilitation potential better than current Clinical Assessment Protocol. *Journal of Clinical Epidemiology*, 60(10), 1015–1021. <https://doi.org/10.1016/j.jclinepi.2007.06.001>
- Zhu, W. (2016). *Shiny Tutorial*. <https://bookdown.org/weicheng/shinyTutorial/>
- Zhuo, Z., Ku, H.-C., Huang, Z., Xing, G., & Xing, C. (2017). Differentiating the Cochran-Armitage trend test and Pearson's χ^2 test: Location and dispersion. *Annals of Human Genetics*, 81(5), 184–189. <https://doi.org/10.1111/ahg.12202>

APPENDICES

Appendix 1: Ethics Approval



College of Agriculture and Environmental Sciences_Health REC

Date: 15/11/2023

Dear: Mr Billy Ogwel

NHREC Registration # : REC-170616-051

Ref #: 2023/CAES_HREC/2192

Name: Mr Billy Ogwel

Student #: 16949145

**Decision: Ethics Approval from
15/11/2023 to 31/10/2028**

Researcher: Mr Billy Ogwel
16949145@mylife.unisa.ac.za +254737225000

Supervisor: Dr Mzazi Vincent homevz@unisa.ac.za

**A MACHINE LEARNING APPROACH TO THE PREDICTION OF POOR DIARRHEAL OUTCOMES
AMONG CHILDREN UNDER FIVE YEARS OF AGE: AN EVALUATION OF LONGER DURATION
DIARRHEA, CHRONIC MALNUTRITION, AND MORTALITY**

Qualification: PhD (Information Systems)

Thank you for the application for research ethics clearance by the College of Agriculture and Environmental Sciences_Health REC for the above mentioned research study. Ethics approval is granted for five years.

The **negligible risk application** was **reviewed** by College of Agriculture and Environmental Sciences_Health REC on 15 November 2023 in compliance with the Unisa Policy on Research Ethics and the Standard Operating Procedure on Research Ethics Risk Assessment.

The proposed research may now commence with the provisions that:

1. The researcher(s) will ensure that the research project adheres to the values and principles expressed in the UNISA Policy on Research Ethics.
2. Any adverse circumstance arising in the undertaking of the research project that is relevant to the ethicality of the study should be communicated in writing to the College of Agriculture and Environmental Sciences_Health REC.
3. The researcher(s) will conduct the study according to the methods and procedures set out in the approved application.
4. Any changes that can affect the study-related risks for the research participants, particularly in terms of assurances made with regards to the protection of participants' privacy and the confidentiality of the data, should be reported to the Committee in writing, accompanied by a progress report.

5. The researcher will ensure that the research project adheres to any applicable national legislation, professional codes of conduct, institutional guidelines and scientific standards relevant to the specific field of study. Adherence to the following South African legislation is important, if applicable: Protection of Personal Information Act, no 4 of 2013; Children's act

Appendix 2: Glossary of Key Terms

Term	Definition
Acute Gastroenteritis	Having at least three looser than normal stools and/or at least one instance of unexplained vomiting followed by loose stool within a 24-hour timeframe, where these symptoms began within the 7 days prior to care-seeking at the facility
Artificial Neural Networks	Computational networks that simulate the human-brain processes thereby establishing empirical relationships between outcome and features
Area Under the Curve	Measure of a model's discriminative ability using the relationship between true and false positive rates across varying decision thresholds.
Disability-Adjusted Life Years	A composite metric of total disease burden that combines years of life lost from early mortality with years lived with disability. It reflects the population impact of health conditions by estimating the number of years of healthy life lost due to illness or premature death.
Explanatory Model Analysis	the process of understanding and interpreting the decisions made by a ML model by giving us visibility into the model's decision-making process including feature importance, model structure, and decision-making processes
Fetal Growth Restriction	A prenatal ultrasound-based estimate of fetal weight that falls below the 10th percentile for gestational age.
Gradient Boosting Machine	An algorithm that builds strong learners from weak learners in a gradual, additive and sequential manner offering lots of flexibility allowing for optimization of different loss functions
Height-for-Age Z-score	Measure of how a child's height (or length for children <2 years) compares with the median height of children of the same age and sex in the World Health Organization child growth standards.
Integrated Management of Childhood Illness guidelines	An integrated strategy for child health that addresses the overall well-being of the child by seeking to reduce mortality, morbidity, and disability while promoting optimal growth and development among infants and children under five years of age. The approach combines preventive and curative interventions, delivered both at the household and community level as well as through healthcare services provided by trained health workers in health facilities.
K-Nearest Neighbors	A classifier that approximates how likely an observation is assigned to a group based on the group membership of observations within close proximity
Longer Duration Diarrhea	A diarrheal illness that extends for seven (7) or more days
Linear Growth Faltering	A decrease of 0.5 HAZ or more ($\Delta \text{HAZ} \geq -0.5$) between enrolment and follow-up
Logistic Regression	A statistical approach in machine learning that approximates the likelihood of an event occurring by using a sigmoid function to

	generate a linear transformation of the predictors into a probability value between 0 and 1.
Medically-Attended Diarrhea	A diarrheal illness (at least three loose stools within 24 hours), starting within the previous week and following at least two diarrhea-free days for which care was sought.
Multiple Imputation by Chained Equations	Framework for performing data imputation and includes several functions for identifying the missing data patterns present in a dataset, which operates by creating several complete datasets with plausible imputations for missing values, employing a chained equations approach where variables are imputed iteratively based on observed values of other variables.
Moderate-to-Severe Diarrhea	Diarrhea (at least three loose stools in 24 hours) that started within the previous week after at least seven diarrhea-free days and displayed at least one of these symptoms: sunken eyes, reduced skin elasticity, bloody diarrhea, need for IV fluids, or being hospitalized.
Machine Learning	Arm of artificial intelligence which builds applications that acquire knowledge from data and continuously improve their performance, without requiring additional programming to achieve such improvements
Naive Bayes	A probability-based classifier which utilizes the likelihood of terms and categories co-occurring in the estimation of the probabilities of distinct classes represented in the data assuming conditional independence between terms within each category and is insensitive to noise
Negative Predictive Value	The percentage of correctly predicted negatives out of all negative predictions.
Persistent Diarrhea	A diarrheal illness that extends for fourteen (14) days or more
Propensity score Mean-Squared-Error	The accuracy of a propensity score model by quantifying the average squared difference between the estimated propensity scores and the true, unknown probabilities of treatment assignment for subjects.
Positive Predictive Value	The percentage of correctly predicted positives out of all positive predictions.
Precision-Recall Area Under the Curve	The measure of a model's performance by evaluating the precision-recall trade-off, offering a more accurate assessment, especially for imbalanced classes.
Person Years of Observation	Metric measuring the actual time-at-risk each individual contributes to a study, combining the number of participants and their follow-up duration
Random Forest	An ensemble learning method that employs bootstrap aggregation technique to train multiple decision trees to get a more accurate and stable prediction
Receiver Operating Characteristic curve	A graphical plot used to evaluate the overall diagnostic accuracy of a test and to compare the performance of two or more diagnostic methods.

Severe Acute Malnutrition	Presence of bilateral pitting oedema of the feet or severe wasting, defined as weight-for-height/length below –3 standard deviations or a mid-upper arm circumference (MUAC) less than 115 mm.
SHapley Additive exPlanations values	A method from explainable artificial intelligence used to quantify the contribution of each feature to an individual prediction made by a machine learning model.
Seattle Index of Comorbidity	A weighted, validated scoring system developed to measure the burden of pre-transplant comorbidities.
Support Vector Machine	A supervised machine learning algorithm used for classification and regression tasks. It works by identifying the optimal decision boundary (hyperplane) that separates data points belonging to different classes with the maximum possible margin.
Years-Lived with-Disability	Quantification the amount of time individuals spend in states of less than full health due to disease or injury.
Years-of-Life Lost	Metric used to quantify the burden of premature mortality

Appendix 3: Methodological Appendix

2.1 Description of Machine Learning Algorithms Used

LR is a supervised learning classification algorithm that predicts the probability of an outcome variable which is dichotomous in nature (Shipe et al., 2019) while NB is a probabilistic classifier that uses joint probabilities of terms and categories to estimate the probabilities of categories given in a dataset working on the assumption that all terms are conditionally independent of each other in a given category and is insensitive to noise (Langarizadeh & Moghbeli, 2016). RF is an algorithm that builds an ensemble of decision trees trained with a bagging approach to get a more accurate and stable prediction (Khalilia et al., 2011), GBM - an algorithm that builds strong learners from weak learners in a gradual, additive and sequential manner offering lots of flexibility allowing for optimization of different loss functions (Zhongheng Zhang, 2019). SVM works through discriminating between two classes by generating a hyperplane which optimally segregates classes after the mathematical transformation of input data into a high dimensional space (Yu et al., 2010). KNN is a classifier that estimates how likely a data point is to be a member of one group or another based on what group the data points nearest to it are in (Shee et al., 2014), ANN are computational networks that simulate the human-brain processes thereby establishing empirical relationships between outcome and features (Shahid et al., 2019). These algorithms have been previously applied in various classification tasks involving disease prediction in healthcare.

2.2 Boruta Feature Selection

The minimal, average and maximum Z score of a shadow feature are represented by blue boxplots. Red, yellow and green boxplots represent Z scores of rejected, tentative and confirmed features, respectively. From the Boruta output, features that were either confirmed or tentative were used in model development. A shadow feature is an artificial feature created to help determine the importance of real features. For each real feature in the dataset, a corresponding shadow feature is created by shuffling the values of the real feature so as to maintain the distribution of the data while destroying any relationship with the outcome. By using shadow features, Boruta provides a robust way to assess the significance of real features, effectively distinguishing between noise and informative features.

2.3 Justification for 75%:25% Data Split

While there is no established guidelines on data split ratios, it is common practice to use 75%-80% of the total sample for training since a large training set reduces bias and will yield better performance (Li, 2022).

2.4 Ensemble Modelling

Ensemble modelling is the technique where different base learners (level-0 models) are created to predict an outcome and the predictions from these models are then aggregated using a meta-learner (Level-1 model) to make a final prediction [9]. Our stacked ensemble used the following algorithms as base learners: RF, ANN, NB, KNN, GBM, SVM, Linear discriminant analysis (LDA), General additive Models (GAM), and Adaboost. While logistic regression was the algorithm of choice in the meta-learner. A simple algorithm was used for the meta-learner because using a complex meta-learner increases the probability of overfitting the predictions from the base-learners [10].

2.5: Hyper parameter evaluation

Model	Hyper parameters
Random Forest (RF)	Number of Trees (ntree): Evaluated values: 100, 200, 500 Number of Variables Tried at Each Split (mtry): Evaluated values: 2, 4, 6, 8 Best Hyper parameters: ntree = 500, mtry = 6
Support Vector Machine (SVM)	Cost (C): Evaluated values: 0.1, 1, 10, 100 Kernel Type: Linear Kernel Gamma (γ): Evaluated values: 0.01, 0.1, 1 Best Hyper parameters: C = 10, γ = 0.1
Artificial Neural Network (ANN)	Number of Hidden Layers: Evaluated values: 1, 2, 3 Number of Neurons per Layer: Evaluated values: 5, 10, 20 Learning Rate: Evaluated values: 0.01, 0.001 Best Hyper parameters: 2 hidden layers, 10 neurons per layer, learning rate = 0.001
Gradient Boosting (GBM)	Number of Trees (n.trees): Evaluated values: 50, 100, 200 Interaction Depth (interaction.depth): Evaluated values: 1, 3, 5 Shrinkage (shrinkage): Evaluated values: 0.01, 0.1 Best Hyper parameters: n.trees = 200, interaction.depth = 3, shrinkage = 0.1
k-Nearest Neighbors (k-NN)	Number of Neighbors (k): Evaluated values: 3, 5, 7, 9 Distance Metric: Euclidean distance Best Hyper parameters: k = 5

2.5 Calibration

Calibration which involves comparing the model's prediction against the real (observed) distribution was assessed using Brier scores (the mean squared error between the actual outcome and the estimated probabilities), Spiegelhalter's z-test (a formal measurement that serves as a proxy for calibration calculated from the decomposition of Brier score) and its accompanying p-value (Huang et al., 2020).

2.6 Business Value Evaluation

To evaluate the business value of the predictive model, modelplotr package (Nagelkerke, 2020) was used to build valuable evaluation plots (cumulative gains, cumulative lift, response and cumulative response plots). The cumulative gains plot was used to visualize the percentage of the target class members that were selected if we decided to select up until percentile X while the cumulative lift plot was used to explain how much better selecting based on our model was compared to taking random selections. The response plot was used to plot the percentage of target class observations per percentile. Lastly, the cumulative response plot was used to show the expected percentage of the target class observations in the selection, when we apply the model and select up until percentile X.

2.7 Sample Size Estimation

Sample size was conducted utilizing a formula developed by Riley et al (Riley et al., 2020).

$$n = P/(S - 1) \ln(1 - \frac{R^2_{cs}}{S}) \text{ (Riley et al., 2020)}$$

Where P= Candidate predictor parameters; S= 1- shrinkage ; R²_{cs}- Cox-Snell R squared statistic ;
P was 25 for mortality prediction; desired shrinkage level was ≤ 10% S=0.9 and R²_{cs} is at least 0.1

No	Participants in the study	Study	Total population (Development cohort)	Sample Size	Temporal Validation cohort
1	Children presenting with Moderate-Severe-Diarrhea and have Diarrhea duration data available in Rural Western Kenya	Vaccine Impact on Diarrhea in Africa (VIDA) Study (2015-2018)	1,484	1,443	685
2	Children presenting with Moderate-Severe-Diarrhea and have Diarrhea duration data available in Rural Western Kenya	VIDA Study (2015-2018)	1,484	1,443	655

3	Children hospitalized with acute gastroenteritis at Siaya County Referral Hospital and discharge outcome available	Rotavirus Inpatient Surveillance (2010-2020)	2,271	2,122	-

The sample size calculation closely aligned with the total population count. Consequently, we opted to include the entire population in the primary studies as participants for each diarrheal outcome in this particular study.

2.8 Definition of Variables

Wasting: Weight for Height Z-score (WHZ) were categorized as: normal if > -2.0 ; Moderate if -3.0 to less than -2.0 ; Severe if less than -3.0 .

Stunting: Height for Age Z-score (HAZ) categorized as: normal if > -2.0 ; Moderate if -3.0 to less than -2.0 ; Severe if less than -3.0 .

Dehydration: WHO classification used:

- Severe Dehydration: Child exhibiting at least two of the following signs: Lethargic or unconscious; sunken eyes; Drinks poorly or is unable to drink; Skin pinch goes back very slowly.
- Some Dehydration: Child exhibiting at least two of the following signs: Restless or irritable; sunken eyes; drinks eagerly or is very thirsty; skin pinch goes back slowly

Fever: Caregiver reported fever/hotness of body.

Weight-loss: Caregiver reported loss of weight with the illness

Vesikari Score: Severity of diarrheal illness was assessed using Vesikari as described in Table below:

Parameter	1	2	3
Diarrhea			
Maximum number stools per day	1–3	4–5	≥ 6
Diarrhea duration (day)	1–4	5	≥ 6
Vomiting			
Maximum number per day	1	2–4	≥ 5
Vomiting duration (day)	1	2	≥ 3
Maximum body temperature ($^{\circ}\text{C}$)	37.1–38.4	38.5–38.9	≥ 39.0
Severity of dehydration (%)	N/A	1–5	≥ 6
Treatment	Rehydration	Hospitalization	N/A
Severity rating scales	<7 (mild)	7–10 (moderate)	≥ 11 (severe)

Adapted from Ruuska T and Vesikari T. Scand J Infect Dis 1990;22:259-67⁴⁾.

Appendix 4: Permissions for Use of Images

RightsLink Printable License

<https://s100.copyright.com/App/PrintableLicenseFrame.jsp?publisherI...>

JOHN WILEY AND SONS LICENSE TERMS AND CONDITIONS

Feb 04, 2026

This Agreement between Billy Ogwel ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	6202221137109
License date	Feb 04, 2026
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Wiley Books
Licensed Content Title	Epidemiology
Licensed Content Author	Yvette J. Johnson-Walker, John B. Kaneene
Licensed Content Date	Mar 9, 2018
Licensed Content Pages	29
Type of use	Dissertation/Thesis
Requestor type	University/Academic
Format	Electronic
Portion	Figure/table

Number of figures/tables	1
Will you be translating?	No
Title of new work	PhD Thesis
Institution name	UNISA
Expected presentation date	Mar 2026
Portions	Figure 1.2
The Requesting Person / Organization to Appear on the License	Billy Ogwel
Requestor Location	Kisumu
	Kisumu, 40100 Kenya
Publisher Tax ID	EU826007151
Total	0.00 USD
Terms and Conditions	

TERMS AND CONDITIONS

This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one of its group companies (each a "Wiley Company") or handled on behalf of a society with which a Wiley Company has exclusive publishing rights in relation to a particular work (collectively "WILEY"). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the billing and payment terms and conditions established by the Copyright Clearance Center Inc., ("CCC's Billing and Payment terms and conditions"), at the time that you opened your RightsLink account (these are available at any time at <http://myaccount.copyright.com>).

Terms and Conditions

- The materials you have requested permission to reproduce or reuse (the "Wiley Materials") are protected by copyright.
- You are hereby granted a personal, non-exclusive, non-sub licensable (on a stand-alone basis), non-transferable, worldwide, limited license to reproduce the Wiley Materials for the purpose specified in the licensing process. This license, **and any CONTENT (PDF or image file) purchased as part of your order**, is for a one-time use only and limited to any maximum distribution number specified in the license. The first instance of republication or reuse granted by this license must be completed within two years of the date of the grant of this license (although copies prepared before the end date may be distributed thereafter). The Wiley Materials shall not be used in any other manner or for any other purpose, beyond what is granted in the license. Permission is granted subject to an appropriate acknowledgement given to the author, title of the material/book/journal and the publisher. You shall also duplicate the copyright notice that appears in the Wiley publication in your use of the Wiley Material. Permission is also granted on the understanding that nowhere in the text is a previously published source acknowledged for all or part of this Wiley Material. Any third party content is expressly excluded from this permission.
- With respect to the Wiley Materials, all rights are reserved. Except as expressly granted by the terms of the license, no part of the Wiley Materials may be copied, modified, adapted (except for minor reformatting required by the new Publication), translated, reproduced, transferred or distributed, in any form or by any means, and no derivative works may be made based on the Wiley Materials without the prior permission of the respective copyright owner.**For STM Signatory Publishers clearing permission under the terms of the [STM Permissions Guidelines](#) only, the terms of the license are extended to include subsequent editions and for editions in other languages, provided such editions are for the work as a whole in situ and does not involve the separate exploitation of the permitted figures or extracts,** You may not alter, remove or suppress in any manner any copyright, trademark or other notices displayed by the Wiley Materials. You may not license, rent, sell, loan, lease, pledge, offer as security, transfer or assign the Wiley Materials on a stand-alone basis, or any of the rights granted to you hereunder to any other person.
- The Wiley Materials and all of the intellectual property rights therein shall at all times remain the exclusive property of John Wiley & Sons Inc, the Wiley Companies, or their respective licensors, and your interest therein is only that of having possession of and the right to reproduce the Wiley Materials pursuant to Section 2 herein during the continuance of this Agreement. You agree that you own no right, title or interest in or to the Wiley Materials or any of the intellectual property rights therein. You shall have no rights hereunder other than the license as provided for above in Section 2. No right, license or interest to any trademark, trade name, service mark or other branding ("Marks") of WILEY or its licensors is granted hereunder, and you agree that you shall not assert any such right, license or interest with respect thereto

- NEITHER WILEY NOR ITS LICENSORS MAKES ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS OR THE ACCURACY OF ANY INFORMATION CONTAINED IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND ITS LICENSORS AND WAIVED BY YOU.
- WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.
- You shall indemnify, defend and hold harmless WILEY, its Licensors and their respective directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.
- IN NO EVENT SHALL WILEY OR ITS LICENSORS BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.
- Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as nearly as possible the same economic effect as the original provision, and the legality, validity and enforceability of the remaining provisions of this Agreement shall not be affected or impaired thereby.
- The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party's right to enforce each and every term and condition of this Agreement. No breach under this agreement shall be deemed waived or excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to any other or subsequent breach by such other party.
- This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.

- Any fee required for this permission shall be non-refundable after thirty (30) days from receipt by the CCC.
- These terms and conditions together with CCC's Billing and Payment terms and conditions (which are incorporated herein) form the entire agreement between you and WILEY concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in writing signed by both parties. This Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.
- In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall prevail.
- WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.
- This Agreement will be void if the Type of Use, Format, Circulation, or Requestor Type was misrepresented during the licensing process.
- This Agreement shall be governed by and construed in accordance with the laws of the State of New York, USA, without regards to such state's conflict of law rules. Any legal action, suit or proceeding arising out of or relating to these Terms and Conditions or the breach thereof shall be instituted in a court of competent jurisdiction in New York County in the State of New York in the United States of America and each party hereby consents and submits to the personal jurisdiction of such court, waives any objection to venue in such court and consents to service of process by registered or certified mail, return receipt requested, at the last known address of such party.

WILEY OPEN ACCESS TERMS AND CONDITIONS

Wiley Publishes Open Access Articles in fully Open Access Journals and in Subscription journals offering Online Open. Although most of the fully Open Access journals publish open access articles under the terms of the Creative Commons Attribution (CC BY) License only, the subscription journals and a few of the Open Access Journals offer a choice of Creative Commons Licenses. The license type is clearly identified on the article.

The Creative Commons Attribution License

The [Creative Commons Attribution License \(CC-BY\)](#) allows users to copy, distribute and transmit an article, adapt the article and make commercial use of the article. The CC-BY license permits commercial and non-

Creative Commons Attribution Non-Commercial License

The [Creative Commons Attribution Non-Commercial \(CC-BY-NC\) License](#) permits use,

distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.(see below)

Creative Commons Attribution-Non-Commercial-NoDerivs License

The [Creative Commons Attribution Non-Commercial-NoDerivs License](#) (CC-BY-NC-ND) permits use, distribution and reproduction in any medium, provided the original work is properly cited, is not used for commercial purposes and no modifications or adaptations are made. (see below)

Use by commercial "for-profit" organizations

Use of Wiley Open Access articles for commercial, promotional, or marketing purposes requires further explicit permission from Wiley and will be subject to a fee.

Further details can be found on Wiley Online Library <http://olabout.wiley.com/WileyCDA/Section/id-410895.html>

Other Terms and Conditions:

v1.10 Last updated September 2015

Questions? customercare@copyright.com.

**SAGE PUBLICATIONS LICENSE
TERMS AND CONDITIONS**

Feb 05, 2026

This Agreement between Billy Ogwel ("You") and Sage Publications ("Sage Publications") consists of your license details and the terms and conditions provided by Sage Publications and Copyright Clearance Center.

License Number	6202701102093
License date	Feb 05, 2026
Licensed Content Publisher	Sage Publications
Licensed Content Publication	Health Informatics Journal
Licensed Content Title	Sources of variation in primary care clinical workflow: Implications for the design of cognitive support
Licensed Content Author	Laura G Militello, Nicole B Arbuckle, Jason J Saleem, Emily Patterson, et al.
Licensed Content Date	Mar 1, 2014
Licensed Content Volume	20
Licensed Content Issue	1
Licensed Content Pages	15
Type of Use	Dissertation/Thesis

Requestor type	I am NOT the author of the requested article
Format	Electronic
Portion	Figure/table
Number of figures/tables	1
Will you be translating?	No, only English
Circulation	7
Title of new work	PhD Thesis
Institution name	UNISA
Expected presentation date	Mar 2026
Portions	Figure 1
The Requesting Person / Organization to Appear on the License	Billy Ogwel
	Mr. Billy Ogwel Kisumu
Requestor Location	Kisumu, 40100 Kenya
Publisher Tax ID	GB232600116
Payment Type	Invoice
Email Address	ogwelbill@gmail.com

	Mr. Billy Ogwel Kisumu
Billing Address	
	Kisumu, Kenya 40100
Total	0.00 USD
Terms and Conditions	

Sage Terms and Conditions for Permissions Administered Through RightsLink

You, the Requestor, are requesting to use the material specified in the permission request (the "Work"). Your agreement to the terms and conditions herein and completion of a permission request constitutes a Permission Request. You are in no way required to use the Work; however, should you decide not to use the Work after you complete this request, you agree to cancel your order through this website. Under the above conditions, and the following terms and conditions, permission to use Work ("Permission") is granted solely to you:

1. Sage reserves the right to revoke any Permission, at Sage's sole discretion, within two (2) business days of the request.
2. The number of copies ("Copies") for print use is defined as the total number of copies made for distribution or for repurposing, and the number of copies ("Copies") for electronic use is defined as the total number of viewers of the Work, recipients of copies of the Work, and individuals or entities who may have access to the Work. The Copies must not exceed the Copies as stated in the Permission Request.
3. Requests to post a full article on a website, internet, intranet, or any publicly accessible site must be submitted to the Publisher through the Sage Permissions Portal (<https://sage-cloud.atlassian.net/servicedesk/customer/portal/10>).
4. Permission is granted only for the type of use specified in the Permission Request. If any information pertaining to your Permission Request changes, you must cancel this request and resubmit a new request with the correct and current permission request information. Sage may exercise its right to revoke Permission, if Sage finds, in Sage's sole opinion, that the context in which you have used or repurposed the Work is considered libelous, in violation of any right of privacy, or otherwise unlawful; infringement or in violation of any copyright or other proprietary right of others; or can be construed to possibly cause harm or injury. You agree that use of Work will be professional, in the context of fact-based and generally acceptable professional practices.
5. Permission does not include the use within Custom Publishing Programs, and all use within such programs is explicitly prohibited.
6. Permission does not include use of the material within Massive Open Online Courses (MOOC's). For permission to include material in a MOOC, please contact Sage directly through the Sage Permissions Portal (<https://sage-cloud.atlassian.net/>)

[servicedesk/customer/portal/10](#)).

7. If your Permission Request includes the right to translate the Work, you agree that the translation of the material shall be made faithfully and accurately, and that abbreviations and/or alterations in the text and/or title of the Work shall be made only with Sage's prior written consent. Requestor shall not own or acquire any copyright or other proprietary rights in the Work or any other material furnished by Sage, including without limitation translations or transcriptions thereof, all of which rights shall be owned by and/or are hereby assigned to Sage. Requestor shall indemnify Sage against any and all claims, including without limitation attorneys' fees and legal costs, that concern or relate to (a) inaccurate translation or transcription of the Work, (b) infringement claims arising out of the inclusion of material not furnished by Sage or (c) infringement or other claims asserted by anyone retained by Requestor to translate the Work. Requestor agrees that the name of the Author(s), Copyright Holder, and Publisher shall appear in due prominence on the title page of every copy of the translation and in all advertisements for the translation, and that the translation shall include: (1) the Work's original copyright notice and original American title, both in English, and (2) notice in granted translated language in identifying the Sage Publications as the original publisher and stating the translation is published by arrangement with Sage. The rights licensed to Requestor are not transferrable. The translated article reprint must include the following disclaimer in English and in the language of the reprint: "While every effort has been made to ensure that the contents of this publication are factually correct, neither the authors nor the publisher accepts, and they hereby expressly exclude to the fullest extent permissible under applicable law, any and all liability arising from the contents published in this Article, including, without limitation, from any errors, omissions, inaccuracies in original or following translation, or for any consequences arising therefrom. Nothing in this notice shall exclude liability which may not be excluded by law. Approved product information should be reviewed before prescribing any subject medications."
8. Permission is granted for prospective use only, and does not apply to any use that has occurred prior to the date of the Permission Request.
9. Permission does not apply to any material (reprints, illustrations, figures, tables, etc.) not controlled by Sage. Requests for permission to re-use third-party material should be submitted to the original copyright holder, as indicated in the credit line for the materials.
10. Full acknowledgment of your source must appear in every copy of your work as follows:

Author(s), Book/Journal Title (Journal Volume Number and Issue Number)
pp. xx-xx. Copyright © YEAR by (Copyright Holder). Reprinted by
Permission of Sage Publications
11. Unless specified in the request or by prior arrangement with Sage, payment is due from you within sixty (60) days after the completion of Permission.
12. It is assumed that the Requester is using the selection in question, and is subject to billing and collections procedures, unless otherwise noted.
13. Permission Requests for reuse of text excerpts shall not exceed 50% of a journal article's content.
14. No more than 20% of any one Sage book or journal issue may be reused at one time.

Other Terms and Conditions:

v4.06

Questions? customer care@copyright.com.

**OXFORD UNIVERSITY PRESS LICENSE
TERMS AND CONDITIONS**

Feb 05, 2026

This Agreement between Billy Ogwel ("You") and Oxford University Press ("Oxford University Press") consists of your license details and the terms and conditions provided by Oxford University Press and Copyright Clearance Center.

License Number	6202710547180
License date	Feb 05, 2026
Licensed content publisher	Oxford University Press
Licensed content publication	International Journal of Epidemiology
Licensed content title	Profile: The KEMRI/CDC Health and Demographic Surveillance System—Western Kenya
Licensed content author	Odhiambo, Frank O; Laserson, Kayla F
Licensed content date	Aug 28, 2012
Type of Use	Thesis/Dissertation
Institution name	
Title of your work	PhD Thesis
Publisher of your work	UNISA
Expected publication date	Mar 2026

Permissions cost	0.00 USD
Value added tax	0.00 USD
Total	0.00 USD
Title of new work	PhD Thesis
Institution name	UNISA
Expected presentation date	Mar 2026
Portions	Figure 1
The Requesting Person / Organization to Appear on the License	Billy Ogwel
Requestor Location	Mr. Billy Ogwel Kisumu
	Kisumu, 40100 Kenya
Publisher Tax ID	GB125506730
Payment Type	Invoice
Email Address	ogwelbill@gmail.com
Billing Address	Mr. Billy Ogwel Kisumu
	Kisumu, Kenya 40100

Total 0.00 USD

Terms and Conditions

**STANDARD TERMS AND CONDITIONS FOR REPRODUCTION OF
MATERIAL FROM AN OXFORD UNIVERSITY PRESS JOURNAL**

1. Use of the material is restricted to the type of use specified in your order details.
2. This permission covers the use of the material in the **English** language in the following territory: *world*. If you have requested additional permission to translate this material, the terms and conditions of this reuse will be set out in clause 12.
3. This permission is limited to the particular use authorized in (1) above and does not allow you to sanction its use elsewhere in any other format other than specified above, nor does it apply to quotations, images, artistic works etc that have been reproduced from other sources which may be part of the material to be used.
4. No alteration, omission or addition is made to the material without our written consent. Permission **must** be re-cleared with Oxford University Press if/when you decide to reprint.
5. The following credit line appears wherever the material is used: author, title, journal, year, volume, issue number, pagination, by permission of Oxford University Press or the sponsoring society if the journal is a society journal. Where a journal is being published on behalf of a learned society, the details of that society **must** be included in the credit line.
6. For the reproduction of a full article from an Oxford University Press journal for whatever purpose, the corresponding author of the material concerned should be informed of the proposed use. Contact details for the corresponding authors of all Oxford University Press journal contact can be found alongside either the abstract or full text of the article concerned, accessible from www.oxfordjournals.org. Should there be a problem clearing these rights, please contact journals.permissions@oup.com
7. If the credit line or acknowledgement in our publication indicates that any of the figures, images or photos was reproduced, drawn or modified from an earlier source it will be necessary for you to clear this permission with the original publisher as well. If this permission has not been obtained, please note that this material cannot be included in your publication/photocopies.
8. While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by Oxford University Press or by Copyright Clearance Center (CCC)) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's

Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and Oxford University Press reserves the right to take any and all action to protect its copyright in the materials.

9. This license is personal to you and may not be sublicensed, assigned or transferred by you to any other person without Oxford University Press's written permission.

10. Oxford University Press reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

11. You hereby indemnify and agree to hold harmless Oxford University Press and CCC, and their respective officers, directors, employs and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

12. Unless otherwise specified, inclusion under a Creative Commons license or any other Open Access license allowing onward reuse is prohibited.

13. Where permission to translate has been agreed, the credit line and any copyright/disclaimer notices provided by OUP shall be included on the same page as the Material, translated into the language of the new work, except for trademark names:

Translated and reproduced by permission of Oxford University Press on behalf of the <Society Name>. Translation Disclaimer: OUP and the <Society Name> are not responsible or in any way liable for the accuracy of the translation. The Licensee is solely responsible for the translation in this publication/reprint.

14. Other Terms and Conditions:

v1.5

Questions? customercare@copyright.com.

**OXFORD UNIVERSITY PRESS LICENSE
TERMS AND CONDITIONS**

Feb 05, 2026

This Agreement between Billy Ogwel ("You") and Oxford University Press ("Oxford University Press") consists of your license details and the terms and conditions provided by Oxford University Press and Copyright Clearance Center.

License Number	6202711153827
License date	Feb 05, 2026
Licensed content publisher	Oxford University Press
Licensed content publication	Clinical Kidney Journal
Licensed content title	External validation of prognostic models: what, why, how, when and where?
Licensed content author	Ramspek, Chava L; Jager, Kitty J
Licensed content date	Nov 24, 2020
Type of Use	Thesis/Dissertation
Institution name	
Title of your work	PhD Thesis
Publisher of your work	UNISA
Expected publication date	Mar 2026

Permissions cost	0.00 USD
Value added tax	0.00 USD
Total	0.00 USD
Title of new work	PhD Thesis
Institution name	UNISA
Expected presentation date	Mar 2026
Portions	Figure 1
The Requesting Person / Organization to Appear on the License	Billy Ogwel
Requestor Location	Mr. Billy Ogwel Kisumu Kisumu, 40100 Kenya
Publisher Tax ID	GB125506730
Payment Type	Invoice
Email Address	ogwelbill@gmail.com
Billing Address	Mr. Billy Ogwel Kisumu Kisumu, Kenya 40100
Total	0.00 USD

Terms and Conditions

STANDARD TERMS AND CONDITIONS FOR REPRODUCTION OF MATERIAL FROM AN OXFORD UNIVERSITY PRESS JOURNAL

1. Use of the material is restricted to the type of use specified in your order details.
2. This permission covers the use of the material in the **English** language in the following territory: *world*. If you have requested additional permission to translate this material, the terms and conditions of this reuse will be set out in clause 12.
3. This permission is limited to the particular use authorized in (1) above and does not allow you to sanction its use elsewhere in any other format other than specified above, nor does it apply to quotations, images, artistic works etc that have been reproduced from other sources which may be part of the material to be used.
4. No alteration, omission or addition is made to the material without our written consent. Permission **must** be re-cleared with Oxford University Press if/when you decide to reprint.
5. The following credit line appears wherever the material is used: author, title, journal, year, volume, issue number, pagination, by permission of Oxford University Press or the sponsoring society if the journal is a society journal. Where a journal is being published on behalf of a learned society, the details of that society **must** be included in the credit line.
6. For the reproduction of a full article from an Oxford University Press journal for whatever purpose, the corresponding author of the material concerned should be informed of the proposed use. Contact details for the corresponding authors of all Oxford University Press journal contact can be found alongside either the abstract or full text of the article concerned, accessible from www.oxfordjournals.org Should there be a problem clearing these rights, please contact journals.permissions@oup.com
7. If the credit line or acknowledgement in our publication indicates that any of the figures, images or photos was reproduced, drawn or modified from an earlier source it will be necessary for you to clear this permission with the original publisher as well. If this permission has not been obtained, please note that this material cannot be included in your publication/photocopies.
8. While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by Oxford University Press or by Copyright Clearance Center (CCC)) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute

copyright infringement and Oxford University Press reserves the right to take any and all action to protect its copyright in the materials.

9. This license is personal to you and may not be sublicensed, assigned or transferred by you to any other person without Oxford University Press's written permission.

10. Oxford University Press reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

11. You hereby indemnify and agree to hold harmless Oxford University Press and CCC, and their respective officers, directors, employs and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

12. Unless otherwise specified, inclusion under a Creative Commons license or any other Open Access license allowing onward reuse is prohibited.

13. Where permission to translate has been agreed, the credit line and any copyright/disclaimer notices provided by OUP shall be included on the same page as the Material, translated into the language of the new work, except for trademark names:

Translated and reproduced by permission of Oxford University Press on behalf of the <Society Name>. Translation Disclaimer: OUP and the <Society Name> are not responsible or in any way liable for the accuracy of the translation. The Licensee is solely responsible for the translation in this publication/reprint.

14. Other Terms and Conditions:

v1.5

Questions? customercare@copyright.com.

**OXFORD UNIVERSITY PRESS LICENSE
TERMS AND CONDITIONS**

Feb 05, 2026

This Agreement between Billy Ogwel ("You") and Oxford University Press ("Oxford University Press") consists of your license details and the terms and conditions provided by Oxford University Press and Copyright Clearance Center.

License Number 6202720141890

License date Feb 05, 2026

Licensed content publisher Oxford University Press

Licensed content publication Clinical Infectious Diseases

Licensed content title The Global Enteric Multicenter Study (GEMS) of Diarrheal Disease in Infants and Young Children in Developing Countries: Epidemiologic and Clinical Methods of the Case/Control Study

Licensed content author Kotloff, Karen L.; Blackwelder, William C.

Licensed content date Dec 15, 2012

Type of Use Thesis/Dissertation

Institution name

Title of your work PhD Thesis

Publisher of your work UNISA

Expected publication date Mar 2026

Permissions cost 0.00 USD

Value added tax 0.00 USD

Total 0.00 USD

Title of new work PhD Thesis

Institution name UNISA

Expected presentation date Mar 2026

Portions Figure 3

The Requesting Person /
Organization to Appear on
the License Billy Ogwel

Requestor Location Mr. Billy Ogwel
Kisumu

Kisumu, 40100
Kenya

Publisher Tax ID GB125506730

Payment Type Invoice

Email Address ogwelbill@gmail.com

Billing Address

Mr. Billy Ogwel
Kisumu

Kisumu, Kenya 40100

Total 0.00 USD

Terms and Conditions

**STANDARD TERMS AND CONDITIONS FOR REPRODUCTION OF
MATERIAL FROM AN OXFORD UNIVERSITY PRESS JOURNAL**

1. Use of the material is restricted to the type of use specified in your order details.
2. This permission covers the use of the material in the **English** language in the following territory: *world*. If you have requested additional permission to translate this material, the terms and conditions of this reuse will be set out in clause 12.
3. This permission is limited to the particular use authorized in (1) above and does not allow you to sanction its use elsewhere in any other format other than specified above, nor does it apply to quotations, images, artistic works etc that have been reproduced from other sources which may be part of the material to be used.
4. No alteration, omission or addition is made to the material without our written consent. Permission **must** be re-cleared with Oxford University Press if/when you decide to reprint.
5. The following credit line appears wherever the material is used: author, title, journal, year, volume, issue number, pagination, by permission of Oxford University Press or the sponsoring society if the journal is a society journal. Where a journal is being published on behalf of a learned society, the details of that society **must** be included in the credit line.
6. For the reproduction of a full article from an Oxford University Press journal for whatever purpose, the corresponding author of the material concerned should be informed of the proposed use. Contact details for the corresponding authors of all Oxford University Press journal contact can be found alongside either the abstract or full text of the article concerned, accessible from www.oxfordjournals.org Should there be a problem clearing these rights, please contact journals.permissions@oup.com
7. If the credit line or acknowledgement in our publication indicates that any of the figures, images or photos was reproduced, drawn or modified from an earlier source it will be necessary for you to clear this permission with the original publisher as well. If this permission has not been obtained, please note that this material cannot be included in your publication/photocopies.
8. While you may exercise the rights licensed immediately upon issuance of the license at

the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by Oxford University Press or by Copyright Clearance Center (CCC)) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and Oxford University Press reserves the right to take any and all action to protect its copyright in the materials.

9. This license is personal to you and may not be sublicensed, assigned or transferred by you to any other person without Oxford University Press's written permission.

10. Oxford University Press reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

11. You hereby indemnify and agree to hold harmless Oxford University Press and CCC, and their respective officers, directors, employs and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

12. Unless otherwise specified, inclusion under a Creative Commons license or any other Open Access license allowing onward reuse is prohibited.

13. Where permission to translate has been agreed, the credit line and any copyright/disclaimer notices provided by OUP shall be included on the same page as the Material, translated into the language of the new work, except for trademark names:

Translated and reproduced by permission of Oxford University Press on behalf of the <Society Name>. Translation Disclaimer: OUP and the <Society Name> are not responsible or in any way liable for the accuracy of the translation. The Licensee is solely responsible for the translation in this publication/reprint.

14. Other Terms and Conditions:

v1.5

Questions? customercare@copyright.com.
