A PROSPECTIVE EVALUATION OF THE COMMUNITY ASSESSMENT TOOLS (CATs) IN CHILDREN AND ADULTS WITH ACUTE LOWER RESPIRATORY TRACT INFECTIONS

AN EXTERNAL VALIDATION STUDY

By

Ruby Armon MBBS, MPH-International Health

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

February 2017

Abstract

Background: CATs is a hospital admission triage tool for influenza and pneumonia, and comprises of 7 criteria (6 objective and 1 subjective); any one of which is expected to trigger hospital admission in adults and children. As the CATs was only recently developed, its performance has only been assessed in the United Kingdom using data from the 2009 A (H1N1) flu pandemic. An external validation study is thus required to assess the generalisability of the CATs in a new patient population before it is recommended for adoption in practice.

Aim: The primary aim of this thesis was to prospectively evaluate the predictive performance of the Community Assessment Tools across a range of acute lower respiratory tract infections in a new geographical setting.

Five studies were included in this thesis, each addressing a specific objective in order to achieve the primary aim of this research:

- The first study examined the predictive performance of individual CATs criteria with regard to the need for hospital based care and mortality,
- The second study examined the performance of CATs as unweighted and weighted scoring indexes for assessing these outcomes in adults and children with ALRTI,
- The third study compared the performance of the unweighted CATs scoring index and the well-established pneumonia severity score –CRB65 in two separate sensitivity analyses involving cases of suspected ALRTI and pneumonia in adults.
- The fourth study sought to investigate the value of clinical intuition in medical decision-making by comparing the performance of the CATs as an

unweighted scoring index with or without the inclusion of its subjective criterion (causing other clinical concern).

- The fifth study was a qualitative exploration of critical care practices in Nigerian hospitals, conducted to provide contextual interpretation of findings related to the performance of the CATs in predicting the need for critical care and mechanical ventilation.

Methods: Data specific to the CATs and CRB65 were recorded at the initial consultation of patients with symptoms suggestive of an acute lower respiratory tract infection. The performance of the tools with regard to predicting the need for hospital admission , acute care interventions (supplemental oxygen, IV antibiotics and IV fluids), or any deaths on admission were assessed using specificity, sensitivity, negative and positive predictive values, Area under Receiver Operating Characteristic Curves (AUROC) with 95% confidence intervals and Hosmer-Lemeshow goodness-of-fit test. The qualitative study was conducted within a pragmatic paradigm and a semi-structured guide was used to for individual interviews with key hospital staff. Thematic analysis was used for analysing interview data.

Results: Data were obtained for 1016 (809 children <16years) consecutive patients from four hospitals in Lagos Nigeria. In the univariate analysis each CATs criterion significantly predicted an outcome which indicated the need for hospital care or mortality, also none of these criteria were redundant in the multivariable analysis. Criterion C representing low oxygen saturation levels of 92% or less and Criterion G –clinicians' intuition or gut feelings ,were the best predictors of the need for hospital based care and risk of death in both children and adults. The predictive accuracy of the individual criteria was generally low in terms of AUROC

values, but this improved when they were combined to create scoring indexes. An unweighted CATs threshold score of three points or more was appropriate for identifying both adults and children who would potentially benefit from hospital based care and therapeutic interventions.

In the comparison between CRB65 and the unweighted CATs score, the latter was seen to have better discrimination and calibration qualities for predicting all outcomes in adults with ALRTI, and although both the unweighted CATs and CRB65 scores demonstrated good predictive ability for the outcome 'in-hospital mortality' in adults with pneumonia, the unweighted CATs score outperformed the CRB65 in terms of AUROC values, however, the difference was marginal and statistically insignificant.

This thesis provided evidence to support considering clinical intuition or gut feelings in the assessment of patients with ALRTI, as the predictive performance of the unweighted CATs score was significantly better with the inclusion of subjective criterion G (causing other clinical concern) than when this criterion was omitted. Due to insufficient data it was not possible to assess the performance of the CATs in regard to the outcome 'need for critical care'. Findings from the qualitative study revealed that ALRTI was not a common indication for ICU admission in these hospitals, also that clinical predictions tools were not generally used in this regard, rather decisions regarding ICU admissions were determined by the ability to pay and bed availability.

Conclusion: the CATs criteria showed geographical generalisability despite differences in case-mix and geographical variations. When combined to create an unweighted scoring system, CATs showed good potential for guiding decisions to admit adults and children with ALRTI to hospital for further care. Although CATs was developed for use during an influenza pandemic, the unweighted CATs score

could be a useful tool for guiding admission decisions in adults with interpandemic pneumonia, as it performed reasonably well when compared with a widely validated and established pneumonia severity score.

Acknowledgements

These past years have been truly life changing, nothing short of an emotional rollercoaster; there were days when I doubted, days when I celebrated small victories, and days when I felt nothing at all. But as I look back on my journey, from deciding to apply for a PhD to completing this thesis, I am most grateful to God for this unique opportunity, for the lessons learned, for growth and for the person I have become – without Him I am indeed nothing.

I gratefully acknowledge the Niger Delta Development Commission (NDDC) for the funding to undertake this PhD.

I do not take full credit for overcoming any challenges experienced in the course of this research, and I would like to express my sincere gratitude to the following people for their invaluable support and contributions over the past years:

To Dr Puja Myles, I would like to say a heartfelt thank you, for her patience, genuine concern and motivation during tough times, her valuable feedback, and flexibility which allowed me to achieve other dreams while writing this thesis; in all she has shown me an excellent example of what a researcher should be.

A special thank you to Professor Jonathan Van-Tam for accepting me into his research group, and for his supervisory role and expert advice on my research.

Many thanks to Dr Calum Semple for taking time to discuss the Community Assessment Tools with me, explaining the use of each criterion in preparation for data collection in Nigeria.

I am particularly grateful to Aamir Khakwani, Sudhir Venkatesan, Kayode Oke, Uche Anyanwagu, Kelly Elimian and Solomon Idinyang for their helpful comments and encouragement throughout my PhD.

I am also grateful to Dr Yemi Osinaike and Professor Ajayi for their corporation and support throughout the data collection phase in Nigeria.

To my support system – my beautiful family, I would like to say a heartfelt thank you to my siblings for believing in me and encouraging me to pursue my dreams, I am blessed and proud to be an Armon.

Lastly but by no means least, no words can express the appreciation, respect and gratitude I feel towards my mother, for her continuous sacrifice and infallible devotion to making sure I succeed in all I do, for tireless prayers and words of wisdom, for being my companion and best friend, this thesis is dedicated to you –thank you Phil.

..*"Veni, vidi, vici"*

Table of contents

TABLE OF CONTENTS **1**
LIST OF FIGURES **8**
LIST OF TABLES **9**
LIST OF ABBREVIATIONS **12**
1: INTRODUCTION **14**
1.1 Thesis background **14**
1.2 The evolution of prognostication in medicine **15**
1.3 Consecutive phases of developing a prediction tool **16**
1.3.1 Step 1: Derivation 17
1.3.2 Step 2: Validation (external) and updating 18
1.3.3 Step 3: Impact analysis 19
1.4 Risk prediction for acute lower respiratory tract infections **20**
1.5 The Community Assessment Tools (CATs) **23**
1.6 An overview of acute lower respiratory tract infections in Nigeria **28**
1.7 Management of ALRTI in Nigeria **29**
1.8 Thesis Rationale and Aims **31**
1.9 Thesis outline **34**
2: RESEARCH METHODOLOGY AND METHODS **35**
2.1 Introduction **35**
2.2 Overall ethics approval for thesis **35**
2.3 Theoretical underpinning for the methodology of the external validation of the CATs **36**
2.4 Study design **37**
2.5 Setting **38**

2.6 Study centres 40

2.6.1. Lagos University Teaching Hospital (LUTH) 40

2.6.2 Lagos State University Teaching Hospital (LASUTH) 40

2.6.3 Massey Street Children's Hospital (MSCH) 40

2.6.4 Surulere General Hospital and Randle Maternity and Child Care Centre 41

2.7 Study participants 41

2.7.1 Inclusion criteria 42

2.7.2 Exclusion criteria 44

2.7.3 Follow-up of patients 44

2.8 Sample size calculation 44

2.9 Data collection pro forma and procedure 46

2.9.1 Pilot phase 48

2.9.2 Data entry 49

2.9.3 Data validation 49

2.9.4 Data storage 50

2.10 Data variables & definition 50

2.10.1 Predictors 50

2.10.2 Outcome variables 51

2.10.3 Covariates 52

2.10.4 Handling missing data 52

2.11 Statistical analyses 52

2.11.1 Descriptive analyses 54

2.11.2 Logistic regression analyses 54

2.11.3 Assessing multicollinearity 55

2.11.4 Creating combined CATs scores 55

2.11.5 Assessing predictive accuracy 56

2.11.5.1 Sensitivity and specificity 57

2.11.5.2 Positive predictive & negative predictive values 58

2.11.5.3 Receiver operating characteristic (ROC) value & Area under ROC (AUROC) curve 58

2.11.5.4 Hosmer-Lemeshow goodness-of–fit test 59

2.12 Ethics Statement 60

2.13 Qualitative Study Methodology 61

2.14 Philosophical underpinning 62

2.15 Study design 62

2.16 Ethics 63

2.17 Study population 63

2.18 Sampling and recruitment strategy 64

2.19 Data collection and interview process 65

2.20 Data analysis 66

2.20.1 Familiarisation with data 67

2.20.2 Generating initial codes 67

2.20.3 Searching for themes 68

2.20.4 Reviewing themes 68

2.20.5 Defining and naming themes 68

3: STUDY CENTRES AND DATA DESCRIPTION 69

3.1 Introduction 69

3.2 Descriptive characteristics 69

3.3 Diagnoses and Study Outcomes 73

4: ASSESSING THE PERFORMANCE OF THE INDIVIDUAL CATS CRITERIA IN PREDICTING HOSPITAL ADMISSION, ACUTE CARE INTERVENTIONS AND IN-HOSPITAL MORTALITY IN ADULTS AND CHILDREN WITH ALRTI 74

4.1 Introduction 74

4.2 Background 74

4.3 Statistical analyses 75

4.4 Results 76

4.4.1 Association between CATs and need for hospital admission 77

4.4.2 Association between CATs and need for supplemental oxygen 77

4.4.3 Association between CATs and need for intravenous antibiotics 78

4.4.4 Association between CATs and intravenous fluids 78

4.4.5 Association between CATs and in-hospital mortality 78

4.4.6 Multivariable analysis: model I including all criteria irrespective of significance in children (< 16years) 85

4.4.7 Multivariable analysis: model I including all criteria irrespective of significance in adults (> 16years) 86

4.4.8 Multivariable analysis model II: forward stepwise selection in adults. 90

4.4.9 Multivariable model II: forward stepwise selection in children 93

4.4.10 Predictive performance of individual CATs criteria for various outcomes in children (< 16 years) 97

4.4.11 Predictive performance of individual CATs criteria for various outcomes in adults 100

4.5 Discussion 103

4.5.1 Key findings 103

4.5.2 Strengths and limitations 103

4.5.3 Comparison with other studies 105

4.5.4 Summary 109

5: ASSESSING THE PERFORMANCE OF THE CATS AS A COMBINED SCORING INDEX FOR PREDICTING HOSPITAL ADMISSION, ACUTE CARE INTERVENTIONS AND IN-HOSPITAL MORTALITY IN ADULTS AND CHILDREN WITH ALRTI 110

5.1 Introduction 110

5.2 Background 110

5.3 Statistical analyses 112

5.4 Results 115

5.4.1 Assessment of multicollinearity between CATs criteria..... 115

5.4.2 Assessment of discriminatory ability of un-weighted CATs index..... 116

5.4.3 Comparing of the predictive performance of un-weighted and weighted CATs combined scores 117

5.4.4 Assessing the goodness of fit and calibration of the unweighted and weighted combined CATs regression models in children and adults 122

5.4.5 Comparison of the predictive performance of CATs as individual criteria and as combined scoring indexes 125

5.4.6 Predictive performance of un-weighted CATs threshold scores for predicting study outcomes in children and adults..... 126

5.4.7 Predictive performance of weighted CATs threshold scores for predicting study outcomes in children and adults..... 132

5.5 Discussion 137

5.5.1 Summary of key findings 137

5.5.2 Strengths and limitations 137

5.5.3 Comparison with existing literature 138

5.5.4 Summary 142

6: COMPARING THE UN-WEIGHTED CATS SCORE AND THE CRB65 FOR SEVERITY ASSESSMENT OF ADULT PATIENTS WITH ALRTI IN NIGERIA 144

6.1 Introduction 144

6.2 Background 144

6.3 Statistical analysis 145

6.4 Results 147

6.4.1 Comparison of un-weighted CATs and CRB65 scores in adults with ALRTI 148

6.4.2 Sensitivity analysis comparing predictive performance of CRB-65 and un-weighted CATS in adults with pneumonia 153

6.4.3 Predictive performance of CRB65 score thresholds in adults with ALRTI 157

6.5 Discussion 160

6.5.1 Summary of key findings 160

6.5.2 Strengths and limitations 160

6.5.3 Comparison with existing literature 162

6.5.4 Summary 165

7: ASSESSING THE VALUE OF CLINICAL INTUITION IN THE PERFORMANCE OF THE UN-WEIGHTED CATS SCORE FOR SEVERITY ASSESSMENT OF PATIENTS WITH ALRTI 167

7.1 Introduction 167

7.2 Decision making in clinical practice 167

7.3 Statistical Analysis 170

7.4 Results 171

7.4.1 Comparing the predictive performance of the unweighted CATs-7 versus CATs-6 scoring indices in children (<16 years) 171

7.4.2 Comparing the performance of unweighted CATs with all 7 criteria versus CATs without criterion G in adults (>16 years) 174
7.5 Discussion 177
7.5.1 Summary of key findings 177
7.5.2 Limitations and strengths 177
7.5.3 Comparison with existing literature 178
7.5.4 Summary 181
8: A QUALITATIVE EXPLORATION OF CRITICAL CARE PRACTICES IN NIGERIA: FINDINGS AND DISCUSSION 182
8.1 Overview 182
8.2 Description of Participants 182
8.3 Overview of themes 184
8.3.1 Definition and structure of critical care 184
8.3.2 Factors affecting critical care admission and discharge 187
8.3.3 ICU organisation and patient profile 192
8.4 Discussion 194
8.4.1: Definition and structure of critical care 194
8.4.2: Factors affecting discharge and admission 197
8.4.3: ICU organisation and patient profile 199
8.5 Generalisability 200
8.6 Strengths 202
8.7 Limitations 203
8.8 Reflexivity 203
8.9 Summary 205
9: THESIS CONCLUSIONS AND RECOMMENDATIONS 207
9.1 Recap of the thesis objectives and key findings 207
9.2 Discussion of the thesis findings 214
9.2.1 Implications for clinical practice and future research 220
9.2.1.1 Implications for practice 221
9.2.1.2 Recommendations for future research 222
REFERENCES 223

APPENDICES 235

APPENDIX I: TRIPOD checklist -methods section 237

APPENDIX II: Data collection Forms 239

APPENDIX III: Qualitative interview questionnaire 254

APPENDIX IV: Distribution of un-weighted CATs scores by patient age group 255

APPENDIX V: Distribution of weighted CATs combined scores in children 256

APPENDIX VI: Distribution of weighted CATs combined scores in adults 257

APPENDIX VII: Lowess plots showing linearity of logit assumption for un-weighted CATs scores in adults 258

APPENDIX VIII: Lowess plots showing linearity of logit assumption for un-weighted CATs scores in children 259

APPENDIX IX: Lowess plots assessing the linearity of logit assumption for CRB-65 scores in adults with ALRTI 260

APPENDIX X: Comparison of ROC curves for CRB65 score thresholds in adults with ALRTI 261

APPENDIX XI: Data extracts from interview transcripts 262

APPENDIX XII: Conference presentations based on this thesis 277

List of Figures

Figure 1: Classification of test performance using the ROC curve 59
Figure 2: Histogram showing age distribution by patient age group 72
Figure 2a: Age distribution in children (<16 years).....72
Figure 2b: Age distribution of adults (> 16 years).....72
Figure 3: Summary of multivariable analysis model II using forward stepwise selection approach in adults.....91
Figure 4: Summary of multivariable analysis model II using forward stepwise selection approach in children.....94
Figure 5: Comparison of ROC curves of unweighted and weighted CATs scores in children.....120
Figure 6: Comparison of ROC curves of unweighted and weighted CATs scores in adults.....121
Figure 7: ROC curves comparing un-weighted CATs score thresholds in children.....130
Figure 8: ROC curves comparing un-weighted CATs score thresholds in adults.....131
Figure 9: Comparison of ROC curves for weighted combined CATs score thresholds in children 135
Figure 10: Comparison of ROC curves for weighted combined CATs score thresholds in adults 136
Figure 11: ROC curves comparing unweighted combined CATs and CRB65 scores in adults with ALRTI.....151
Figure 12: ROC curves comparing unweighted CATs and CRB65 scores in adults with pneumonia 156
Figure 13: Pattern of CATs criteria selection in children <16years 171
Figure 14: ROC curves comparing unweighted CATs with all 7 criteria versus CATs without criterion G in children.....173
Figure 15: Pattern of CATs criteria selection in adults >16 years 174
Figure 16: ROC curves comparing unweighted CATs with all 7 criteria versus CATs without criterion G in adults.....176

List of Tables

Table 1.1: Clinical prediction tools for acute lower respiratory tract infections.....22

Table 1.2: Children/paediatric community assessment tools 26

Table 1.3: Adult community assessment tools.....27

Table 3.1: Distribution and characteristics of children (< 16years.....70

Table 3.2: Distribution and characteristics of adults (≥ 16years) 71

Table 4.1: Association between CATs criteria and hospital admission...80

Table 4.2: Association between CATs criteria and need for supplemental oxygen 81

Table 4.3: Association between CATs criteria and need for Intravenous antibiotics 82

Table 4.4: Association between CATs criteria and need for Intravenous Fluid.....83

Table 4.5: Association between CATs criteria and In-hospital mortality 84

Table 4.6: Multivariable analysis (model I) to determine independent predictors of acute care interventions and mortality in children 88

Table 4.7: Multivariable analysis (model I) to determine independent predictors of acute ca re interventions and mortality in adults 89

Table 4.8: Predictive values of individual CATs criterion for predicting admission, acute care interventions and in-hospital mortality in children 98

Table 4.9: Predictive values of individual CATs criteria for predicting admission, acute care interventions and in-hospital mortality in adults 101

Table 5.1: Assessment of multi-collinearity between CATs criteria and in-patient mortality.....116

Table 5:2: Logistic regression analyses involving unweighted CATs scoring indexes as continuous variables in adults and children 117

Table 5.3: Comparison of AUROC values of unweighted and weighted CATs scores derived in children 118

Table 5.4: Comparison of AUROC values of unweighted and weighted CATs scores derived in adults 119

Table 5.5: Hosmer-Lemeshow goodness of fit test for assessing the regression models of unweighted and weighted CATs combined scores in children and adults 122

Table 5.6a: Hosmer-Lemeshow test for calibration of un-weighted CATs index in children 123

Table 5.6b: Hosmer-Lemeshow test for calibration of weighted CATs index in children 123

Table 5.7a: Hosmer-Lemeshow test for calibration of un-weighted CATs index in adults 124

Table 5.7b: Hosmer-Lemeshow test for calibration of weighted CATs index in adults 124

Table 5.8: comparison of the predictive performance of individual CATs criteria, un-weighted and weighted CATs combined scores in children 125

Table 5.9: comparison of predictive performance of individual CATs criteria, un-weighted and weighted CATs combined scores in adults... 126

Table 5.10: Predictive performance of combined un-weighted CATs score thresholds for predicting admission, acute care interventions and in-hospital mortality in children 128

Table 5.11: Predictive performance of combined un-weighted CATs score thresholds for predicting admission, acute care interventions and in-hospital mortality in adults 129

Table 5.12: Predictive performance of weighted CATs combined score thresholds in adults 133

Table 5.13: Predictive performance of weighted CATs combined score thresholds in children 134

Table 6.1: Logistic regression analyses involving competing scoring indexes as continuous variables in adults with ALRTI.....148

Table 6.2: Assessing the goodness of fit of the regression models of CRB65 and un-weighted CATs scores in adults with ALRTI 149

Table 6.3: Comparison of AUROC values of CRB65 and un-weighted CATs scores in adults with ALRTI 150

Table 6.4: Hosmer-Lemeshow test for assessing the calibration of crb65 scores in adults with ALRTI 152

Table 6.5: Hosmer-Lemeshow goodness of fit test for calibration of un-weighted CATs index in adults with ALRTI 152

Table 6.6: AUROC (95%CI) values for CRB65 and un-weighted CATs scores in adults with pneumonia 154

Table 6.7: Assessing the goodness of fit of the regression models for CRB65 and CATs scores in adults with pneumonia 154

Table 6.8a: Assessing calibration of CRB scores in adults with pneumonia 155

Table 6.8b: Assessing calibration of Un-weighted CATs scores in adults with pneumonia 155

Table 6.9: Predictive performance of CRB65 score thresholds in adults with ALRTI 158

Table 6.10: Predictive performance of CATs score thresholds in adults with ALRTI 159

Table 7.1: comparison of AUROC values of unweighted CATs with all 7 criteria versus CATs without criterion G in children 172

Table 7.2: comparison of AUROC values of unweighted CATs with all 7 criteria versus CATs without criterion G in adults 175

Table 8.1: Demographics of interview respondents (n=14) 183

List of Abbreviations

ALRTI	Acute Lower Respiratory Tract Infection
ARTI	Acute Respiratory Tract Infection
AUROC	Area under Receiver Operating Curve
CATs	Community Assessment Tools
CI	Confidence Interval
COPD	Chronic Obstructive Pulmonary Disease
CAP	Community Acquired Pneumonia
ED	Emergency Department
FLU-CIN	Pandemic Influenza Clinical Information Network
H-L	Hosmer-Lemeshow
ICU	Intensive Care Unit
LASUTH	Lagos State University Teaching Hospital
LUTH	Lagos University Hospital
MSCH	Massey Street Children's Hospital
NPV	Negative Predictive Value
OR	Odds Ratio
PPV	Positive Predictive Value
PSI	Pneumonia Severity Index
PMEWS	Pandemic Modified Early Warning Score
QUIPS	Quality in Prognostic Studies
ROC	Receiver Operating Curve
TRIPOD	Transparent Reporting of a multivariable Prediction model for Individual Prognosis or Diagnosis
WHO	World Health Organisation

It appears to me a most excellent thing for the physician to cultivate Prognosis; for by foreseeing and foretelling, in the presence of the sick, the present, the past, and the future, and explaining the omissions which patients have been guilty of, he will be the more readily believed to be acquainted with the circumstances of the sick; so that men will have confidence to intrust (sic) themselves to such a physician. …
Thus a man will be the more esteemed to be a good physician, for he will be the better able to treat those aright who can be saved, having long anticipated everything; and by seeing and announcing beforehand those who will live and those who will die, he will thus escape censure.

-Hippocrates, 400 BCE: The Book of Prognostics

1: Introduction

1.1 Thesis background

The timely recognition and assessment of patients with severe illness is a major step towards minimising their risk of developing adverse outcomes, however in times of extreme medical surge such as pandemic events, health care demand often exceeds clinical capacity to provide usual standard of care. Such situations will require the use of carefully selected and clearly defined clinical criteria for making objective assessments of illness severity. When combined using statistical methods, such criteria have been referred to as clinical prediction tools or rules(1). Clinical prediction tools have become increasingly popular in recent times as they not only aid in decision making for hospital triage, but have also been found useful in screening for diseases, designing therapeutic research studies, and developing policies for managing high-risk patients and resource allocation. With the increasing popularity of such tools in clinical practice, it is important that they are properly developed and that their performance is also assessed in terms of accuracy (clinical validity) and impact on clinical practice (clinical utility)(2). According to the literature on predictive research, there are three steps involved in the development of a clinical prediction tool: an initial derivation stage which involves criteria or predictor variable selection using logistic regression analysis and the assessment of the performance of this new model in the derivation population (internal validation); the second stage involves an evaluation of the performance of the tool in a new population (external validation) in order to assess its generalisability ;and finally an impact analysis of the prediction tool's usability and impact on patient management and outcomes(1, 3-9). It has however been reported that many clinical prediction tools are developed without proper

validation and the lack of such evidence presents a risk of poor decisions and patient management(1). Furthermore, prediction tools which are only applicable in the setting in which they were developed hold no value in practice as clinicians are more inclined to adopt a prediction tool with the potential to perform similarly across plausibly related patient groups and geographical or clinical settings.

As the purpose of this thesis is an external validation of a recently developed clinical prediction tool –Community Assessment Tools (CATs) which was proposed for hospital triage of both adults and children with influenza and severe respiratory illnesses, this opening chapter will provide a compendious discussion on the evolution and development of clinical prediction tools. Also included, is a section on some widely used prediction tools for severe respiratory tract illnesses including a discussion on the development of the CATs. In order to contextualise this thesis to the geographical setting in which the research was conducted an overview of the burden and management of ALRTI in Nigeria has also been included. This chapter finally concludes with a thesis rationale and outline summarising the content of the subsequent chapters.

1.2 The evolution of prognostication in medicine

According to the history of medicine, the science and practice of providing care and preventing death is based on a trichotomous relationship between diagnosis; prognostication and treatment(10). The writings of historical physicians such as Hippocrates and Galen reveal that prognostication which involves predicting or foretelling the course or outcome of an illness was the strongest of the three principles of medicine(11). However, medical and scientific breakthroughs brought

about improved diagnosis and treatment which led to a paradigm shift and the decline of prognostication(12). It has been suggested that the reasons for the decline in prognostication was that it was believed to be unscientific(13); it was not until the twentieth century when a string of articles on prognostication were published in the Lancet(10), that opinions started to change and the scientification of prognostication began with the development of the first prognostic index for estimating mortality from myocardial infraction in 1953(14). Over the next couple of decades a renewed interest in prognostication emerged, such that the number of publications on developing prognostic indexes for quantifying mortality risk and other management outcomes was seen to rise steadily. Clinicians are now faced with the difficulty of deciding which tool is best from the pool of existing tools, especially as most of these tools report varying levels of predictive accuracy in different patient populations. The next section will provide a detailed description of the stages involved in the development of clinical prediction tools.

1.3 Consecutive phases of developing a prediction tool

The importance of accurate medical decision making cannot be overemphasised; by applying the statistical approach of risk and probability estimation, clinical prediction tools provide a solution to some of the subjectivities and uncertainties of medical decision making(3). With the growing popularity of prediction research, there has been an increase in the number of clinical prediction tools such that a specific clinical domain might have several different tools for the same purpose(15). For clinicians to confidently adopt a prediction tool for use in practice, it must perform similarly across different patient demographics and clinical settings(16). In predictive research this is referred to as internal and

external validity; the internal validity of a tool is when its performance is in accordance with the purpose for which it was designed within the derivation population, while external validity is when its performance is generalisable to an entirely different group or population (5, 17). It is therefore of utmost importance that a prediction tool should possess both internal and external validity, as this is proof of its generalisability across different demographics (17). Moons and colleagues (6-9) have extensively discussed the principles of predictive research in a series of four articles published in the British Medical Journal, highlighting three stages for developing prediction tools which are described below:

1.3.1 Step 1: Derivation

This stage involves identifying relevant predictors from pre-selected variables identified from clinical evidence and subject knowledge, followed by model selection using multivariable logistic regression analysis(6). Two approaches have been suggested for model selection; a full model approach where all predictor variables are included and a stepwise regression approach using either backward elimination or forward selection(7). These approaches have their pros and cons, and there is no consensus on which method is best; the full model approach may avoid selection bias, but it requires extensive knowledge of the subject area and including too many predictors in the final model would be impractical for use. On the other hand, either backward elimination or forward selection may result in a more parsimonious model by removing redundant variables based on specified significance levels, however the risk of excluding clinically relevant predictors is high. Next, it is important to assess the predictive performance of the model in terms of its discrimination and calibration qualities in the development data or population. Put simply, discrimination is the ability of the model or tool to

accurately distinguish between people who are or are not at risk of an event, while calibration is the probability that the predicted events match the actual observed events. The issue with the performance of the new prediction tool/model at this stage is that of optimism, which means that such a model is usually found to have very good predictive accuracy when tested in the development data, which may not be the case when tested in new data. Optimism is caused by overfitting i.e. the prediction model was designed to fit the development data(5). Optimism is often investigated by randomly splitting the data into two such that part of the data becomes a development sample and the other a validation sample for comparison, or by using bootstrapping– this process is referred to as internal validation(7).

1.3.2 Step 2: Validation (external) and updating

As previously stated, a model which is only applicable in one setting will not be considered very useful in practice, which is why the next stage in the development of a prediction model is an external validation in a new and different population(8). The difference between internal and external validation is that the former uses the same development data to assess the performance of the model while external validation is conducted using entirely different data(17). There are three types of external validation: temporal; geographical and domain validation (15, 17). Temporal validation as the name suggests relates to time and involves using data similar to the derivation cohort but collected at a different time; although some degree of variation may exist, this new sample population still shares similar characteristics with the derivation cohort. Geographical validation on the other hand assesses the transportability of the prediction model by assessing its

performance in a different country or location thus allowing for variations in case-mix and measurement of the predictors and outcomes. In the third form of external validation -domain validation, the performance of the prediction model is evaluated in a new clinical setting i.e. primary versus secondary care or adults versus children. In order of stringency, domain validation ranks highest while temporal validation is the least rigid of the three validation methods(17). In most cases, prediction models tend to perform poorly when externally validated, which is understandable given the existing variations between the development and validation data in terms of sample size, case-mix and measurements of the outcome and predictor variables(8, 17). When this occurs, model adjustment is either attempted in the new data using recalibration (adjustments of model's intercept) or model revision (addition of new predictors or adjustment of regression coefficients) or the model is discarded and a new one is developed(18). Model adjustment is usually the preferred solution, as discarding the old rule to develop new ones only results in an illogical amount of prediction rules studies which are never used.

1.3.3 Step 3: Impact analysis

Prediction tools are developed with an aim to guide clinical decisions towards improving management outcomes. A tool which does not influence practice or result in improved outcomes has no value, as such it is important to assess the impact of using or not using the tool in terms of cost effectiveness, patient outcomes and influencing decisions made by the clinician. Impact analysis are usually conducted as randomised clinical trials, or as before and after studies(9).

1.4 Risk prediction for acute lower respiratory tract infections

Acute respiratory tract infections are the most prevalent indications for hospital attendance in both developed and developing countries irrespective of age. The clinical manifestation of acute respiratory tract infections could range from minor self-limiting viral illnesses to severe and life-threatening complicated bacterial infections; as such, decisions regarding site of care and extent of investigation and treatment should be based on illness severity (19, 20). Quite a number of prediction tools have been developed for this purpose and with pneumonia being the most common ALRTI and most associated with mortality, it is of no surprise that the majority of clinical prediction rules for respiratory illness were developed for the severity assessment of pneumonia (21, 22). A summary and description of the most established and validated prediction rules for pneumonia are presented in Table 1.1.

Influenza is another common acute respiratory tract infection resulting in severe illness and increased hospital attendance. Influenza is a viral illness which typically occurs in seasonal epidemics during the winter season or unexpectedly as a pandemic outbreak. Although usually self-limiting, complicated influenza may result in patients presenting with pneumonia, infective exacerbations of chronic pulmonary illnesses such as COPD and respiratory failure. Although there are limited epidemiological data on the seasonality of influenza in the tropics, there is still some evidence of a year-round occurrence in these parts of the world (23, 24). During an influenza pandemic, demand for healthcare usually outweighs available resources and clinical staff, such that the capacity to provide normal standard of care is exceeded. Such situations will require the use of triage tools to distinguish between high risk patients in need of prompt hospital referral and low risk patients who can be safely managed out of hospital(25). In spite of the

fact that a number of tools have been developed for use in influenza pandemics, the emergence of novel viruses and the virulence of such new virus strains make it quite difficult to develop prediction tools for future outbreaks (26, 27). In the absence of well validated influenza prediction tools, previous pandemic preparedness guidance plans in the United Kingdom(28) adopted the CURB-65 (21, 29)and the PMEWS(30) for use in the event of an influenza pandemic. It is likely that these tools were chosen because influenza is commonly complicated by pneumonia and vice versa, however the major drawback of some these tools is that they were originally developed for use in adult patients. The need for age-appropriate tools which are able to predict the need for clinical interventions and critical care services was highlighted in the 2011 UK influenza pandemic preparedness strategy which reported severe morbidity in younger adults and children during the recent 2009 influenza pandemic(31). In a bid to bridge the existing age-gap in the prediction tools previously recommended for triaging patients with influenza, the Community Assessment Tools (CATs) was developed as an age-appropriate hospital triage tool for use in both adults and children in times of extreme medical surge(32). The CATs which is made up of six objective and one subjective criteria is yet to be used, however its performance has been assessed using data collected in the United Kingdom during the first and second waves of 2009 H1N1 influenza pandemic. The next section provides historical context with regard to the development of the CATs and discusses findings from the UK validation studies.

Table 1.1: clinical prediction tools for acute lower respiratory tract infections

Prediction tool	Clinical context & criteria	Outcome	Author (reference)
CURB -65 **CRB-65**	6 point scoring system or 5 point in the case of CRB65 for stratifying adults with community acquired pneumonia into 3 management groups based on mortality risk. **C**-confusion, **U**- urea>7mmol/l, **R**-respiratory rate ≥30cycles/min, **B**- systolic blood pressure <90mmHg or diastolic <60mmHg, age**≥65 years**	30- day mortality	Lim et al. (2003) (21)
Pneumonia Severity Index (PSI)	Consists of 20 variables consisting of demographic variables (age, sex and nursing home residence), existing comorbid illness, physical examination findings (pulse rate, systolic blood pressure, temperature, respiratory rate, mental status) and laboratory investigation findings. Categorises patients with community acquired pneumonia into five risk groups (I-V)	30- day mortality	Fine et al. (1997) (22)
Pandemic Medical Early Warning Score (PMEWS)	Modification of a simple physiological scoring system (MEWS) used for in-hospital assessment of patients to include oxygen saturation measurements, age ≥65 years, social isolation, existing comorbidity and performance status. Intended for use in severity assessment of patients with influenza	Hospital admission and the need for higher levels of care	Challen et al. (2007) (30)
American Thoracic Society Criteria	Consists of 10 criteria -6 baseline clinical parameters available at admission and 4 parameters assessed during admission for assessing patients with severe community acquired pneumonia defined as the need for ICU admission.	ICU admission	Ewig et al. (1998)(33)

1.5 The Community Assessment Tools (CATs)

CATs was proposed by a group at Alder Hey Children's Hospital as a package of care to work in sync with National Health Service hospital pathways (34, 35). The CATs consists of 7 (6 objective and 1 subjective) criteria with physiological thresholds specific to adults and children <16 years; meeting any one or more of these criteria was expected to trigger hospital referral or admission(36). Detailed definitions for each of the assessment criteria within the paediatric and adult CATs are provided in Tables 1.2 and 1.3. The CATs criteria are yet to be used in routine clinical practice as they were developed for use in 'severe and exceptional circumstances when surge demand calls for strict triage' (35). Semple at al.(2013) conducted an evaluation of the CATs criteria using data from the FLU-CIN cohort (35), FLU-CIN was an emergency sentinel surveillance system set up by the Department of health England to record clinical, epidemiological and outcome data during the 2009 pandemic from patients hospitalised with PCR confirmed A(H1N1)pdm09 flu (37). All seven criteria were associated with the need for higher levels of care, hospital admission and death in the univariate analysis involving adults and children below the age of 16 and none of the criteria remained redundant in the multivariable analysis thus justifying the choice of criteria and also the potential for creating a composite scoring system (35). The performance of the CATs criteria as a combined score has been examined in three studies; Myles et al. (2012)(36) and Goodacre et al. (2010)(38) compared the performance of a CATs score which was calculated by awarding each of the seven criteria a single point on an unweighted basis with two well established scoring systems – CURB65 and PMEWS. The third study which was a masters dissertation

compared the performance of the unweighted CATs score with a new score calculated using weights based on the regression coefficients derived from the multivariable logistic regression analysis including all seven criteria(39). In the comparative validation study conducted using the UK FLU-CIN data, Myles et al. (2012) reported that of the three scores, the unweighted CATs score had the best performance for predicting the need for higher levels of care and mortality in both children and adults(36). The performance of the unweighted CATs for predicting death in adults was similar to that of the CURB-65 which is an established mortality predictor in adults with pneumonia. Goodacre and colleagues compared the unweighted CATs, CURB-65 and PMEWS scores using data prospectively collected data from 481 cases across three hospitals in the UK during the first wave of the 2009 influenza pandemic (38). These authors reported that the three scores were unable to discriminate between patients with severe illness who were at risk of poor outcomes such as death within 30 days of follow up and those who less ill and could be managed out of hospital. These findings may have been affected by the mild nature of the 2009 pandemic, as there were very few deaths and their study was limited by insufficient data (38). The third study which was a dissertation conducted in the division of epidemiology and public health at the University of Nottingham by a master's student involved a comparison of the unweighted CATs score and with a new CATs score calculated using weights based on the regression coefficients obtained from the multivariable logistic regression analysis including all seven criteria. It was concluded from this study that the CATs scores performed better when used as combined scores than as stand-alone criteria, with the unweighted CATs been the better of the two(39). A recent study 'the FLU-CATs study' which was conducted by Venkatesan et al. used UK

primary care data to assess the validity and utility of the CATs criteria and the Pandemic Medical Early Warning score (PMEWS) which is a pandemic hospital admission triage tool(40). In addition to comparing the two triage tools, this study also intended to develop surveillance systems which could capture and allow near real time analysis of data on the clinical assessment of adults and children presenting with influenza-like illness in primary care. One of the findings from this study was that many clinical parameters were not routinely measured by GPs, such that there were a limited number of participants with complete data for PMEWS, but the CATs criteria were recorded for all participants. Of the 863 (704 adults and 159 children) study participants, only 17 people were referred to hospital and only 15 deaths were recorded, as such it was not possible to conduct inferential analyses on the CATs and PMEWS(40) .

As these studies have all been conducted using data collected in the United Kingdom, the performance of the CATs criteria and scores are yet to be validated or used in a different geographical setting, or in patients with other respiratory conditions with similar presentation to influenza. Furthermore, the limitations of using retrospective data in some of these studies are worth noting; the FLU-CIN data were based on information extracted from case notes and physicians' clinical assessments which were not always complete and the possibility of bias in the classification and measurement of the study variables could not be ruled out. However these limitations can be overcome by using prospectively collected study-specific data in an external validation study.

Table 1.2: Children/Paediatric Community Assessment Tools

Criteria labels	Criteria Definitions
A	SEVERE RESPIRATORY DISTRESS Lower chest wall indrawing, sternal recession, grunting/ noisy breathing when calm.
B	INCREASED RESPIRATORY RATE (measured over at least 30 seconds) ≥50 breaths/minute if under 1 year, or ≥40 breaths/minute if ≥1 year.
C	OXYGEN SATURATION ≤92% ON PULSE OXIMETRY, BREATHING AIR/ OXYGEN Absence of cyanosis is a poor discriminator for sever illness
D	RESPIRATORY EXHAUSTION OR APNOEIC EPISODE Apnoea defined as ≥20 second pause in breathing
E	EVIDENCE OF SEVERE CLINCAL DEHYDRATION OR CLINICAL SHOCK Sternal capillary refill time >2 seconds, reduced skin turgor, sunken eyes/ fontanelle.
F	ALTERED CONSCIOUS LEVEL Strikingly agitated or irritable, seizures or floppy infant
G	CAUSING OTHER CLINCAL CONCERN TO THEIR OWN GP OR CLINICAL TEAM E.g. a rapidly progressive or unusually prolonged illness.

Table 1.3: Adult Community Assessment Tools

Criteria labels	Criteria Definitions
A	SEVERE RESPIRATORY DISTRESS Severe breathlessness, e.g. unable to complete sentences in one breath. Use of accessory muscles, supra-clavicular recession, tracheal tug or feeling of suffocation.
B	INCREASED RESPIRATORY RATE (measured over at least 30 seconds) Over 30 breaths per minute.
C	OXYGEN SATURATION <92% ON PULSE OXIMETRY, BREATHING AIR/ OXYGEN Absence of cyanosis is a poor discriminator for sever illness
D	RESPIRATORY EXHAUSTION New abnormal breathing pattern, e.g. alternating fast and slow rate or long pauses between breaths.
E	EVIDENCE OF SEVERE CLINCAL DEHYDRATION OR CLINICAL SHOCK Systolic blood pressure <90mmHg and/or diastolic blood pressure <60mmHg Sternal capillary refill time >2 seconds, reduced skin turgor.
F	ALTERED CONSCIOUS LEVEL New confusion, striking agitation or seizures
G	CAUSING OTHER CLINCAL CONCERN TO THEIR OWN GP OR CLINICAL TEAM E.g. a rapidly progressive or unusually prolonged illness.

1.6 An overview of acute lower respiratory tract infections in Nigeria

According to the WHO global burden of disease statistics for 2012, acute lower respiratory tract infections rank as one of the top three leading causes of death across all age groups globally and the major cause of death in low income countries(41). The actual burden of acute lower respiratory tract infections in terms of incidence and prevalence is a poorly researched topic in most developing countries especially in Nigeria(42); Cooper et al. (1998) conducted a study on disease burden in Sub-Saharan Africa and highlighted the absence of basic population denominators in these countries which make it impossible to actually calculate important health statistics(43). Furthermore a WHO review on studies reporting incidence of ALRTIs in low income countries has also revealed that only a few studies met the minimum methodological requirements(44). Therefore it can be inferred that the available figures with regards to disease burden for most sub-Saharan countries including Nigeria might actually be 'guesstimates'. Notwithstanding, recent WHO evidence on the epidemiology of ALRTIs in children shows that there are over 150 million new cases annually in developing countries, with approximately 13% severe enough to need hospital admission(45). Although the epidemiology of ALRTIs in adults in this part of the world is not as well documented, the estimated burden of disease in adults is thought to mirror that in children.

According to country health statistics by the WHO, ALRTIs were the leading cause of mortality accounting for almost 300 thousand deaths in the year 2012(46) and in 2013 about 34% of children less than 5 years were reported to have presented to a health facility with symptoms of an acute lower

respiratory tract infection (47).These figures suggest that acute lower respiratory tract infections are an immense burden on the already sparsely available medical resources and manpower which presents a suitable setting for conducting an external validation of the Community assessment tools. The common

1.7 Management of ALRTI in Nigeria

To a large extent the management of ALRTI is dependent on the severity of the illness, as high risk or severely ill patients would require hospital based care and interventions, while less severely ill patients could be managed out of hospital on oral therapy(48). Management of Acute lower respiratory tract infections is also determined by aetiology which could either be bacterial or viral. Viral infections are known to cause acute bronchitis, influenza and exacerbation of chronic obstructive pulmonary disease (COPD)(49). Such infections are usually self-limiting requiring supportive measures such as intravenous fluids and supplemental oxygen administration. Influenza is the only viral ALRTI for which there is licensed antiviral therapy (oseltamivir - Tamiflu and Zanamivir), however, the high cost of these medications limits their use in developing countries(50). The aetiology of pneumonia could be either viral or bacterial but mostly bacterial, requiring the use of oral or intravenous antibacterial medications depending on severity. COPD may be also be exacerbated by bacterial infections and treatment in such cases would involve the use of antibiotics. The main issue in making a diagnosis of ALRTI is distinguishing between pneumonia and other ALRTI due to the overlap in their presentations. The established the gold standard for making a diagnosis of pneumonia is the use of radiographic evidence, however chest x-rays are not

typically performed in the community and such investigations may be not be readily available or are too expensive to be used routinely in resource-poor settings. British Thoracic Society and World Health Organisation guidelines recommend the use of clinical features for diagnosing ALRTI in the settings where chest x-rays are not readily available such as in the community or in low and middle income countries(51, 52).

Despite differences in patient characteristics, available resources and local hospital guidelines, it is expected that the same universal principle is followed for the management of ALRTI irrespective of geographical setting. In developing countries including Nigeria where healthcare is funded by out-of-pocket payments, ALRTI and the pneumonia are usually diagnosed based on clinical features, furthermore antibiotics are the most commonly prescribed medications for managing ALRTI as several studies conducted in Nigeria have reported that the bacteria are the major pathogens causing ALRTI. Also, evidence on the epidemiology of influenza viruses is lacking in Nigeria, which may be due to the seasonality of such infections; as such antivirals are not commonly used in the management of ALRTI. It is also worth mentioning at this point that ICU admissions of patients with severe ALRTI is not well documented in studies conducted in Nigeria, and this has implications for the use of a clinical prediction tool for making decisions regarding the need for ICU admission and mechanical ventilation in patients with ALRTI.

1.8 Thesis Rationale and Aims

The initial premise for using CATs, was that each criterion could serve as a stand-alone trigger for urgent hospital referral and admission of adults and children with severe respiratory illnesses. However the use of stand-alone criterion as triggers for hospital care is not without certain resource allocation implications especially in resource limited settings and during periods of extreme medical surge. When combined into a score the CATs offers an opportunity for possible risk stratification of patients into high and low risk groups which will guide in choosing appropriate interventions for care and foster more efficient use of limited resources. The performance of the CATs criteria and its combined unweighted and weighted scores have been investigated in the United Kingdom, however evidence of its generalisability to a non-UK population is required for recommending its use in practice and to provide direction for further validation or updating if required. The CATs includes a subjective criterion 'causing concern to the clinician' which has been shown to be associated with outcomes studied in the CATs validation studies conducted in the UK. In externally validating the CATs, it will also be worth investigating the performance of this outcome particularly its contribution to the performance of CATs as a combined score –if it retains its validity or not and if it remains relevant when the CATs score is applied in new patients and in a different geographical setting.

Furthermore, there is a paucity of evidence on the validation and use of clinical prediction tools or severity scores in low and middle income countries, Nigeria included. Majority of the existing severity scores are developed and validated in high income countries with no evidence of their applicability in other settings. Resource-limited countries with a high burden of ALRTI stand to gain from the

use of such clinical prediction tools. The findings from thesis will therefore generate evidence for recommending the use of a clinical prediction tool for the management of ALRTI in Nigeria.

The primary aim of this thesis therefore, was to assess the external validity of the Community Assessment Tools (criteria and combined scores) for the severity assessment of adults ($\geq$ 16years) and children with acute lower respiratory tract infections (ALRTI) in an entirely new geographical and clinical setting. This thesis which included five studies was conducted in Nigeria – a resource-poor country with a high ALRTI burden, and these studies specifically investigated in turn:

1. The performance of the individual CATs criterion for predicting the need for hospital admission, acute care interventions such as supplemental oxygen, intravenous antibiotics or fluids and the risk of in-hospital mortality in adults and children with ALRTI.
2. The performance of the CATs as un-weighted or weighted combined scoring indexes for predicting the need for hospital admission, acute care interventions and risk of in-hospital mortality in adults and children with ALRTI.
3. The predictive performance of the un-weighted combined CATs score in comparison with a well validated pneumonia severity assessment tool – the CRB65 score; first in adults with ALRTI and then specifically in those diagnosed with pneumonia.
4. The benefit if any, of considering clinicians' gut feelings or intuition as part of the unweighted CATs score for the severity assessment of adults and children with ALRTI.

5. Critical care practice and its utilization in Nigerian hospitals by interviewing doctors and nurses from the participating hospitals.

While the first four studies were conducted using a quantitative approach, the fifth was conducted as a qualitative study to provide contextual insights into the provision of critical care in Nigeria in relation to the possible use of CATs for predicting ICU admissions in this setting.

1.9 Thesis outline

As mentioned earlier, this thesis is made up of five linked studies which have each been presented as individual chapters and are outlined below:

Chapter 2: describes the general methodology and methods used in this thesis.

Chapter 3: describes the data for this thesis in terms of the demographic characteristics of the participants.

Chapter 4: considers the first objective of this thesis which evaluates the CATs based on the individual performance of each criterion.

Chapter 5: focuses on the second objective and describes the creation and comparison of the unweighted and weighted CATs scores or indexes in adults and children with ALRTI.

Chapter 6: deals with the third objective and compares the predictive performance of the unweighted CATs and the CRB65 scores in adults with ALRTI and pneumonia separately.

Chapter 7: examines the role of clinical judgement or intuition in the severity assessment of ALRTI cases by comparing the performance of the unweighted CATs score with and without its subjective criterion.

Chapter 8: focuses on the fifth and final objective which is the qualitative study examining critical care practices in Nigeria.

Chapter 9: provides an overall summary of all the findings from this thesis in line with the aims and objectives; it also considers the implications of these findings for practice and offers recommendations for further research.

2: Research methodology and Methods

2.1 Introduction

This chapter is presented in two parts; the first will focus on the quantitative aspect of this thesis, presenting a comprehensive description of the methods and statistical analyses used in the first four studies. The second part of this chapter discusses the philosophical framework within which the qualitative study on critical care practice was undertaken and also describes the methods used for data collection and analysis.

2.2 Overall ethics approval for thesis

The protocol for this thesis was approved by the University of Nottingham Medical school ethics committee prior to commencement of field work in Nigeria. Ethics approval for data collection was also obtained from the health research and ethics committee of Lagos university teaching hospital (LUTH), Lagos state university teaching hospital health research ethics committee (LASUTH) and the health service commission Lagos state (RGH and MSCH). See appendix X for copies of ethics approval documents. All ethics approval letters have been included in the appendix section.

2.3 Theoretical underpinning for the methodology of the external validation of the CATs

Consistency in the conduct and reporting of prediction tool studies is imperative for assessing the validity and the potential application of such tools in practice. The increasing popularity in the use of prediction tools for informing decisions within clinical practice calls for an improvement in the methodological quality of prognostic factor studies (53, 54). The majority of published studies on prediction tools fail to adhere to an established principle in terms of design, statistical analysis and reporting (1, 55, 56). Observational studies with the most reliable results are those which have been designed as much as possible to mimic the exemplary methodological standards of randomised controlled trials(57) Various authors in a bid to improve on the quality of prognostic factor research have proposed frameworks for quality assessment (58-61). Hayden et al. (2013) developed the 'Quality in Prognostic Studies (QUIPS), Tool for assessing six biases ; 'participation, attrition, prognostic factor measurement, outcome measurement, confounding, statistical analysis and reporting of results' which could affect the quality of prognostic factor studies (59). A more recent guideline for reporting of prognostic factor studies which covers all aspects of reporting from the title and abstract of the study to the discussion of findings was developed in 2015 by Collins et al. (61); the TRIPOD (Transparent Reporting of a multivariable Prediction model for Individual Prognosis Or Diagnosis) statement is basically a 22 item checklist which is applicable to both development and validation (internal and external) multivariable prediction model studies (61). The methodology for this validation study will be informed by the checklist from the TRIPOD statement as it offers a concise, systematic and transparent approach for reporting

prognosis research(54). This chapter will be reported in line with recommendations made in the methods section of the TRIPOD checklist in Appendix I.

2.4 Study design

This study was conducted using a prospective cohort design which is regarded as the most suitable observational study design for estimating the risk of outcomes in a chosen population and arguably the best design for prognostic model studies (6, 54). Although retrospective and case-cohort designs are also commonly used in prognostic model studies, as they provide a nominal avenue for obtaining quick results using readily available data (6, 55), they are not without important shortcomings which ultimately compromise the validity of their results(62). Missing data, though a common problem in most observational studies, remains an inescapable disadvantage of using retrospectively collected data invariably leading to biased estimates(63). While potential confounders which might be of interest in the index study can be easily collected using a prospective cohort design, such variables might not be readily available using retrospectively collected data. In addition, although retrospectively collected data may present a wide variety of potentially useful information, certain study-specific variables might be unavailable and the abundance of variables in such data sets could result in possible data dredging, as researchers might be tempted to test for extemporary hypotheses. It is also more desirable that an 'inception cohort' is identified prospectively, as this will ensure accuracy in collection and measurement of study variables as well as allow for measurement of other additional factors which could confound associations or predictions (54, 64).

2.5 Setting

A multicentre study involving secondary and tertiary care hospitals in Lagos, Nigeria, was conducted to assess the CAT's generalisability. The decision to conduct the study in this region was based on the fact that Lagos with a population of 21 million people is the largest city in Nigeria as well as Africa (65). Lagos is a metropolitan city as well as a major cultural and economic hub being home to indigenes from almost all ethnic groups in Nigeria thus ensuring the generalisability of our sample to the entire Nigerian population. Healthcare in Nigeria is provided predominantly via the public (government owned) hospitals, although there has been a steady growing contribution from the private (non-government owned) healthcare market since the mid-80s(66). The Nigerian public healthcare sector is a decentralised system which is run by the three tiers of government in Nigeria. The Federal Ministry of Health which essentially provides all health policy guidance oversees the activities of the tertiary care hospitals, while the State Ministry and Local Government Health Department run the secondary or general hospitals and primary health care centres respectively(67). Following a bottom-up approach, the Primary health centres (PHC) act as the entry point into the health system providing both basic preventive and curative services as well as referrals; the General hospitals act as the intermediary between the lowest and highest levels of care and offer both out and in-patient services; while the highest and specialised level of care is provided by the tertiary centres(68). Though it may appear as if a structured referral pattern and division of labour exists within the Nigeria healthcare system, overlap and confusion between the primary and secondary levels in terms service delivery and their respective roles and responsibilities have been reported (69).

The Federal government provides at least one tertiary care hospital in every state in Nigeria; Lagos state is an exception having two tertiary care hospitals, Lagos University Teaching Hospital (LUTH) Idiaraba and Lagos State University Teaching Hospital (LASUTH) Ikeja which were both chosen for this study. These tertiary care hospitals were chosen because though Nigeria operates a decentralised healthcare system, most patients seeking healthcare directly access the tertiary care centres even for illnesses which are usually considered the remit of the primary and sometimes secondary care centres(70). Several explanations have been offered for this behaviour, the most common being a shortage of accessible PHC facilities and even when accessible, the inadequate staffing and financing of these PHCs results in a lack of confidence among people who resort to visiting the tertiary centres in the expectation of better care from specialist doctors (70, 71).

After three months of data collection, it was imperative to include secondary care hospitals as the anticipated throughput of patients at the tertiary centres was not being achieved. This was possibly due to the recurrent strike actions by the association of resident doctors during the data collection period which crippled all hospital activities including admission of patients from emergency departments. The two secondary care hospitals were strategically chosen based on the geography of Lagos state. The state is divided into two geographical regions; a group of islands and the mainland connected by three bridges, one of the two secondary care hospitals was located on the island while the other was on the mainland. A concise description of each hospital is presented in the next section.

2.6 Study centres

2.6.1. Lagos University Teaching Hospital (LUTH)

LUTH is a federal government owned tertiary hospital located in the mainland region of Lagos state; it is classified as a tertiary health centre based on its capacity to provide specialised medical care and also function as a university teaching hospital to the College of Medicine of the University of Lagos. LUTH is the largest teaching Hospital in Nigeria with a capacity of 761 beds, 17 operating theatres and a 7 bed ICU; it also has the highest number of specialist doctors in any Nigerian hospital.

2.6.2 Lagos State University Teaching Hospital (LASUTH)

LASUTH is the second tertiary health centre in this study, but unlike LUTH, it is owned and managed by the Lagos state government. Also located on the mainland of Lagos, LASUTH was initially a general (secondary care) hospital serving Ikeja the capital of Lagos state and its environs, before it was converted to a 600 bed tertiary health care facility with a 5 bed ICU in the year 2001. In terms of structure, LASUTH and LUTH are quite similar; as they both provide postgraduate and undergraduate medical training, and patients presenting to these hospital are either self-referred or referred from other hospitals (private or government owned).

2.6.3 Massey Street Children's Hospital (MSCH)

Massey Street hospital is a Lagos state government owned specialised service hospital; it is a secondary care facility located in a densely populated area in Lagos Island and it is the oldest children's hospital in the West African sub

region. Though classed as a secondary care hospital, ''Massey' as it is popularly called, functions in the capacity of a primary as well as tertiary level facility providing paediatric training to both medical students and resident doctors. Having both in-patient and out-patient departments, Massey receives self-referred patients and patients from other health facilities around metropolitan Lagos. This hospital does not have an Intensive care unit and refers all critically ill patients to the two tertiary care hospitals in Lagos (LUTH and LASUTH).

2.6.4 Surulere General Hospital and Randle Maternity and Child Care Centre

Surulere or Randle general hospital is a secondary care facility also owned and managed by the Lagos state government health management board. Surulere is a commercial and residential area on Lagos mainland, with a population of over 500, 000 people This hospital provides both in and out-patient medical and surgical services at the main site while the newly built Maternal and Child centre at Gbaja provides paediatric emergency and ward services. Surulere General Hospital has over 150 beds and serves as referral centre for primary health centres and other private hospitals in Surulere district, but refers critically ill patients to the nearby tertiary care centre (LUTH) as it does not have an ICU.

2.7 Study participants

Adults and children presenting to the emergency departments of the four study hospitals whose symptoms met the study inclusion criteria were consecutively recruited between December 2013 and July 2014. Although this recruitment

strategy is a form of non-probability sampling and therefore prone to selection bias, randomisation, which is the core of probability sampling might have been impractical for this study. As the most superior form of non-probability sampling techniques, consecutive sampling improved the representativeness of this thesis's study population by recruiting every possible patient based on carefully selected inclusion and exclusion criteria based on the essential characteristics and typical presenting pattern of patients of all ages with ALRTI (44, 72, 73). The geography of Lagos was also considered in choosing the study centres so as to avoid any differences that might exist between patients domiciled in either of the two zones (Lagos Island and Mainland).

2.7.1 Inclusion criteria

The case definition of acute lower respiratory tract infections in literature has remained an area of debate as there appears to be a lack of a single consistent label in the definitions adopted (73). Although patients with ALRTI tend to present with a variety of physical signs and symptoms; cough and at least one other lower respiratory tract symptom with abnormalities on auscultation are almost always present when making a diagnosis of acute lower respiratory tract infection (44, 73). It has also been reported that making a distinction between pneumonia and other types of acute lower respiratory tract infections will require radiological evidence of chest infiltrates, however the use of extensive investigations to confirm or eliminate pneumonia in every patient who presents with cough and/or difficulty in breathing is neither cost-effective nor practical, especially in developing countries where resources are limited or expensive(74).

As mentioned earlier in the introductory chapter, ALRTI in Nigeria are usually diagnosed on the basis of clinical features as such the case definition of ALRTI for inclusion of patients in this thesis was based on clinical features (44). The following inclusion criteria are based on the WHO- CHERG recommendation for ALRTI case definition in children <5 years:

- Cough or breathing difficulty with at least one of the other signs below,
- increased respiratory rate or fast breathing of $\geq$60 breaths per minute in children less than 2 months of age, $\geq$50 breaths per minute in children age between 2 to 11 months and $\geq$40 breaths per minute between the ages of 1 to 5 years
- Evidence of lower chest in-drawing observed at rest as 'inward movement of the bony structures but not soft tissues of the lower chest wall during inspiration'.

The same inconsistencies exist with regards to the case definition of ALRTI in adults(75); as a result, this thesis adopted an approach based on a review by Macfarlane(76) which states that the clinical definition of ALRTI should include

- cough productive or non- productive of sputum and one other lower respiratory tract symptom such as dyspnoea/difficulty in breathing, tachypnoea/ fast breathing, chest discomfort/pleural pain
- abnormal chest signs on auscultation
- at least one systemic feature (i.e. a temperature of 38 degrees or more, chills, malaise)
- no other explanation for the illness

2.7.2 Exclusion criteria

Due to the overlap in the symptoms of asthma and bronchiolitis, children who presented with just wheezing or a diagnosis of asthma without other signs of lower respiratory tract illness as per the specified inclusion criteria were excluded from the study. Patients who had recently been hospitalised in the 14 days preceding index presentation were also excluded as this would suggest a hospital acquired infection.

2.7.3 Follow-up of patients

Patients who were admitted from the emergency department were followed up till discharge or death in hospital. There were few cases where patients self-discharged either as a result of inability to continue paying for their stay in hospital or because of the strike actions by doctors which occurred during the course of data collection. Although few in number, these patients were excluded from the data as it was not possible to ascertain recovery or death and additional follow-up was beyond the scope of the study.

2.8 Sample size calculation

The validation of a prognostic model involves determination of its predictive accuracy, usually done by calculating the Area under receiver operating characteristic curve (AUROC). A detailed approach for calculating the sample size of studies involving the use of AUROC to determine predictive accuracy has been described by Obuchowski et al. (2004)(77). Using this approach,

sample sizes for objectives 3 and 4 using the outcome variable mortality are calculated below:

1. Test hypothesis: The area under the ROC curve exceeds 0.5 i.e. the unweighted combined CATs index has the ability to discriminate between patients who died (cases) and those who survived (controls). The following formula was applied:

$$n_D = \frac{[\, Z_\alpha \sqrt{0.0792 \times (1 \times 1/\kappa) + Z_\beta \sqrt{V(\hat{\theta})}]^2}}{(\theta - 0.05)^2}$$

Where $V(\hat{\theta})$ is the variance function of $\hat{\theta}$

The following assumptions have been used:

- $k = 4$[1] (i.e. 4 times as many survivors as compared to deaths)
- The type I error set at 0.05
- The type II error rate of 0.10 (power = 0.90)
- Conjectured accuracy of CATs as a predictive tool (θ) using area under the ROC curve as a measure of accuracy = 0.70
- One sided test

The estimated sample size according to these calculations is 84 survivors (controls) and 21 deaths (cases). If the conjectured accuracy was changed to an AUROC of 0.90 (high accuracy), while leaving all other parameters unchanged, the estimated sample size would be 40 survivors (controls) and 10 deaths (cases).

2. Test hypothesis: The area under the ROC curve obtained for the un-weighted combined CATs and CRB65 with respect to mortality are significantly different. The tables provided in Obuchowski et al. (2000)(78) for detecting a difference in accuracy between two diagnostic tests using area under ROC curve have

been used in estimating the required sample size using the following conservative assumptions(78):

- Up to 4 observers with moderate intra-observer (0.0025) and inter-observer (0.05) differences in applying the CATs criteria
- Moderate accuracy of each of the predictive tools being compared i.e. AUROCs in the range of 0.75
- Small differences between the accuracy of the two predictive tools i.e. an absolute difference between the two AUROCs of 0.05
- Ratio of survivors/deaths (i.e. controls/cases) of 4
- The type I error set at 0.05
- The type II error set 0.20 (power =0.80)
- Two tailed test

The estimated sample size calculated for this study was: 983 patients i.e. approximately 788 survivors (controls) and 197 deaths (cases).

2.9 Data collection pro forma and procedure

Data were collected prospectively from time of presentation until exit from study (discharge or death in hospital) between December 2013 and July 2014 using a pro forma to capture patient data in line with the study objectives, see Appendix II.

The data collection pro forma was designed using the forms from the UK FLU-CATS development study as prototype, and incorporating appropriate changes where necessary to fit the Nigerian context. The form was designed in two parts for ease of use; the first part clearly titled 'initial assessment in the emergency department', captured data on patient demographics, inclusion criteria and

CATs criteria, examination findings and working diagnoses on first contact with the patients. The initial intention for the data collection process was that all the CATs assessments were to be conducted as part of the patients' case notes either by a medically qualified staff in the ED (i.e. medical intern or registrar), however, this was not practical due to the busy workload and after he pilot phase, the forms were ultimately completed by trained data collectors. A training session was held for the data collectors who were mostly final year medical students who had successfully completed a year of clinical rotations in the major specialities, and had an acceptable level of medical knowledge to enable them understand the contents of patients' case notes and the medical terms used in the forms. In the weeks preceding commencement of data collection, special sessions were organised with the emergency department doctors at each study centre to discuss the aims of the project, inclusion and exclusion criteria and clarifications on clinical signs for identifying patients with ALRTI. Age-specific pulse oximeters were also made available in consulting rooms to ensure recording of oxygen saturation measurements.

Trained data collectors were stationed in the ED with the attending clinical officers, who would identify cases of ALRTI and notify the data collectors or pass on case notes to them for data extraction. The second part of the form which was designed for recording death and non-death events occurring during follow-up of patients admitted for observation in the emergency departments or moved to the ward, was completed by the primary investigator of the study with the assistance of the data collectors. Information recorded included length of stay on admission (on the wards, ICU or ED), comorbidities, types of medication administered, duration (days) on supplemental oxygen, death or discharge. A section with guidance note explaining the variable fields was also

provided; this was to help data collectors in case there were any queries during the process of entering information into the forms.

2.9.1 Pilot phase

An initial pilot phase of data collection was conducted in LUTH and LASUTH between December 2013 and January 2014 to assess the applicability of the pro forma and the practicality of the data collection process. After the pilot phase, a few minor adjustments were made to the data collection forms to suit the context of the study setting. In the updated form, the 'date of onset' field for presenting symptoms was changed to 'duration of symptom'; as it seemed patients found it easier to estimate duration of illness in days than the actual date of onset. The 'ALRTI not specified' option in the provisional diagnosis field was also taken out as this was not commonly entered as a diagnosis in the case notes. As earlier mentioned, after the pilot phase it was decided that the high- pressure, time-critical ED working environment may not allow doctors enough time to complete the CATs assessment forms, and also that having a data collector assigned to each ED, would reduce that forms were completed and patients were not being missed.

A further modification to the initial study plan which occurred following the pilot phase, was the decision to recruit additional patients from secondary care hospitals in order to achieve the estimated sample size and ensure the representativeness of the cohort. This was because the patient throughput at LUTH and LASUTH during the pilot was lower than the presumed rate based on the available literature on ALRTI incidence in Nigeria.

2.9.2 Data entry

A database was created using Microsoft Excel for manual data entry; and as data were being collected for children and adults, two separate spreadsheets were created for this purpose. Data were entered on an ongoing basis alongside data collection data by the study's principal investigator and a data entry assistant who received training from the PI to ensure standardised data entry and coding practices.

2.9.3 Data validation

Data entry errors are inevitable, and therefore, the following steps were taken to minimise the occurrence of such errors and ensure data quality:

While the database was being designed an initial data validation plan was put in place to ensure accuracy and consistency of the terms being entered in each variable field. This was done using a range of data validation features in Microsoft Excel such as message prompts to indicate field requirements and error alerts to avoid entry of incorrect values.

Initial entries by the data entry assistant were checked against entries made by the PI (for the same patients, identified with their ID/ hospital numbers) to detect incorrectly entered data and ensure consistency in entry.

Although double-data entry was not possible at all times during the course of this project, the PI performed random periodic checks of the data entered by the assistant for errors and data completeness. Any errors found or queries raised by the data entry assistant during this process were clarified by contacting the data collectors for that study centre.

2.9.4 Data storage

Data collected was stored in a password protected hard drive and transferred to a secure University of Nottingham backup server that was only assessable via an approved login account for a period of 7 years as per University protocol. Patient details were anonymised during data cleaning by dropping personal identifiers. During the course of data collection, completed follow-up data of patients who had been enrolled in the study were periodically sent to thesis supervisors using a shared online password-secured study folder created within the University of Nottingham workspace.

2.10 Data variables & definition

A total of twelve predictor and five outcome variables were involved in the various statistical analyses. These variables will be explained in the subsequent sub-sections.

2.10.1 Predictors

All CATs criteria were recorded during initial consultation in the emergency department. As the CRB-65 tool shared similar criteria with the CATs tool and blood pressure was a regularly collected physiological parameter in the adult group, it was not necessary to separate data collection for the tools. All CATs and CRB-65 criteria were considered individually as binary categorical variable (Yes/No) for all analyses except for the comparison of weighted and un-weighted score indices where combined to yield a continuous score (this is discussed further in the data analysis section).

2.10.2 Outcome variables

The outcome variables in this thesis were chosen to reflect the clinical course and the interventions involved in the management of ALRTI (see chapter 1). Irrespective of the aetiology of ALRTI, antibiotics are the mainstay of treatment for ALRTI in Nigeria. Other supportive therapies used in the management of ALRTI depending on clinical presentation and findings include intravenous fluid and supplemental oxygen administration. As most of these interventions will necessitate hospital admission, a major decision in the management of ALRTI is the site of care; that is whether to manage on outpatient or inpatient basis. Based on this, the five outcomes considered are listed below:

- need for hospital admission
- need for supplemental oxygen
- need for intravenous antibiotics
- need for intravenous fluids
- In-hospital mortality

All study outcomes were treated as binary categorical variables (Yes/No), and in-hospital mortality was recorded as all-cause mortality during hospital stay. Mechanical ventilation and admission into the intensive care unit were initially considered as outcomes, but were later excluded as it was observed that patients with ALRTI were rarely admitted to the ICU in the study hospitals, and the reason for this was explored in the qualitative study conducted as part of this thesis.

2.10.3 Covariates

These were variables which were extracted from patients' case notes, after the patient had been assessed and admitted or discharged from the ED. If information for any of these variables was not documented by the attending physician, they were considered to be negative or not present. Examples of these variables include; comorbidities, birth and immunisation history for paediatric patients, social history such as alcohol consumption and smoking.

2.10.4 Handling missing data

There were instances of patients who self-discharged during the recurrent strike actions as there were no doctors in work to continue care. It was not possible to follow up these patients to determine management outcomes i.e. death which lead to their exclusion from the final analyses and as such complete case analysis was conducted for patients where complete follow up was possible. There were no missing data for the CATs criteria as the importance of data completeness was emphasised to the data collectors, however in the event that a missing criterion was discovered during data entry there was always an opportunity to trace patients' case notes for retrieval.

2.11 Statistical analyses

All analyses were conducted separately for adults and children, and Stata version 12 was used for data management and statistical analyses. Statistical significance for the analyses was set at the conventional level of 0.05 (P value< 0.05 considered significant). Given the number of variables and the various

statistical tests required for the analyses, there was a possibility that issues of multiple comparisons resulting in a type I error may occur (79). The Bonferroni correction is a popular method of solving issues arising from conducting multiple comparisons by adjusting the P-value to allow for the number of statistical tests performed(80); using the formula below:

$$\alpha/n$$

Where α = 0.05 and n = the total number of statistical tests, which in this case is the adjusted P value becomes: 0.05/164 = 0.0003

Although this newly adjusted P value will reduce the occurrence of type I errors, it also increases the likelihood of type II errors which is one of the major pitfalls of the Bonferroni correction method (81, 82). The decision to use the conventional P value was upheld, based on the fact that all associations studied in the analyses were decided *a priori*, and a detailed statistical plan was included in the protocol submitted for ethics approval (82, 83). Also, the use of such corrections and adjusted P values may contradict clinical evidence; particularly in regard to the CATs criteria, where established physiological evidence exists on why there is bound to be some association between the criteria themselves and some of the study outcomes i.e. a patient with low oxygen saturation levels, severe respiratory distress, increased respiratory rate will in a clinical setting will require supplemental oxygen administration, however, such associations may be considered false by using adjusted p-values.

Effect sizes (ORs, AUROC etc.) were reported using their corresponding confidence intervals (95%CI) and P values; as this has been found to be a more prudent style of reporting providing information on both the statistical and

clinical relevance of the findings(84). Details of statistical analyses conducted are described in subsequent sub-sections.

2.11.1 Descriptive analyses

Patient characteristics by study outcomes of interest (as described in earlier sections) were presented as percentages and compared using the chi square test for categorical variables, while continuous variables like age were summarised using the median and interquartile range.

2.11.2 Logistic regression analyses

Logistic regression was used to investigate the crude association between each of the seven CATs criteria and the five study outcomes (the need for hospital admission, supplemental oxygen, intravenous antibiotics or fluids and in-hospital mortality) in turn. Results were presented separately for each patient group as unadjusted odds ratios (ORs) and 95% confidence intervals (95% CI). The Odds ratios in this case represented the 'odds' of an outcome occurring in the presence of a given predictor relative to an occurrence in the absence of that same predictor. The association between the CATs criteria and outcomes were considered statistically significant if the P values were ≤ 0.05.

Multivariable logistic regression analyses were then conducted to determine which CATs criteria would independently predict the outcomes of interest. Two multivariable models were created for this purpose: the first model included all seven CATs criteria irrespective of significance attained in the univariate analysis, while the second multivariable model adopted the forward stepwise selection approach to determine the criteria to be included in the model.

2.11.3 Assessing multicollinearity

Multicollinearity is a phenomenon observed in multiple regression modelling which arises when the correlation between predictor variables is so strong that it becomes extremely difficult to interpret their individual effects on the outcome (85). It is therefore important to address any issues of multicollinearity before constructing a multivariable regression model. Most physiological parameters share similarities or are derived from each other making multicollinearity an unavoidable occurrence e.g. a model which has Body Mass Index (BMI) and weight as variables. As such tackling multicollinearity goes beyond 'the effect of presence' but becomes one of 'degree of presence' (86). In order to determine the degree of multicollinearity present in the regression model, the variance inflation factor (VIF) and tolerance of each CATs criterion were calculated. VIF and tolerance (1/VIF) values are a measure of how much the degree of collinearity between the predictor variables has affected the variance of the regression coefficient of the outcome variable. A rule of thumb used for assessing multicollinearity and usually favoured by statistic textbooks is that a VIF value of more than 10 suggests a red flag signalling a serious issue of multicollinearity with the regression (87-89).

2.11.4 Creating combined CATs scores

As mentioned in the introductory chapter (see chapter 1), the performance of the CATs criteria as combined un-weighted and weighted scoring systems has been previously assessed in three independently conducted studies using data

from the H1N1 pandemic(36, 38, 39); the CATs scores in this thesis were created using the approaches described in these studies. The un-weighted CATs score was created by awarding a single point to each CATs criterion: severe respiratory distress, increased respiratory rate, oxygen saturation levels less than 92%, respiratory exhaustion, clinical dehydration and other clinical concern raised by the clinician. Scores for each patient were calculated based on the summation of points for any criteria present. Weighted scores were also calculated for comparison with the un-weighted system, and weighting was achieved by multiplying the beta regression coefficients of each criterion obtained from the multivariable logistic regression by 10 and rounding up to the nearest integer. As before, patients' scores were calculated by adding the weighted scores to yield a weighted CATs index score. Regression coefficients were favoured over using odds ratios for assigning weights as scoring systems developed using odds ratios have been said to be algebraically incorrect(90).

2.11.5 Assessing predictive accuracy

Before accepting a tool for use in practice, clinicians require evidence that it correctly identifies the presence of disease or outcome in people who are truly affected and people without the disease as unaffected. Predictive accuracy is assessed in terms of discrimination and calibration. Discrimination is the ability of a prediction tool to perform binary categorisation of patients in terms of the risk or presence of an outcome(91). Calibration is defined as the ability of a prediction tool to accurately judge the probability of an event occurring across a range of estimated values(92); in simpler terms, if the values predicted by the tool agree with the actual observed values. Several approaches have been

suggested for assessing predictive accuracy, however this thesis will adopt the two most commonly used approaches which are the receiver operating characteristic (ROC) curves and area under the ROC (AUROC) values for discrimination and the Hosmer-Lemeshow goodness of fit test for calibration. Other measures used in describing accuracy as well as usefulness are sensitivity, specificity, negative and positive predictive values; which are usually expressed as proportions of the population under study. These measures are also needed to calculate AUROC values and their uses are discussed in detail below.

2.11.5.1 Sensitivity and specificity

Sensitivity is the proportion of people with the disease who are identified by a test/tool as being positive, while specificity is the exact opposite (proportion of people unaffected with the disease identified as negative). For example, if a tool has a sensitivity of 90%, it can be interpreted that 90% (45/50) people with the disease were identified by the tool, which is important as this could improve management outcomes as affected patients will receive the needed medical attention. On the other hand if the specificity of the tool was 90% for identifying patients who were negative for the disease in a healthy population, the tool is regarded as equally useful as this will encourage efficient use of hospital facilities and manpower. Although not unusual, it is rare for a test to have both high sensitivity and specificity as the relationship between the two parameters is inversely proportional such that one decreases as the other increases, in which case a trade-off between the two will depend on the purpose for which the tool is being used.

2.11.5.2 Positive predictive & negative predictive values

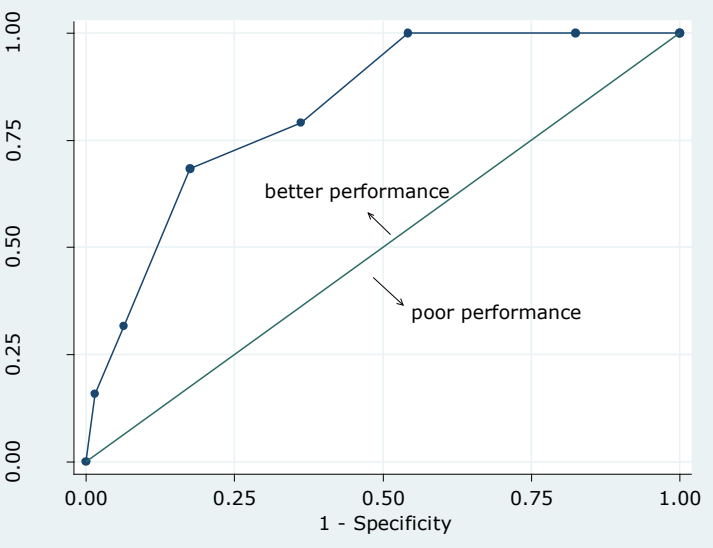
Positive predictive value is the proportion of people who were identified by the tool as positive (needing clinical intervention as in the case of triage tools) who are actually positive (in need for clinical care). With that being said, negative predictive value is thus the proportion of people who were identified as negative (not requiring clinical intervention as in the case of triage tools) who are actually negative and not in need for clinical care. However these values must be interpreted with caution as they are highly dependent on prevalence of the outcome being predicted (in this case mortality or need for clinical intervention); the higher the prevalence of the outcome, the higher the positive predictive values and vice versa for negative predictive values

2.11.5.3 Receiver operating characteristic (ROC) value & Area under ROC (AUROC) curve

The Receiver operating characteristic curve shown in Figure 1, uses the sensitivity and specificity (1- specificity) properties of a tool to plot a curve for all possible thresholds of the scoring system. The area under receiver operating characteristic curve is a summary of the ROC plot which represents the probability that if a person with the outcome of interest is chosen at random, such a person would be ranked higher in terms of their risk of that outcome by a prediction tool in comparison to a person without the outcome (93). The concept of perfect discrimination represented as an AUROC value of 1 is purely theoretical and unattainable; it is expressed as a ROC plot with perfect sensitivity and perfect specificity for use as a guide or marker of predictive accuracy. A test with no discriminative ability will have a ROC plot expressed

as a 45^{o} diagonal line and an AUROC of 0.5, meaning the tool is not able to separate patients into groups(93). Tools with AUROC values closer to 1.00 ($\geq$ 0.7) are said to have good to excellent predictive accuracy, while AUROC values < 0.70 are said to have poor accuracy.

Figure 1: Classification of test performance using the ROC curve



2.11.5.4 Hosmer-Lemeshow goodness-of–fit test

The Hosmer-Lemeshow test follows the principle of hypothesis testing by using a significance level of 0.05 to determine if the observed values match the predicted values. Calibration and the goodness of fit of the logistic models were assessed using the Hosmer-Lemeshow goodness-of-fit test. The model under investigation is said to have good calibration if the H-L test statistic is greater than 0.05 meaning that there was no difference between the models prediction and the observed values. The reverse is the case when the test statistic is found to be significant (< 0.05) as this suggests discrepancies between model

prediction and observed values(94). The same principle applies to assessing the 'goodness of fit' of a prediction tool, and in this case the predictions of the tool in the development sample are compared with predictions in the validation sample. Although this was an external validation study, the performance of the CATs scores was being assessed in new data and as such that its performance may be affected by overfitting (see chapter 1 for explanation). It was therefore important to conduct an internal validation of the findings from this new data by splitting the study population into two equal halves to create a hypothetical development and validation sample to compare the accuracy of the CATs score predictions using the Hosmer-Lemeshow test.

2.12 Ethics Statement

Data were collected in the background during consultation and did not interfere with the attending physicians' clinical decisions. Approval was given from the ethics board of each participating hospitals for collection of data without individual patient consent as the research objectives were considered to be non-therapeutic and data required were typically routinely collected as part of patient assessment. This decision was based on a WHO research ethics committee publication which advised that informed consent can be waived if the risks posed by the study on participants was minimal in terms of influencing clinical decisions and patient confidentiality(95) . Evidence of all ethics approval provided by the ethics committees of the participating hospitals are presented in the appendix section.

2.13 Qualitative Study Methodology

The need for mechanical ventilation which would necessitate admission into an intensive care unit was considered an important outcome in the initial study protocol for the CATs external validation study. A preliminary search for studies conducted in Nigeria on the management of patients with ALRTI or pneumonia in an ICU did not yield many results, however two things were evident: the few studies which investigated the use of severity scoring systems for predicting ICU admissions in such patients reported low ICU admission rates and poor performance by the tools in question; secondly, research on critical or intensive care practice in Nigeria was lacking which led to the decision to conduct a qualitative study to provide some insight into the critical care system in Nigeria. Coincidentally, the ICU admissions rates for ALRTI cases in this present study were low, thus providing further support for conducting this study. Also, as dissimilarities in the definition and practice of critical care may exist across study settings, the information gathered from this study is expected to facilitate contextually relevant interpretations of the findings from the CATs validation study and also contribute to the literature on critical care practice in Nigeria. The overarching aim of this section is to present the paradigmatic framework within which the qualitative study was conducted as well as the methods applied in answering this research question.

2.14 Philosophical underpinning

The overall aim of this study was to elicit an understanding of critical care practice and utilization in Nigerian health facilities, by answering the following questions: (A) what does critical care mean to healthcare professionals in Nigeria (B) how is critical care practised in terms of structure what factors affect critical care delivery in Nigeria and (C) how is critical care practiced in Nigerian ICUs in relation to ICU admissions, discharge and patient profile? This study was conducted within the pragmatic paradigm of scientific enquiry which Greene and Caracelli (2003) described as a bridge between philosophical paradigms and research methodology(96). The core merit of pragmatism is its context-driven and philosophically unaligned approach of focusing on practicality when considering methodological options for achieving research objectives (97, 98). The potential downside of this methodological-flexibility is that the researcher is solely responsible for judging the apparent utility of the methods applied and for interpreting emergent findings (99). In order to ensure the quality and validity of this study, a reflexive statement was included in the discussion section on how the researcher's background, experiences and interests may have influenced the research process, relationship with participants, analysis and findings (99, 100)

2.15 Study design

The exploratory nature of the research question supports the use of a qualitative methods to achieve the study's objectives; face to face interviews were the chosen method of data collection as they provided an opportunity to actively elicit responses from the participants about a range of topics (101).

Individual interviews allowed participants to freely discuss their views without interruptions or influence from other respondents.

2.16 Ethics

The main ethical issues considered in the conduct of this study were informed consent, confidentiality, and anonymity (102). Ethics approval was received from all participating institutions as detailed in the first part of this chapter, the purpose of the study was explained to the participants by the researcher (RA) who was also the only interviewer. Verbal informed consent was audio recorded and obtained from each participant before starting the interviews.

As the findings of this study were intended for dissemination and future publication offering full confidentiality to participants would defy the purpose of using the information obtained, however ensuring anonymity during transcription and analysis by substituting any personal identifiers with pseudonyms such as LSR1 (LUTH senior registrar 1) and LASR1
(LASUTH senior registrar 1) guaranteed that confidentiality would be maintained(103). Confidentiality was further maintained by taking care to avoid using complete extracts of quotes by participants while writing up the results to prevent them from being identified from their words.

2.17 Study population

Study participants were chosen by the researcher (RA) from the four hospitals used in the CATs validation study; two were tertiary care centres with functioning ICUs and the other two were secondary care centres with no ICU. Based on the capacity of the health facilities, study participants in the tertiary

hospitals were typically doctors, nurses and three anaesthetist, while participants in the secondary care hospitals with no ICUs were mainly doctors. The researcher was also able to capture a wide range of perspectives and deviant views by choosing clinical staff with different levels and lengths of experience, this was so that they would have both background knowledge and skills to answer interview questions. Respondents had to have at least 2-3 years of experience working in that facility, and the doctors also had to be at the level of a junior registrar or higher (speciality registrar grade).

2.18 Sampling and recruitment strategy

The heads of the emergency departments, ICU and the major medical specialties (medicine, paediatric, surgery and obstetrics and gynaecology) at each of the four hospitals were approached by the researcher for a formal introduction to discuss the study and the recruitment strategy. After the introductory meeting, the researcher was shown around the emergency units and ICU; using the snowball method of purposive (non-probability) sampling, the first respondents were senior doctors who then suggested other doctors and nurses whom they felt would have the relevant knowledge and experience to participate in the interviews.

Quite a number of doctors and nurses declined stating lack of free time as an excuse, while some were just uncomfortable with having their voice recorded; in the end a total of 14 interviews were conducted.

There are no explicit guidelines for estimating the sample size of qualitative studies utilising interviews, however several authors have suggested data saturation which is the point at which no extra information is gained from data

collection or analysis as a guide to determining the sample size when using purposive samples (104-107). Using data from sixty interviews conducted in Ghana and Nigeria, Guest et al. (2006) conducted a study to estimate the sample size required to reach saturation in studies involving purposive sampled interviews, and reported that twelve interviews were sufficient for developing significant themes and achieving data saturation(107). These authors however emphasise that the choice of an adequate sample size is to a large extent dependent on the purpose of the study and have cautioned against adhering strictly to this rule of twelve as this finding will not be applicable to studies with 'a relatively heterogeneous population and ill-defined objectives'(107). With its well-defined and explicit aim of understanding critical care practice and utilization, among a group of individuals with relatively similar background by virtue of being healthcare professionals, responding to a similar set of questions, this qualitative study conducted an appropriate number of interviews according to the criteria stated by Guest et al. (2006).

2.19 Data collection and interview process

The interview guide which was informed by knowledge gaps in literature was designed and refined during a tripartite meeting held between supervisors and the principal investigator (appendix III). Semi-structured face to face interviews were chosen for data collection for the following reasons:

- The conversational style interview process allows the respondents express themselves freely in their own words which provides very rich data for analysis(108)

- It also allows the researcher/interviewer some discretion with regards to the interview process, offering opportunities for probing using open ended question which generates in-depth understanding of the topic(101).

All interviews were audio recorded using a portable digital recorder, and were conducted during working hours in locations chosen by the participants; mostly in consulting rooms and ICU side rooms as this ensured an element of comfort and security throughout the interview process. Informed consent was obtained from each participant before the start of the interview, after which respondents were asked to introduce their roles (without giving names). All participants were asked to explain what they understood by critical care; participants in tertiary care facilities were asked about ICUs admission and discharge criteria and protocols, while those in secondary care hospitals with no ICUs were asked questions on referral and management protocols for critically ill patients. The questions were all open-ended encouraging a natural two way communication/ conversation between the researcher and respondents.

2.20 Data analysis

Data were analysed using the thematic analysis approach; this method of qualitative data analysis is simple, relatively easy to use and able to generate understanding by identifying similarities and differences across the data(109). In line with the study's pragmatic framework, thematic analysis was chosen based on its theoretical flexibility and its ability to work across all philosophical paradigms(110). Data were analysed according to steps outlined by Braun and Clarke (2006), and themes were identified based on semantic content focusing

mainly on literal interpretations of the data (109). The six steps by Braun and Clarke as used in this study have been described in the following subsections:

2.20.1 Familiarisation with data

The researcher RA familiarised herself with the data by listening to the audio recordings repeatedly before transcribing the interviews verbatim. Pseudonyms such as LSR1 (LUTH senior registrar 1) and LASR1 (LASUTH senior registrar 1) were used to identify participants and copies of transcripts were stored in a password protected hard drive. RA reviewed transcripts while listening to audio recordings to further familiarise herself with the data and ensure accuracy of transcription.

2.20.2 Generating initial codes

Data were manually coded independently by the researcher RA and supervisor PM, after which a set of probable codes related to the interview questions (analyst-driven) were agreed on by (RA) and PM using the deductive thematic approach (111, 112). This approach helped identify and categorise instances within the texts which provided insights specific to the research purpose of describing critical care practice. Codes and relevant sections of texts were collated and entered into an Excel spreadsheet by RA.

2.20.3 Searching for themes

Codes were organised into broader groups known as themes and sub-themes to represent patterns within the data by RA. All relevant participant quotes were also included within the themes and sub-themes at this stage.

2.20.4 Reviewing themes

Coded text extracts under each candidate theme were first reviewed to determine if they were a good fit and a consistent representation of the theme in question. Extracts which did not fit in were taken out for possible reclassification into other existing or new themes. Sub-themes were then reviewed to determine if they correctly represented the data and this also led to some sub-themes being dropped or merged into themes.

2.20.5 Defining and naming themes

The themes were considered in relation to the study aims, as well as the other study themes and in order to ensure coherence in presenting the findings. Themes were defined and named according to what they represented.

3: Study centres and Data description

3.1 Introduction

In order to contextualise the findings from this thesis, this chapter describes the data collected across the four study hospitals including the distribution of the variables and participants.

3.2 Descriptive characteristics

Between December 2013 and July 2014, 1071 consecutive patients with symptoms suggestive of an acute lower respiratory tract infection were recruited across four large hospitals in Lagos; after 55 patients who opted for early discharge were dropped from the data, a total of 1016 patients (809 children <16 years and 207 adults aged $\geq$ 16 years) with complete follow up data were studied. 714 cases were identified from the two tertiary hospitals (LUTH= 454 and LASUTH =260) while the other 302 cases were identified from March 2014 at Randle General Hospital (130) and Massey Street Children's Hospital (172). The majority of patients self-referred (635/1016; 63%), while most of the referred cases were from primary health centres (132/1016; 13%). There was a slight male preponderance among the participants (n=628; 62%), the median (IQR) age in years was 1 (1-2) for children and 42 (31-58) for adults; histograms showing the age distribution of the study population are presented in Figures 4 and 5. The distribution and characteristics of the participants in relation to the study outcomes have been presented separately for adults and children in Tables 3.1 and 3.2.

Table 3.1: Distribution and characteristics of children (< 16years)

	Total sample (n=809) Frequency (%)	Hospital admission (661/809)	Supplemental oxygen (376/809)	IV Antibiotics (577/809)	IV Fluid (546/809)	In-hospital Mortality (73/809)
Age (years): median (IQR)	1 (1-2)					
Sex (%)						
Male	409 (50.6%)	334 (50.5%)	181 (48.4%)	285 (49.4%)	290 (53.1%)	41 (56.2%)
Female	400 (49.4%)	327 (49.5%) NS	195 (51.9%) NS	292 (50.6%) NS	256 (46.9%) †	32 (43.8%) NS
Ethnicity (%)						
Yoruba	388 (48%)	317 (48%)	187 (49.7%)	279 (48.4%)	265 (48.5%)	33 (45.2%)
Ibo	213 (26.3%)	174 (26.3%)	95 (25.3%)	146 (25.3%)	146 (26.7%)	20 (27.4%)
Hausa	34 (4.2%)	30 (4.5%)	18 (4.8%)	29 (5%)	23 (4.2%)	4 (5.5%)
Others	170 (21%)	137 (20.7%)	74 (19.7%)	120 (20.8%)	109 (20%)	15 (20.6%)
n/a	4 (0.5%)	3 (0.5%) NS	2 (0.5%) NS	3 (0.5%) NS	3 (0.6%) NS	1 (1.4%) NS
Comorbidities (%)						
0	362 (44.8%)	273 (41.3%)	140 (37.2%)	228 (39.5%)	229 (41.9%)	14 (19.2%)
1	372 (46%)	322 (48.7%)	191 (50.8%)	289 (50.1%)	252 (46.2%)	46 (63%)
≥2	75 (9.3%)	66 (10%) ǂ	45 (12%) ǂ	60 (10.4%) ǂ	65 (11.9%) ǂ	13 (17.8%) ǂ
Hospital referred (%)						
No	495 (61.2%)	380 (57.5%)	190 (50.5%)	322 (55.8%)	297 (54.4%)	44 (60.3%)
Yes	314 (38.8%)	281 (42.5%) ǂ	186 (49.5%) ǂ	255 (44.2%) ǂ	249 (45.6%) ǂ	29 (39.7%) NS
General hos.	105 (33.4%)	95 (33.8%)	64 (34.4%)	83 (32.6%)	85 (34.1%)	7 (24.4%)
Private hos.	88 (28%)	81 (28.3%)	57 (30.7%)	77 (30.2%)	69 (27.7%)	13 (44.8%)
PHC	121 (38.5%)	105 (37.4%) NS	65 (35%) NS	95 (37.3%) NS	95 (38.2%) NS	9 (31%) NS
Provisional diagnosis (%)						
Pneumonia	626 (77.4%)	545 (82.5%)	346 (92%)	511 (88.6%)	440 (80.6%)	72 (98.6%)
Acute bronchitis	99 (12.2%)	54 (8.2%)	13 (3.5%)	37 (6.4%)	48 (8.8%)	0 (0%)
Bronchiolitis	84 (10.4%)	62 (9.4%) ǂ	17 (4.5%) ǂ	29 (5%) ǂ	58 (10.6%) ǂ	1 (1.4%) ǂ

NS=p value >0.05; †= p-value <0.05; ǂ= p-value <0.001
P values relate to the comparison of characteristics in those who experienced a given outcome VS those who did not

Table 3.2: Distribution and characteristics of adults (≥ 16years)

	Total sample (n=207) Frequency (%)	Hospital admission (135/207)	Supplemental oxygen (75/207)	IV Antibiotics (106/207)	IV Fluid (96/207)	In-hospital Mortality (19/207)
Age (years): median (IQR)	42 (31- 58)					
Sex (%)						
Male	119 (57.5%)	76 (56.3%)	43 (57.3%)	63 (59.4%)	55 (57.3%)	13 (68.4%)
Female	88 (42.5%)	59 (43.7%) NS	32 (42.7%) NS	43 (40.6%) NS	41 (42.7%) NS	6 (31.6%) NS
Ethnicity						
Yoruba	102 (49.3%)	62 (45.9%)	31 (41.3%)	49 (46.2%)	43 (44.8%)	5 (26.3%)
Ibo	42 (20.3%)	29 (21.5%)	19 (25.3%)	25 (23.6%)	26 (27.1%)	8 (42.1%)
Hausa	11 (5.3%)	8 (5.9%)	6 (8%)	6 (5.7%)	3 (3.1%)	2 (10.5%)
n/a	52 (25.1%)	36 (26.7%) NS	19 (25.3%) NS	26 (24.5%) NS	24 (25%) NS	4 (21.1%) N
Comorbidities (%)						
0	79 (38.2%)	14 (10.3%)	10 (13.3%)	12 (11.3%)	7 (7.3%)	3 (15.8%)
1	110 (53.1%)	82 (60.7%)	45 (60%)	65 (61.3%)	62 (64.6%)	13 (68.4%)
≥2	18 (8.7%)	39 (28.9%) **ǂ**	20 (26.7%) †	29 (27.4%) †	27 (28.1%) †	3 (15.8%) NS
Hospital referred (%)						
No	140 (67.6%)	78 (57.8%)	37 (49.3%)	61 (57.6%)	58 (60.4%)	2 (10.5%)
Yes	67 (32.4%)	57 (42.2%) ǂ	38 (50.7%) ǂ	45 (42.5%) **†**	38 (39.6%) **†**	17 (89.5%) ǂ
General hos.	21 (31.3%)	19 (33.3%)	16 (42.1%)	17 (37.8%)	12 (31.6%)	8 (47.1%)
Private hos.	35 (52.4%)	29 (50.9%)	17 (44.7%)	20 (44.4%)	19 (50%)	9 (52.9%)
PHC	11 (16.4%)	9 (15.8%) NS	5 (13.2%) NS	8 (17.8%) NS	7 (18.4%) NS	0 (16.4%) NS
Provisional diagnosis (%)						
Pneumonia	129 (62.3%)	111(81%)	64 (92.8%)	84 (84%)	73 (81.1%)	17 (88.9%)
Acute bronchitis	76 (36.7%)	22 (17.5%)	3 (4.4%)	14 (14%)	16 (17.8%)	0 (0%)
Inf. Exacerbation of COPD	2 (1%)	2 (1.6%) **ǂ**	2 (2.9%) **ǂ**	2 (2%) ǂ	1 (1.1%) **ǂ**	2 (1%) ǂ

NS=p value >0.05; †= p-value <0.05; ǂ= p-value <0.001
P values relate to the comparison of characteristics in those who experienced a given outcome VS those who did not

Figure 2: Histogram showing age distribution by patient age group

Figure 2a: Age distribution in children (<16 years)

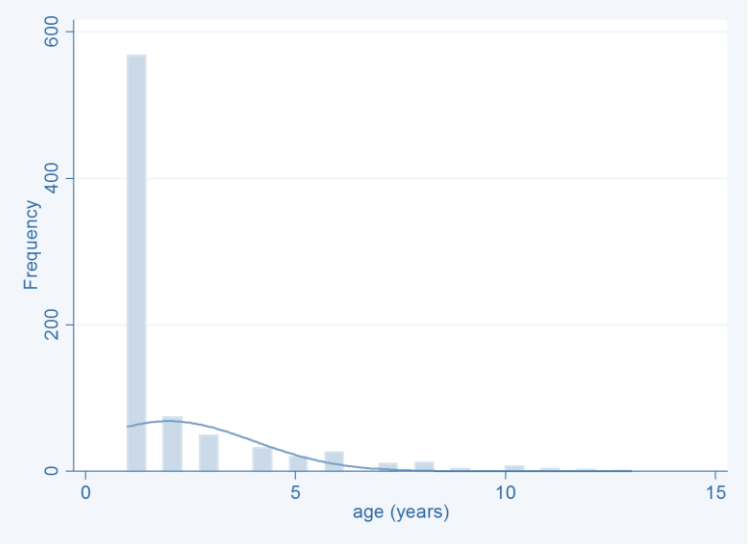
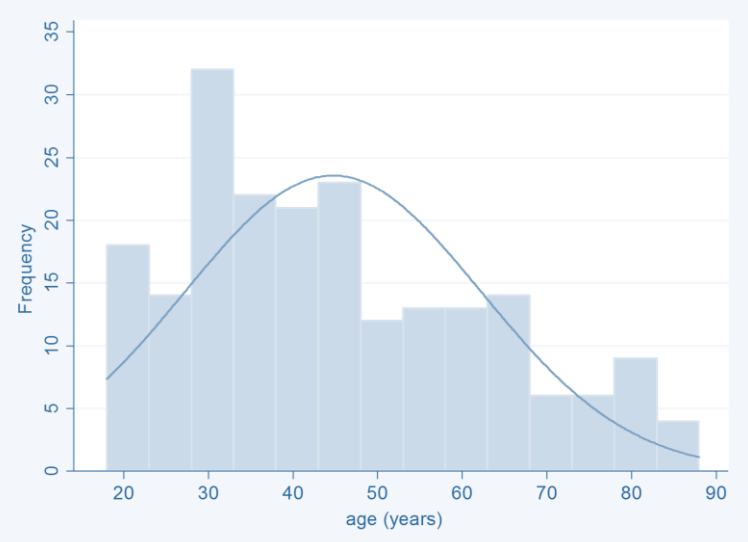


Figure 2b: Age distribution of adults (≥ 16 years)



3.3 Diagnoses and Study Outcomes

Of the 1016 patients recruited into this study, 796 (78 %) were admitted (661 children and 135 adults) while the rest were managed as outpatients, pneumonia was the most common diagnosis in this data (n=755; 74.3%) and these diagnoses were made on clinical grounds; 77.4% (626/809) of children and 62.3% (129/207) of the adult population were diagnosed with pneumonia. After pneumonia, the next common diagnosis in this data was acute bronchitis (n= 175; 17.2%), there were 84 cases of children with bronchiolitis and only 2 cases of adults with acute infectious exacerbation of chronic obstructive pulmonary disease. More than half of the population of adults (140/207; 68%) and children (498/809; 62%) had self-medicated prior to presentation to any hospital, Of the 796 patients admitted, 86% (683/796) received intravenous antibiotics, 81% (642/796) received intravenous fluid and 441/796 (55%) required supplemental oxygen. Patients received other medications such as antimalarial, anticonvulsant, anti-hypertensive medications depending on co-existing illnesses.

The total mortality reported among admitted patients was 92/796 (12%) of which 96% (89/93) of the reported in-patient deaths were in patients who had pneumonia. 79% (73/92) of the reported deaths occurred the younger participants and 19 in the adult participants.

4: Assessing the performance of the individual CATs criteria in predicting hospital admission, acute care interventions and in-hospital mortality in adults and children with ALRTI

4.1 Introduction

The first stage in the external validation of the Community Assessment Tools is presented in this chapter. The predictive potential of each CATs criterion in relation to the severity assessment of a range of acute lower respiratory tract infections has been prospectively assessed in a population of adults and children presenting to tertiary and secondary care hospitals in Nigeria. This chapter describes the specific statistical analyses employed in investigating the performance of the criteria and also discusses the implications of the findings in the concluding sections.

4.2 Background

The seven criteria Community Assessment Tool was proposed as a guide to improve clinical decision making during periods of extreme medical surge or pressure on healthcare capacity and resources (32). The first six CATs criteria (A to F) were selected based on established clinical evidence for identifying a range of lower respiratory tract and severe feverish illness in patients of all ages(35), while the seventh criterion represents the role of a clinician's instinct or 'judgement' in recognising a potentially severe situation which may require hospital admission or referral. The initial concept when the CATs was

developed, was that meeting any one or more of its criteria would serve as a cue for referral or hospital admission. Evidence for using the CATs criteria as prompts for making clinical decisions was further confirmed by the findings of an evaluation study conducted using retrospective data from the 2009 UK A/H1N1 influenza pandemic (35, 113); where each CATs criterion was seen to be independently capable of predicting hospital based interventions and severe outcomes (35).

Although the CATs were not intended for use as illness-specific prediction tools, they have only been studied in cases of influenza and pneumonia. This present study aims to investigate the predictive ability and accuracy of the CATs criteria, in a multi-centre study involving a prospective cohort of adults and children presenting with symptoms suggestive of acute lower respiratory tract infections. The reported burden of acute lower tract respiratory infections, coupled with the disproportionate distribution of healthcare resources and manpower in relation to demand in Nigeria presents a presumably high-bar triage setting for validating a tool developed for use in a pandemic situation.

4.3 Statistical analyses

The predictive ability of each CATs criterion was initially assessed using logistic regression analyses; univariate logistic regression was initially used to investigate the associations between the individual CATs criterion and each study outcome: need for acute care interventions (supplemental oxygen, intravenous antibiotics or fluids), hospital admission and in-patient mortality. The CATs criteria were then included in two separate multivariable logistic models to determine which criteria would emerge as independent predictors of

the study outcomes. Model I included all criteria irrespective of statistical significance achieved in the unadjusted analysis, while model II used the forward stepwise regression method to include criteria based on a significance threshold of 0.05 (p value < 0.05). Results are presented as unadjusted and adjusted odds ratios (OR) and 95% confidence intervals (95%CI) for the univariate and multivariable analyses respectively.

Area under receiver operating curve values AUROC (95% CI), as well as other measures of predictive performance such as sensitivity, specificity, negative and positive predictive values were then used to assess predictive accuracy.

All analyses were conducted separately for adults (≥ 16years) and children (< 16years), using Stata version 12.0 (StataCorp. 2011. *Stata Statistical Software: Release 12.* College Station, TX: StataCorp LP).

4.4 Results

The data used for the analyses have been described previously in chapter 3, and in the following sections, the results for the univariate analyses have been presented for each of the study outcome in separate tables (Table 4.1 to 4.5). Each table includes p-values for the regression analysis and the distribution of the patients meeting each CATs criterion for that specific outcome.

As independent predictors of the study outcomes were identified using two different multivariable regression modelling approaches; the results according to patient age group are presented in Tables 4.6 and 4.7 for model I and in Figures 3 and 4 for model II.

4.4.1 Association between CATs and need for hospital admission

The univariate analyses investigating the relationship between the CATs criteria and need for hospital admission are presented in table 4.1 for adults and children. With the exception of CATs criterion D (respiratory exhaustion) in children where the numbers were insufficient for the analysis, all unadjusted odds ratios and p values show significant associations between each individual CATs criterion and hospital admission for both patient groups. The criteria with the strongest likelihood for admission in the adult group were criterion C (peripheral oxygen saturation $\leq$ 92%) [OR (95%CI):20.02(6.91- 57.99)] and F (altered consciousness) [OR (95%CI):10.23(1.33-78.52)]. In children, the strongest associations were observed for criterion F (altered consciousness) [OR (95%CI):30.64(4.24-221.21)] and criterion E (clinical dehydration or shock) [OR (95%CI):20.82(2.87-150.81)].

4.4.2 Association between CATs and need for supplemental oxygen

CATs C which is the criterion indicating an oxygen saturation value of 92% or less (a marker of hypoxemia) was excluded from this analysis as it is a direct indication for the administration of supplemental oxygen(114). All other CATs criteria were significantly associated with an increased need for supplemental oxygen in both adults and children ($p< 0.001$). The strongest associations observed were for criterion D (respiratory exhaustion) [OR (95%CI):17.71(8.47-37.02)] in children (< 16 years) and criterion A (severe respiratory distress) [OR (95%CI):11.37(5.05-25.56)] in the adult group. These findings are summarised in table 4.2.

4.4.3 Association between CATs and need for intravenous antibiotics

The unadjusted odds ratios presented in table 4.3 show significant associations between each of the seven CATs criterion and need for intravenous antibiotics in both adults and children. Criterion D (respiratory exhaustion) [OR (95%CI):15.81(4.96-50.40)] had the strongest association with the need for IV antibiotics use in the children's group, while criterion C (oxygen saturation $\leq$ 92%) [OR (95%CI):8.08(4.13-15.80)] had the strongest association in the adult group.

4.4.4 Association between CATs and intravenous fluids

Results presented for the younger patient group (< 16years) in table 4.4 show that all CATs criteria except criterion D (respiratory exhaustion) were significantly associated with the outcome 'need for intravenous fluid'; each one showing an increase in the likelihood of requiring intravenous fluids. Criterion D on the other hand, showed a non-significant decrease in the likelihood of needing IV-fluids [OR (95%CI):0.83(0.54-1.27)]. In the adult patient group, all CATs criteria except criterion F (altered consciousness) and criterion G (causing other clinical concern) were each seen to be associated with the need for IV-fluid administration.

4.4.5 Association between CATs and in-hospital mortality

Table 4.5 summarises the results of the unadjusted analyses involving the outcome in-hospital mortality and each criterion. In both adults and children, Criterion B (increased respiratory rate) was not associated with in-hospital

mortality; all of the other criteria except criterion E (clinical dehydration) in the children and D (respiratory exhaustion) in the adults were significantly associated with in-hospital mortality.

Table 4.1: Association between CATs criteria and hospital admission

CATs Criteria	No. of children <16years in need of admissions (%)	Unadjusted odds ratio OR (95%CI)	No. of adults ≥ 16years in need of admissions (%)	Unadjusted odds ratio OR (95%CI)
A: severe respiratory distress				
No	115 (17.4%)	1.00	34 (25.2%)	1.00
yes	546 (82.6%)	6.77 (4.61-9.94) ǂ	101 (74.8%)	6.75 (3.58-12.73) ǂ
B: increased respiratory rate				
No	75 (11.4%)	1.00	31 (23%)	1.00
Yes	586 (88.7%)	3.31 (2.16-5.06) ǂ	104 (77%)	4.97 (2.68-9.23) ǂ
C: oxygen saturation. ≤ 92%				
No	361 (54.6%)	1.00	62 (45.9%)	1.00
Yes	300 (45.4%)	19.67 (8.57-45.15) ǂ	73 (54.1%)	20.02 (6.91-57.99) ǂ
D: respiratory exhaustion				
No	559 (84.6%)	**	107 (79.3%)	1.00
Yes	102 (15.4%)		28 (20.7%)	9.16 (2.11-39.67) †
E: clinical dehydration				
No	579 (87.6%)	1.00	116 (85.9%)	1.00
Yes	82 (12.4%)	20.82 (2.87-150.81) †	19 (14.1%)	3.77 (1.08-13.20) †
F: altered consciousness				
No	547 (82.8%)	1.00	118 (87.4%)	1.00
Yes	114 (17.3%)	30.64 (4.24-221.21) ǂ	17 (12.6%)	10.23 (1.33-78.52) †
E: causing other clinical concern	464 (70.2)	1.00	94 (69.6%)	1.00
No	197 (29.8%)	2.88 (1.73-4.80) ǂ	41 (30.4%)	7.41 (2.54-21.68) ǂ
Yes				

**** Odds ratio not calculated due to insufficient numbers**
ǂ P-value <0.001
† P-value <0.05
All values are presented as the odds of having the outcome if CATs criteria are present.

Table 4.2: Association between CATs criteria and need for supplemental oxygen

Triage Criteria	No. of children <16years in need of suppl. O2 (%)	Unadjusted odds ratio OR (95%CI)	No. of adults ≥ 16years in need of suppl. O2 (%)	Unadjusted odds ratio OR (95%CI)
A: severe respiratory distress				
No	46 (12.2%)	1.00	8 (10.7%)	1.00
yes	330 (87.8%)	4.04 (2.80-5.82) ǂ	67 (89.3%)	11.37 (5.05-25.56) ǂ
B: increased respiratory rate				
No	32 (8.5%)	1.00	11 (14.7%)	1.00
Yes	344 (91.5%)	2.70 (1.76-4.16) ǂ	64 (85.3%)	5.31 (2.57-10.97) ǂ
C: oxygen saturation ≤ 92%				
No	115 (30.6%)		13 (17.3%)	
Yes	261 (69.4%)	**X**	62 (82.7%)	**X**
D: respiratory exhaustion				
No	282 (75%)	1.00	51 (68%)	1.00
Yes	94 (25%)	17.71 (8.47- 37.02) ǂ	24 (32%)	9.88 (3.81-25.60) ǂ
E: clinical dehydration				
No	322 (85.6%)	1.00	60 (80%)	1.00
Yes	54 (14.4%)	2.34 (1.45-3.75) ǂ	15 (20%)	4.46 (1.73-11.53) †
F: altered consciousness				
No	293 (77.9%)	1.00	60 (80%)	1.00
Yes	83 (22.1%)	3.55 (2.30-5.48) ǂ	15 (20%)	10.75 (3.00-38.54) ǂ
G: causing other clinical concern	245 (65.2%)	1.00	48 (64%)	1.00
No	131 (34.8%)	2.19 (1.59-3.01) ǂ	27 (36%)	3.56 (1.80-7.07) ǂ
Yes				

**** Odds ratio not calculated due to insufficient numbers**
X this criterion directly predicts the outcome
ǂ P-value <0.001
† P-value <0.05
All values are presented as the odds of having the outcome if CATs criteria are present.

Table 4.3: Association between CATs criteria and need for Intravenous antibiotics

Triage Criteria	No. of children <16years in need of IV Antibiotics (%)	Unadjusted odds ratio OR (95%CI)	No. of adults ≥ 16years in need of IV Antibiotics (%)	Unadjusted odds ratio OR (95%CI)
A: severe respiratory distress				
No	101 (17.5%)	1.00	27 (25.5%)	1.00
yes	476 (82.5%)	3.63 (2.59-5.09) ǂ	79 (74.5%)	3.79 (2.11-6.82) ǂ
B: increased respiratory rate				
No	66 (11.4%)	1.00	21 (19.8%)	1.00
Yes	511 (88.6%)	2.29 (1.54-3.42) ǂ	85 (80.2%)	4.47 (2.41-8.28) ǂ
C: oxygen saturation ≤92%				
No	298 (51.7%)	1.00	44 (41.5%)	1.00
Yes	279 (48.4%)	7.11 (4.61-10.96) ǂ	62 (58.5%)	8.08 (4.13-15.80) ǂ
D: respiratory exhaustion				
No	478 (82.8%)	1.00	81 (76.4%)	1.00
Yes	99 (17.2%)	15.81 (4.96-50.40) ǂ	25 (23.6%)	5.93 (2.17-16.18) †
E: clinical dehydration				
No	502 (87%)	1.00	87 (82.1%)	1.00
Yes	75 (13%)	4.18 (1.98-8.82) ǂ	19 (17.9%)	7.13 (2.04-24.94) †
F: altered consciousness				
No	476 (82.5%)	1.00	91 (85.9%)	1.00
Yes	101 (17.5%)	3.30 (1.85-5.91) ǂ	15 (14.2%)	5.38 (1.51-19.21) †
G: causing other clinical concern	401 (69.5%)	1.00	68 (64.2%)	1.00
No	176 (30.5%)	2.11 (1.43-3.09) ǂ	38 (35.9%)	7.50 (3.16-17.81) ǂ
Yes				

**** Odds ratio not calculated due to insufficient numbers**
ǂ P-value <0.001
† P-value <0.05
All values are presented as the odds of having the outcome if CATs criteria are present.

Table 4.4: Association between CATs criteria and need for Intravenous Fluid

Triage Criteria	No. children <16years in need of IV fluid (%)	Unadjusted odds ratio OR (95%CI)	No. of adults ≥ 16years in need of IV fluid (%)	Unadjusted odds ratio OR (95%CI)
A: severe respiratory distress				
No	100 (18.3%)	1.00	27 (28.1%)	1.00
yes	446 (81.7%)	2.83 (2.03-3.93) ‡	69 (71.9%)	2.70 (1.51-4.82) †
B: increased respiratory rate				
No	71 (13%)	1.00	23 (24%)	1.00
Yes	475 (87%)	1.49 (1-2.23) †	73 (76%)	2.70 (1.48-491) †
C: oxygen saturation ≤ 92%				
No	317 (58.1%)	1.00	46 (47.9%)	1.00
Yes	229 (41.9%)	1.75 (1.27-2.39) †	50 (52.1%)	3.38 (1.87-6.10) ‡
D: respiratory exhaustion				
No	481 (88.1%)	1.00	72 (75%)	1.00
Yes	65 (11.9%)	0.83 (0.54-1.27) NS	24 (25%)	5.83 (2.27-14.99) ‡
E: clinical dehydration				
No	469 (85.9%)	1.00	79 (82.3%)	1.00
Yes	77 (14.1%)	7.03 (3.02-16.37) ‡	17 (17.7%)	4.56 (1.61-12.89) †
F: altered consciousness				
No	456 (83.5%)	1.00	85 (88.5%)	1.00
Yes	90 (16.5%)	1.88 (1.17-3.01) †	11 (11.5%)	1.92 (0.71-5.17) NS
G: causing other clinical concern				
No	379 (69.4%)	1.00	70 (72.9%)	1.00
Yes	167 (30.6%)	1.92 (1.34-2.76) ‡	26 (27.1%)	1.80 (0.92-3.51) NS

**** Odds ratio not calculated due to insufficient numbers**
‡ P-value <0.001
† P-value <0.05
All values are presented as the odds of having the outcome if CATs criteria are present.

Table 4.5: Association between CATs criteria and In-hospital mortality

Triage Criteria	Number of deaths (%) in children <16years	Unadjusted odds ratio OR (95%CI)	Number of deaths (%) in adults ≥ 16years	Unadjusted odds ratio OR (95%CI)
A: severe respiratory distress				
No	7 (9.6%)	1.00	2 (10.5%)	1.00
yes	66 (90.4%)	3.40 (1.53-7.53) †	17 (89.5%)	6.58 (1.48-29.27) †
B: increased respiratory rate				
No	6 (8.2%)	1.00	5 (26.3%)	1.00
Yes	67 (91.8%)	2.03 (0.86-4.78) NS	14 (73.7%)	1.62 (0.56-4.70) NS
C: oxygen sat. ≤ 92%				
No	10 (13.7%)	1.00	1 (5.3%)	1.00
Yes	63 (86.3%)	12.78 (6.45-25.35) ‡	18 (94.7%)	39.36 (5.13-301.79) ‡
D: respiratory exhaustion				
No	45 (61.6%)	1.00	15 (79%)	1.00
Yes	28 (38.4%)	5.57 (3.28-9.45) ‡	4 (21.1%)	1.66 (0.51-5.40) NS
E: clinical dehydration				
No	61 (83.6%)	1.00	14 (73.7%)	1.00
Yes	12 (16.4%)	1.84 (0.95-3.58) NS	5 (26.3%)	3.59 (1.15-11.19)†
F: altered consciousness				
No	53 (72.6%)	1.00	13 (68.4%)	1.00
Yes	20 (27.4%)	2.55 (1.46-4.45) †	6 (31.6%)	6.77 (2.19-20.96)†
E: causing other clinical concern	36 (49.3%)	1.00	8 (42.1%)	1.00
No	37 (50.7%)	3.20 (1.96-5.21) ‡	11 (57.9%)	6.23 (2.33-16.65)‡
Yes				

**** Odds ratio not calculated due to insufficient numbers**
‡ P-value <0.001
† P-value <0.05
All values are presented as the odds of having the outcome if CATs criteria are present.

4.4.6 Multivariable analysis: model I including all criteria irrespective of significance in children (< 16years)

Criterion C (oxygen saturation $\leq$92%) independently positively predicted all outcomes except need for supplemental oxygen which was removed a priori, and was the strongest independent predictor of in-hospital mortality [OR (95%CI): 8.36 (3.98-17.54)]. Criterion G (causing other clinical concern) was the only CATs criterion to independently predict the need for hospital admission, all acute care interventions (need for supplemental oxygen, IV antibiotics and fluids) and was the strongest predictor of in-hospital mortality after criterion C [OR (95%CI): 2.95 (1.73-5.02)]. Criterion D was significantly associated with all outcomes except hospital admission as it lacked sufficient entries to be included in the analysis for this outcome. Criterion D which was the strongest independent positive predictor for supplemental oxygen and IV antibiotic, independently predicted the need for IV fluids negatively [OR (95%CI): 0.43 (0.26-0.72)]. Criterion F (altered consciousness) which was the strongest positive independent predictor of admissions [OR (95%CI): 14.90 (2-110.92)] independently predicted other outcomes except in-hospital mortality. Criterion A (severe respiratory distress) and criterion B (increased respiratory distress) both independently predicted the need for admission and supplemental oxygen but not in-hospital mortality. Criterion E (clinical dehydration) was the strongest positive independent predictor of need for IV fluid.

In summary, after adjustment: the need for IV antibiotics was predicted by all criteria except A (severe respiratory distress), hospital admission was predicted by all criteria except criterion D (respiratory exhaustion) due to insufficient

numbers, need for supplemental oxygen was predicted by all except criterion E (clinical dehydration), all criteria except B (increased respiratory rate) were independently associated with the need for IV fluid, and lastly only criteria C (oxygen saturation ≤92%), D(respiratory exhaustion) and G (causing other clinical concern) independently predicted in-hospital mortality. Results are presented in Table 4.6.

4.4.7 Multivariable analysis: model I including all criteria irrespective of significance in adults (> 16years)

The results for this analysis are presented in Table 4.7, and unlike in the younger patient group, no one CATs criterion independently predicted all study outcomes in adults (≥ 16years). Criterion C (oxygen saturation ≤92%) independently positively predicted all outcomes, except the need for IV fluid and supplemental oxygen (removed a priori). Criterion G (causing clinical concern) in this patient group had an independent positive association with all outcomes except need for IV fluid and was the strongest positive predictor of admission [OR (95%CI): 8.37 (2.27-30.91)] and IV antibiotic use [OR (95%CI): 7.37 (2.73-19.90)]. Criterion B (increased respiratory distress) had a negative independent association with in-hospital mortality while positively independently predicting all other outcomes except need for IV fluid. Criterion A (severe respiratory distress) only independently predicted admission and supplemental oxygen while Criterion E (clinical dehydration) was not an independent predictor of any of the outcomes. Criterion F (altered consciousness) which was only independently associated with supplemental oxygen [OR (95%CI): 9.12 (1.47-8.86)] was the strongest predictor after criterion A (severe respiratory distress) [OR (95%CI): 9.17 (3.43-24.55)].

Although criterion D (respiratory exhaustion) had sufficient entries for the outcome hospital admission in the adult group it was not an independent predictor of the need for admission, it however, positively and independently predicted the need for supplemental oxygen and was the only independent predictor for IV fluid [OR (95%CI): 3.26 (1.18-8.89)].

In summary, after adjustment, the need for hospital admission was independently predicted by criteria A (severe respiratory distress), B (increased respiratory distress) and G (causing clinical concern). The need for supplemental oxygen administration was predicted by all the criteria except criteria E (clinical dehydration) and C (oxygen saturation $\leq$92%) which was removed a priori. IV antibiotic administration was only predicted by criteria B (increased respiratory distress), C (oxygen saturation $\leq$92%) and G (causing clinical concern), while in-hospital mortality was only independently predicted by criteria C (oxygen saturation $\leq$92%) and G (causing clinical concern). Lastly, the need for IV fluid administration was predicted by criterion D (respiratory exhaustion) alone.

Table 4.6: Multivariable analysis (model I) to determine independent predictors of acute care interventions and mortality in children (n=809)

CATS Criteria	Admission	Supplemental oxygen	IV antibiotic	IV fluid	In-hospital mortality
A: severe resp. distress					
No	1.00	1.00	1.00	1.00	1.00
Yes	3.80 (2.47-5.85) **ǂ**	1.77 (1.12-2.78) **†**	2.08 (1.44-3.02) NS	2.49 (1.73-3.58) ǂ	1.14 (0.47-2.77) NS
B: increased resp. rate					
No	1.00	1.00	1.00	1.00	1.00
Yes	2.11 (1.25-3.56) **†**	2.05 (1.15-3.66) **†**	1.60 (1.01-2.54) †	1.15 (0.74-1.79) NS	1.51 (0.57-3.97) NS
C: oxygen sat. ≤ 92%					
No	1.00	1.00	1.00	1.00	1.00
Yes	8.11 (3.43-19.15) **ǂ**	X	3.86 (2.43-6.14) ǂ	1.47 (1.00-2.17) **†**	8.36 (3.98-17.54) ǂ
D: resp. exhaustion					
No	1.00	1.00	1.00	1.00	1.00
Yes	*****	5.13 (2.27-11.62) ǂ	5.40 (1.62-17.95) **†**	0.43 (0.26-0.72) **†**	2.26 (1.26-4.05) **†**
E: clinical dehydration					
No	1.00	1.00	1.00	1.00	1.00
Yes	11.39 (1.51-86.05) †	1.41 (0.77-2.56) NS	2.73 (1.24-5.99) †	5.69 (2.41-13.41) ǂ	1.36 (0.66-2.80) NS
F: altered consciousness					
No	1.00	1.00	1.00	1.00	1.00
Yes	14.90 (2.00-110.92) **†**	2.31 (1.35-3.96) **†**	1.88 (1.01-3.50) **†**	1.44 (1.32-2.82) **†**	1.74 (0.94-3.21) NS
G: other clinical concern					
No	1.00	1.00	1.00	1.00	1.00
Yes	3.07 (0.25-5.49) **†**	2.50 (1.65-3.78) **ǂ**	2.00(1.31-3.06) †	1.93 (1.32-2.82) †	2.95 (1.73-5.02) ǂ

NS=p value >0.05; †=p<0.05 and ǂ<0.001;

*** Criterion had insufficient numbers to calculate OR**

Table 4.7: Multivariable analysis (model I) to determine independent predictors of acute care interventions and mortality in adults (n=207)

CATS Criteria	Admission	Supplemental oxygen	IV antibiotic	IV fluid	In-hospital mortality
A: severe resp. distress **No** **Yes**	1.00 3.69 (1.68-8.09) **†**	1.00 9.17 (3.43-24.55) **ǂ**	1.00 1.74 (0.82-3.73) NS	1.00 1.45 (0.74-2.83) NS	1.00 3.49 (0.61-19.97) NS
B: increased resp. rate **No** **Yes**	1.00 2.81 (1.32-5.99) **†**	1.00 3.69 (1.54-8.87) **†**	1.00 2.68 (1.28-5.58) **†**	1.00 1.79 (0.92-3.48) NS	1.00 0.14 (0.03-0.71) †
C: oxygen sat. ≤ 92% **No** **Yes**	1.00 8.12 (2.51-26.20) ǂ	X	1.00 3.88 (1.71-8.79) **†**	1.00 1.72 (0.83-3.57) NS	1.00 68.26 (5.99-778.35) †
D: resp. exhaustion **No** **Yes**	1.00 3.08 (0.50-18.90) NS	1.00 5.83 (2.03-16.78) †	1.00 2.48 (0.76-8.13) NS	1.00 3.26 (1.18-8.98) **†**	1.00 0.73 (0.19-2.84) NS
E: clinical dehydration **No** **Yes**	1.00 0.52 (0.09-3.10) NS	1.00 0.93 (0.28-3.12) NS	1.00 2.44 (0.54-10.96) NS	1.00 2.58 (0.79-8.44) NS	1.00 0.80 (0.19-3.37) NS
F: altered consciousness **No** **Yes**	1.00 2.07 (0.17-24.45) NS	1.00 9.12 (1.47-8.86) **†**	1.00 1.13 (0.25-5.02) NS	1.00 0.80 (0.24-2.62) NS	1.00 3.11 (0.80-12.09) NS
G: causing other clinical concern **No** **Yes**	1.00 8.37 (2.27-30.91) †	1.00 3.60 (1.47-8.86) †	1.00 7.37 (2.73-19.90) ǂ	1.00 1.36 (0.64-2.88) NS	1.00 7.84 (2.18-28.19) †

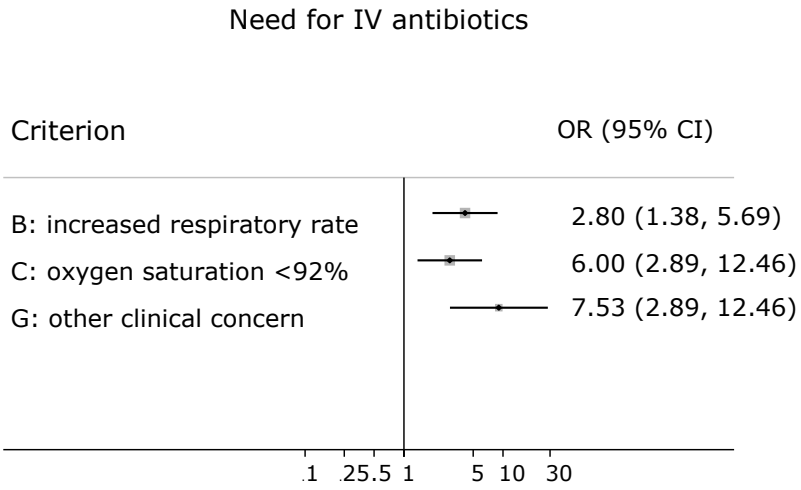
NS=p value >0.05; †=p<0.05; ǂ<0.001;
X = criterion directly predicts this outcome

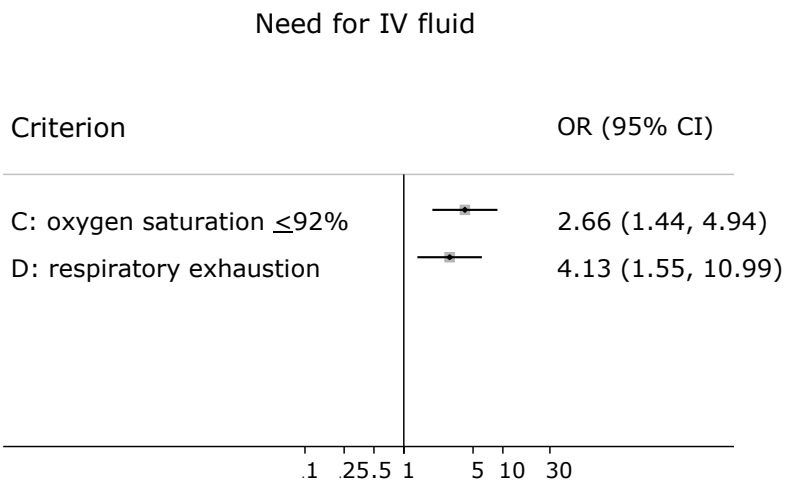
4.4.8 Multivariable analysis model II: forward stepwise selection in adults

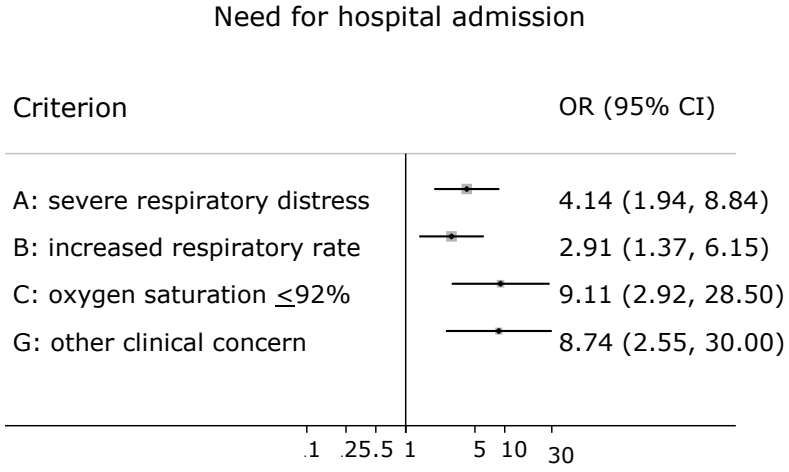
The independent predictors of the study outcomes (need for hospital admission, acute care interventions and in-hospital mortality) using forward stepwise selection in the older patient group ($\geq$16 years) are presented in Fig 3.

The likelihood of study participants in the adult category requiring intravenous antibiotics, supplemental oxygen and hospital admission were predicted by the same CATs criteria regardless of which multivariable regression approach was used. Criteria B (increased respiratory rate), C (oxygen saturation $\leq$ 92%) and G (causing other clinical concern) independently predicted the need for intravenous antibiotics; and these same criteria also predicted the need for hospital admission along with criterion A (severe respiratory distress). The 'need for supplemental oxygen' was independently predicted by criterion A (severe respiratory distress), B (increased respiratory rate), D (respiratory exhaustion), F (altered consciousness) and G (causing other clinical concern). Criteria C (oxygen saturation $\leq$ 92%) was the only predictor of in-hospital mortality common to both regression models, other predictors of this outcome were Criterion D (respiratory exhaustion) in model II and G (causing clinical concern) in model I. Lastly, while criterion D was the only independent predictor of IV fluid requirement in the first model which included all CATs criteria, Criteria C (oxygen saturation $\leq$ 92%) was also associated with this outcome using forward stepwise selection.

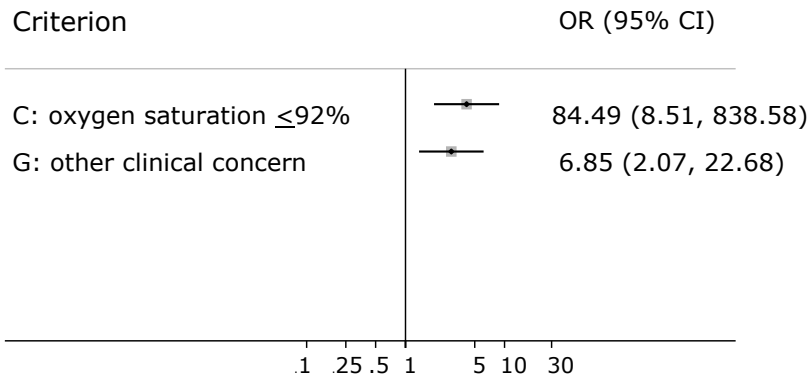
Figure 3: Summary of multivariable analysis model II using forward stepwise selection approach in adults



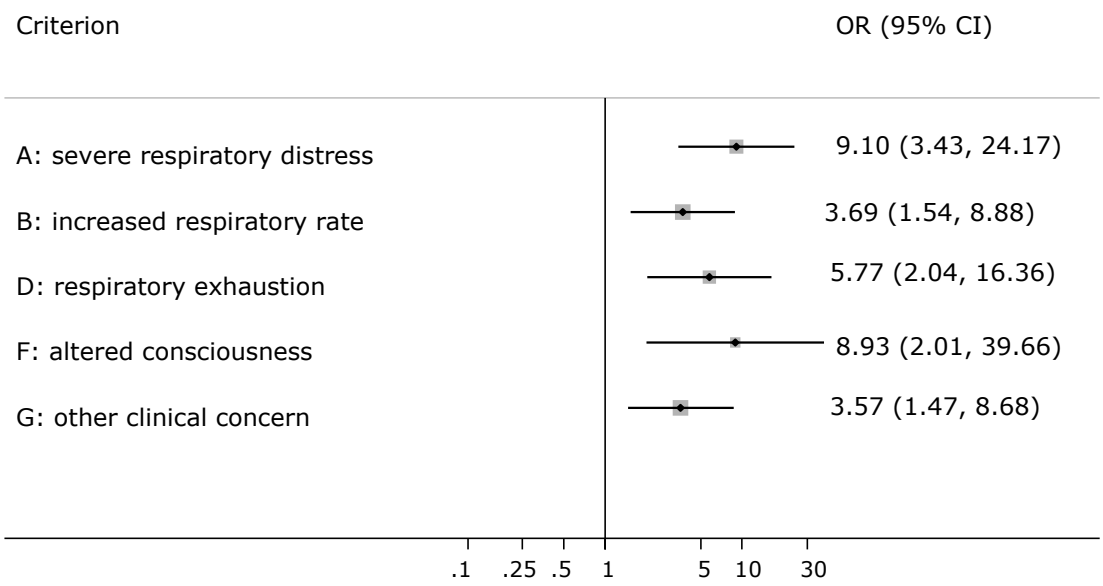




In-patient mortality



Need for supplemental oxygen



4.4.9 Multivariable model II: forward stepwise selection in children

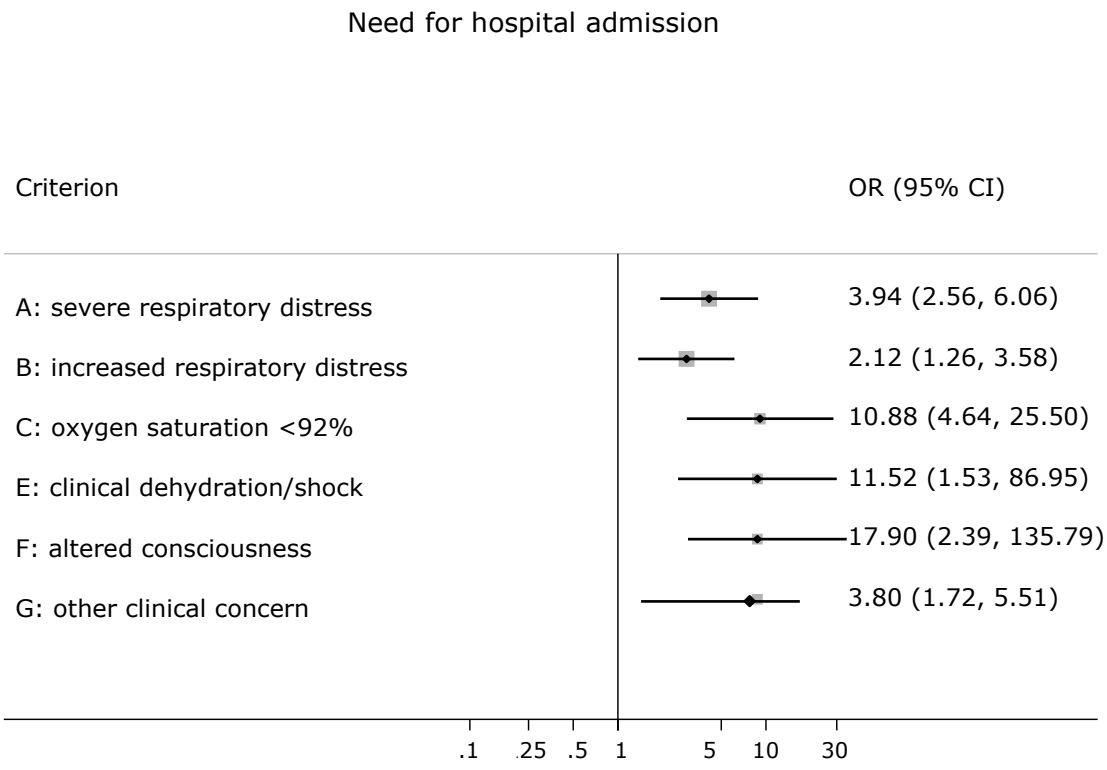
The CATs criteria which independently predicted admission and in-hospital mortality using a forward stepwise approach (model II) in the younger (< 16 years) study participants were similar to the results obtained using the approach in model I. Admission was independently predicted by all the CATs criteria except criterion D (respiratory exhaustion) which had insufficient numbers for the analysis, while in-hospital mortality was only predicted by criteria C (oxygen saturation $\leq$92%), D (respiratory exhaustion) and G (causing other clinical concern).

There were only slight differences between the two models in terms of predicting the need for supplemental oxygen and intravenous antibiotics ;with the exclusion of criterion C (oxygen saturation $\leq$92%) from both models, all CATs criteria in model II predicted the need for supplemental oxygen administration, while criterion E (clinical dehydration) was the only non-predictor in model I. Using a stepwise approach (model II), need for intravenous antibiotics was independently predicted by all criteria except criterion B (increased respiratory rate), while model I included all the criteria except A(severe respiratory distress).

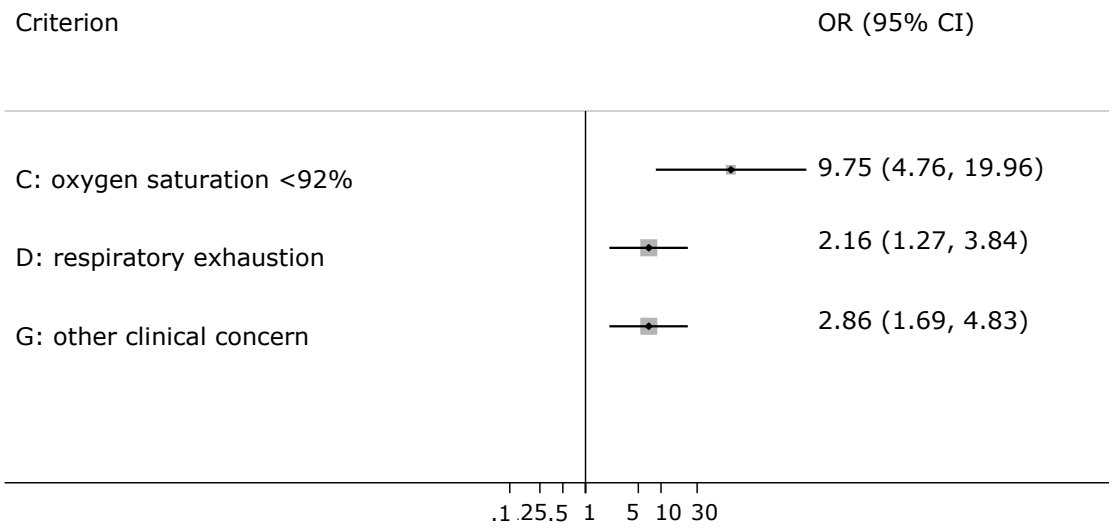
The major differences between the two multivariable modelling approaches were seen for outcome IV fluid, which was only predicted by criteria A (severe respiratory distress), E (clinical dehydration) and G (causing clinical concern) using the forward stepwise approach (model II), compared to model I where all criteria except criteria B (increased respiratory rate) and D (respiratory

exhaustion) independently predicted this outcome. These findings are presented separately for each outcome in figure 4.

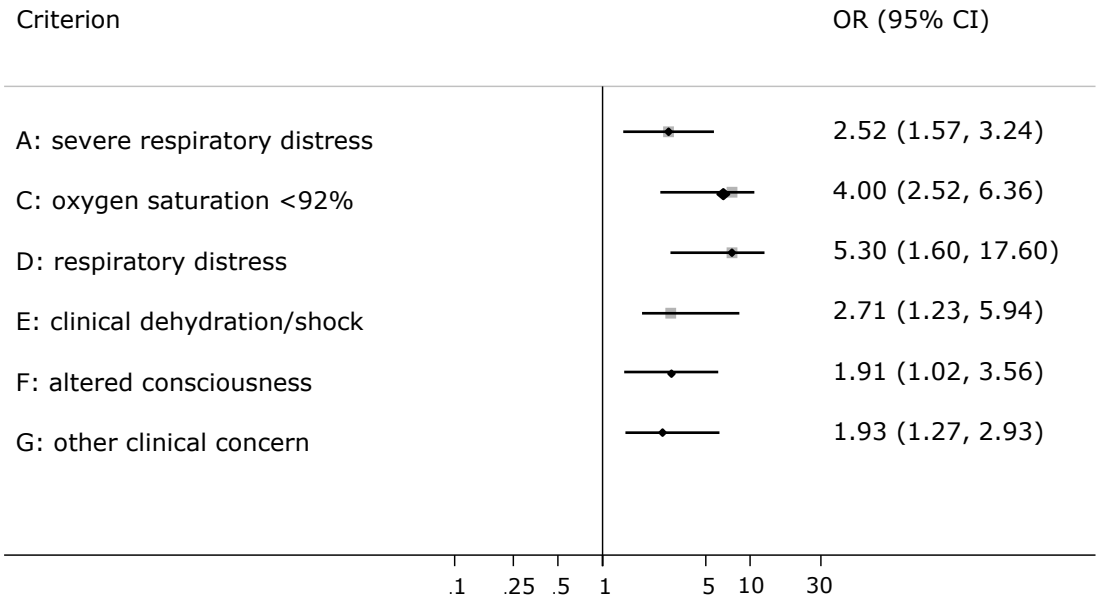
Figure 4: Summary of multivariable analysis model II using forward stepwise selection approach in children



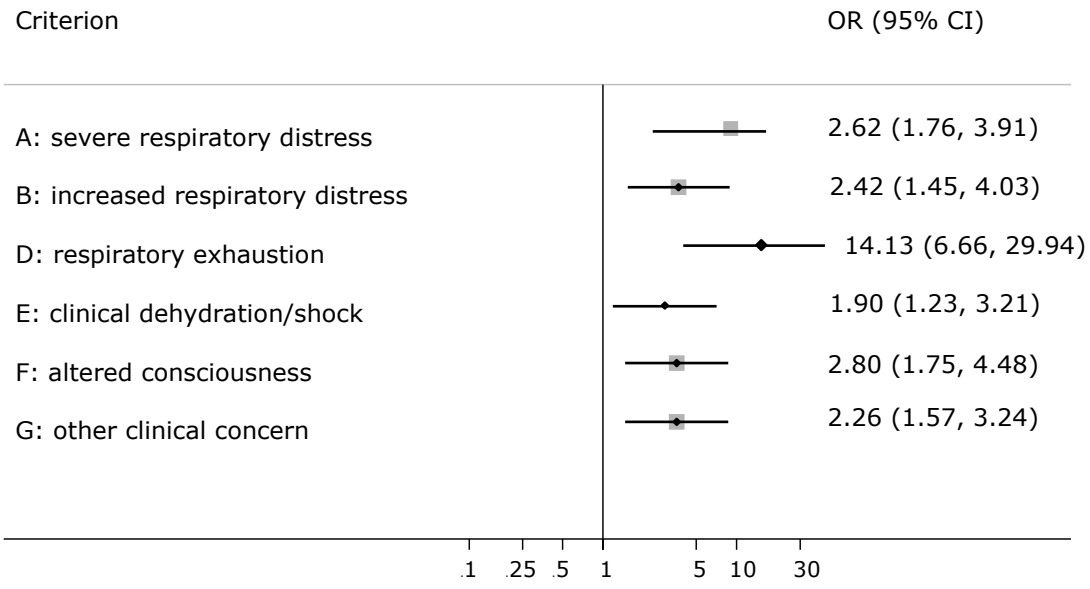
In-patient mortality



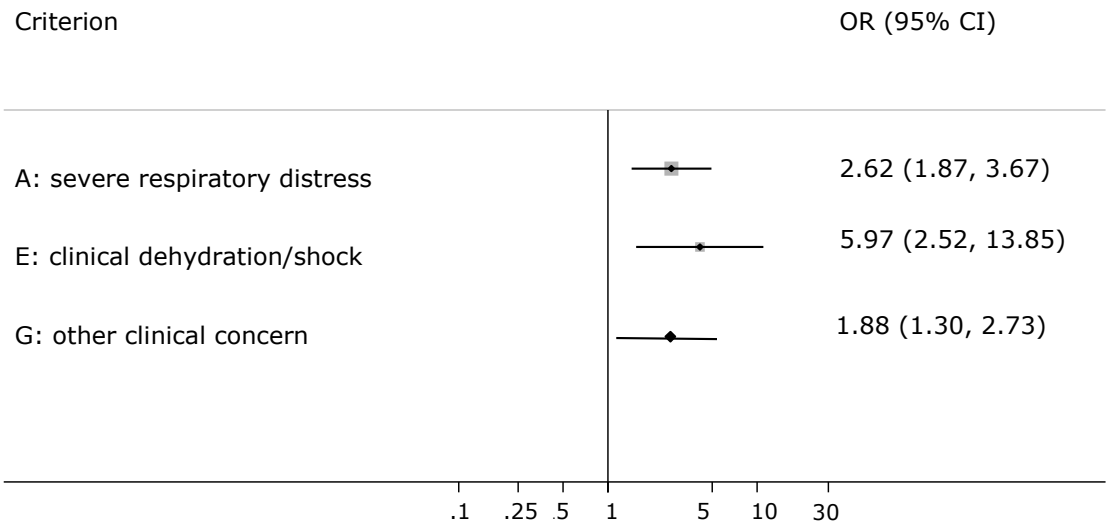
Need for IV antibiotics



Need for supplemental oxygen



Need for IV fluid



4.4.10 Predictive performance of individual CATs criteria for various outcomes in children (< 16 years)

The predictive performance of each CATs criterion in terms of their individual area under receiver operating characteristic, sensitivity, specificity, positive and negative predictive values have been summarised for the younger study population in table 4.8.

The AUROC values for the CATs criteria were generally poor, with values as low as 0.49 and none over 0.80. Criterion D had the lowest AUROC value of 0.49 while C (oxygen saturation $\leq$ 92%) had the highest at 0.77. Overall, 'Severe respiratory distress' criterion A (0.58-0.71) and 'oxygen saturation $\leq$92%' criterion C (0.56-0.77) had the highest range of AUROC values across the five study outcomes.

Sensitivity values were generally high across all the study outcomes for criteria A and B ranging between 81.7% -91.8%, compared to sensitivity values for the other CATs criteria which were mostly below 50%.

On the other hand, while criteria C to G had better specificity (67%- 100%) for predicting all the outcomes, criteria A and B had the lowest range of values for specificity (15.4%-58.8%). The negative predictive values of the criteria were generally low for all study outcomes except 'in-hospital mortality' (91.6% - 98%) and 'need for supplemental oxygen' (57.8% - 77.2%). The reverse was the case for positive predictive values as the CATs criteria were seen to have better values for outcomes: need for admission (90% -100%), intravenous antibiotics (74.1% -97.1%) and fluids (63.7% - 92.8%), while values for 'need for supplemental oxygen' and particularly in-hospital mortality and were not as high.

Table 4.8: Predictive values of individual CATs criterion for predicting admission, acute care interventions and in-hospital mortality in children (n=809)

OUTCOMES	CATs CRITERION	AUROC (95%CI)	SENSITIVITY (%) (95%CI)	SPECIFICITY (%) (95%CI)	PPV (%) (95%CI)	NPV (%) (95%CI)
Admission	A	**0.71 (0.66-0.75)**	82.6% (79.5%-85.4%)	58.8% (50.4%-66.8%)	90% (87.3%-92.2%)	43.1% (36.1%-50.2%)
	B	0.59 (0.55-0.63)	88.7% (86%-91%)	29.7% (22.5%-37.8%)	84.9% (82%-87.5%)	37% (28.3%-46.3%)
	C	**0.71 (0.68-0.73)**	45.4% (41.5%-49.3%)	95.5% (91.4%-98.5%)	98% (95.5%-99.3%)	28.2 %(24.3%-32.4%)
	D	0.58 (0.56-0.59)	15.4% (12.8%-18.4%)	100% (97.5%-100%)	100% (96.4%-100%)	20.9% (18%-24.1%)
	E	0.56 (0.54-0.57)	12.4% (10%-15.2%)	99.3% (96.3%-100%)	98.8% (93.5%-100%)	20.2% (17.4%-23.4%)
	F	0.58 (0.57-0.60)	17.2% (14.4%-20.3%)	99.3% (96.3%-100%)	99.1% (95.3%-100%)	21.2% (18.2%-24.4%)
	G	0.58 (0.55-0.62)	29.8% (26.3%-33.5%)	87.2% (80.7%-92.1%)	91.2% (86.6%-94.6%)	21.8% (18.5%-25.3%)
Supplemental oxygen	A	**0.62 (0.59-0.65)**	87.8% (84%-90.9%)	36% (31.5%-40.7%)	54.4% (50.3%-58.4%)	77.2% (70.8%-82.8%)
	B	0.56 (0.53-0.58)	91.5% (88.2%-94.1%)	20.1% (16.4%-24.2%)	49.9% (46.1%-53.7%)	73.1% (64.2%-80.8%)
	D	**0.62 (0.59-0.64)**	25% (20.7%-29.7%)	98.2% (96.4%-99.2%)	92.2% (85.1%-96.6%)	60.1% (56.4%-63.7%)
	E	0.54 (0.52-0.56)	14.4% (11%-18.3%)	93.3% (90.5%-95.5%)	65.1% (53.8%-75.2%)	55.6% (51.9%-59.3%)
	F	0.57 (0.55-0.60)	22.1% (18%-26.6%)	92.6% (89.7%-94.9%)	72.2% (63%-80.1%)	57.8% (54%-61.5%)
	G	0.58 (0.55-0.61)	34.8 %(30%-39.9%)	80.4% (76.3%-84%)	60.6% (53.8%-67.2%)	58.7% (54.6%-62.7%)
IV antibiotic	A	0.63 (0.59-0.67)	82.5% (79.1%-85.5%)	43.5% (37.1%-50.2%)	78.4% (74.9%-81.6%)	50% (42.9%-57.1%)
	B	0.56 (0.53-0.59)	88.6% (85.7%-91%)	22.8% (17.6%-28.8%)	74.1% (70.6%-77.3%)	44.5% (35.4%-53.9%)
	C	**0.68 (0.65-0.71)**	48.4% (44.2%-52.5%)	88.4% (83.5%-92.3%)	91.2% (87.4%-94.1%)	40.8% (36.4%-45.2%)

	D	0.58 (0.56-0.60)	17.2% (14.2%-20.5%)	98.7% (96.3%-99.7%)	97.1% (91.6%-99.4%)	32.4% (28.9%-36%)
	E	0.55 (0.53-0.57)	13% (10.4%-16%)	96.6% (93.3%-98.5%)	90.4% (81.9%-95.7%)	30.9% (27.5%-34.4%)
	F	0.56 (0.54-0.58)	17.5% (14.5%-20.9%)	94% (90.1%-96.7%)	87.8% (80.4%-93.2%)	31.4% (28%-35%)
	G	0.57 (0.54-0.60)	30.5% (26.8%-34.4%)	82.8% (77.3%-87.4%)	81.5% (75.6%-86.4%)	32.4% (28.6%-36.3%)
IV fluid	A	**0.60 (0.57-0.64)**	81.7% (78.2%-84.8%)	38.8% (32.9%-45%)	73.5% (69.8%-76.9%)	50.5% (43.4%-57.6%)
	B	0.53 (0.50-0.55)	87% (83.9%-89.7%)	18.35 (13.8%-23.5%)	68.8% (65.2%-72.3%)	40.35 (31.4%-49.7%)
	C	0.56 (0.53-0.60)	41.9% (37.8%-46.2%)	70.7% (64.8%-76.2%)	74.8% (69.6%-79.6%)	37% (32.7%-41.4%)
	D	0.49 (0.46-0.51)	11.9% (9.3%-14.9%)	85.9% (81.1%-89.9%)	63.7% (53.6%-73%)	32% (28.5%-35.5%)
	E	0.56 (0.54-0.58)	14.15 (11.3%-17.3%)	97.7% (95.1%-99.25)	92.8% (84.9%-97.3%)	35.4% (31.9%-39%)
	F	0.53 (0.51-0.56)	16.5% (13.5%-19.9%)	90.5% (86.3%-93.8%)	78.3% (69.6%-85.4%)	34.3% (30.8%-38%)
	G	0.56 (0.53-0.59)	30.6% (26.7%-34.6%)	81.4% (76.1%-85.9%)	77.3% (71.1%-82.7%)	36.1% (32.2%-40.1%)
In-hospital mortality	A	0.58 (0.55-0.62)	90.4% (81.2%-96.1%)	26.5% (23.3%-29.8%)	10.9% (8.5%-13.6%)	96.5% (93%-98.6%)
	B	0.54 (0.50-0.57)	91.8% (83%-96.9%)	15.4% (12.8%-18.2%)	9.7% (7.6%-12.2%)	95% (89.3%-98.1%)
	C	**0.77 (0.72-0.81)**	86.3% (76.2%-93.2%)	67% (63.5%-70.4%)	20.6% (16.2%-25.6%)	98% (96.4%-99%)
	D	0.64 (0.58-0.70)	38.4% (27.2%-50.5%)	89.9% (87.5%-92%)	27.5% (19.1%-37.2%)	93.6% (91.6%-95.3%)
	E	0.53 (0.49-0.58)	16.4% (8.8%-27%)	90.4% (88%-92.4%)	14.5% (7.7%-23.9%)	91.6% (89.3%-93.5%)
	F	0.57 (0.52-0.63)	27.4% (17.6%-39.1%)	87.1% (84.5%-89.4%)	17.4% (11%-25.6%)	92.4% (90.1%-94.2%)
	G	0.63 (0.57-0.69)	50.7% (38.7%-62.6%)	75.7% (72.4%-78.7%)	17.1% (12.4%-22.8%)	93.9% (91.7%-95.7%)

4.4.11 Predictive performance of individual CATs criteria for various outcomes in adults

AUROC values of the CATs criteria in adults were averagely low for all outcomes (0.53-0.86) and within a similar range as values seen in children. CATs A (severe respiratory distress) and C (oxygen saturation $\leq$ 92%) were the best predictors of outcomes: admission (0.72; 0.74), need for supplemental oxygen (0.73; 0.86), IV antibiotics (0.65; 0.72), and IV fluids (0.62; 0.64), while CATs C and G had the highest values for in-hospital mortality (0.82; 0.70).

Although criteria A (severe respiratory distress) and B (increased respiratory rate) appeared to be the most sensitive of the CATs criteria, criteria C (oxygen saturation $\leq$ 92%) had the highest sensitivity (94.7%) for predicting 'in-hospital mortality' in adults. Specificity values were similar to the values in children with criteria A and B having the lowest range of values for all five study outcomes.

Lastly, while positive predictive values of the CATs criteria were lowest for in-hospital mortality (10.5%-24.2%), and highest for 'admission' (78.2% - 94.8%), negative predictive values were highest for in-hospital mortality (91.5%-99.2%) and lowest for admission (37.3% -59.5%). These results have been presented in Table 4.9.

Table 4.9: Predictive values of individual CATs criteria for predicting admission, acute care interventions and in-hospital mortality in adults (n=207)

OUTCOMES	CATs CRITERION	AUROC (95%CI)	SENSITIVITY (%) (95%CI)	SPECIFICITY (%) (95%CI)	PPV (%) (95%CI)	NPV (%) (95%CI)
Admission	**A**	0.72 (0.66-0.79)	74.8% (66.6%-81.9%)	69.4% (57.5%-79.8%)	82.1% (74.2%-88.4%)	59.5% (48.3%-70.1%)
	B	0.68 (0.62-0.75)	77% (69%-83.8%)	59.7% (47.5%-71.1%)	78.2% (70.2%-84.9%)	58.1% (46.1%-69.5%)
	C	**0.74 (0.69-0.79)**	54.1% (45.3%-62.7%)	94.4% (86.4%-98.5%)	94.8% (87.2% - 98.6%)	52.3% (43.4%-61.1%)
	D	0.59 (0.55-0.63)	20.7% (14.2%-28.6%)	97.2% (90.3%-99.7%)	93.3% (77.9%-99.2%)	39.5% (32.3%-47.2%)
	E	0.55 (0.51-0.59)	14.1% (8.7%-21.1%)	86.4% (65.1%-97.1%)	86.4% (65.1%-97.1%)	37.3% (30.3%-44.7%)
	F	0.56 (0.52-0.59)	12.6% (7.5%-19.4%)	98.6% (92.5%-100%)	94.45 (72.7%-99.9%)	37.6% (30.6%-44.9%)
	G	0.62 (0.58-0.67)	30.4% (22.8%-38.9%)	94.4% (86.4%-98.5%)	91.1% (78.8%-97.5%)	42% (34.3%-50%)
Supplemental oxygen	**A**	**0.73 (0.68-0.79)**	89.3% (80.1%-95.3%)	57.6% (48.7%-66.1%)	54.5% (45.2%-63.5%)	90.5% (82.1%-95.8%)
	B	0.67 (0.61-0.72)	85.3% (75.35-92.4%)	47.7% (39%-56.6%)	48.1% (39.4%-56.9%)	85.1% (75%-92.3%)
	D	0.64 (0.58-0.69)	32% (21.7%-43.8%)	95.5% (90.4%-98.3%)	80% (61.4%-92.3%)	71.25 (63.9%-77.7%)
	E	0.57 (0.52-0.62)	20% (11.6%-30.8%)	94.7% (89.4%-97.8%)	68.2% (45.1%-86.1%)	67.6% (60.3%-74.3%)
	F	0.59 (0.54-0.64)	20% (11.6%-30.8%)	97.8% (93.5%-99.5%)	83.3% (58.6%-96.4%)	68.3% (61.1%-74.8%)
	G	0.61 (0.55-0.67)	36% (25.2%-47.9%)	86.4% (79.3%-91.7%)	60% (44.3%-74.3%)	70.4% (62.7%-77.3%)
IV antibiotic	**A**	0.65 (0.59-0.72)	74.5% (65.1%-82.5%)	56.4% (46.2%-66.3%)	64.2% (55.1%-72.7%)	67.9% (56.8%-77.6%)
	B	0.66 (0.60-0.73)	80.2% (71.3%-87.3%)	52.5% (42.3%-62.5%)	63.9% (55.1%-72.1%)	71.6% (59.9%-81.5%)
	C	**0.72 (0.66-0.78)**	58.5% (48.5%-68%)	85.1% (76.7%-91.4%)	80.5% (69.9%-88.7%)	66.2% (57.3%-74.2%)

	D	0.59 (0.55-0.64)	23.6% (15.9%-32.8%)	95% (88.8%-98.4%)	83.3% (65.3%-94.4%)	54.25 (46.6%-61.7%)
	E	0.57 (0.53-0.62)	17.9% (11.2%-26.6%)	97% (91.6%-99.4%)	86.4% (65.1%-97.1%)	53% (45.5%-60.3%)
	F	0.56 (0.52-0.59)	14.2% (8.1%-22.3%)	97% (91.6%-99.4%)	83.3% (58.6%-96.4%)	51.9% (44.5%-59.2%)
	G	0.64 (0.59-0.70)	35.8% (26.8%-45.7%)	93.1% (86.2%-97.2%)	84.4% (70.5%-93.5%)	58% (50%-65.7%)
IV fluid	**A**	0.62 (0.55-0.68)	71.9% (61.8%-80.6%)	51.4% (41.7%-61%)	56.1% (46.9%-65%)	67.9% (56.8%-77.6%)
	B	0.61 (0.55-0.67)	76% (66.35-84.2%)	45.9% (36.4%-55.7%)	54.9% (46%-63.5%)	68.9% (57.1%-79.2%)
	C	**0.64 (0.57-0.70)**	52.2% (41.6%-62.4%)	75.7% (66.6%-83.3%)	64.9% (53.2%-75.5%)	64.6% (55.8%-72.8%)
	D	0.60 (0.55-0.65)	25% (16.7%-34.9%)	94.6% (88.6%-98%)	80% (61.4%-92.3%)	59.3% (51.7%-66.6%)
	E	0.57 (0.52-0.61)	17.7% (10.7%-26.8%)	95.5% (89.8%-98.5%)	77.3% (54.6%-92.2%)	57.3% (49.8%-64.5%)
	F	0.53 (0.49-0.57)	11.5% (5.9%-19.6%)	93.7% (87.4%-97.4%)	61.1% (35.7%-82.7%)	55% (47.6%-62.3%)
	G	0.55 (0.49-0.61)	27.1% (18.5%-37.1%)	82.9% (74.6%-89.4%)	57.8% (42.2%-72.3%)	56.8% (42.2%-72.3%)
In-hospital mortality	**A**	0.67 (0.59-0.74)	89.5% (66.9%-98.7%)	43.6% (36.4%-51%)	13.8% (8.3%-21.2%)	97.6% (91.7%-99.7%)
	B	0.55 (0.44-0.66)	73.7% (48.8%-90.9%)	36.7% (29.8%-44%)	10.5% (5.9%-17%)	93.2% (84.9%-97.8%)
	C	**0.82 (0.76-0.88)**	94.7% (74%-99.9%)	68.6% (61.5%-75.2%)	23.4% (14.5%-34.4%)	99.2% (95.8%-100%)
	D	0.54 (0.44-0.63)	21.1% (6.1%-45.6%)	86.2% (80.4%-90.8%)	13.3% (3.8%-30.7%)	91.5% (86.4%-95.2%)
	E	0.59 (0.48-0.69)	26.3% (9.1%-51.2%)	91% (85.9%-94.6%)	22.7% (7.8%-45.4%)	92.4% (87.6%-95.8%)
	F	0.63 (0.52-0.73)	31.6% (12.6%-56.6%)	93.6% (89.1%-96.7%)	33.3% (13.3%-59%)	93.1% (88.5%-96.3%)
	G	0.70 (0.58-0.82)	57.9% (33.5%-79.7%)	81.9% (75.5%-87.1%)	24.2% (12.9%_39.5%)	95.1% (90.5%-97.8%)

4.5 Discussion

4.5.1 Key findings

This study has investigated the associations between each CATs criterion and important interventions and outcomes associated with the management of acute lower respiratory tract infections. The findings from this study provide evidence of the generalisability of the CATs to the Nigerian context by showing which criteria were significantly associated with and capable of predicting the risk of mortality and need for hospital care in both adults and children with a range of ALRTIs.

The overall predictive performance of the CATs criteria in terms of AUROC values was generally poor across the study outcomes in both adults and children. Criterion A (severe respiratory distress) and criterion C (oxygen saturation <92%) had the highest AUROC values for all outcomes in children and adults, except in-hospital mortality where criterion G (causing clinical concern) had the best AUROC value.

4.5.2 Strengths and limitations

One of the key strengths of this study lies in its methodological approach, which involved using well defined inclusion and exclusion criteria to collect data from consecutive patients presenting to multiple hospitals representing the different tiers of the health system in Nigeria. This multicentre approach provides a reasonably strong justification for the study's findings being generalisable to the wider population of patients with ALRTI presenting to other Nigerian hospitals. A further argument in favour of this study's strengths, is the

prospective approach to data collection which helped minimise occurrences of missing data and selection bias.

On the other hand, using clinical signs for diagnosing acute lower respiratory tract infections without microbiological or radiographic diagnostic tests confirmation might be considered as a possible limitation of this study. However, according to recommendations made by the American Thoracic Society (ATS) and Infectious Disease Society of America (IDSA) joint guidelines for managing CAP in adults (115); chest radiographs taken at the initial stages of pneumonia are mostly inconclusive and make the distinction between the pneumonia and other ALRTI in adults quite difficult. Furthermore, two separate studies conducted in developing countries (Nigeria and Pakistan) reported that chest x-rays did not offer any significant improvements over clinical signs and symptoms with regard to the diagnosis of pneumonia in children (116, 117). With regards to microbiological tests, the joint IDSA/ATS guidelines also state that these investigations have low specificity for diagnosing pneumonia and tend to result in false negatives(115). Asides from these diagnostic uncertainties, the major issue with subjecting every patient with a suspicion of ALRTI to diagnostic tests is the cost implication in low resource settings, especially as lack of funds may delay the process of getting results in time to make urgent clinical decisions. Therefore the choice of using highly sensitive clinical features rather than diagnostic tests for diagnosing ALRTI in this study presents a pragmatic and accurate reflection of real life practice in low resource or heavily burdened healthcare settings.

Finally, although the effect sizes reported were statistically significant ($p < 0.05$), most of the 95% confidence values were quite broad; as such, caution is advised when interpreting the strength of association between the criteria

and the outcomes as these confidence intervals suggest that a larger sample size could have increased the precision of the findings.

4.5.3 Comparison with other studies

This is the first external validation of the community assessment tools; and the findings of this study were encouraging in terms of establishing the generalisability of the CATs criteria. The main difference between this study and the other CATs studies conducted in the United Kingdom was in the patient cases studied. While the UK studies included confirmed cases of pandemic influenza (FLU-CIN cohort) (35, 37, 39, 118), the study population in Nigerian included patients with pneumonia, acute bronchitis, acute exacerbation of COPD and bronchiolitis in children. Most prediction tools tend to perform differently and sometimes disappointingly when validated in new subjects (35, 113, 119); however, this was not the case for the CATs, as the criteria maintained their ability to predict important interventions and outcomes in the management of children and adults with ALRTI in a new population despite differences in setting and patients.

Overall, criteria C (oxygen saturation $\leq$92%) and G (causing clinical concern) were the most common independent predictors of all outcomes in both adults and children, more importantly they were the only criteria to predict mortality in both patient groups. This finding is quite important in regard to the use of CATs in low resource settings, where clinicians often face several diagnostic uncertainties when assessing patients without the adequate investigations. Pulse oximeters which are used in measuring oxygen saturation are less expensive than chest radiographs, and low oxygen saturation levels have been

shown to be an accurate predictor of radiographic pneumonia (35, 120, 121). As hospitals in LMIC lack the resources to perform chest x-rays on every patients, they stand to benefit from the use of pulse oximeters which are less expensive. This study's findings on criterion G (causing other clinical concern) are in agreement with those reported in the UK-CATs study by Semple et al., which also reported that this criterion was associated with mortality in children and severe outcomes in adults(35). As the risk of under-treating is worse than over-treating, this finding could empower clinicians to rely on their gut feelings rather than on blind algorithmic compliance with prediction rules which may devalue the process of clinical reasoning. However, as this is only apparent evidence of the validity of this criterion, the study conducted in the seventh chapter of this thesis, will investigate the actual value of clinical intuition by comparing the performance of a combined CATs score with criterion G vs one without .

It was not possible to assess the performance of criterion D in this study as very few children met the criterion for respiratory exhaustion. While this was only in relation to predicting the need for hospital admission in this present study, the UK study lacked sufficient numbers for this criterion to perform the analysis across all outcomes in children. The mechanism of respiratory exhaustion or respiratory muscle fatigue is an important clinical feature which indicates a patient's progression towards ventilatory failure, however more research is needed to establish how best to diagnose it and how frequently it occurs (122). In this present study, there were 102/809 (12%) observations for respiratory exhaustion in participants aged below 16 years compared to almost 15% in the adult population ($\geq$16 years). A probable explanation for why respiratory exhaustion was more frequently observed in the adult patients

than in children is that children are able to mask hypoventilation by maintaining an appearance of normal breathing even while in respiratory distress(123). Respiratory exhaustion is a late sign, and it is likely that patients in this present study may have been more presented much sooner than in a pandemic when the patients are more likely to be severely ill at presentation, however it is more than a coincidence that very few patients met this criterion in two independently conducted studies. If the CATs is to have face validity and be successfully adopted in practice, it is important that it includes the most clinically relevant and routinely used clinical parameters; criterion D may be an important clinical sign for identifying a deteriorating patient, but the reported missing values suggest that it is not used very often in practice (even across geographical settings) and the exclusion of this criterion from the CATs should be considered.

The criterion for clinical dehydration/shock (criterion E) did not predict any of the outcomes in adults while altered consciousness (criterion F) only predicted the need for supplemental oxygen. These criteria were seen to perform better in children, predicting all study outcomes except mortality. Semple et al. (2013) reported that clinical dehydration (criterion E) and altered consciousness (criterion F) were both independent predictors of severe outcomes and the need for high levels of care in both adults and children(35). In developing countries, most childhood illness are complicated by dehydration, making it one of the top reasons for hospital admission, also confusion of recent onset is a major indication of severe illness in the older population, which could also be as a result of sepsis or dehydration such that the poor performance of these criteria may be as a result of less severe cases in the Nigerian adult population compared to the patients with pandemic influenza in the UK data.

It is therefore worth retaining these criteria for further validation in a new population with more severe cases of ALRTI.

Severe respiratory distress -criterion A and increased respiratory rate –criterion B, each had a role in predicting one or more of the study outcomes particularly the need for hospital admission, supplemental oxygen and intravenous antibiotics in both adults and children. It was also reported in the dissertation study which evaluated the predictive performance of the individual CATs criteria using FLU-CIN data, that criterion A predicted both severe outcomes and the need for higher levels of care in children; and criterion B also predicted these outcomes in adults(39). These findings are in agreement with WHO recommendations which advise the use of elevated respiratory rate and lower chest wall indrawing in developing countries for the diagnosis of severe ALRTI and pneumonia in children(124). When these recommendations were evaluated by Mulholland et al. (1992) in both developing and developed countries, these authors reported that the sensitivities of both clinical signs remained the same across settings(125).

The reported AUROC values of the criteria in these study were within the range of 0.49 to 0.77 in children and 0.53 to 0.82 in adults which suggests low to moderate levels of predictive accuracy. Similar AUROC values (0.52 -0.73 in adults and 0.52-0.68 in children) were reported in the dissertation study using FLU-CIN data;, however when the criteria were combined into weighted and unweighted scores in this study, the predictive performance of the CATs improved which suggests that the initial premise to use the CATs as stand-alone for assessing severity of illness may be flawed (39). This thesis will also investigate the validity of these findings in the next chapter which compares the performance of combined CATs scores with the individual criteria.

4.5.4 Summary

In summary, this study provides preliminary evidence of the generalisability of the CATs criteria to patients in a new geographical setting, however their use as stand-alone predictors may not be ideal as individual AUROC values indicated poor to moderate accuracy. Respiratory exhaustion does not appear to be a routinely recorded clinical finding in both adults and children; as such the CATs should be revised to exclude this criteria as it is unlikely to be used during time constrained assessments which occur during pandemics. This study also provides further evidence supporting the use of clinical intuition and oxygen saturation levels in the assessment of patients of all ages with respiratory symptoms.

5: Assessing the performance of the CATs as a combined scoring index for predicting hospital admission, acute care interventions and in-hospital mortality in adults and children with ALRTI

5.1 Introduction

This chapter describes the creation of two composite clinical scores using the CATs criteria and goes on to compare their performances with regards to predicting the need for hospital care and interventions as well as in-hospital mortality in adults and children with acute lower respiratory tract infections. It also includes a comparison of the predictive performance of the newly created scoring indexes with that of the individual CATs criteria, and then concludes by discussing the pros and cons of using the criteria as single predictors or as a combined scoring index for assessing patients with ALRTI.

5.2 Background

The individual CATs criteria have each been shown to independently predict important outcomes associated with the management as well as the prognosis of acute lower respiratory tract infections in adults and children in the previous chapter. However, in terms of predictive accuracy, the AUROC values of the individual criteria tell a different story; most of the criteria had AUROC values which fell within the range of 0.5 to 0.79, signifying moderate to no discriminatory value at all(126).These findings therefore suggest that as

standalone predictors, the criteria were not good at discriminating between people with the outcomes (i.e. in-patient mortality or admission), from people without.

A plausible explanation for this is that even when diagnosed with the same medical condition patients tend to vary in terms of presenting clinical features, as well as illness aetiology and trajectory thus making it highly impractical for doctors to prognosticate using a singular criterion or predictor variable (6). By combining individual prognostic variables into scoring indexes, doctors are better able to categorise patients using appropriate thresholds into risk groups which ensure that the patients get interventions best suited to them; a phenomenon known as 'risk-tailored care'(5, 127).

There are no hard and fast rules when it comes to creating scoring indexes as authors of studies on prediction models have been seen to adopt different methods. Some authors have created unweighted scoring systems by simply awarding equal points to each predictor (21, 128), while others have assigned weighted points using the odds ratios (129), or regression coefficients from the multivariable models (130, 131). Although weighting appears to be more logical given that some variables will be prioritised over others based on severity at presentation and knowledge of the subject area , it is actually quite complex to achieve, which makes it less popular among physicians(132).

The performance of the CATs criteria as unweighted and weighted scores have been assessed previously using UK data, the aim of this study therefore is to compare the performance of the CATs criteria as un-weighted and weighted scoring indexes in a new population of patients. Additionally this study will compare the two combined scoring indexes with the individual CATs criteria to determine which have better predictive performance.

5.3 Statistical analyses

This study was conducted using prospectively collected data on 1016 children and adults with symptoms suggestive of ALRTI, who presented to the emergency department of four hospitals in Lagos, Nigeria between December 2013 and July 2014. The data collection process has been previously described in chapter 3; study outcomes which were recorded during follow-up included hospital admission, administration of supplemental oxygen, intravenous antibiotics or fluids and death on admission described as in-patient mortality.

Before creating the scoring indices, a test for multicollinearity was conducted to investigate the degree of correlation existing between the CATs criteria; the more correlated variables are in a regression model, the more unstable the coefficient estimates(85, 133). Multicollinearity is most popularly investigated using the variance inflation factor (VIF) and tolerance (1/VIF) values, and as the name suggests VIF tells us how much the variance of the regression coefficient has been inflated by the collinearity between the predictor variables; as a rule of thumb, VIF and tolerance values should generally be below 10 and above 0.1 respectively (88, 89, 133).

In order to create the unweighted CATs scoring indices (children and adults), a single point was allotted to each of the seven criteria and the total score for each patient was calculated by summing up the points for all criteria present. The predictive potential of the unweighted CATs scoring index was investigated by conducting a logistic regression analysis including the scoring index as a continuous variable to determine if outcome prediction by this scoring system followed a linear trend, i.e. a higher likelihood of an adverse outcome with higher scores. Results are presented separately for each scoring index as odd ratios (ORs) with 95% confidence intervals (95%CI) and p values. Because the

unweighted CATs index was included in the models as a continuous predictor variable, the 'linearity assumption' of logistic regression which requires that a linear relationship exists between the independent variable and the logit transformation of the outcome was assessed using Lowess algorithm (94, 134, 135). Lowess curves showing the relationship between the continuous CATs index and each outcome are presented in appendix X.

For the weighted CATs index, individual weights were created and assigned by multiplying the regression coefficients of each criterion from the fully adjusted multivariable model (Model I as described in chapter 4) by ten, and rounding to the nearest whole number. Unlike the unweighted CATs index where the criteria were assumed to be equal and assigned a point each (irrespective of the outcome), the points assigned to each criterion in the weighted index varied with each outcome and the range of score values were different for each outcome.

In prognostic research, the two most commonly used qualities for describing the performance of predictive models are 'discrimination' and 'calibration' (5, 56), and for the purpose of this analysis receiver operating characteristic (ROC) curves were used to assess and compare the discriminatory ability of the unweighted and weighted combined CATs indices (as continuous variables) for each of the outcomes. Results were presented separately for children and adults using AUROC values with 95% confidence intervals (95%CI). In addition to this, the calibration of the predictions of these combined scores i.e. how well the outcome predictions of the scoring index in a development sample match predictions in a new sample (5, 56) was tested using the Hosmer-Lemeshow goodness-of-fit test by splitting the data into two separate equal halves to represent a hypothetical development and validation sample. The overall fit of

the regression models was also tested using the Hosmer-Lemeshow goodness of fit test which uses the principles of chi square testing to assess the closeness of the values in the model. A significant p-value will indicate that predicted values are significantly different from the observed values also meaning that the model does not fit the data and vice versa if the p-values were insignificant. The sole premise behind using prediction rules or scores in clinical practice is to improve patients' outcomes by aiding clinicians in decision making. This is usually achieved by using such scores as classifiers to put the patients in categories which reflect their actual risk of having an adverse outcome (136). In order to further boost the utility of the CATs as a clinical decision rule, each combined CATs scoring index was divided into three tiers using suitable threshold score values which would distinguish between alternative outcomes in patients. This was achieved in the unweighted CATs index by creating separate binary variables using cut-off values of $\geq$3 (yes/no), $\geq$4(yes/no), $\geq$5(yes/no), just like in the UK CATs comparative validation study by Myles et al.(2012) (119). In the case of the weighted CATs score, three scoring tiers for each outcome were derived from the median and interquartile range of the weights of the individual criteria, i.e. the lower tier represented scores above the 25th percentile, the middle tier included scores above the median, and scores above the 75th percentile made up the upper tier. This approach was used because the range of potential score values of the weighted CATs score were seen to vary for each outcome, such that even the threshold scores for each of the study outcomes also differed between children and adults. Histograms showing the distribution of the un-weighted and weighted scores, median (IQR) are presented in Appendices IV, V and VI.

The performance of the individual threshold scores of the combined CATs indexes for each of the study outcomes were compared using ROC curves, area under receiver operating characteristic curve (AUROC), sensitivity, specificity, negative and positive predictive values. All analyses were conducted using Stata version 12 (StataCorp LP).

5.4 Results

5.4.1 Assessment of multicollinearity between CATs criteria

Given that the composite variance inflation factor (VIF) and tolerance values for the individual CATs criterion will remain the same when assessed in relation to the different outcomes, only values for a single outcome (in-patient mortality) have been presented in table 5.1. The degree of multicollinearity between the CATs criteria is negligible as per widely accepted threshold values (VIF <10 or tolerance > 0.1).

Table 5.1: Assessment of multi-collinearity between CATs criteria and in-patient mortality

CATS Criteria	Adults ≥ 16years		Children < 16years	
	Variance inflation factor (VIF)	Tolerance	Variance inflation factor (VIF)	Tolerance
Criterion A: severe respiratory distress	1.27	0.7889	1.16	0.8606
Criterion B: increased respiratory rate	1.18	0.8488	1.08	0.9249
Criterion C: oxygen saturation ≤ 92%	1.50	0.6645	1.29	0.7723
Criterion D: respiratory exhaustion	1.17	0.8521	1.19	0.8424
Criterion E: clinical dehydration	1.22	0.8181	1.03	0.9726
Criterion F: altered consciousness	1.22	0.8202	1.04	0.9594
Criterion G: causing other clinical concern	1.09	0.9187	1.02	0.9814

Mean VIF for adults = 1.24 Mean VIF for children = 1.12

5.4.2 Assessment of discriminatory ability of un-weighted CATs index

Table 5.2 shows the unadjusted odds ratios and 95% confidence intervals for the association between the unweighted combined CATs scores and the study outcomes in both children and adults. All the P values (P trend <0.001), and 95% confidence intervals (95%CI) for the odds ratios suggest the presence of a linear trend in outcome prediction i.e. every unit increase on the scoring scale increases the likelihood of an adverse outcome.

Although these findings are akin to simple logic that a patient with a higher score will have an increased likelihood of an adverse outcome compared with

a patient with a lower score, the validity of the linearity of logit assumption in the regression models is further reinforced by the Lowess plots presented in Appendices VII and VIII. In both patient groups, the smoothed curves appear fairly linear with only slight plateauing towards the higher scores thus showing that a positive relationship exists between the unweighted combined CATs scores and the outcomes.

Table 5.2: Logistic regression analyses involving unweighted CATs scoring indexes as continuous variables in adults and children

Outcomes	Un-weighted adults CATs score OR (95% CI) P trend	Un-weighted children CATs score OR (95% CI) P trend
Hospital admission	3.55 (2.47 – 5.10) <0.001	4.03 (3.16 – 5.13 <0.001
Supplemental oxygen	4.62 (3.09 – 6.91) <0.001	3.49 (2.93 -4.17) <0.001
Intravenous antibiotics	2.66 (2.02 – 3.49) <0.001	2.43 (2.07 – 2.85) <0.001
Intravenous fluid	1.69 (1.38 – 2.07) <0.001	1.49 (1.32 – 1.68) <0.001
In-patient mortality	2.21 (1.56 – 3.15) <0.001	2.58 (2.07 – 3.23) <0.001

5.4.3 Comparing of the predictive performance of un-weighted and weighted CATs combined scores

The AUROC values comparing the unweighted and weighted scoring indices in children and adults are presented in Tables 5.3 and 5.4 respectively. The two competing CATs indices had similar performance in predicting all five outcomes in children and adults as their AUROC values were within the same range. The ROC curves for children and adults displayed in Figures 5 and 6 also show similarities between the unweighted and weighted indices as the lines

representing the two scoring systems are seen to overlap for all outcomes. However, the AUROC values comparing the unweighted and weighted scores presented in Tables 5.3 and 5.4 and the p-values suggest that the weighted CATs index was significantly better than the unweighted CATs at predicting supplemental oxygen, intravenous antibiotics and in-patient mortality in adults and all outcomes in children.

Table 5.3: Comparison of AUROC values of unweighted and weighted CATs scores derived in children (n=809)

	CATs Scores	AUROC (95%CI)	P-values
Admission	Unweighted CATs	0.85 (0.82-0.88)	
	Weighted CATs	0.86 (0.83-0.89)	**0.0394**
Suppl. oxygen	Unweighted CATs	0.83 (0.80-0.86)	
	Weighted CATs	0.86 (0.83-0.88)	**<0.0001**
IV antibiotic	Unweighted CATs	0.77 (0.73-0.80)	
	Weighted CATs	0.78 (0.75-0.81)	**0.0203**
IV fluid	Unweighted CATs	0.65 (0.61-0.69)	
	Weighted CATs	0.67 (0.63-0.71)	**0.0039**
In-patient mortality	Unweighted CATs	0.81 (0.76-0.85)	
	Weighted CATs	0.84 (0.79-0.88)	**0.0081**

Significant p-values have been highlighted

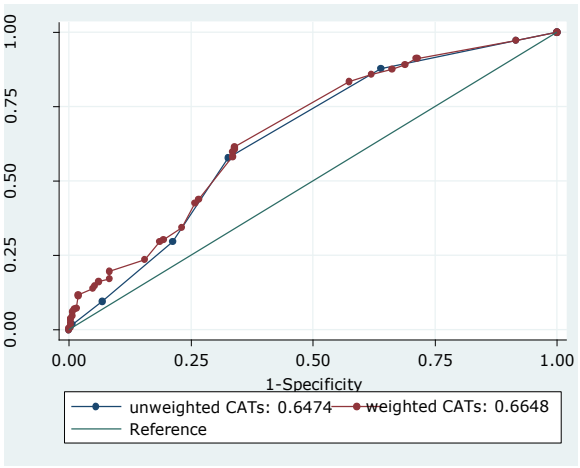
Table 5.4: Comparison of AUROC values of unweighted and weighted CATs scores derived in adults

	CATs Scores	AUROC (95%CI)	P-values
Admission	Unweighted CATs	0.86 (0.81-0.91)	
	Weighted CATs	0.87(0.82-0.92)	0.5809
Supplemental oxygen	Unweighted CATs	0.92 (0.88-0.95)	
	Weighted CATs	0.94 (0.91-0.97)	**0.0155**
IV antibiotic	Unweighted CATs	0.83 (0.77-0.88)	
	Weighted CATs	0.85 (0.79-0.90)	**0.0325**
IV fluid	Unweighted CATs	0.72(0.65-0.79)	
	Weighted CATs	0.73 (0.66-0.80)	0.3630
In-patient mortality	Unweighted CATs	0.82 (0.74-0.91)	
	Weighted CATs	0.87 (0.80-0.93)	**0.0478**

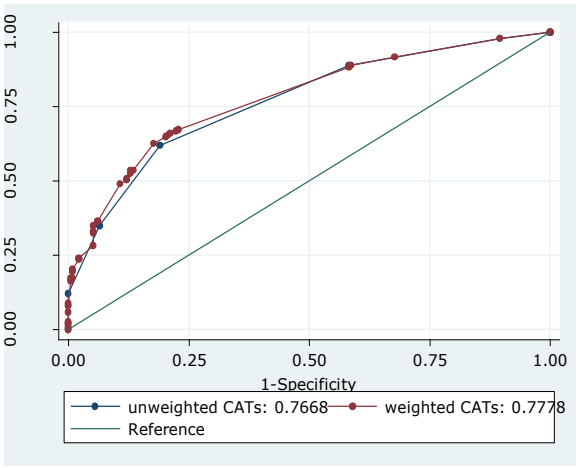
Significant p-values have been highlighted

Figure 5: Comparison of ROC curves of unweighted and weighted CATs scores in children

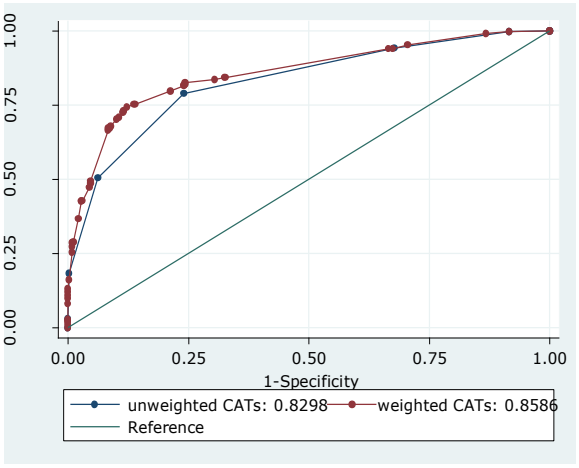
IV fluids



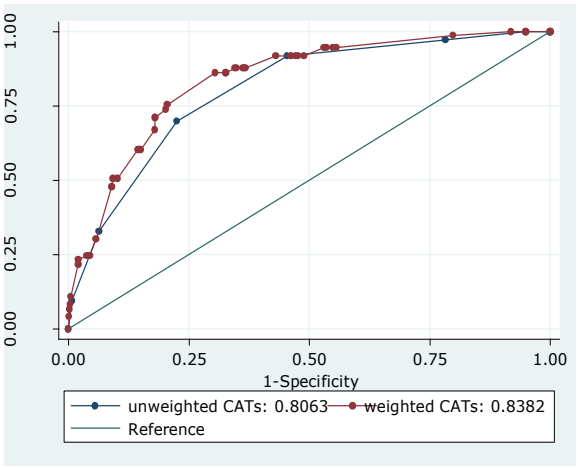
IV antibiotics



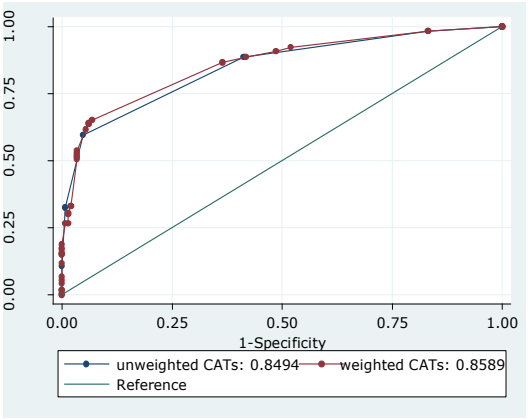
Supplemental oxygen



In-patient mortality



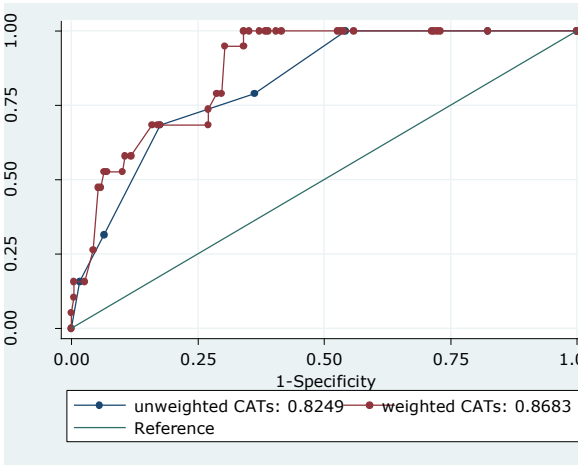
Hospital admisison



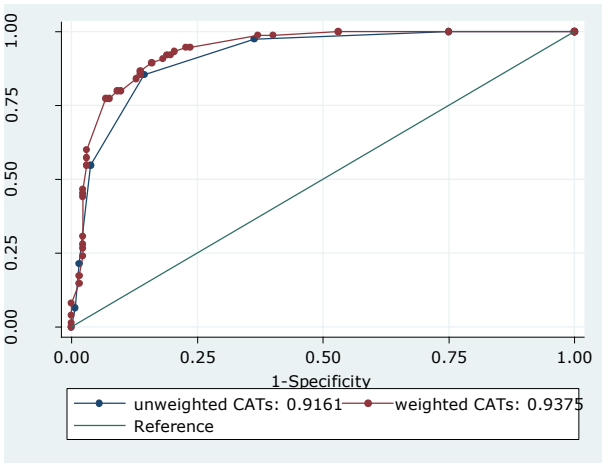
Unweighted CATs	Blue line
Weighted CATs	Red line
Reference	Light green line

Figure 6: Comparison of ROC curves of unweighted and weighted CATs scores in adults

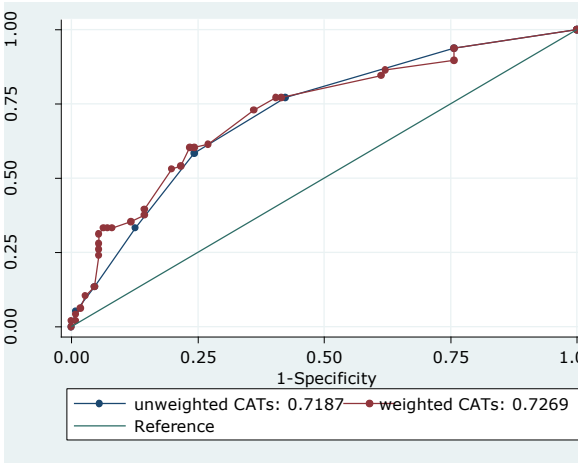
In-patient mortality



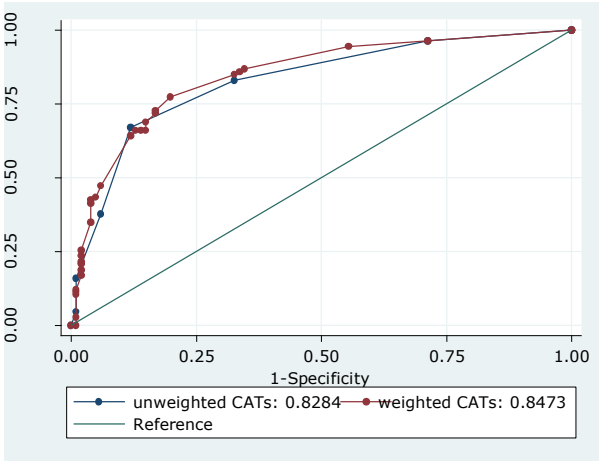
Supplemental oxygen



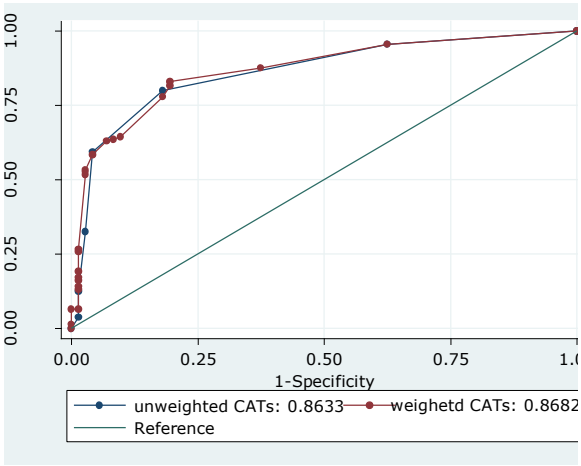
IV fluid



IV antibiotics



Hospital admission



Unweighted CATs	Blue line
Weighted CATs	Red line
Reference	Light green line

5.4.4 Assessing the goodness of fit and calibration of the unweighted and weighted combined CATs regression models in children and adults

According to the P values presented in table 5.5, the regression models for the un-weighted and weighted CATs scores in adults had a good fit as the p-values for all outcomes were non-significant (>0.05). On the other hand, the models in children were a good fit except for those evaluating the un-weighted score with regard to the outcome supplemental oxygen (0.0160) and those evaluating the combined CATs scores (both unweighted and weighted indices) for the outcome intravenous fluid.

Table 5.5: Hosmer-Lemeshow goodness of fit test for assessing the regression models of unweighted and weighted CATs combined scores in children and adults

Outcomes	Un-weighted CATs (P-value) Children	Un-weighted CATs (P-value) Adults	weighted CATs (P-value) Children	weighted CATs (P-value) Adults
Admission	0.2377	0.0538	0.4849	0.0554
Supplemental Oxygen	0.0160*	0.1368	0.6883	0.1221
Intravenous antibiotic	0.2344	0.4611	0.7256	0.7321
Intravenous fluid	0.0000*	0.6524	0.0000*	0.4832
In-hospital mortality	0.4087	0.1658	0.7229	0.8835

*P-values <0.005 signifying model with poor fit

According to tables 5.6 (a, b) and 5.7 (a, b), when the study sample was split in half, the performance of both unweighted and weighted scores in the first column (development sample) matched the findings in table 5.5. When compared with predictions in the second column (validation sample) the predictions of both unweighted and weighted scores for all the outcomes in adults were well calibrated (p value >0.05). On the other hand the unweighted and weighted CATs in children were poorly calibrated for predicting the need for both intravenous antibiotics and fluids, in adults and children (Tables 5.7 a, b). In addition to this the un-weighted CATs was not well calibrated for predicting the need for supplemental oxygen administration in children (Table 5.6a).

Table 5.6a: Hosmer-Lemeshow test for calibration of un-weighted CATs index in children

outcomes	Development sample p-values	Validation sample p-values
Admission	0.5316	0.2362
Supplemental oxygen	0.0910	**0.0187**
Intravenous antibiotic	0.7037	**0.0261**
Intravenous fluid	0.0004	**0.0008**
In-patient mortality	0.5139	0.7049

Table 5.6b: Hosmer-Lemeshow test for calibration of weighted CATs index in children

Outcomes	Development sample p-values	Validation sample p-values
Admission	0.0455	0.3040
Supplemental oxygen	0.3757	0.2807
Intravenous antibiotic	0.3801	**0.0281**
Intravenous fluid	0.0047	**0.0130**
In-patient mortality	0.3069	0.4768

Table 5.7a: Hosmer-Lemeshow test for calibration of un-weighted CATs index in adults

outcomes	Development sample p-values	Validation sample p-values
Admission	0.0345	0.4979
Supplemental oxygen	0.3178	0.7992
Intravenous antibiotic	0.7496	0.0620
Intravenous fluid	0.6148	0.0689
In-patient mortality	0.4726	0.4926

Table 5.7b: Hosmer-Lemeshow test for calibration of weighted CATs index in adults

Outcomes	Development sample p-values	Validation sample p-values
Admission	0.1859	0.4497
Supplemental oxygen	0.5828	0.1526
Intravenous antibiotic	0.0953	0.6707
Intravenous fluid	0.9383	0.2150
In-patient mortality	0.9806	0.6894

5.4.5 Comparison of the predictive performance of CATs as individual criteria and as combined scoring indexes

The range of AUROC values and 95% CI for the individual CATs criteria, alongside the values for each of the combined scoring indexes have been presented according to the study outcomes in Tables 5.8 and 5.9 for children and adults respectively. Outcome prediction was generally better when the CATs were used as either an un-weighted or weighted index compared which suggests that the performance of the individual criteria was inferior to that of using CATs as a combined scoring index.

Table 5.8: comparison of the predictive performance of individual CATs criteria, un-weighted and weighted CATs combined scores in children

Outcomes	Range of AUROC values for individual CATs criteria (95% CI)	AUROC values for un-weighted combined CATs scores (95%CI)	AUROC values for weighted combined CATs scores (95%CI)
Admission	0.56 (0.54-0.57) – 0.71 (0.66-0.75)	0.81 (0.76-0.85)	0.84 (0.79-0.88)
Supplemental oxygen	0.54 (0.52-0.56) – 0.62 (0.59-0.65)	0.83 (0.80-0.86)	0.86 (0.83-0.88)
Intravenous antibiotics	0.55 (0.53-0.57) – 0.68 (0.75-0.71)	0.77 (0.73-0.80)	0.78 (0.75-0.81)
Intravenous fluids	0.49 (0.46-0.51) – 0.60 (0.57-0.64)	0.65 (0.61-0.69)	0.67 (0.63-0.71)
In-patient mortality	0.53 (0.49-0.58) – 0.64 (0.58-0.70)	0.81 (0.76-0.85)	0.84 (0.79-0.88)

Table 5.9: comparison of predictive performance of individual CATs criteria, un-weighted and weighted CATs combined scores in adults

Outcomes adults	Range of AUROC values for individual CATs criteria (95% CI)	AUROC values for un-weighted combined CATs scores (95%CI)	AUROC values for weighted combined CATs scores (95%CI)
Admission	0.55 (0.51-0.59) – 0.74 (0.69-0.79)	0.86 (0.81-0.91)	0.87 (0.82-0.92)
Supplemental oxygen	0.57 (0.52-0.62) – 0.73 (0.68- 0.79)	0.92 (0.88-0.95)	0.94 (0.91-0.97)
Intravenous antibiotics	0.56 (0.52-0.59) – 0.72 (0.66-0.78)	0.83 (0.77-0.88)	0.85 (0.79-0.90)
Intravenous fluids	0.53 (0.49-0.57) – 0.64 (0.57-0.70)	0.72 (0.65-0.79)	0.73 (0.66-0.80)
In-patient mortality	0.54 (0.44-0.63) – 0.82 (0.76-0.88)	0.82 (0.74-0.91)	0.87 (0.80-0.93)

5.4.6 Predictive performance of un-weighted CATs threshold scores for predicting study outcomes in children and adults

The predictive performance of the threshold scores for the unweighted CATs index in terms of AUROC, sensitivity, specificity, positive and negative predictive values are presented in Tables 5.10 and 5.11 for children and adults respectively. In terms of overall predictive accuracy of the unweighted CATs index, the threshold score of $\geq$ 3 appeared to be the best predictor of need for hospital admission, supplemental oxygen, IV antibiotics and fluids in both adults and children, while a CATs score of $\geq$ 4 was the best predictor of in-patient mortality in adults [AUROC (95% CI: (0.75 (0.64-0.87)] and children [AUROC (95% CI: (0.74 (0.68-0.79)].

The sensitivity and negative predictive values (NPV) appeared to be lower than specificity and positive predictive values (PPV) in both adults ($\geq$16 years) and

children (<16 years). CATs score threshold of $\geq$ 3 had the best sensitivity and negative predictive values in both adults and children, while an unweighted thresholds score of $\geq$5 had the best specificity for predicting all the outcomes in both patient groups. Except for outcomes 'need for IV fluids' in children and 'hospital admission' in adults, positive predictive values were seen to improve with increase in score threshold.

The Receiver Operating Characteristic (ROC) curves for children and adults displayed in Figures 7 and 8 respectively show that the lines representing unweighted CATs threshold score of $\geq$3 are farthest from the diagonal for all outcomes except for 'in-patient mortality' where threshold score of $\geq$4 had the best performance .

Table 5.10: Predictive performance of combined un-weighted CATs score thresholds for predicting admission, acute care interventions and in-hospital mortality in children

OUTCOMES	SCORE THRESHOLD	AUROC (95%CI)	SENSITIVITY (%) (95%CI)	SPECIFICITY (%) (95%CI)	PPV (%) (95%CI)	NPV (%) (95%CI)
Admission	**CAT ≥3**	**0.77 (0.75-0.80)**	**59.6% (55.8%-63.4%)**	95.3% (90.5%-98.1%)	98.3% (96.4%-99.3%)	**34.6% (29.9%-39.4%)**
	CAT ≥4	0.66 (0.64-0.68)	32.7% (29.1%-36.4%)	99.3% (96.3%-100%)	99.5% (97.5%-100%)	24.8% (21.4%-28.5%)
	CAT ≥5	0.55 (0.54-0.56)	10.6% (8.3%-13.2%)	**100% (97.5%-100%)**	**100% (94.9%-100%)**	20% (17.2%-23.1%)
Suppl. oxygen	C**AT ≥3**	**0.77 (0.75-0.80)**	**79% (74.5%-83%)**	76% (71.7%-79.9%)	74.1 %(69.5%-78.3%)	**80.6% (76.5%-84.4%)**
	CAT ≥4	0.72 (0.69-0.75)	50.5% (45.4%-55.7%)	93.8% (91.1%-95.9%)	87.6% (82.4%-91.6%)	68.6% (64.7%-72.3%)
	CAT ≥5	0.59 (0.57-0.61)	18.4% (14.6%-22.6%)	**99.8% (98.7%-100%)**	**98.6% (92.3%-100%)**	58.5% (54.8%-62%)
IV antibiotic	**CAT ≥3**	**0.71 (0.68-0.75)**	**61.9% (57.8%-65.9%)**	81% (75.4%-85.9%)	89% (85.6%-91.9%)	**46.1% (41.2%-51.1%)**
	CAT ≥4	0.64 (0.62-0.67)	35% (31.1%-39.1%)	93.5% (89.6%-96.3%)	93.1% (88.9%-96.1%)	36.7% (32.8%-40.7%)
	CAT ≥5	0.56 (0.55-0.57)	12.1% (9.6%-15.1%)	**100% (98.4%-100%)**	**100% (94.9%-100%)**	31.4% (28.1%-34.9%)
IV fluid	**CAT ≥3**	**0.63 (0.59-0.66)**	**57.9% (53.6%-62.1%)**	67.7% (61.7%-73.3%)	**78.8% (74.5%-82.7%)**	**43.6% (38.8%-48.6%)**
	CAT ≥4	0.54 (0.51-0.57)	29.5% (25.7%-33.5%)	78.7% (73.3%-83.5%)	74.2% (67.8%-79.9%)	35% (31.1%-39%)
	CAT ≥5	0.51 (0.49-0.53)	9.5% (7.2%-12.3%)	**93.2% (89.4%-95.5%)**	74.3% (62.4%-84%)	33.2% (29.8%-36.7%)
In-hospital mortality	CAT ≥3	0.73 (0.70-0.77)	**91.8% (83%-96.9%)**	54.6% (50.9%-58.3%)	16.7% (13.2%-20.7%)	**98.5% (96.8%-99.5%)**
	CAT ≥4	**0.74 (0.68-0.79)**	69.9% (58%-80.1%)	77.4% (74.3%-80.5%)	23.5% (18%-29.7%)	96.3% (94.4%-97.7%)
	CAT ≥5	0.63 (0.58-0.69)	32.9% (22.3%-44.9%)	**93.8% (91.8%-95.4%)**	**34.3% (23.3%-46.6%)**	93.4% (91.3%-95.1%)

The best predictive performance (AUROC, NPV, PPV, sensitivity, specificity) values for each outcome have been highlighted for the score thresholds

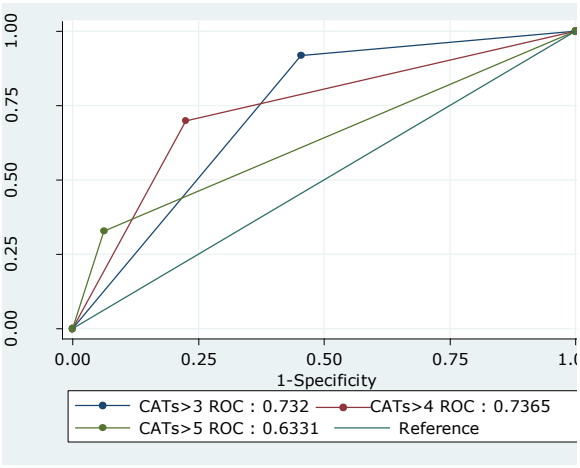
Table 5.11: Predictive performance of combined un-weighted CATs score thresholds for predicting admission, acute care interventions and in-hospital mortality in adults

OUTCOMES	SCORE THRESHOLD	AUROC (95%CI)	SENSITIVITY (%) (95%CI)	SPECIFICITY (%) (95%CI)	PPV (%) (95%CI)	NPV (%) (95%CI)
Admission	CAT ≥3	**0.78 (0.73-0.82)**	**59.3% (50.5%-67.6%)**	95.8% (88.3%-99.1%)	**96.45 (89.8%-99.2%)**	**55.6% (46.5%-64.6%)**
	CAT ≥4	0.65 (0.61-0.69)	32.6% (24.85-41.2%)	97.2% (90.3%-99.7%)	95.7% (85.2%-99.5%)	43.5% (35.7%-51.5%)
	CAT ≥5	0.56 (0.52-0.59)	12.6% (7.5%-19.4%)	**98.6% (92.5%-100%)**	94.4% (72.7%-99.9%)	37.6% (30.6%-44.9%)
Suppl. oxygen	CAT ≥3	**0.85 (0.80-0.90)**	**85.3% (75.3%-92.4%)**	85.6% (78.4%-91.1%)	77.1% (66.6%-85.6%)	**91.1% (84.7%-95.5%)**
	CAT ≥4	0.75 (0.70-0.81)	54.7% (42.7%-66.2%)	96.2% (91.4%-98.8%)	89.1% (76.4%-96.4%)	78.9% (71.8%-84.9%)
	CAT ≥5	0.60 (0.55-0.65)	21.3% (12.7%-32.3%)	**98.5% (94.6%-99.8%)**	**88.9% (65.3%-98.6%)**	68.8% (61.7%-75.3%)
IV antibiotic	CAT ≥3	**0.78 (0.72-0.83)**	**67% (57.2%-75.8%)**	88.1% (80.2%-93.7%)	85.5% (76.1%-92.3%)	**71.8% (63%-79.5%)**
	CAT ≥4	0.66 (0.61-0.71)	37.7% (28.5%-47.7%)	94.1% (87.5%-97.8%)	87% (73.7%-95.1%)	59% (51%-66.7%)
	CAT ≥5	0.58 (0.54-0.61)	16% (9.6%-24.4%)	**99% (94.6%-100%)**	**94.4% (72.7%-99.9%)**	52.9% (45.5%-60.2%)
IV fluid	CAT ≥3	**0.67 (0.61-0.73)**	**58.3% (47.8%-68.3%)**	75.7% (66.6%-83.3%)	67.5% (56.3%-77.4%)	**67.7% (58.8%-75.9%)**
	CAT ≥4	0.60 (0.55-0.66)	33.3% (24%-43.7%)	87.4% (79.7%-92.9%)	69.6% (54.2%-82.3%)	60.2% (52.2%-67.9%)
	CAT ≥5	0.55 (0.51-0.58)	13.5% (7.4%-22%)	**95.5% (89.8%-98.5%)**	**72.2% (46.5%-90.3%)**	56.1% (48.75-63.3%)
In-hospital mortality	CAT ≥3	0.71 (0.61-0.81)	**78.9% (54.4%-93.9%)**	63.8% (56.5%-70.7%)	18.1% (10.5%-28%)	**96.8% (10.5% (28%)**
	CAT ≥4	**0.75 (0.64-0.87)**	68.4% (43.4%-87.4%)	82.4% (76.2%-87.6%)	28.3% (16%-43.5%)	96.3% (92.1%-98.6%)
	CAT ≥5	0.63 (0.52-0.73)	31.6% (12.6%-56.6%)	**93.6% (89.1%-96.7%)**	**33.3% (13.3%-59%)**	93.1% (88.5%-96.3%)

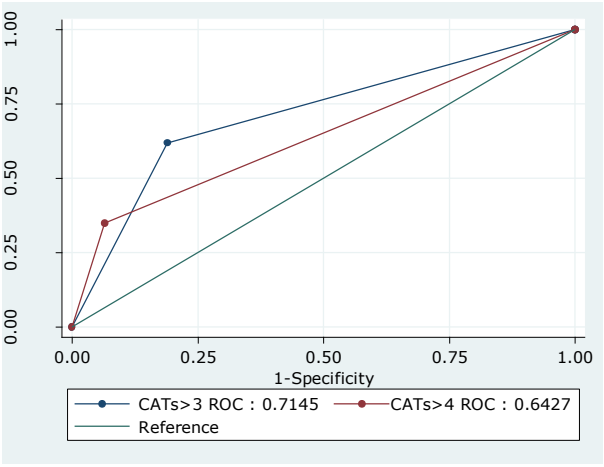
The best predictive performance (AUROC, NPV, PPV, sensitivity, specificity) values for each outcome have been highlighted for the score thresholds

Figure 7: ROC curves comparing un-weighted CATs score thresholds in children

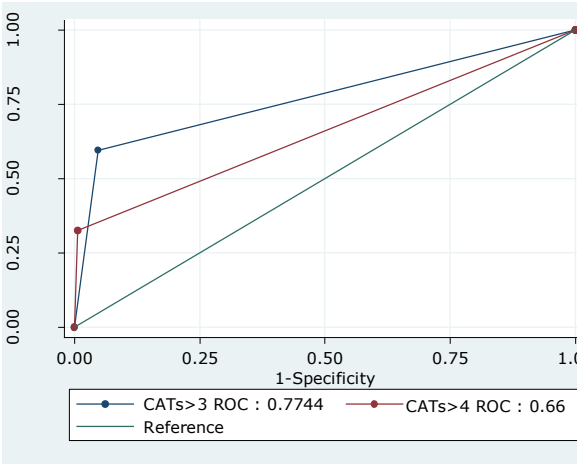
In-hospital mortality



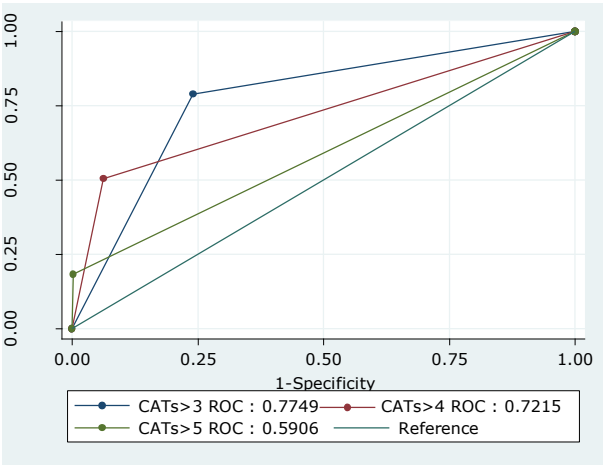
IV antibiotics



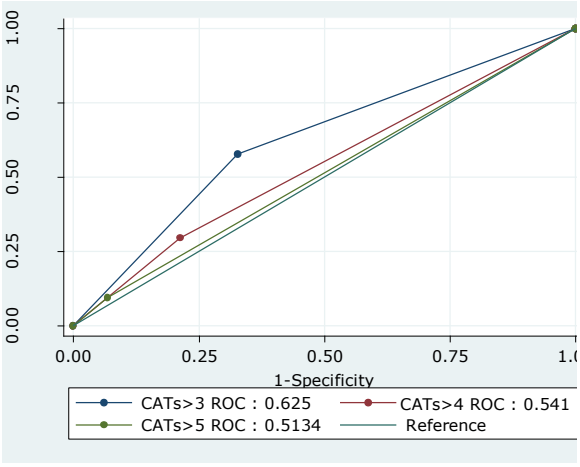
Hospital admission



Supplemental oxygen



IV fluid



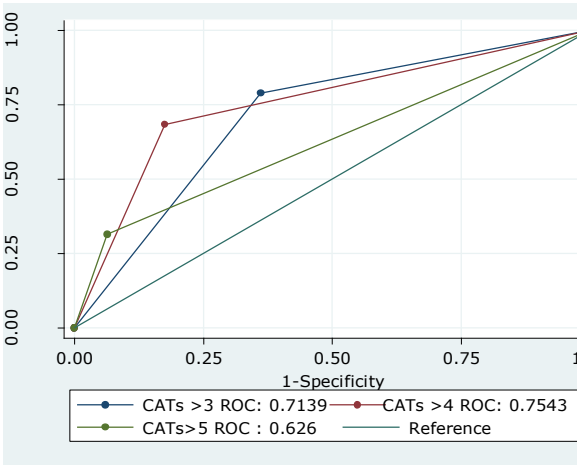
CATs score ≥3 Blue line

CATs score ≥4 Red line

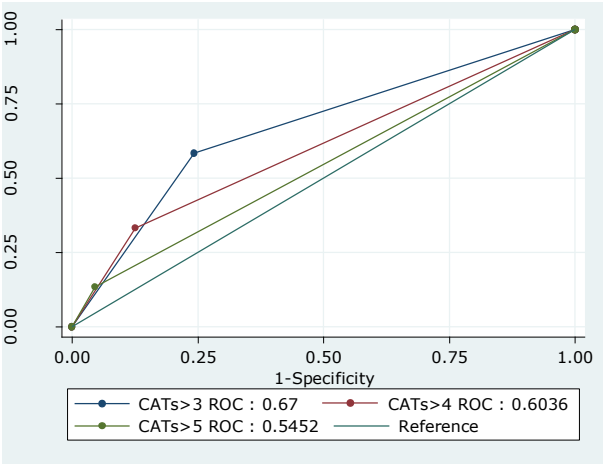
CATs score ≥5 Light green line

Figure 8: ROC curves comparing un-weighted CATs score thresholds in adults

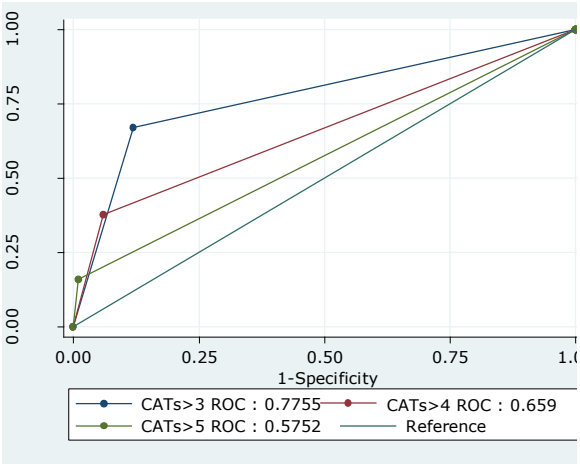
In-hospital mortality



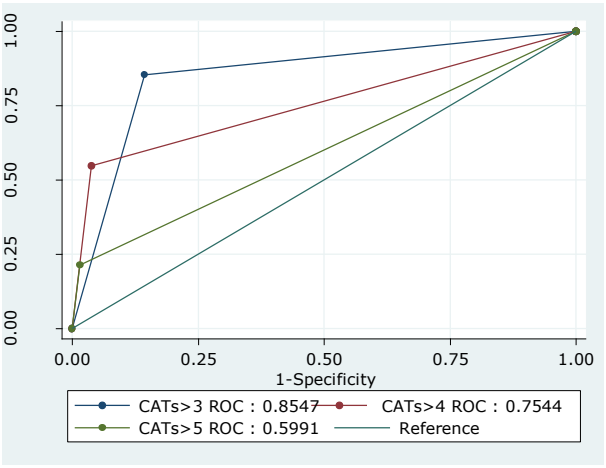
Supplemental oxygen



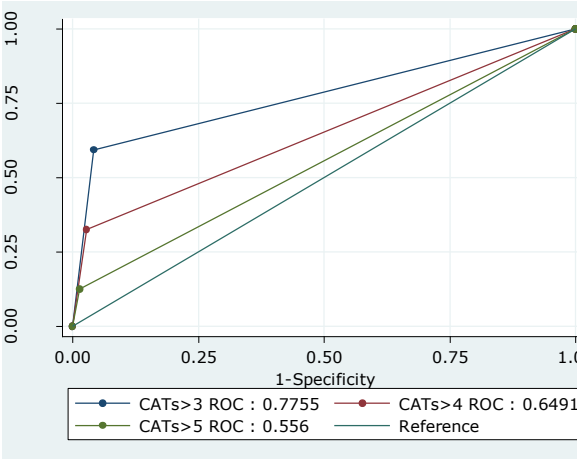
IV antibiotics



IV fluid



Hospital admission



CATs score ≥3 Blue line

CATs score ≥4 Red line

CATs score ≥5 Light green line

5.4.7 Predictive performance of weighted CATs threshold scores for predicting study outcomes in children and adults

Specificity and positive predictive values appeared to get better as the score threshold values increased while sensitivity and negative predictive values decreased with higher threshold scores in both patient groups. Low threshold scores were seen to have the poorest AUROC values in both adults and children but improved as score thresholds increased. A visual representation of the performance of the threshold scores in predicting each outcome using ROC curves is presented in Figures 9 and 10.

Table 5.12: Predictive performance of weighted CATs combined score thresholds in adults

OUTCOMES	SCORE THRESHOLD	AUROC (95%CI)	SENSITIVITY (%) (95%CI)	SPECIFICITY (%) (95%CI)	PPV (%) (95%CI)	NPV (%) (95%CI)
Admission	CAT ≥10	0.67 (0.61-0.72)	**95.6 (90.6-98.4)**	37.5 (26.4-49.7)	74.1 (67.0-80.5)	**81.8(64.5-93.0)**
	CAT ≥23	**0.80 (0.74-0.86)**	77.8 (69.8-84.5)	81.9 (71.1-90)	89.0 (81.9-94.0)	66.3(55.5-76.0)
	CAT ≥44	0.75 (0.70-0.79)	51.9 (43.1-60.5)	**97.2 (90.3-99.7)**	**97.2 (90.3-99.7)**	51.9(43.1-60.5)
Suppl. oxygen	CAT ≥6	0.63 (0.59-0.66)	**100 (95.2-100)**	25.0 (17.9-33.3)	43.1 (35.6-50.8)	**100 (89.4-100)**
	CAT ≥24	0.81 (0.76-0.85)	98.7 (92.8-100)	62.9 (54.0-71.1)	60.2 (50.9-68.9)	98.8 (93.5-100)
	CAT ≥54	**0.85 (0.80-0.90)**	77.3 (66.2-86.2)	**93.2 (87.5-96.8)**	**86.6 (76-93.7)**	87.9 (81.3-92.8)
IV antibiotic	CAT ≥10	0.69 (0.64-0.75)	94.3 (88.1-97.9)	44.6 (34.7-54.8)	64.1 (56.0-71.6)	**88.2 (76.1-95.6)**
	CAT ≥16	0.62 (0.58-0.67)	**96.2 (90.6-99.0)**	28.7 (20.1-38.6)	58.6 (50.9-66.0)	87.9 (71.8-96.6)
	CAT ≥31	**0.71 (0.65-0.76)**	47.2 (37.4-57.1)	**94.1 (87.5-97.8)**	**89.3 (78.1-96.0)**	62.9 (54.7 -70.6)
IV fluid	CAT ≥4	0.57 (0.52-0.62)	**89.6 (81.7 -94.9)**	24.3 (16.7-33.4)	50.6 (42.8-58.3)	73 (55.9-86.2)
	CAT ≥10	**0.68(0.62- 0.75)**	72.9(62.9 -81.5)	**64.0 (54.3-72.9)**	63.6 (53.9-72.6)	**73.2(63.2-81.7)**
	CAT ≥17	0.62 (0.56- 0.67)	37.5 (27.8 -48)	**64.0 (54.3-72.9)**	**69.2 (54.9-81.3)**	61.3 (53.1-69)
In-hospital mortality	CAT ≥15	0.64 (0.60-0.67)	**100 (82.4-100)**	27.1 (20.9-34.1)	12.2 (7.5-18.4)	**100 (93.0-100)**
	CAT ≥33	**0.74 (0.70-0.77)**	**100 (82.4-100)**	47.3 (40.0-54.7)	16.1 (10.0-24.0)	**100 (95.9-100)**
	CAT ≥75	0.71 (0.59-0.82)	68.4 (43.4-87.4)	**72.9 (65.9-79.1)**	**20.3 (11.3-32.2)**	95.8(91.1-98.4)

The best predictive performance (AUROC, NPV, PPV, sensitivity, specificity) values for each outcome have been highlighted for the score thresholds

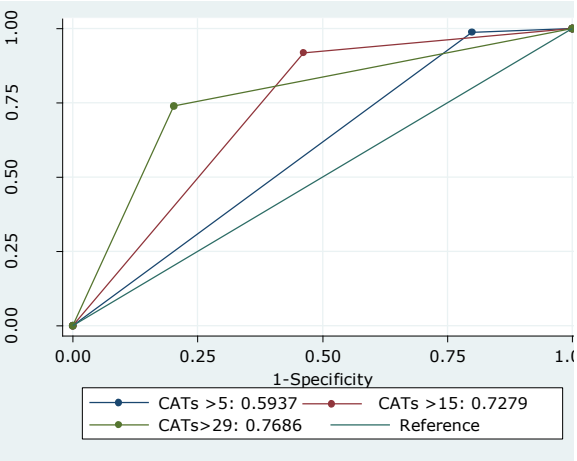
Table 5.13: Predictive performance of weighted CATs combined score thresholds in children

OUTCOMES	SCORE THRESHOLD	AUROC (95%CI)	SENSITIVITY (%) (95%CI)	SPECIFICITY (%) (95%CI)	PPV (%) (95%CI)	NPV (%) (95%CI)
Admission	CAT ≥21	0.75 (0.71-0.79)	**86.7 (83.9-89.2)**	63.5 (55.2-71.3)	91.4 (88.9-93.5)	**51.6 (44.1-59.1)**
	CAT ≥32	**0.78 (0.75-0.81)**	61.6 (57.7-65.3)	94.6 (89.6-97.6)	98.1 (96.2-99.2)	35.5 (30.8-40.5)
	CAT ≥40	0.65 (0.63-0.67)	30.4 (26.9-34.1)	**98.6 (95.2-99.8)**	**99.0 (96.5-99.9)**	24.1(20.7-27.7)
Suppl. oxygen	CAT ≥13	0.64 (0.61-0.66)	**93.9 (91.0-96.1)**	33.5 (29.1-38.1)	55.1 (51.1-59.0)	**86.3 (80.2-91.1)**
	CAT ≥21	**0.79 (0.76-0.81)**	81.4 (77.1-85.2)	76.0 (71.7-79.9)	74.6 (70.1-78.8)	82.5 (78.4-86.1)
	CAT ≥40	0.72 (0.69-0.75)	48.7 (43.5-53.8)	**95.4 (93.0-97.2)**	**90.1 (85.2-93.9)**	68.2 (64.3-71.8)
IV antibiotic	CAT ≥12	0.65 (0.62-0.68)	**88.2 (85.3-90.7)**	41.8 (35.4-48.4)	74.2 (75.7-82.1)	**58.8 (50.9-66.4)**
	CAT ≥18	**0.72 (0.69-0.76)**	65.0 (60.9-68.9)	79.7 (74.0-84.7)	88.9 (85.5-91.7)	47.8 (42.7-52.9)
	CAT ≥31	0.64(0.62-0.66)	33.1 (29.3-37.1)	**94.8 (91.1-97.3)**	**94.1 (89.9-96.9)**	36.3 (32.5 -40.3)
IV fluid	CAT ≥9	0.62 (0.59-0.65)	**85.7 (82.5 -88.5)**	38.0 (32.1-44.2)	74.2 (70.6-77.5)	**56.2 (48.6-63.6)**
	CAT ≥14	**0.63(0.59- 0.66)**	58.4 (54.2 -62.6)	66.9 (60.9-72.6)	**78.6 (74.3-82.5)**	43.7 (38.8-48.7)
	CAT ≥21	0.56 (0.52- 0.59)	29.7 (25.9 -33.7)	**81.4 (76.1-85.9)**	76.8 (70.5-82.3)	35.8 (31.9-39.8)
In-patient mortality	CAT ≥5	0.59 (0.57-0.61)	**98.6(92.6-100)**	20.1 (17.3-23.2)	10.9 (8.6-13.5)	**99.3 (96.3-100)**
	CAT ≥15	0.73 (0.69-0.76)	91.8 (83.0-96.9)	53.8 (50.1-57.5)	16.5 (13.0-20.4)	98.5 (96.8-99.5)
	CAT ≥29	**0.77 (0.72-0.82)**	74.0 (62.4-83.5)	**79.8 (76.7-82.6)**	**26.6 (20.7-33.2)**	96.9(95.1-98.1)

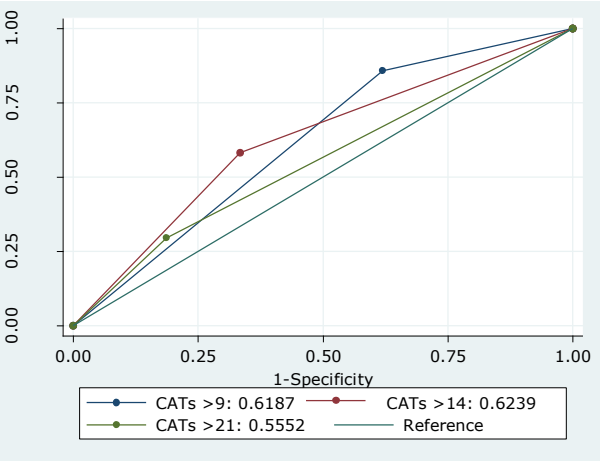
The best predictive performance (AUROC, NPV, PPV, sensitivity, specificity) values for each outcome have been highlighted for the score thresholds

Figure 9: Comparison of ROC curves for weighted combined CATs score thresholds in children

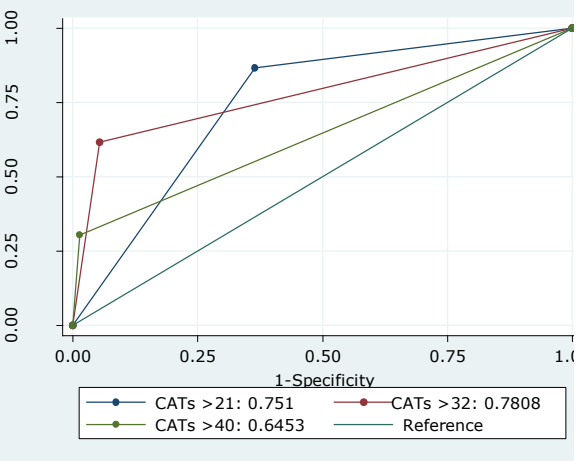
In-hospital mortality



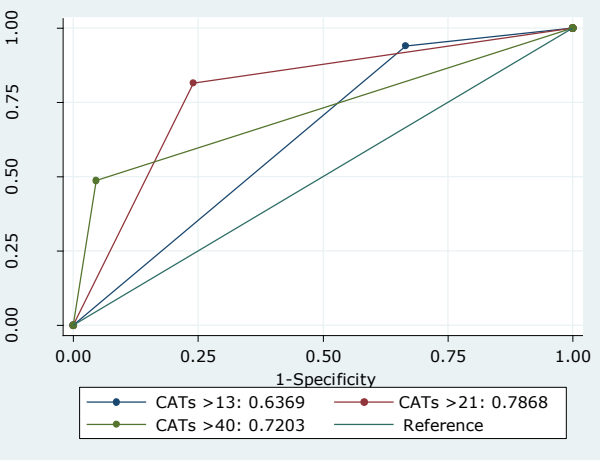
IV antibiotics



Hospital admission



IV fluid



Supplemental oxygen

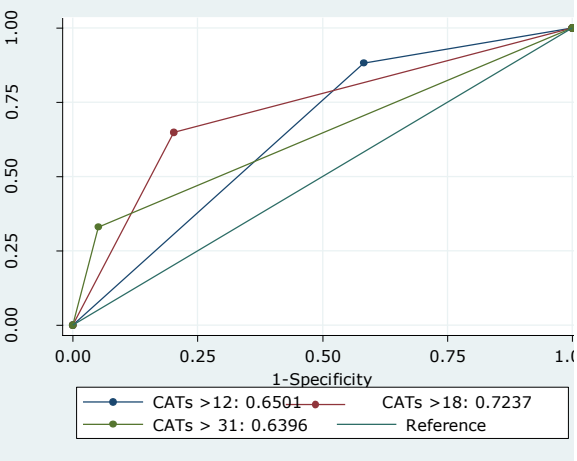
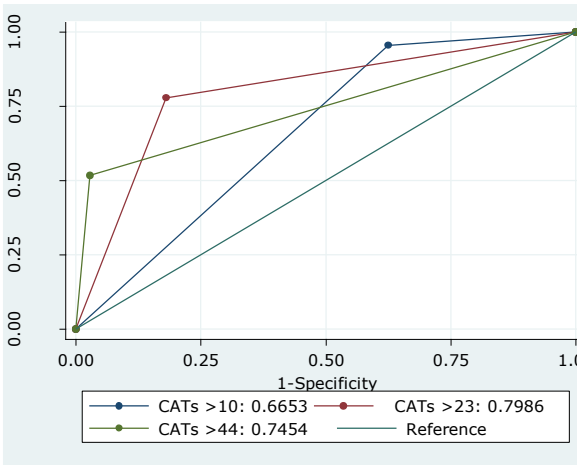
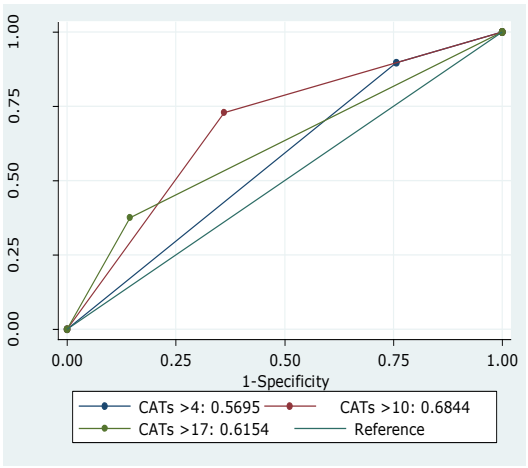


Figure 10: Comparison of ROC curves for weighted combined CATs score thresholds in adults

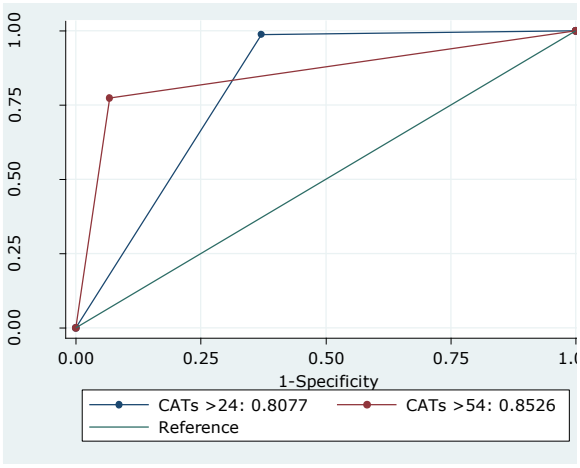
Hospital admission



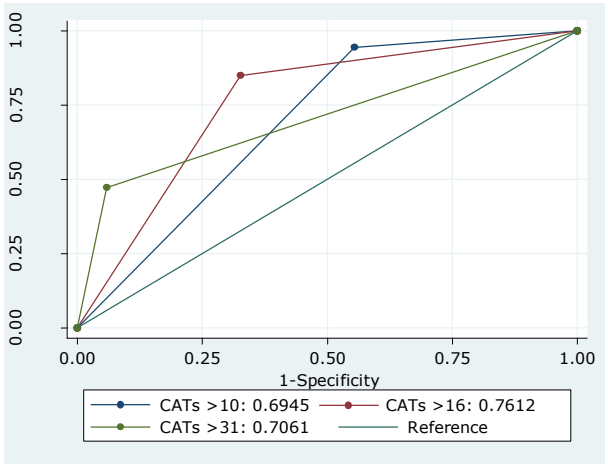
IV fluid



Supplemental oxygen



IV antibiotics



* It was not possible to plot a ROC curve for in-patient mortality in the adult group as there were not enough numbers in the threshold groups.

5.5 Discussion

5.5.1 Summary of key findings

This study demonstrates that the Community assessment tools (CATs) perform better as predictors of hospital admission, administration of acute care interventions and death on admission, when used as combined scoring indices than as individual stand-alone criteria. AUROC values for the weighted CATs scoring index were higher than those for the un-weighted index and in most instances this difference was statistically significant. However, the AUROC values of the weighted index were only marginally better (differences ranging between 0.01 -0.05) and the 95% CI values were overlapping which suggests that these differences may not necessarily translate into a clinically meaningful result. The weighted CATs score also poses several potential difficulties in terms of ease of utility; one of which was the complexities of calculating the weighted scores and identifying sensible score thresholds. While this study's findings suggest that an unweighted CATs score of 3 or more is a suitable threshold for making a decision on the need for hospital care and interventions in both adults and children, it was less straightforward to identify clinically relevant thresholds using the weighted scoring index.

5.5.2 Strengths and limitations

As the overarching strengths and limitations relating to this research have already been outlined in Chapter 4, this section will focus on further strengths and limitations that are specific to the second study in this thesis. The major strength of this study lies in the use of beta regression coefficients over odds ratios for generating weighted scores. Although odds ratios (OR) have been

used for generating scores in a number of studies, one of the flaws of this approach is that odds ratios are heavily influenced by sample size such that larger effect estimates are reported in studies with smaller samples (137), which is not the case with regression coefficients (138). More importantly, it has been pointed out that it is mathematically incorrect to assigned a score of '1' to a predictor having an OR of 1, as there is actually no association between this predictor and the outcome in question as an OR of 1 is synonymous with a regression coefficient of 0 (90).

5.5.3 Comparison with existing literature

As the CATs was only recently developed (139) and validated using data from the 2009 A/H1N1 flu pandemic in the United Kingdom, there are just a handful of published papers on this prediction tool which limits the number of studies available for direct comparison(35, 113, 118). The performance of CATs as a scoring index was first assessed by Myles et al. (2012) in patients with PCR confirmed influenza in the UK (119). They assessed the discriminatory value of an unweighted CATs index using logistic regression and just as in this present study, the authors also found a linear relationship between the CATs scores and clinical outcomes i.e. significant test for trend. This further supports the combination of the CATs criteria into a scoring index for severity assessment. Although the methods used for creating the weighted scores in this study were similar to those used in the UK study which assessed the performance of the CATs as a weighted scoring system in the FLU-CIN cohort, the weights obtained for each CATs were very different, thus highlighting a major issue with the use of statistically derived weights.

The evaluation of any prediction tool involves the assessment of two important concepts which are clinical validity and utility(136). A tool is said to have clinical validity if its predictions correspond with the actual risk of the outcome (136) , and this quality can be assessed by determining the discrimination and calibration of such a tool(140). The un-weighted and weighted combined CATs scoring indices in this study, demonstrated good clinical validity in terms of discrimination for both children and adults, but had a less satisfactory performance in children in terms of their calibration qualities. According to Steyerberg et al. (2010) in their study on assessing the performance of prognostic models, it is not uncommon to find a model with good discrimination also having poor calibration(141), and this does not give enough reasons to condemn the CATs. Moreover, the Hosmer-Lemeshow test which was used in this study to assess the calibration qualities of the scoring indices has been found to be affected by sample size(142). This test follows the principles of hypothesis testing by using a significance level of 0.05 to either accept or reject the null hypothesis i.e. that there is no difference between the observed and predicted estimates of an outcome by a prediction rule. Therefore calibration will most likely be better in a small sample which will lack the power to identify poor calibration i.e. reject the null hypothesis while the smallest conflict between predicted and observed estimates in a large sample will result in a significant Hosmer-Lemeshow test(143, 144). This offers an explanation for why the calibration of the CAT s indices was better in adults than in children who make up about almost 80% (809/1016) of the patient population.

Findings from the dissertation study which compared the predictive performance of the CATs scoring indices were very similar to those in this study as the weighted scores also had higher AUROC values than unweighted scores

in both adults and children. However just like in this present study the differences between the CATs indices in the dissertation study were marginal with overlapping 95% CI and therefore statistically insignificant. Additionally, the AUROC values in the dissertation study were also higher when CATs was used as a combined scoring index than when its criteria are used as stand-alone predictors. Although these studies provide cogent statistical evidence that clinical outcome prediction is better when the CATs criteria are used as combined scoring indexes than as stand-alone criterion, the application of a prediction tool is highly dependent on the clinical scenario, and to reiterate, these tools are not replacements for clinical judgement instead they should serve as guides and to improve the specificity of clinical decisions(127).

The second aspect of prediction tool evaluation involves assessing clinical utility, which has been described as the ability of a tool to sufficiently inform management decisions that will result in better patient outcomes(136). This entails defining suitable score 'cut-offs' or thresholds capable of reflecting the severity of a patient's condition, in order to guide clinicians in making appropriate management decisions(127). Threshold values are usually decided by using sensitivity and specificity values, bearing in mind that the concept of a prediction tool having both high sensitivity and specificity is purely theoretical and does not occur in real life (145). Concessions are made in favour of whichever of the two measures best fits the purpose of the tool in practice and whatever threshold value achieves the most favourable value of these two measures (118). ROC curves and AUROC values are also used to decide on appropriate thresholds and in this study an unweighted threshold score of $\geq$3 had the highest AUROC values for predicting the need for admission, respiratory support and intravenous therapy in both adults and children. A

similar performance was also observed in the dissertation study and by Myles et al. (2012), an unweighted score of three or more was also identified as the best predictor of several clinical outcomes in patients of all ages(118).

In terms of predicting 'in-patient mortality' in this present study, a threshold score of ≥4 had the highest AUROC values in children [score ≥4: 0.74 (0.68-0.79) vs score ≥3 0.73 (0.70-0.77)] and in adults [score ≥4: 0.75 (0.64-0.87) vs score ≥3 0.71 (0.61-0.81)]. Once again, the overlapping confidence intervals and marginal differences (0.01- 0.05) suggests that the performance of these thresholds were comparable. Moreover, a score of ≥3 had the best sensitivity and negative predictive values of the three score thresholds. Even though acute lower respiratory tract infections are severe enough to result in an adverse outcome such a death, early recognition of signs of deterioration and timely interventions usually result in good outcome. Therefore an unweighted score threshold of ≥3 with its high sensitivity will allow prompt identification of high risk patients (true positives), and its good negative predictive value also ensures accurate identification of low risk patients.

On the other hand, it can be argued that the weighted CATs scoring index might be less clinically useful than its un-weighted counterpart, seeing as it was quite difficult to identify a suitable weighted score threshold. Clinical usefulness though rarely considered in a lot of prediction studies, relies less on statistical validity but deals with the practicalities of using the tool in real life. Vergouwe et al. have stated that even though a prognostic tool has good calibration and discriminative properties, such a tool might still be unsatisfactory in terms of clinical usefulness(140). Despite having high AUROC values and good calibration, specificity values of the weighted index threshold scores were lower than that of the unweighted index which meant that using the weighted CATs

index will result in more false positives. False positives lead to the admission of low risk patients and inefficient use of hospital resources. Another potential pitfall of using statistically derived weighted scores is that the beta regression coefficients vary according to variability in case-mix and patients' characteristics(138), which means that the total scores and threshold of the weighted index will vary in different populations. For instance the individual weights of the CATs criteria for the adults were different from the weights for children, which were also very different from weights reported in the CATs dissertation study. Notwithstanding, the difficulties associated with calculating the different weights in real life clinical setting, one of the advantages of weighting is that it gives a more realistic estimate of the strength and clinical relevance of each criterion. Instead of using statistically derived weights which might not be clinically sensible, the utility of the weighted CATs score can be potentially improved through expert clinical consensus on appropriate weight scales.

5.5.4 Summary

It can be concluded that the CATs perform better as combined scores than when used as stand-alone individual criteria, especially as the use of individual criterion as triggers for hospital admission and acute care interventions assumes a more resource-rich environment with adequate staff and hospital beds which is unlike the healthcare system in developing countries, Nigeria included. However, the use of the CATs criteria as combined unweighted scores presents an opportunity for creating trigger-threshold scores which could be incorporated into practice to indicate a scale of response best suited to the clinical setting, capacity or scenario. For example, a score of one or more in

the community could indicate the need for prompt referral for urgent assessment in hospital, while a threshold score of three or more on presentation in the emergency department will indicate the need for hospital based admission and intravenous therapy, and a score of four indicates critical illness requiring escalated response. Although there are potential gains associated with such adaptations, they are more likely to work in well organised health systems, and its generalisability may be affected by differences in local practices, patient characteristics and disease incidence. A typical example is the Nigerian health sector with its three tiers system of primary, secondary and tertiary care hospitals; the primary health centres are supposedly the Nigerian version of primary care in the UK and are expected to act as entry points or referral centres. However the primary health centres in Nigeria are usually bypassed by patients who present first to tertiary centres, and this may defeat the purpose of adopting adjusted score thresholds. Therefore such adjustments must take into consideration the individual health care systems, capacity and health seeking pattern of the population, also it will require government involvement and planning as the necessary infrastructure must be put in place. For now such adjustments may only be achievable using the un- weighted CATs score as threshold adjustments may be quite complicated using statistically derived weights, which tend to vary with population and data.

6: Comparing the un-weighted CATs score and the CRB65 for severity assessment of adult patients with ALRTI in Nigeria

6.1 Introduction

In this chapter, the un-weighted CATs scoring index and the British Thoracic Society's (BTS) CRB-65 score have been compared in terms of their ability to predict the need for hospital admissions, lifesaving clinical interventions and also the risk of mortality among adults with a range of acute lower respiratory tract infections in Nigerian hospitals. This chapter also reports on a sensitivity analysis comparing the performance of the two scoring indices in patients with a clinical diagnosis of pneumonia, and finally concludes with a discussion on the clinical validity and utility of the scoring indices.

6.2 Background

Assessment of illness severity is a crucial step in the management of acute lower respiratory tract infections (20, 146). It has also been noted that severity assessment of ALRTIs based solely on clinical judgement is usually susceptible to inaccuracies (147, 148). In some cases the clinical experience of the doctor has a huge impact on clinical judgement(22). Thus several international guidelines have recommended the use of severity scoring systems in the management of acute lower respiratory tract infections to improve and standardise decision making across the spectrum of clinical experience (51, 115, 149-151). With pneumonia being the most common and serious type of

ALRTI, the majority of severity scoring systems for respiratory illnesses are for the management of community acquired pneumonia (21, 22, 33, 152). In the absence of a flu-specific prediction score, certain well validated pneumonia prediction scores have been formerly used in triaging patients with influenza (see chapter one) (21) (30).

There is a limited amount of information on the use of clinical prediction tools in the management of ALRTI in Nigeria despite being the leading cause of mortality in this region. This chapter will compare the performance of the unweighted CATs scoring index with that of a well validated pneumonia severity assessment tool among a cohort of Nigerian adult patients. Although the CATs has been previously compared with the CURB-65 score in a UK study, the CRB-65 score was selected for this present study as it does not require extensive laboratory investigations, making it better suited for use in a resource-limited setting (51, 153-155). Despite the omission of serum urea from the CRB-65 score, findings from several large prospective studies have shown that its performance is comparable to that of the CURB-65 in patients with CAP (21, 153, 154, 156, 157).

6.3 Statistical analysis

The data used in this study were from three of the four participating hospitals in this doctoral research. Younger patients were pragmatically excluded from this study as blood pressure measurement which is one of the CRB-65 criteria is not a routinely recorded parameter among children in this setting except there was an underlying cardiac condition.

Data were collected on CRB-65 and CATs criteria, diagnosis based on presenting symptoms and planned therapeutic interventions were all recorded at the time of initial clinical assessment using a standard pro forma. The methods used in creating the un-weighted CATs score have been described in chapter 5, and a similar approach was used in creating the CRB-65 score by awarding a single point to each of its criteria (confusion, respiratory rate $\geq$30mins, systolic Bp <90mmHg/ diastolic Bp $\leq$90mmHg and age $\geq$65 years). Patients in this study could therefore have a minimum of 0 and a maximum of 4 points for the CRB-65 and 7 points for the unweighted CATs scores. The predictive performance of both scores was first assessed by conducting separate logistic regression analysis involving the individual study outcomes in turn: need for hospital admission, supplemental oxygen therapy, in-patient mortality, intravenous antibiotics and fluids and each score fitted as a continuous variable. The linearity of logit assumption of the logistic regression models was assessed using the Lowess algorithm and displayed as graphs in appendix X (94, 135), while Hosmer-Lemeshow goodness-of-fit test was used to assess the fit of the regression models (144).

The predictive ability of the two scoring systems was investigated and compared in terms of discrimination and calibration. Discrimination was assessed using the area under receiver operating characteristic curves (AUROC) and 95% confidence intervals (CI) and ROC curves. (158).

Hosmer-Lemeshow (HL) goodness of fit statistic was used in assessing calibration by first randomly and equally splitting the total study population into a separate development and validation sample, after which the Pearson's chi square test was used to compare the predicted and observed values for each study outcome within the two newly created populations. A poorly

calibrated tool will have a p-value of < 0.05 and vice versa in the validation sample(94). These analyses were repeated in a sensitivity analysis to compare the predictive performance of the CRB-65 and un-weighted CATs scores in the patients who had a diagnosis of pneumonia.

The scoring systems were each re-categorised into binary variables and thresholds for the un-weighted CATs scores were recoded as $\geq$3, $\geq$4 and $\geq$5 as in the previous chapter, while the CRB-65 score was categorised as '0' (low risk), '1/2' (intermediate risk) and '3/4' (high risk) according to previously specified score threshold values used in published studies and guidelines (21, 51, 159). The predictive accuracy of the thresholds for each of the scoring indexes was assessed using sensitivity, specificity, area under ROC curve (AUROC) values, positive predictive (PPV) and negative predictive values (NPV). All analyses were conducted using Stata version 12 (StataCorp LP).

6.4 Results

A total of 207 consecutive patients aged 16 years and above presenting with symptoms of ALRTIs to the emergency departments of Lagos University Teaching Hospital (LUTH), Lagos State University Teaching Hospital (LASUTH) and Randle General Hospital Surulere (RGH) were prospectively identified based on the study's inclusion criteria (see chapter 2). 62.3% (129/207) of the adult population were diagnosed with pneumonia, 135 (65%) were admitted, and in-patient mortality rate was 14.1% (19/135). A more detailed discussion of the data collection process and descriptive statistics is presented in chapter 3.

6.4.1 Comparison of un-weighted CATs and CRB65 scores in adults with ALRTI

Table 6.1 shows the unadjusted odds ratios (95% CI) and p-values for the logistic regressions analyses between the two scoring systems as continuous variables and each study outcome. All p-values were highly statistically significant ($p < 0.001$) which indicated that the likelihood of being admitted, receiving supplemental oxygen and intravenous therapy or dying on admission increased directly with every unit increase in either the CRB-65 and CATs scores. Examining the Lowess plots for each scoring system versus the individual study outcomes (presented in appendix VIII and IX), the linearity of logit assumption was maintained in the regression models as the shape of the smoothed curves for both the unweighted CATs and CRB65 scores appeared to be relatively linear.

Table 6.1: Logistic regression analyses involving competing scoring indexes as continuous variables in adults with ALRTI

Outcomes	CRB65 score OR (95% CI) P-values	Unweighted adults CATs score OR (95% CI) P-values
Hospital admission	4.08 (2.46 – 6.78) <0.001	3.55 (2.47 – 5.10) <0.001
Supplemental oxygen	3.72 (2.36 – 5.86) <0.001	4.62 (3.09 – 6.91) <0.001
Intravenous antibiotics	4.10 (2.53 – 6.64) <0.001	2.66 (2.02 – 3.49) <0.001
Intravenous fluid	2.50 (1.69 – 3.70) <0.001	1.69 (1.38 – 2.07) <0.001
In-patient mortality	2.83 (1.70 – 4.71) <0.001	2.21 (1.56 – 3.15) <0.001

According to the p-values of the HL goodness-of-fit test presented in Table 6.2, all regression models involving the un-weighted CATs and the five study outcomes were a good fit for the data ($p \geq 0.05$), while the models for the CRB-65 score fit reasonably well except for outcomes admission (p= 0.0356) and intravenous antibiotic (p= 0.0056).

Table 6.2: Assessing the goodness of fit of the regression models of CRB65 and un-weighted CATs scores in adults with ALRTI

Outcomes	CRB65 p-values	Un-weighted CATs p-values
Hospital admission	**0.0356**	0.0538
Supplemental oxygen	0.5024	0.1368
Intravenous antibiotic	**0.0056**	0.4611
Intravenous fluid	0.1115	0.6524
In-patient mortality	0.4377	0.1658

**** Highlighted p-values are significant and suggest poor model fit**

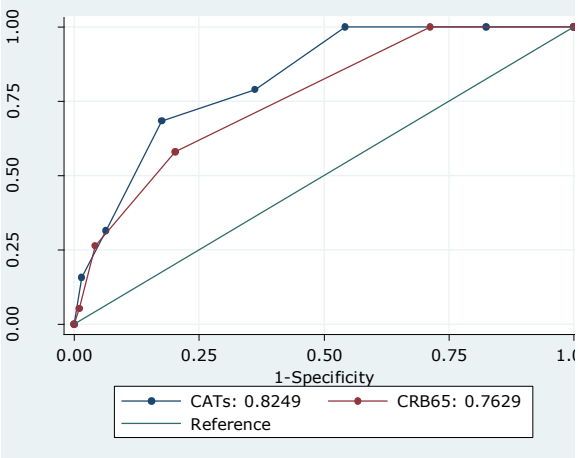
The un-weighted CATs score as seen by its higher AUROC (95% CI) values in Table 6.3, appears to be better than the CRB-65 score at predicting all five study outcomes. The difference between the performance of the two scores was only statistically significant (<0.05) for predicting the need for admission, supplemental oxygen and IV antibiotics administration. ROC curves for the comparisons are displayed in Figure 11 and the CATs score represented by the blue line consistently appears to outperform the CRB-65 (red line).

Table 6.3: Comparison of AUROC values of CRB65 and un-weighted CATs scores in adults with ALRTI

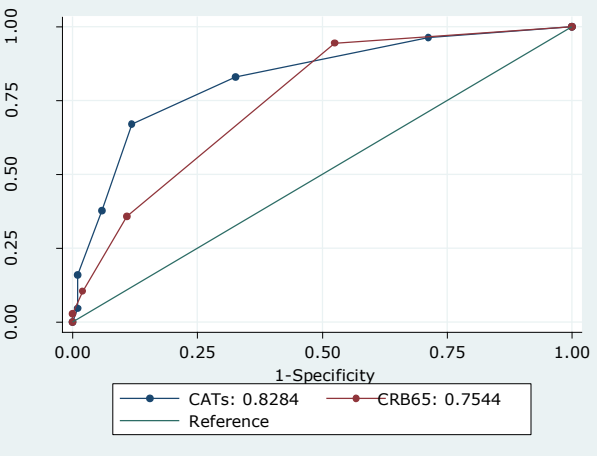
Outcomes	Scores	AUROC (95%CI)	P-values
Admission	CRB65	0.75 (0.68-0.81)	**<0.0001**
	Unweighted CATs	0.86 (0.81-0.91)	
Suppl. oxygen	CRB65	0.74 (0.68-0.80)	**<0.0001**
	Unweighted CATs	0.92 (0.88-0.95)	
IV antibiotic	CRB65	0.75 (0.70-0.81)	**0.0070**
	Unweighted CATs	0.83 (0.77-0.88)	
IV fluid	CRB65	0.68 (0.62-0.75)	0.2420
	Unweighted CATs	0.72 (0.65-0.79)	
In-patient mortality	CRB65	0.76 (0.67-0.86)	0.1543
	Unweighted CATs	0.82 (0.74-0.91)	

Figure 11: ROC curves comparing unweighted combined CATs and CRB65 scores in adults with ALRTI

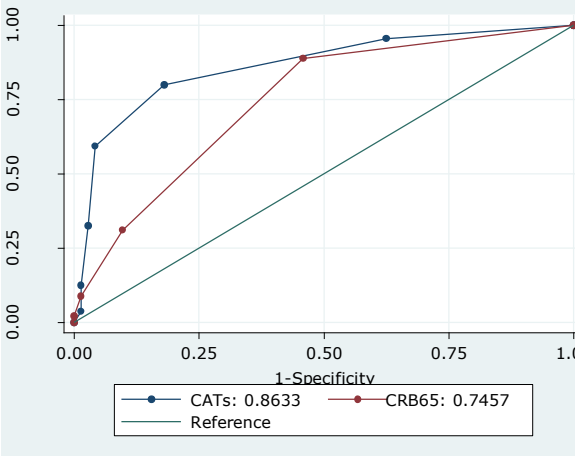
In-patient mortality



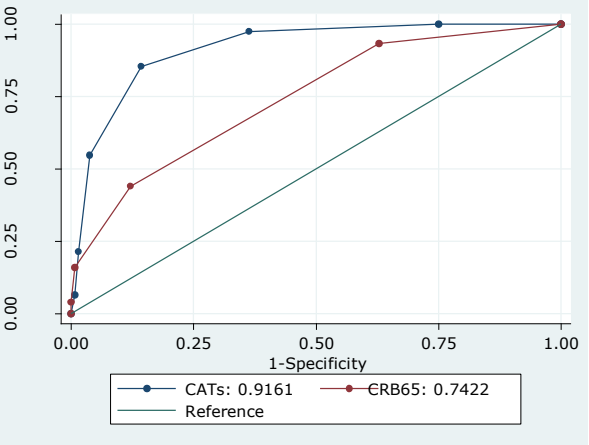
IV fluids



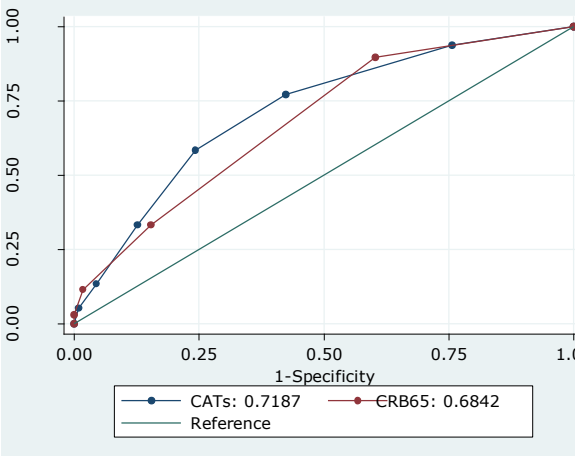
Hospital admission



IV antibiotics



Supplemental oxygen



CATs score	Blue line
CRB65	Red line
Reference	Light green line

Finally, the calibration of the CRB-65 and un-weighted CATs scores as seen by the Hosmer-Lemeshow test p-values have been presented in Tables 6.4 and 6.5 respectively. In the validation sample CRB-65 showed good calibration for all outcomes except intravenous antibiotic (0.0005), compared to the un-weighted CATs score which was well calibrated for all five outcomes.

Table 6.4: Hosmer-Lemeshow test for assessing the calibration of crb65 scores in adults with ALRTI

outcomes	Development sample p-values	Validation sample p-values
Hospital admission	**0.0331**	0.1505
Supplemental oxygen	0.1732	0.8070
Intravenous antibiotic	**0.0049**	**0.0005**
Intravenous fluid	**0.0401**	0.0618
In-patient mortality	0.5039	0.9362

** Highlighted p values are significant and suggest poor calibration

Table 6.5: Hosmer-Lemeshow goodness of fit test for calibration of un-weighted CATs index in adults with ALRTI

outcomes	Development sample p-values	Validation sample p-values
Hospital admission	**0.0345**	0.4979
Supplemental oxygen	0.3178	0.7992
Intravenous antibiotic	0.7496	0.0620
Intravenous fluid	0.6148	0.0689
In-patient mortality	0.4726	0.4926

** Highlighted p values are significant and suggest poor calibration

6.4.2 Sensitivity analysis comparing predictive performance of CRB-65 and un-weighted CATS in adults with pneumonia

A total of 129 patients with a diagnosis of pneumonia were studied and the discriminatory ability of each tool was assessed using ROC curves and the AUROC (95% CI) values, while model fit was assessed using the Hosmer-Lemeshow test. The predictive performance of the scores was not different from the previous findings in the complete patient group, however AUROC (95%CI) values for the two scoring systems in Table 6.6 were much lower in patients with pneumonia. The un-weighted CATs score again appeared to perform better than the CRB-65 at predicting all five study outcomes, but the differences were not statistically significant for IV fluid and in-patient mortality. In terms of mortality prediction, there was a marginal difference of 0.02 in the AUROC values of the unweighted CATs [0.73 (0.59- 0.86)] and CRB-65 [0.71 (0.59-0.84)] scores which was not statistically significant. ROC curves comparing the CRB-65 and unweighted CATs scores are displayed in Figure 12. A visual inspection of the ROC curves shows that the red line representing the CRB-65 score is closer to the green-diagonal line for outcomes admission, supplemental oxygen and IV antibiotics indicating a poorer performance in comparison to the CATs score, while the CATs score appears to be only marginally superior to the CRB-65 for outcomes in-patient mortality and IV fluid.

The p-values from the Hosmer-Lemeshow goodness of fit test presented in Table 6.7 were all statistically insignificant ($\geq$0.05), which indicates that the regression models of both CRB-65 and CATs are a good fit for all the outcomes in this patient group. When the sample was split in two in Tables 6.8a and 6.8b, the CATs score was seen to lose some of its calibration qualities as the Hosmer-Lemeshow P-values for outcomes hospital admission and in-patient mortality were significant,

while CRB-65 maintained calibration for all study outcomes except admission and need for supplemental oxygen.

Table 6.6: AUROC (95%CI) values for CRB65 and un-weighted CATs scores in adults with pneumonia

Outcomes	CRB65 AUROC (95% CI)	Un-weighted CATs AUROC (95% CI)	p- value
Admission	**0.64 (0.50 – 0.77)**	**0.87 (0.78 – 0.95)**	**<0.0001**
Supplemental oxygen	**0.67 (0.79 – 0.75)**	**0.85 (0.79 – 0.92)**	**<0.0001**
Intravenous antibiotics	**0.64 (0.55 – 0.74)**	**0.79 (0.71 – 0.87)**	**0.0003**
Intravenous fluids	0.60 (0.51 – 0.69)	0.63 (0.53 -0.72)	0.4857
In-patient mortality	0.71 (0.59 – 0.84)	0.73 (0.59 – 0.86)	0.8265

**statistically significant values are highlighted

Table 6.7: Assessing the goodness of fit of the regression models for CRB65 and CATs scores in adults with pneumonia

outcomes	Un-weighted CATs p-values	CRB65 p-values
Admission	0.4439	0.6029
Supplemental oxygen	0.9121	0.9206
Intravenous antibiotics	0.1952	0.7376
Intravenous fluids	0.5656	0.9148
In-patient mortality	0.2252	0.8118

Table 6.8a: Assessing calibration of CRB scores in adults with pneumonia

outcomes	Development sample p-values	Validation sample p-values
Admission	0.4769	**0.0004**
Supplemental oxygen	0.1702	**<0.0001**
Intravenous antibiotic	0.7664	0.9497
Intravenous fluid	0.4556	0.1257
In-patient mortality	0.2222	0.3286

**** Highlighted p values are significant and suggest poor calibration**

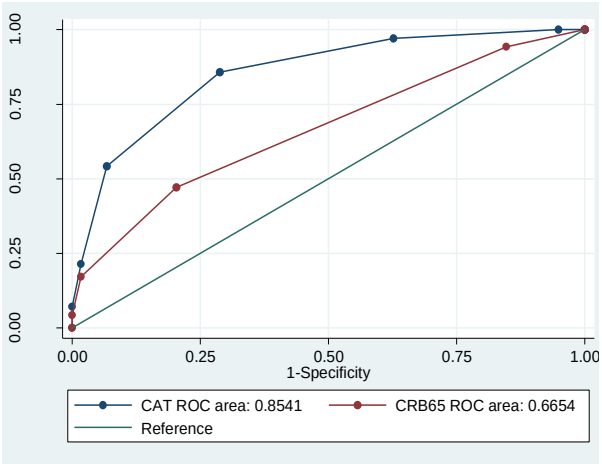
Table 6.8b: Assessing calibration of Un-weighted CATs scores in adults with pneumonia

outcomes	Development sample p-values	Validation sample p-values
Admission	0.9412	**0.0063**
Supplemental oxygen	**0.0172**	0.1234
Intravenous antibiotic	0.4712	0.8083
Intravenous fluid	0.7786	0.4269
In-patient mortality	0.9442	**<0.0001**

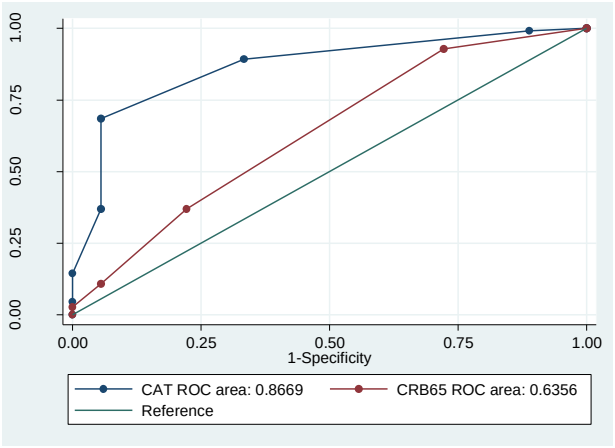
**** Highlighted p values are significant and suggest poor calibration**

Figure 12: ROC curves comparing unweighted CATs and CRB65 scores in adults with pneumonia

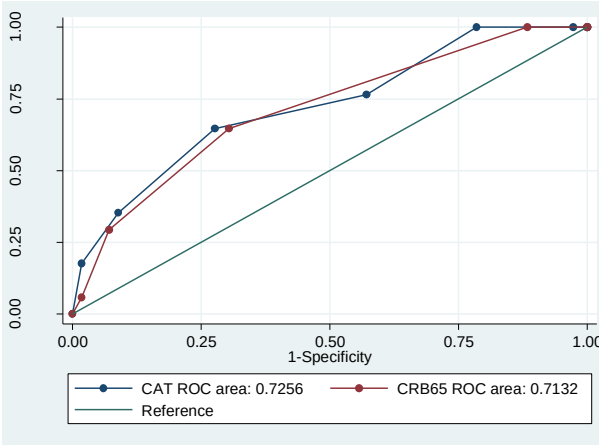
Supplemental oxygen



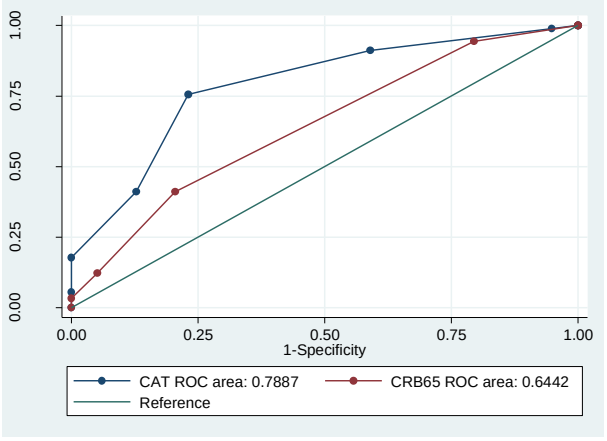
Hospital admission



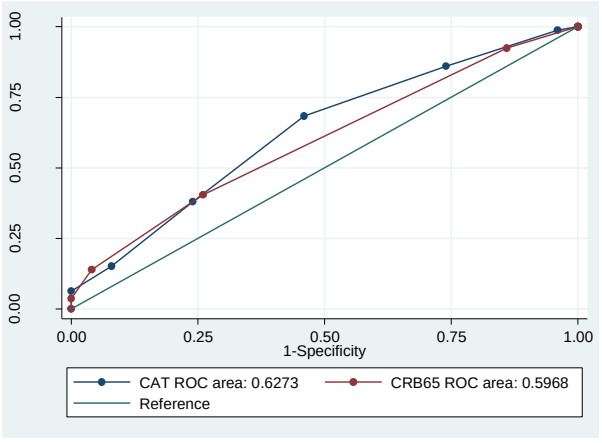
In-patient mortality



IV antibiotics



IV fluid



CATs score	Blue line
CRB65	Red line
Reference	Light green line

6.4.3 Predictive performance of CRB65 score thresholds in adults with ALRTI

In terms of overall predictive accuracy, AUROC values for the CRB-65 presented in Table 6.9 were generally poor [0.29 (0.22- 0.35) – 0.68 (0.61-0.74)] across all thresholds and much lower than values for the un-weighted CATs index [0.51 (0.49 – 0.53) – 0.74 (0.68 – 0.71)] in table 6.10. The need for hospital admission, supplemental oxygen therapy, intravenous antibiotics or fluids were best predicted by a CRB-65 score of either 1/2 and an un-weighted CATs score of $\geq$3, while in-patient mortality was best predicted by a CRB-65 score of 3/ 4, and by an un-weighted CATs score of $\geq$4.

A CATs score of $\geq$3 was seen to have the highest sensitivity and negative predictive values for each of the five outcomes. CRB-65 score of 1/2 had the best sensitivity values for all study outcomes, and was also seen to have the best NPV for all outcomes except in-patient mortality where a threshold score of 3/4 was seen to perform marginally better. Specificity and positive predictive values (PPV) were seen to improve with increase in un-weighted CATs and CRB-65 threshold values, even so CRB-65 threshold scores still had lower positive predictive and specificity values compared to the CATs threshold scores. The ROC curves of the comparison between score thresholds in adults with ALRTI have been presented for the unweighted CATs in the previous chapter, curves for the CRB65 are presented in appendix X.

Table 6.9: Predictive performance of CRB65 score thresholds in adults with ALRTI

OUTCOMES	SCORE THRESHOLDS	AUROC (95%CI)	SENSITIVITY (%) (95%CI)	SPECIFICITY (%) (95%CI)	PPV (%) (95%CI)	NPV (%) (95%CI)
Admission	CRB65 =0 (low risk)	0.28 (0.22-0.35)	11.1 (6.4-17.7)	45.8 (34.0-58.0)	27.8 (16.5-41.6)	21.6 (15.3-28.9)
	CRB65=1/2 (moderate risk)	**0.68 (0.61-0.74)**	**80.0 (72.3-86.4)**	55.6 (43.4-67.3)	77.1 (69.3-83.8)	**59.7 (47.0-71.5)**
	CRB65=3/4 (high risk)	0.54 (0.51-0.57)	8.9 (4.7-15.0)	**98.6 (92.5-100.0)**	**92.3 (64.0-99.8)**	36.6 (29.8-43.8)
Suppl. Oxygen	CRB65 =0 (low risk)	0.35 (0.30-0.40)	6.7 (2.2-14.9)	62.9 (54.0-71.1)	9.3 (3.1-20.3)	54.2 (46.0-62.3)
	CRB65=1/2 (moderate risk)	**0.58 (0.51-0.64)**	**77.3 (66.2-86.2)**	37.9 (29.6-46.7)	41.4 (33.2-50.1)	**74.6 (62.5-84.5)**
	CRB65=3/4 (high risk)	0.58 (0.53-0.62)	16.0 (8.6-26.3)	**99.2 (95.9-100)**	**92.3 (64.0-99.8)**	67.5 (60.4-74.1)
IV antibiotics	CRB65 =0 (low risk)	0.29 (0.24-0.34)	5.7 (2.1-11.9)	52.5 (42.3-62.5)	11.1 (4.2-22.6)	34.6 (27.1-42.7)
	CRB65=1/2 (moderate risk)	**0.67 (0.61-0.73)**	**84.0 (75.6-90.4)**	49.5 (39.4-59.6)	63.6 (55.0-71.5)	**74.6 (62.5-84.5)**
	CRB65=3/4 (high risk)	0.54 (0.51-0.57)	10.4 (5.3-17.8)	**98.0 (93.0-99.8)**	**84.6 (54.6-98.1)**	51.0 (43.8-58.3)
IV fluids	CRB65 =0 (low risk)	0.35 (0.30-0.41)	10.4 (5.1-18.3)	60.4 (50.6-69.5)	18.5 (9.3-31.4)	43.8 (35.8-52.0)
	CRB65=1/2 (moderate risk)	**0.60 (0.54-0.66)**	**78.1 (68.5-85.9)**	41.4 (32.2-51.2)	53.6 (45.0-62.0)	**68.7 (56.2-79.4)**
	CRB65=3/4 (high risk)	0.55 (0.51-0.58)	11.5 (5.9-19.6)	**98.2 (93.6-99.8)**	**84.6 (54.6-98.1)**	56.2 (48.9-63.3)
In-hospital mortality	CRB65 =0 (low risk)	0.36 (0.32-0.39)	0.0 (0.0-17.6)	71.3 (64.2-77.6)	0.0 (0.0-6.6)	87.6 (81.3-92.4)
	CRB65=1/2 (moderate risk)	0.53 (0.43-0.64)	**73.7 (48.4-90.9)**	33.0 (26.3-40.2)	10.0 (5.6-16.2)	92.5 (83.4-97.5)
	CRB65=3/4 (high risk)	**0.61 (0.51-0.71)**	26.3 (9.1-51.2)	**95.7 (91.8-98.1)**	**38.5 (13.9-68.4)**	**92.8 (88.2-96.0)**

The best predictive performance values (AUROC, sensitivity, specificity, PPV and NPV) have been highlighted for each score threshold

Table 6.10: Predictive performance of CATs score thresholds in adults with ALRTI

OUTCOMES	SCORE THRESHOLDS	AUROC (95%CI)	SENSITIVITY (%) (95%CI)	SPECIFICITY (%) (95%CI)	PPV (%) (95%CI)	NPV (%) (95%CI)
Admission	**CAT ≥3**	**0.77 (0.75-0.80)**	**59.6% (55.8%-63.4%)**	95.3% (90.5%-98.1%)	98.3% (96.4%-99.3%)	**34.6% (29.9%-39.4%)**
	CAT ≥4	0.66 (0.64-0.68)	32.7% (29.1%-36.4%)	99.3% (96.3%-100%)	99.5% (97.5%-100%)	24.8% (21.4%-28.5%)
	CAT ≥5	0.55 (0.54-0.56)	10.6% (8.3%-13.2%)	**100% (97.5%-100%)**	**100% (94.9%-100%)**	20% (17.2%-23.1%)
Suppl. oxygen	C**AT ≥3**	**0.77 (0.75-0.80)**	**79% (74.5%-83%)**	76% (71.7%-79.9%)	74.1 %(69.5%-78.3%)	**80.6% (76.5%-84.4%)**
	CAT ≥4	0.72 (0.69-0.75)	50.5% (45.4%-55.7%)	93.8% (91.1%-95.9%)	87.6% (82.4%-91.6%)	68.6% (64.7%-72.3%)
	CAT ≥5	0.59 (0.57-0.61)	18.4% (14.6%-22.6%)	**99.8% (98.7%-100%)**	**98.6% (92.3%-100%)**	58.5% (54.8%-62%)
IV antibiotic	**CAT ≥3**	**0.71 (0.68-0.75)**	**61.9% (57.8%-65.9%)**	81% (75.4%-85.9%)	89% (85.6%-91.9%)	**46.1% (41.2%-51.1%)**
	CAT ≥4	0.64 (0.62-0.67)	35% (31.1%-39.1%)	93.5% (89.6%-96.3%)	93.1% (88.9%-96.1%)	36.7% (32.8%-40.7%)
	CAT ≥5	0.56 (0.55-0.57)	12.1% (9.6%-15.1%)	**100% (98.4%-100%)**	**100% (94.9%-100%)**	31.4% (28.1%-34.9%)
IV fluid	**CAT ≥3**	**0.63 (0.59-0.66)**	**57.9% (53.6%-62.1%)**	67.7% (61.7%-73.3%)	**78.8% (74.5%-82.7%)**	**43.6% (38.8%-48.6%)**
	CAT ≥4	0.54 (0.51-0.57)	29.5% (25.7%-33.5%)	78.7% (73.3%-83.5%)	74.2% (67.8%-79.9%)	35% (31.1%-39%)
	CAT ≥5	0.51 (0.49-0.53)	9.5% (7.2%-12.3%)	**93.2% (89.4%-95.5%)**	74.3% (62.4%-84%)	33.2% (29.8%-36.7%)
In-hospital mortality	CAT ≥3	0.73 (0.70-0.77)	**91.8% (83%-96.9%)**	54.6% (50.9%-58.3%)	16.7% (13.2%-20.7%)	**98.5% (96.8%-99.5%)**
	CAT ≥4	**0.74 (0.68-0.79)**	69.9% (58%-80.1%)	77.4% (74.3%-80.5$)	23.5% (18%-29.7%)	96.3% (94.4%-97.7%)
	CAT ≥5	0.63 (0.58-0.69)	32.9% (22.3%-44.9%)	**93.8% (91.8%-95.4%)**	**34.3% (23.3%-46.6%)**	93.4% (91.3%-95.1%)

The best predictive performance values (AUROC, sensitivity, specificity, PPV and NPV) have been highlighted for each

6.5 Discussion

6.5.1 Summary of key findings

This study provides evidence of the predictive abilities of CRB65 and unweighted combined CATs scores for severity assessment of adults with acute lower respiratory tract infections (ALRTI) in Nigeria. On comparing the performance of these two prediction scores, the unweighted CATs was a better tool in terms of calibration and discrimination for predicting the need for hospital admission and hospital based clinical interventions in adults with ALRTI and pneumonia, while the CRB65 was a better at mortality prediction than at predicting outcomes associated with the management of adults with ALRTI and pneumonia.

6.5.2 Strengths and limitations

In addition to the other strengths of this research which have been discussed in previous chapters, it is worth mentioning that unlike a large number of clinical prediction studies where mortality prediction is often the main outcome of interest, this study incorporates other clinically relevant outcomes such as the need for supplemental oxygen and intravenous therapy which are just as important as predicting survival (160-162). Scoring systems which are only able to predict mortality risks may not be particularly helpful in medical surge scenarios or economically disadvantaged countries where clinicians face the ethical dilemma of deciding which patients are more deserving of life sustaining interventions when every patient deserves a right to live. The ethical theory of 'distributive justice' which suggests giving priority to the patients with the worst prospects i.e. highest mortality risks is open to interpretation as the worst cases do not necessarily make

significant recovery and result in inefficient use of limited resources(163) . In emergency triage it is now common practice for doctors to allocate scarce resources based on maximising the number of patients who will realistically make good recovery, such that a prediction rule which is able to identify patients who stand a greater chance of benefitting from the use of such resources may be more useful than a rule which only identifies high mortality risk patients (164, 165).

The lack of radiological evidence supporting a diagnosis of ALRTI or for differentiating between pneumonia and other ALRTIs in this study may be viewed as a limitation. In order to ensure generalisability of findings, it was important that this study was conducted to reflect real life practice in a resource-poor setting. Medical treatment in Nigeria is financed by out-of-pocket payments(166, 167), and because most people attending hospitals are unable to afford the cost of chest x-rays, doctors tend to rely on clinical features and symptoms when diagnosing ALRTI(168, 169). Furthermore, a recent publication from Nigeria conducted by Onyedum et al. (2011) provides support for the use of clinical diagnosis of ALRTI in this present study as the authors reported that CAP is usually diagnosed and treated on clinical grounds in Nigeria (170).

Lastly, this study did not report on the performance of the CATs and CRB-65 in terms predicting the use of mechanical ventilation and admission into intensive care units. Of the three participating hospitals in this study, only two had ICUs (LUTH and LASUTH), and there were only three ALRTI related ICU admissions recorded in the study population. Low ICU admission rates for patients with CAP have also been reported by other studies conducted in developed countries (21, 157), a possible reason for this might be that practice varies from hospital to hospital and ICU bed spaces are only allocated to the most critical cases based on the clinical judgement and local practices (171, 172). The qualitative study

findings in chapter 8 provide further insight into critical care practice in these hospitals and expand upon the factors which affect ICU admissions.

6.5.3 Comparison with existing literature

This is the first comparative validation of the CATs and CRB65 scores, the other studies on CATs were conducted in the United Kingdom and they have compared the CATs with other well validated prediction tools such as CURB-65 which is the version of the CRB65 which includes serum urea as a criterion in cases of pandemic flu (113, 118). Despite the fact that ALRTI is a major public health problem and a huge contributor to morbidity and mortality in Nigeria, there are a limited number of published studies from Nigeria on the use of clinical prediction tools in the management of ALRTI (170, 172-175).

To start off it is worth commenting that the mortality rate of ALRTI, more specifically pneumonia in this present study, is not far off from rates reported in the other Nigerian studies. Desalu et al. (2011) found that the leading cause of ALRTI visits to the ED in south western Nigeria was pneumonia (176) which was in agreement with findings from this present study. Onyedum et al. (2011) reported a mortality rate of 11.9% among patients with CAP in eastern Nigerian (170), which is comparable with this study's mortality rate of 12.6%.

The findings from this study have shown that both the CATs and the CRB-65 possess clinical validity as described in the study by Simon (2010) (136), as a unit increase in score for either of the tools was seen to be associated with greater risks of the study outcomes i.e. being admitted into a hospital, requiring intravenous and supplemental oxygen therapy or having an adverse outcome such as death. Myles et al (2012) reported a similar linear trend between

unweighted combined CATs scores and in-patient death in adult patients with PCR confirmed A/HINI influenza in the United Kingdom(118). Capelastegui et al. also reported that the risk of hospital admission and 30-day mortality increased with a unit rise in CRB-65 scores(157). Mbata et al. (2014) conducted the only study from Nigeria on the CRB-65 which compared the performance of the CRB-65 and CURB-65 in terms mortality prediction and ICU admissions, reporting an increase in the risk of mortality with higher CRB-65 scores among adults with CAP (174).

Clinical validity which can be further described in terms of the discriminatory and calibration qualities of a tool (136, 140) was further assessed in this study and on average AUROC values for CRB-65 in patients with ALRTI were lower than that of unweighted CATs score for all study outcomes. A similar pattern in performance was observed in the sensitivity analysis conducted in patients with pneumonia; however there was only a marginal difference in performance between unweighted CATs (0.73) and CRB-65 (0.71) in terms of AUROC values. These findings agree with the evidence reported in the systematic reviews conducted by Chalmers et al.(2010) and Loke et al. (2010) which concluded that CRB-65 was a better at mortality prediction than at predicting other outcomes associated with the management of CAP (177, 178). Authors like Marrie et al. (2005) and Bont et al. (2007) have suggested that mortality risk alone should not be the sole factor for deciding on hospital admission (161, 179) especially as a patient's functional state could decline without being initially classed as high risk(155). Based on this, the CATs score which provides more information on other relevant clinical outcomes might be a more useful tool for deciding on hospital admissions.

Another important and commonly overlooked measure of clinical validity is the calibration of a tool which represents the accuracy of its predictions in terms of how well they match the actual observed effects (127). The unweighted CATs

score was well calibrated for all outcomes in patients with ALRTI while the CRB-65 score had good calibration for all outcomes except intravenous antibiotics in this group. CATs has also shown good calibration for death, admission to higher levels of care, and a combination of these two outcomes in the FLU-CIN cohort as reported by Myles et al. (2013) (118). On the other hand both tools were less well calibrated in the sensitivity analysis including only pneumonia cases. It is not unusual to find a tool performing poorly when tested in a new group of patients (180) , and in this study there is a possibility that the patients with pneumonia were too similar in terms of severity seeing as majority of such cases required admissions and other interventions thus introducing a degree of confounding which might have affected calibration. Also calibration has been known to be affected by sample size(144), as such it is expected that the performance of the two scoring systems in terms of calibration will be poorer as the smaller number of patients included in the sensitivity analysis would result in lack of power to reject the null hypothesis. One thing remained certain; the role of CRB-65 as a validated tool for mortality prediction in adults with CAP was evident in the sensitivity analysis as it was seen to be well calibrated for this outcome while the unweighted CATs score was not.

Clinical utility defined by Simon (2010) as the ability of the prediction tool to influence clinical decisions towards improving management outcomes (136) was also assessed by comparing the performance of the threshold values of the two scores for predicting study outcomes. An unweighted CATs score of $\geq$3 and a CRB-65 score of 1 or 2 had the highest sensitivity values and were the best predictors of all outcomes in terms of AUROC values except for in-patient mortality which was best predicted by an unweighted CATs threshold score of $\geq$4 and a CRB65 of 3 or 4 . When Myles et al. (2012) compared threshold scores of the unweighted

CATs, PMEWS and CURB-65, the unweighted CATs threshold score of three or more was the best predictor of the need for higher levels of care and death in adults.(118). Although there may be good enough reason to suggest an unweighted CATs score of ≥3 as a cut-off for identifying the need for hospital based care and mortality, the CATs threshold scores have not been validated in the community, and it is possible that other thresholds of the CATs score may be better suited to different settings or scenarios which may require a higher triage bar. In the context of a high-bar triage scenario such as a pandemic, resource rationing and allocation is a major ethical dilemma and patients will need to be prioritised according to their likelihood of benefiting from an intervention. If the CATs is to be used in a pandemic scenario, a reassessment of the unweighted CATs score thresholds may be necessary as a threshold score of 3 or more appears to indicate a global idea of severity without indicating which patients are too critically ill to benefit from aggressive interventions. Although a CATs threshold score of ≥4 was the best predictor of in-hospital mortality in this study, future research could explore if patients with high CATs score were less likely to survive despite receiving acute care interventions, this could be done retrospectively as a prospective evaluation may present ethical issues.

6.5.4 Summary

The CRB65 and the Unweighted CATs demonstrated good performance for predicting the need for hospital based care and mortality in a non-UK adult population with a range of acute lower respiratory tract infections. Overall, the unweighted CATs score performed better at predicting the need for hospital based care and mortality in adults with ALRTI. The CRB65 was shown to be a better

mortality predictor in adults with pneumonia, while the unweighted CATs score was better at identifying patients in need for hospital admission and acute care interventions.

7: Assessing the value of clinical intuition in the performance of the un-weighted CATs score for severity assessment of patients with ALRTI

7.1 Introduction

Preliminary findings from this thesis have shown the CATs criterion G (causing clinical concern), to be an independent predictor of important clinical outcomes in both adults and children with ALRTI. This chapter seeks to explore what added prognostic value a clinician's gut feeling or intuition contributes to the performance of the unweighted combined CATs scoring index for severity assessment of patients with ALRTI.

This chapter will begin by reviewing pertinent literature on the processes involved in medical decision making in order to contextualise the study's aims. The other sections will include the statistical analysis and findings from comparing the performance of an unweighted CATs score excluding the subjective criterion G with that of an unweighted CATs score with all seven criteria (including criterion G) in predicting hospital care, therapeutic interventions and death on admission.

7.2 Decision making in clinical practice

With clinical judgement being the mainstay of medical practice, the processes employed by clinicians in arriving at diagnostic and prognostic decisions has been an actively researched topic in medical literature(181). Understanding these processes appears to be a key step towards improving the performance of clinicians by reducing diagnostic error rates in practice(182), and over time two

viewpoints with origins in the field of psychology have been seen to dominate the literature on medical decision-making. Paired under the term 'dual-process theory', the two processes by which clinicians decide on a clinical course of action have been described as non-analytical (intuitive) and analytical (conscious) or in neutral terms, as system 1 and system 2 (182-184).

System 1 or the intuitive process represents a fast and unconscious response to a situation, while system 2 or analytic thinking is conscious and much slower in comparison. Most decisions made using the system 1 process have been associated with an emotional component which has earned it the moniker 'gut-feeling' and in medical practice it is has been linked with the use of heuristics such as pattern-recognition and rules of thumb (182, 183, 185, 186). The system 2 analytical process as the name suggests employs deliberate thinking and mental effort making it more objective (182, 184). By virtue of this, common sense will naturally consider the more evidence driven system 2 (analytical) as more accurate, compared to the intuitive system 1 approach which favours the use of mental shortcuts. However there are studies which suggest that both models of thinking are equally prone to cognitive biases and in some cases intuition is a stronger and better predictor of accuracy in medical decision making especially in the early stages of patient consultation (187-192). There is also evidence from studies comparing the reasoning of experienced clinicians (experts) and less experienced junior doctors (novices), that experts engage in less analytical reasoning, and make faster and successful decisions based on their gut feelings or intuitions(184, 192, 193). While these findings might fuel a debate on superiority, it is worth noting that experts were once novices who over years of clinical practice acquired a vast amount of knowledge and experience to create a knowledge bank which formed the basis of their analytic reasoning thus enabling

them to use the intuitive system 1 approach to achieve a better success than novices at recognising disease patterns; a phenomenon described as tacit knowledge (182, 192, 194, 195).

Some authors have also suggested a bidirectional relationship between these two systems of reasoning with neither being superior to the other, and propose that analytical and intuitive reasoning can be applied holistically to arrive at a clinical decision under the cognitive continuum theory (186, 191, 196-198). The expert group behind the development of the Community Assessment Tools appeared to be thinking along these lines when they included a subjective criterion (G: causing clinical concern) as one of the seven CATs criteria. According to the findings from the multivariable analysis involving all the CATs criteria in both adults and children; this criterion remained an independent predictor of important outcomes such as the need for hospital admissions, acute care interventions and in-hospital mortality in this thesis (see chapter 4), as well as length of stay $\geq$6 days, mechanical ventilation and combined severe outcomes (level 2/3 admission or death)in the CATs evaluation study conducted by Semple et al. (2013) using FLU-CIN data (35).

Although these findings are in agreement with the initial premise for using the CATs in times of extreme medical surge which proposed that meeting any one of the seven criteria would prompt urgent referral or hospital admission, the exact value of this subjective criterion in comparison to the other six objective criteria remains unknown when all CATs criteria are combined into a composite unweighted scoring system. This study aims to explore if considering gut feelings or intuitions adds any value to clinical decision making by comparing the performance of unweighted CATs score excluding the criterion G to that of an

unweighted CATs score inclusive of this criterion for predicting the need for hospital admission, interventions and in-hospital mortality.

7.3 Statistical Analysis

As a result of the limitations associated with using statistically derived weights to create scores (see chapter 5), this study will only assess the value of the subjective criterion in unweighted CATs scores. The methods and statistical analyses employed in this study are also similar to those used in chapter 5 for comparing unweighted and weighted combined CATs scores.

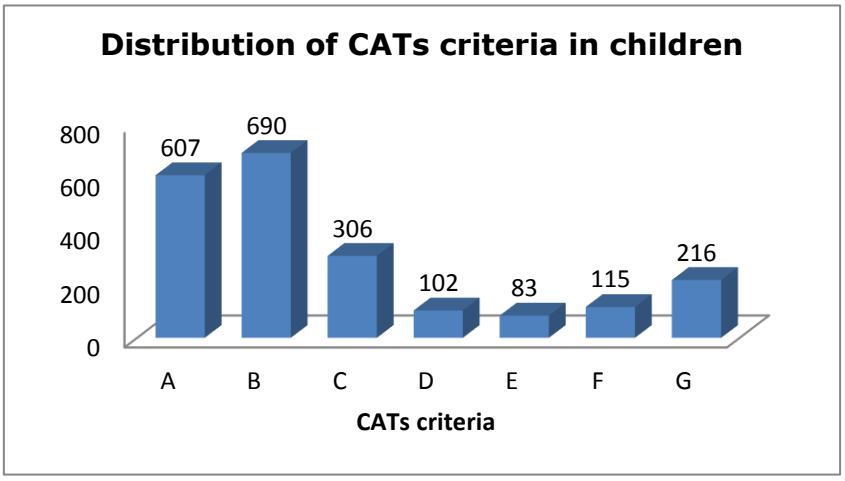
An unweighted CATs scoring index with all of its 7 (6 objective and one subjective) criteria was created by allotting a single point to each criterion and the total score for each patient was calculated by summing up the points for all criteria present. Similar steps were followed to create a second unweighted CATs score with only six objective criteria A to F, and for clarity these two scoring systems will be referred to as unweighted CATs-7 and unweighted CATs-6 scores. Using receiver operating characteristic (ROC) curves and Area under ROC (AUROC) with 95% confidence intervals values, the performance of each variant of the unweighted CATs scoring index in predicting the need for hospital admissions and acute care interventions, or the risk of in-hospital mortality were investigated and compared in children and adults with ALRTI.

7.4 Results

7.4.1 Comparing the predictive performance of the unweighted CATs-7 versus CATs-6 scoring indices in children (<16 years)

There were a total of 216 instances where criterion G was selected in children (< 16 years) in this study, making it the fourth most commonly selected CATs criterion after criteria A: severe respiratory distress, B: increased respiratory rate and C: low oxygen saturation $\leq$ 92% (Figure 13). Of these 216 cases, there were 15 instances where criterion G alone was selected, with 2 of such cases dying on admission.

Figure 13: Pattern of CATs criteria selection in children <16years



AUROC values presented in Table 7.1 for the unweighted CATs-7 and CATs-6 scores, suggest that their predictive powers were virtually within the same range. The range of values for the unweighted CATs-7 scores [AUROC (95%): 0.65 (0.61-0.69) – 0.85 (0.82-0.88)], were slightly higher than the unweighted CATs-6 scores [AUROC (95%): 0.63 (0.59-0.68) – 0.84 (0.81-0.87)], with only a marginal difference of 0.01-0.03 between their values. The difference in the discriminative ability was only significant for predicting the need for intravenous fluid and the

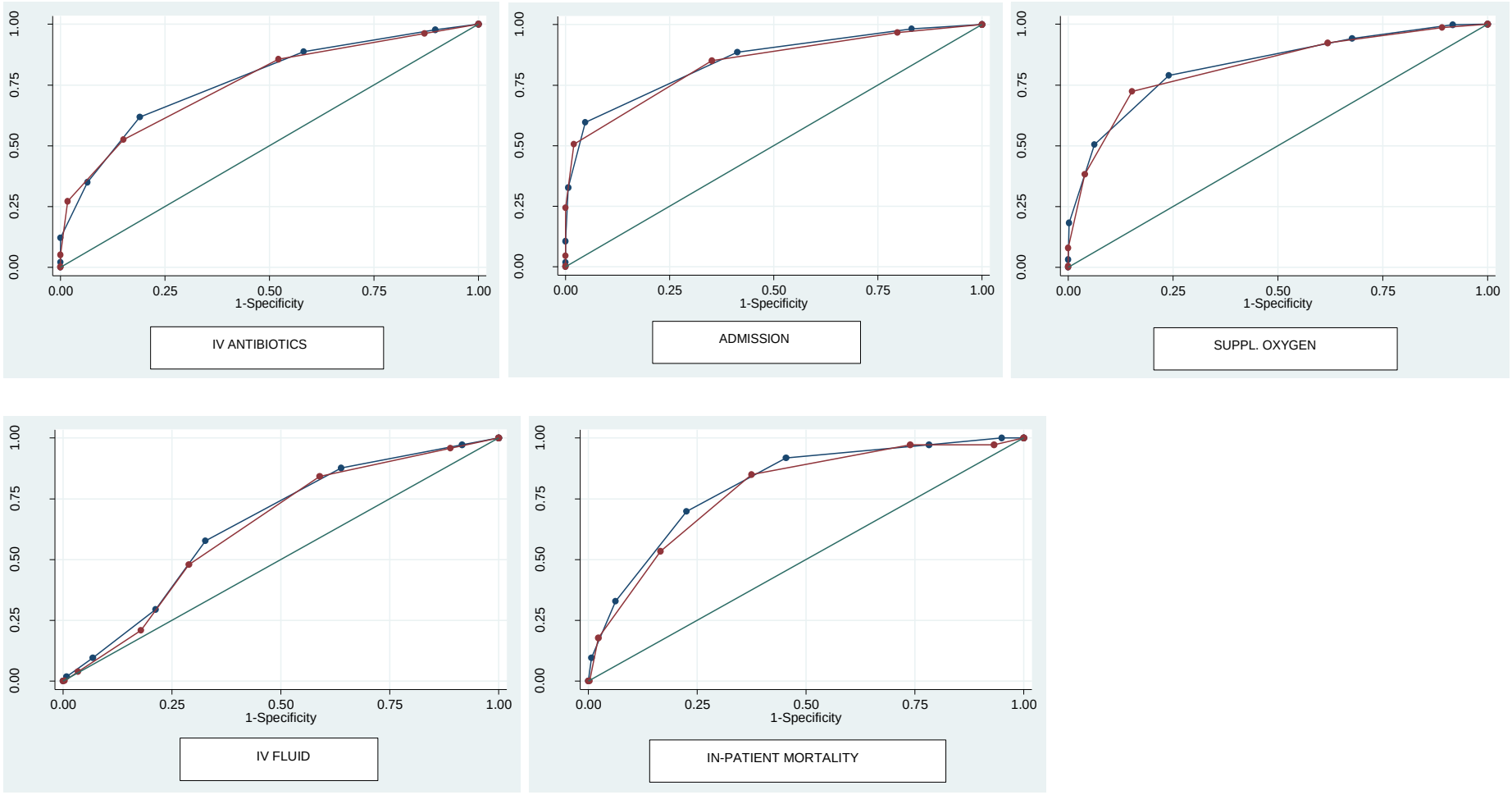
risk of in-patient mortality, with CATs -7 which included a subjective criterion being better than the CATs-6 score with only objective criteria.

Table 7.1: comparison of AUROC values of unweighted CATs with all 7 criteria versus CATs without criterion G in children

OUTCOMES	CATs scores	AUROC (95% CI)	P-value
Admission	CATs-7	0.85 (0.82-0.88)	0.068
	CATs-6	0.84 (0.81-0.87)	
Suppl. Oxygen	CATs-7	0.83 (0.80-0.86)	0.297
	CATs-6	0.82 (0.80-0.85)	
IV antibiotic	CATs-7	0.77 (0.73-0.80)	0.111
	CATs-6	0.76 (0.72-0.77)	
IV fluid	CATs-7	0.65 (0.61-0.69)	**0.026**
	CATs-6	0.63 (0.59-0.68)	
In-patient mortality	CATs-7	0.81 (0.76-0.85)	**0.038**
	CATs-6	0.78 (0.73-0.83)	

The performance of the two unweighted CATs scores are presented for comparison using ROC curves in Figure 14, the blue lines which represent the CATs-7 score appear to be farther away from the diagonal line intersecting the plot than the red lines (CATs-6) for all study outcomes. This suggests that the CATs-7 score was slightly better than the CATs-6 without the subjective criterion G for predicting admissions, acute care interventions and in-hospital mortality.

Figure 14: ROC curves comparing unweighted CATs with all 7 criteria versus CATs without criterion G in children

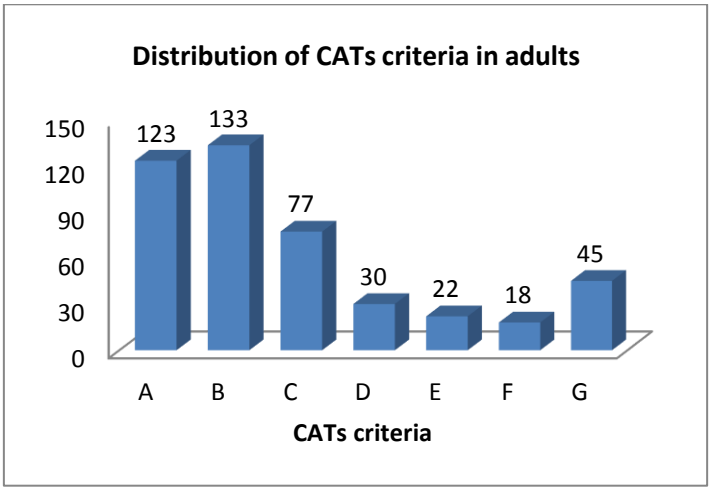


Blue line = unweighted CATs with all 7 criteria ——— Red line = unweighted CATs without subjective criterion G ———

7.4.2 Comparing the performance of unweighted CATs with all 7 criteria versus CATs without criterion G in adults (>16 years)

The distribution of the individual CATs criteria in adults shown in figure 15 is similar to the pattern in children (Figure 15), with criterion G (causing clinical concern) being the fourth most selected criterion for predicting the study outcomes after A: severe respiratory distress, B: increased respiratory rate and C: low oxygen saturation $\leq$ 92%. Of the 45 instances were criterion G was selected, the other six objective criteria (A – F), were not selected in 4 of these cases and although these four adults were admitted, none died on admission.

Figure 15: Pattern of CATs criteria selection in adults >16 years



Both variants of the unweighted CATs scoring index appear to be comparable in terms of their AUROC values presented in Table 7.2 as their AUROC values were within the same range. The seven-criteria CATs-7 score [AUROC (95%): 0.72 (0.65-0.79) – 0.91 (0.88-0.95)], appears to have a slightly better range of AUROC values than the unweighted CATs-6 score [AUROC (95%): 0.71 (0.64-0.78) – 0.90 (0.86-0.95)], without a subjective element. However, the difference between the

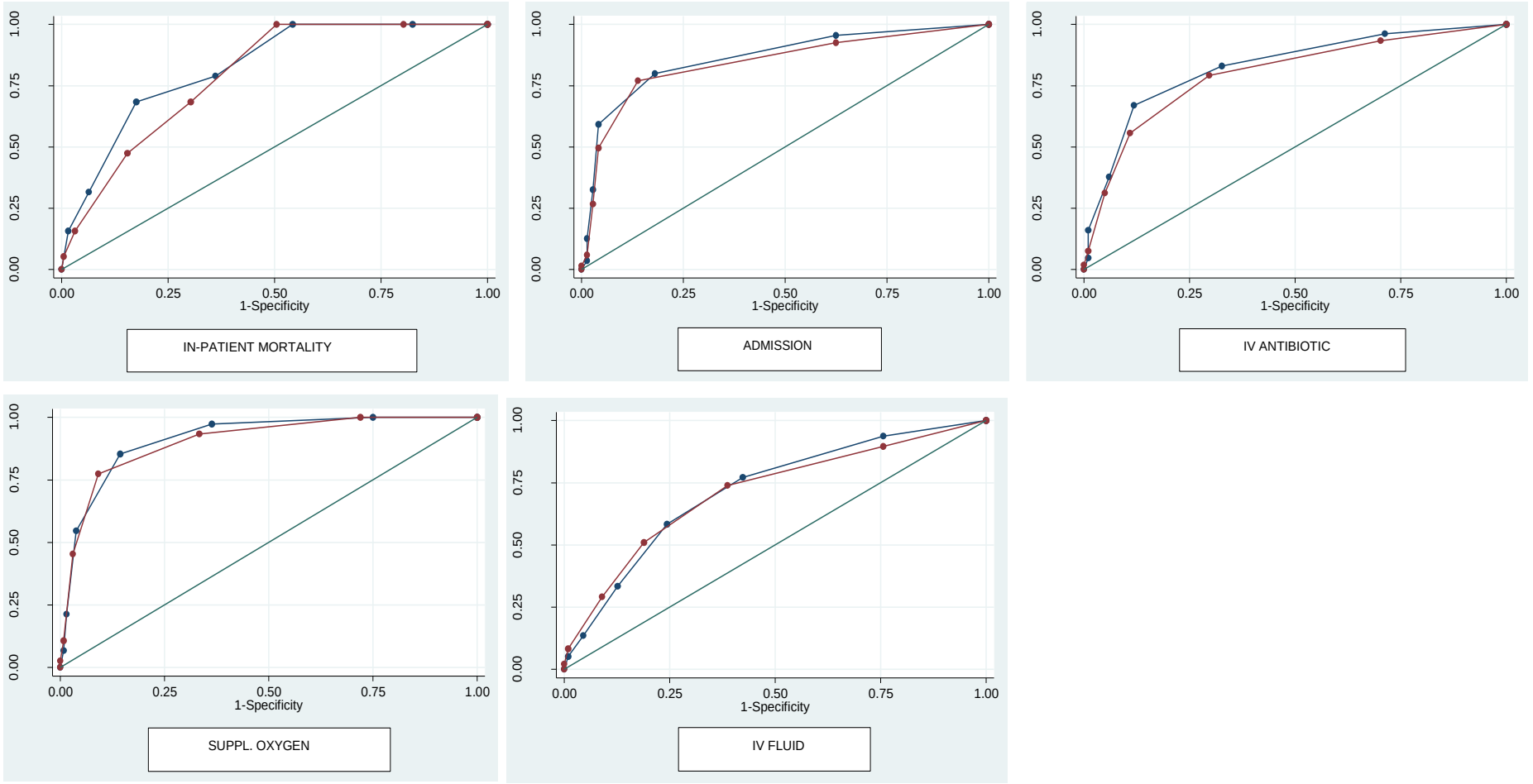
two scores was only significant for predicting hospital admission [CATs-7: 0.86 (0.81-0.91) vs CATs-6: 0.84 (0.79-0.90)], needing IV antibiotics [CATs-7: 0.91 (0.88-0.95) vs CATs-6: 0.90 (0.86-0.95)] and in-patient mortality [CATs-7: 0.82 (0.74-0.91) vs CATs-6: 0.79 (0.71-0.87)].

Table 7.2: comparison of AUROC values of unweighted CATs with all 7 criteria versus CATs without criterion G in adults

OUTCOMES	CATs scores	AUROC (95% CI)	P-value
Hospital admission	CATs-7	0.86 (0.81-0.91)	**0.02**
	CATs-6	0.84 (0.79-0.90)	
Suppl. Oxygen	CATs-7	0.91 (0.88-0.95)	0.218
	CATs-6	0.90 (0.86-0.95)	
IV antibiotic	CATs-7	0.83 (0.77-0.88)	**0.001**
	CATs-6	0.80 (0.74-0.86)	
IV fluid	CATs-7	0.72 (0.65-0.79)	0.533
	CATs-6	0.71 (0.64-0.78)	
In-patient mortality	CATs-7	0.82 (0.74-0.91)	**0.057**
	CATs-6	0.79 (0.71-0.87)	

The ROC curves in Figure 16 comparing the performance of the two variants of unweighted CATs scores show that the unweighted CATs-7 score represented by the blue line was better than the unweighted CATs-6 score (red lines) for all the study outcomes, as the CATs-6 was slightly closer to the diagonal line in green intersecting the plot.

Figure 16: ROC curves comparing unweighted CATs with all 7 criteria versus CATs without criterion G in adults



Blue line = unweighted CATs with all 7 criteria

Red line = unweighted CATs without subjective criterion G

7.5 Discussion

7.5.1 Summary of key findings

Based on the reported distribution of the subjective CATs criterion G in both children and adults, this study shows that intuitions are a common element of decision making in medical practice. Furthermore, the study findings show that considering intuitions or gut feelings in addition to the assessment of a patient based on clinical signs and physical examinations, will improve the identification of adults and children with acute lower respiratory tract infections who are likely to benefit from hospital based care and clinical interventions or were at risk of death. This study therefore provides some justification for the inclusion of a subjective criterion in the community assessment tools.

7.5.2 Limitations and strengths

As earlier discussed in the methods chapter of this thesis, most of the data for this thesis were extracted from case notes by trained data collectors, and it is likely that certain instances where doctors may have acted on their intuitions may have been missed as 'clinical intuitions' are usually difficult to explain or put into words (199). Although the occurrence of such unrecorded instances of subjective decision making in the assessment of the patients may have biased the findings from this study, there is also a possibility that the added value of subjective clinical judgement in the performance of the CATs-7 score may have been underestimated. It was also not possible to ascertain if the level of expertise of the doctors assessing each patient had an impact on the accuracy of their subjective judgement, and as such, it is unclear if the CATs-7 score performed

better than CATs-6 because more senior doctors may have assessed the patients. However, it is unlikely that this was the case as junior doctors tend to have the initial contact with patients in most EDs, and it is safe to assume that the instances where criterion G was selected in this study were not biased by the expert reasoning of more senior doctors. However, as the last statement was purely conjectural, further research is required to compare patient assessments made by senior doctors using the CATs score versus those made by junior to mid-level doctors.

While most studies have evaluated the use of clinical intuition in medical decision making by simply reviewing case studies or interviewing clinicians, the major strength of this study is the use of a prediction tool in examining the value of clinical intuition which allowed for an unbiased comparison in a large number of unselected ALRTI cases and provided more robust results. Furthermore, this was a multi-centre study involving different tiers of hospitals i.e. tertiary and secondary care, which makes the study findings more generalisable.

7.5.3 Comparison with existing literature

The study has shown that the presence of criterion G (causing clinical concern) albeit marginal, improved the performance (albeit marginally) of the unweighted CATs score for predicting mortality in children and adults with ALRTI. Brabrand and colleagues (2014) investigated the accuracy of in-hospital mortality predictions using clinical intuition in adult patients ($\geq$15 years) presenting to the medical assessment unit of large teaching hospital in Denmark, and reported that predictions made by hospital staff using clinical intuitions were accurate and also improved when used in conjunction with an objective scoring system. Although most of the predictions in the study by Brabrand et al. (2014) were made by

nursing staff, these findings are comparable to this present study as the authors have also assessed predictive accuracy in terms of discrimination using AUROC and 95% CI values. A study by Van den Bruel et al.(2012) also reported that acting on gut feelings could potentially prevent diagnostic errors in 33% of children who were clinically assessed as having non-serious illness(200). Marcin et al. (2000) also compared clinician's subjective predictions of mortality with a physiological- mortality based score in children using AUROC and reported that the combined use of these two prediction methods was superior to the use of either methods. On the other hand the inclusion of clinical intuition in the assessment of patients in this study did not result in a significant difference in the performance of the unweighted CATs score for predicting other outcomes such as admission, and the need for IV antibiotics and supplemental oxygen therapy. A possible reason for this in terms of predicting the need for supplemental oxygen is that clinical signs of hypoxemia such as increased respiratory rate or bluish discoloration of the lips (cyanosis) which might trigger clinical concern are less specific for predicting reduced oxygen saturation compared to the use of a pulse oximeter (201-203). Another reason for this might be that most of the reasons given for the clinicians' intuitions in this study were based on clinical findings and the general appearance of the patients. The most commonly cited reasons for selecting criterion G in the assessment of the patients in this study were based on auscultation findings such as crepitation or reduced air entry, rapid deterioration of the patient's clinical status or an unusually prolonged illness, which were consistent with the suggestion made by Van den Bruel et al.(2012) that gut feelings were not entirely devoid of input from physical examinations and clinical findings(200). Based on this it can be said that pattern recognition a system 1 approach might have been used by doctors in this present study in conjunction

with clinical findings based on system II analytical processes to arrive at a medical decision. Clinicians would have previously learnt during medical training which symptoms or signs are associated with worsening of a patient's condition, and over time this knowledge is applied in similar situations almost without deliberate thinking. Djulbegovic et al. (2012) identified a number of factors involved in the system I or intuitive reasoning and these authors believe that intuitive decision making is rooted in 'over-learned processes' or knowledge gained from cognitive learning strategies typical of system II processes(204).

There are two opposing views on the correlation between level of clinical experience and the ability to make decisions using intuition; while some authors are of the belief that more experienced clinicians compared to novices or junior doctors are more adept at using their intuitions(181, 205-207), there is also evidence which disproves this theory(200, 208). As this study did not report on the level of experience of the doctors who assessed the patients, it is difficult to ascertain the contribution of the clinical experience of the doctors to the performance of CATs-7. However, it is highly unlikely that senior doctors or consultants would have been involved in the initial assessment of patients in this study; rather it is more likely that junior level to midgrade (registrar) level doctors were responsible for the initial assessments of the patients from whom data were collected. If this assumption holds then it bodes well for the use of CATs which was designed for use by front line medical staff who would most likely be less experienced medics in the event of a surge or pandemic. In line with the different views on intuition and clinical experience, further research is needed to determine if the level of experience of doctors using the CATs score has any effect on the validity of their intuitions for the assessment of illness severity.

Notwithstanding, this study shows that there are benefits in including a subjective criterion in the Community assessment tool, particularly in the context of a resource limited setting such as Nigeria or in the event of an extreme surge scenario, where time and resource constraints may limit the extent of diagnostic investigations available, doctors will have to base their judgements on cues, intuitive insights and gut feelings. Furthermore, blind algorithmic compliance with clinical prediction tools and evidence based medicine may devalue the process of clinical reasoning and deprive clinicians of applying their 'individual clinical expertise'. In conclusion, criterion G should be retained in the CATs as its inclusion not only adds value (albeit minimal) to the performance of score, but will also increase clinician's willingness to use it.

7.5.4 Summary

The seven criteria CATs score was significantly better than the six criteria CATs score without criterion G (causing clinical concern) at predicting the need for hospital admission and also in-hospital mortality. As such clinicians ought to pay more attention to early clues and gut-feelings as clinical decisions which incorporate both intuitions and clinical evidence are less liable to errors of cognitive bias.

8: A Qualitative Exploration of Critical Care practices in Nigeria: Findings and Discussion

8.1 Overview

This chapter reports on the findings of the qualitative study exploring critical care practice in Nigeria (the methods were described in Chapter 2). It begins by describing the study participants and the core themes generated from the analysis of the interview transcripts substantiated by illustrative quotes from the participants' responses. The second part of this chapter reviews and discusses these findings in relation to relevant published literature and critical care systems in other settings. A reflexive statement on how the researcher's background and perceptions might have influenced the study is also included in the discussion.

8.2 Description of Participants

Fourteen healthcare professionals were interviewed for this study between September and October 2014; 10 were from the two tertiary care hospitals (LUTH=6 and LASUTH=4) while the other 4 were from the secondary care hospitals (RGH=3 and MSCH=1). There was only one nurse in this study while the other thirteen participants were doctors, with the majority of the participants working in LUTH (n=6) and LASUTH (n=4). Doctors at all levels of practice, representing all the major medical specialities except obstetrics and gynaecology participated in this study. The demographic characteristics of the participants are summarised in Table 8.1.

Table 8.1: Demographics of interview respondents (n=14)

Number of respondents in each Hospital (14)	LUTH	6
	LASUTH	4
	MSCH	1
	RGH	3
Healthcare level of hospitals (4)	Tertiary care	2
	Secondary care	2
Speciality of respondents	Medicine	2
	Surgery	3
	Anaesthesia	4
	Accident & Emergency	2
	Paediatrics	2
	Nursing	1
Gender	Male	10
	Female	4
Practice level	Consultant	1
	Senior registrar	4
	Junior registrar	4
	Senior medical officer	2
	Medical officer	2
	Senior nursing officer	1

8.3 Overview of themes

Besides some necessary adaptations to allow for hospital specific differences and level in the healthcare system, participants' responses indicated that the study centres were comparable in structure and delivery of critical care services. The three overarching themes identified from analysing participants' responses to questions regarding critical care practice in the four study hospitals are presented along with corroboratory text extracts in the next few sub-sections. A supplementary table with the themes and all relevant data extracts or quotes from the analysis can be found in appendix XI.

8.3.1 Definition and structure of critical care

The first theme to emerge was actively elicited during interviews, as all participants were asked at the start of the interview to explain what they understood by critical care. Participants defined critical care in terms of a patient's clinical stability i.e. critically ill; patients were 'critically ill' if they needed single or multiple organ support and as such respondents associated critical care with supporting such patients who had lost function of one or more organ systems.

"It is basically taking care of critically ill patients who need multi organ support" (ASR-LASUTH)

"*…means intensive care provided in specialized units for patients that are critically ill or have one or more organ system failure, and is essentially the same for adults and children" (SJR-LUTH)*

Participants also described critical care in terms of the level of monitoring received by critically ill patients, as compared to less critically ill patients in terms of

patient-health care worker proportion i.e., critical care involved a higher nurse to patient ratio.

"*It equally means intensive care, where patients are being cared for or given a high level of care in terms of patient care giver proportion whereby there is a nurse to one patient, the monitors; the level of monitoring is more than what is found in an ordinary ward" (AJR-LUTH)*

"Well on the general ward the support or care usually is about 2 nurses to 24 patients while the HDU and ICU operate on a 1:1 (a nurse for a patient)" (SSR-LUTH).

"We try to sort out patients that are critically ill , which means haemodynamically unstable patients or with severe organ dysfunction and would benefit from closer monitoring than other patients, based on that recommendation we can get them into critical care" (MJR-LASUTH)"

The structure and delivery of critical care was quite distinct between the tertiary and secondary care hospitals; only the two tertiary care centres (LUTH and LASUTH) had functioning intensive care units. Although participants were aware that there were different levels of critical care, it was not clear if they knew what these levels were: staff in LASUTH had conflicting ideas about the structure of critical care unit in their hospitals, and it appeared as though neither of the hospitals with ICUs had a designated high dependency unit.

"The levels are the ICU[1], the HDU[2] or the recovery room, and the operating room/theatre" (AJR-LUTH)

[1] Throughout the interview respondents referred to the intensive care unit using the abbreviation ICU
[2] They also referred to the high dependency unit as the HDU

"Yes there are different levels, even in the main emergency room we have some patients we keep in the resuscitation room where they have dedicated oxygen administration with hourly monitoring of their vitals. They also have access to suctioning and are typically kept closer than other patients to the nurses' station. Then there is the ICU unit where patients require mechanical ventilation and will have a dedicated anaesthetist on call and also specially trained ICU nurses in addition to the primary team handling that patient." (MJR-LASUTH)

"*In LASUTH we do not have different levels of CC; we have just the intensive care unit. We have a 4 bedded ICU, but even though it is 4 bedded we have rapid turnovers" (ASR-LASUTH)*

"*In the theatre we have a recovery room which performs the functions of an HDU and we have a 4 bedded ICU" (SSR-LASUTH)*

The participants in LUTH described a recovery room in the theatre that was used mostly by the surgery department when they needed to keep an eye on a patient following surgery. This recovery room which was situated beside the ICU functioned as a high dependency unit as patients who did not meet the criteria for ICU admission were kept there for close monitoring.

"We have the ICU which is shared by all departments, however the surgical specialties have access to a portion of the recovery room which functions as a HDU, it is just beside the ICU but not within the ICU and sometimes patients who are critically ill but have not met the criteria for being in the ICU might stay there and may benefit from some level of intensive care from anaesthetists, and nurses. Sometimes they might have access to ventilators. If they recover they might be taken to the open ward but if they deteriorate then they are taken into the ICU" (SJR-LUTH)

"In LUTH presently we have HDU, that is the high dependency unit, so that is one level lower than that of the ICU; well in the ICU , we have five beds, but it is not limited to adults alone, beds are allocated as the patients come, so it's first come first serve" (ASR-LUTH)

It was also evident that these hospitals had modified their delivery of critical care to suit the level of skills and resources available to them. The secondary care hospitals did not have the capacity or manpower to manage critically ill patients and referred such patients to one of the tertiary centres after stabilising them.

"*Unfortunately we do not have an ICU in this hospital so the best thing we can do here is that we stabilise any critically ill child and refer to a tertiary institution for further management" (SMO-MASSEY)*

"No we don't have an ICU. We refer such patients to LUTH, LASUTH and general hospital Lagos" (SR-RGH)

"We have an oxygen concentrator and tanks, suction but when you talk about ICU care you are referring to nurse patient ratio and doctor patient ratio which we can't provide" (SR-RGH).

8.3.2 Factors affecting critical care admission and discharge

It was evident that the demand for critical care outweighed the facilities available and with both tertiary centres having approximately 10 ICU beds between them (LUTH=6, LASUTH=4), the most common factor affecting admission was availability of bed space.

"Okay, well firstly, the commonest factor is availability of bed space. There are limited beds in the ICU" so if there is no space available it becomes difficult to admit patients." (MJR-LASUTH)

"*We only have 6 beds here and one is for isolation. Most times we have the issue of lack of bed space and we try to use our HDU from time to time" (SNO-LUTH).*

In situations where there are more patients than ICU beds, a decision was made to determine who would be admitted, and most times this was a joint decision between the consultant anaesthetist and the head of the managing unit with clinical judgement playing a significant role:

"It's a joint decision between the consultant anaesthetist and the consultant of the managing team" (SSR-LUTH)

"Both the managing team and the anaesthetist are involved; the managing team decides who needs critical care based on assessment and the anaesthetist also reviews prior to accepting the patients" (MJR-LASUTH)

These decisions were dependent on a number of patient-related factors; several participants mentioned the prognosis of the patient as an important criterion for determining which patient would be admitted. Based on clinical judgement, patients who had a better prognosis and were more likely to recover were given precedence over patients with terminal illnesses or other irreversible conditions. Besides prognosis, age was the other patient factor, as younger patients were considered stronger and more likely to benefit from ICU admission than older patients.

"Well, one thing is that at times we have more people trying to come into the ICU than the number of bed spaces available, so what we usually do is first of all see cases that

are reversible, so not terminal stage cases. So the younger people that we know can benefit from the ICU care are given preference before the patients with terminal illnesses or with bad prognosis". (ASR-LUTH)

"There are certain things we consider, we look at patient's age, the illness if it is possible that the patient will recover, we do a quick triage" (SNO-LUTH)

Participants also mentioned that critical care admission was more expensive than being admitted on the wards, one of the participants gave the cost of ICU admission for a week as two hundred and fifty thousand naira (approximately seven hundred and fifty pounds), which most patients were unable to afford.

"Yes you see finance and bed space are indeed major factors in LASUTH, as I said we have only just 4 beds. And we may have 6 people requiring admission into the ICU at the same time so we need to triage and find out who will benefit from ICU and who will not. Not all the patients who want to come into the ICU will benefit from the admission so we have to triage. Also if they do not have enough money how do you manage them? This is a general hospital and patients buy all they need for care, it's out of pocket payment they pay for everything from syringes to drugs. So if they cannot procure all that is required it will be impossible to manage such patients. So the monetary aspect is very important it is a major factor for admission into ICU in LASUTH, not even LASUTH alone; everywhere any ICU" (ASR-LASUTH).

"Yes definitely, you find patients requiring admission into the ICU where the patient is not able to afford the fee, also bed space is a factor, here we have a five bed ICU, and if the beds are filled that becomes an issue as well" (AJR-LUTH)

There were no restrictions on length of stay in an ICU, and once admitted the patients stayed in the ICU for as long as they could afford or until they were

deemed fit for discharge; participants attempted to give examples of how long a patient could stay in an ICU to show variations in length of stay.

"We have patients staying longer than 4/5 months." (SNO-LUTH)

"I have seen patients stay from as little as a few days and others up to 120 days…." (SJR-LUTH)

Participants suggested that the reason why some patients continued to stay in the ICU was because of the absence of a functioning HDU. Most of the patients in the ICU were out of the critical state and could be managed in a high dependency unit, but in the absence of this they either stayed longer than required in the ICU or were taken straight to the wards.

"*Most of our patients stay for long periods, because we do not have the HDU unit to keep them in when their condition is less critical, so we make sure they are fine and fit in the ICU before we send them to the ward"(ASR-LASUTH)*

"*If we had a functioning HDU we might be able to step down care for patients in the ICU. But I am afraid most of the patients here go straight from the ICU to the regular wards or from wards to the ICU where there might not be any other form of monitoring but I am aware that some wards now have electronic monitors" (SSR-LASUTH)*

"*…they go straight to the ward most times" (SSR-LUTH)*

As in the case of admission, the decision to discharge a patient rested on the managing consultant and anaesthetists; however, the patient's family also shared some of the responsibility for example in the event when a patient did not show signs of improving and needed to be taken off the ventilator:

"I have seen patients stay from as little as a few days and others up to 120 days. But the decision to move a patient out rests on the admitting consultant, the anaesthetists and the patient's family" (SJR-LUTH)

According to the participants, the two main criteria for discharge were obvious improvement in the patient's condition and the patient's ability to continue paying the hospital bills. Just as in the case of admission, if a patient was unable to keep up with payment an alternative arrangement was sought.

"Obvious clinical improvement, or in the case where the bed is needed for a more critical patient, we can step down the intensity of care of one patient to allow for another patient" (SJR-LUTH)

"The patient can stay as long as they can afford the bills or as long as is needed for care" (MJR-LASUTH)

In situations where the ICU was full and could not take any more patients, depending on the route of admission, patients in the emergency department who required ICU admissions would receive conservative management in the resuscitation room until there was a free bed in the ICU, while post-op surgical patients were managed in the theatre-recovery room. The other options would be referral to another tertiary institution, or a privately run ICU.

"It depends on the kind of patient it is; if it is a surgical case, requiring post up ICU care, we actually keep them in the recovery room, or use our ventilators. Actually in the theatre, that's the best we can do when the bed spaces are full, but they are usually managed on the wards and as soon as we get a free bed then we can quickly bring the patient in." (AJR-LASUTH)

"Most times we carry out conservative management and intubate the patients wherever they are." (SNO-LUTH)

"Well we do give them the option of referral but most of these patients are managed conservatively and i am afraid I cannot say what the outcome of such cases will be." (SSR-LASUTH)

8.3.3 ICU organisation and patient profile

According to the participants, post-op surgical patients requiring stabilisation or patients with head injury made up a large majority of the adult patients seen in critical care unit while paediatric patients in need of critical care were usually burn victims or children infected with tetanus. Acute lower respiratory tract infection was not a common indication for critical care admission as patients were usually managed on the wards. The participants believed that such patients were more suited for a HDU rather than an ICU:

"*The spectrum includes both medical and surgical patients… the common denominator for these patients will be that their cardiopulmonary function cannot be sustained unsupported." (CO-LUTH)*

"Most are post-op, but we still get people from outside the hospital for ICU care. They could be post -op or medically ill patients" (ASR-LASUTH)

"Yes we do get ALRTI patients, but not often. In a situation where we have a more critical case such as post-op patient from an open heart surgery, the surgical patient takes precedence over the ALRTI." (ASR-LASUTH)

"If the ALRTI patient is deemed to require ventilatory support, yes; but we don't not get them often" (SSR-LUTH)

"Yes I am aware patients with ALRTI may be in need of respiratory support, but the ICU here is pressed for space so I am not sure that it will admit such patients. Ideally every patient who needs admission should go to the ICU , however they will be admitted based on priority because of space constraints" (SSR-LASUTH)

As most of the patients requiring critical care were post-surgery, they usually came in from the theatre, while the others were admitted from the emergency department or the hospital ward.

"Patients come in either from the ward, theatre or the ER, for patients from the ER these are patients with severe head injury people who cannot maintain their airway, people with respiratory failure and burns or patients who had a stroke and are unconscious. From the ward, patients who have deteriorated are taken into the ICU." (SJR-LUTH)

"Most are post-op, but we still get people from outside the hospital for ICU care; they could be post -op or medically ill patients" (ASR-LASUTH).

8.4 Discussion

This section will discuss the themes in relation to critical care practices in other settings specifically more developed countries, as such comparisons have been known to provide meaningful insight into the various factors which affect critical care delivery and also drive future research on ways to improve on current practices(209).

8.4.1: Definition and structure of critical care

A recent systematic review on ICU capacity in low-income countries by Murthy et al. (2015), reported disparities in the definitions of critical care within this setting; this was following earlier work by the same authors in which they that noted that there was no universal generic definition of critical care(209, 210).

Wunsch et al. (2008) studied variations in critical care services across eight countries in Western Europe and America and reported that there was no established definition of intensive care or ICU beds in these countries. With the exception of Wales, critical care in the UK referred to care of patients at risk of multi-organ failure and requiring organ support; definitions in France and Belgium, also focused on illness severity and the need for organ support; Americans focused on intensity of monitoring, making reference to the nurse: patient ratio and equipment involved ; the Canadian definition was more focused on monitoring patients in specialised units according to the type of illness either surgical or medical; Wales, Netherland and Germany mentioned treatment with intensive care beds while Spain did not have any clear cut definition(211). Overall it appeared as though country specific variations in critical care delivery and practice were reflected in their adopted definitions and this might result in certain

challenges for comparing critical care delivery across different settings. Participants in this present study based their definitions of critical care on the severity of illness and the required level of monitoring as well as the need for organ support; as such their definitions were comparable with those used in other settings as reported in the study by Wunsch et al. (211). Most of the Nigerian studies did not provide explicit definitions of critical care; instead these studies included definitions and descriptions of intensive care units which did not allow for comparison with the definitions provided by the participants in this present study (212-215). A possible explanation for this is that critical care is still in its primitive stages in developing countries and mostly associated with sophisticated technology found in ICUs which ultimately reflects on the way their critical care system is structured and defined (216).

Similar to descriptions found in studies from Nigeria, the participants in this study described a critical care system where only tertiary care centres had ICUs (213, 216-219). Compared to high income countries like the United Kingdom where different levels of critical care are provided according to the patient's need and clinical stability (see table 8.1) (220), there were no clearly defined levels of critical care or intermediate care units between ICUs and the wards in the participating tertiary care centres; a typical scenario being the modification of surgical recovery rooms into ad hoc HDUs for managing patients who had been stepped down from ICU care. A likely reason for this might be the fact that being a developing country, Nigeria lacks the funds and manpower required for a more advanced critical care system (213, 221).

Interestingly, the system described by this study's participants where critically ill patients from less specialised hospitals were referred to tertiary care centres is not unlike an emerging trend in critical care seen in more developed settings

known as 'regionalisation'. Regionalisation is system which reorganises hospitals into tiers according to the level of critical care they are capable of providing, such that the most severely ill patients are referred to and managed in the most specialised hospitals (222-224). Low income countries with inadequate health budget allocations and manpower stand to benefit from regionalisation of critical care as concentrating expenditure in specific centres will result in cost-savings and more efficient use of manpower (222, 225, 226). Regionalisation also has the potential to result in improved outcomes as patients will be managed in specialised centres with highly skilled staff, as there is evidence in the critical care literature that patients managed in high volume ICUs had better outcomes than those managed in low volume ICUs (226-228). Besides its advantages, regionalisation has some pitfalls with serious implications for developing countries, one of which is the possibility that the hospitals receiving all the referred patients might exceed their capacity to manage the entire patient load(222). Secondly, the success of regionalisation requires appropriate government and stakeholder involvement; however the political instability and lack of basic infrastructure seen in most developing countries may make it harder to achieve the full benefits of this process. A typical example is the provision of good roads and affordable transportation which must be available for transporting patients to specialised centres with ICUs; in developing countries ambulance services are not up to par and geographical disparities exists in the location of ICUs with majority situated far away from rural areas where mortality from critical illness is high (222, 226, 229, 230).

8.4.2: Factors affecting discharge and admission

With only two ICUs serving a population of 9 million people in Lagos state where this study was conducted, it was not surprising that participants identified bed availability as one of the major factor affecting ICU admissions. According to the study participants, patients were usually managed on the wards or in a resuscitation room in the emergency department when there was a shortage of beds. Bed shortage is a major issue even in developed countries; according to a review by Murthy et al. (2012), bed shortage could either lead to patients being denied ICU admissions or their admission being delayed(209). A study in the Lancet by Metcalfe and colleagues (1997), found that over half of the patients refused admissions into ICUs in the UK were due to bed shortages(231), while another study by Goldhill et al. (2004) conducted in the UK found that patients who had to spend longer periods on the wards awaiting ICU admission had higher mortality rates than those whose admissions were not delayed(232).

The other factor determining ICU admission mentioned by the participants was the ability of the patient to pay for their treatment, as ICU care was generally more expensive compared to the cost of care on acute wards. In Nigeria and other developing countries where patients cover the cost of their own care, patients who are unable to afford the cost of ICU admission lose their bed space to patients who can (221, 233, 234). In some countries like Spain and the United Kingdom where the National Health Service covers the cost of healthcare for citizens, inability to pay is never a reason to deny a patient ICU care(235); however, in countries such as the United States where citizens are required to have their own health insurance, people without any health insurance are not likely to be admitted into a hospital and much less an ICU(235, 236). It has also been reported that people without health insurance tend to have poorer prognosis as they typically present

at later (more severe) stages, and receive less services and treatment while on admission as they tend to struggle with the cost of their treatment(237, 238).

Concurrently these two factors also determine discharge from ICU in the participating hospitals. According to study participants patients who are no longer able to afford ICU treatment might be discharged on request from family members or the managing clinician to reduce costs, while some patients might have their care stepped down to accommodate new ICU admissions because of bed shortages. It is common practice across various health settings to step down care of patients who show improvements so that new critical cases can be admitted, however in terms of post ICU care, this study confirms reports from other Nigerian studies that tertiary care centres do not have HDUs and patients were either discharged straight to the wards or remained in the ICU for longer than necessary (219, 233). Not all patients are suitable for ward care which means ICU readmission rates could increase if patients deteriorate due to inadequate monitoring and care on the wards. A study conducted in a tertiary care hospital in Nigeria stated that lack of HDUs and fear of readmission result in prolonged ICU stay in this setting(233).

Lastly, participants were not aware of the existence of any hospital specific guidelines for making admission or discharge decisions, rationing was purely based on the clinicians' judgements. This appeared to be common practice in both developed and developing countries as studies from have shown that clinician's tend to make inappropriate decisions with regards to ICU admissions (214, 233, 239, 240).

8.4.3: ICU organisation and patient profile

The type of patients admitted and the way ICUs are organised tends to vary across different settings; in less developed countries, medical and surgical cases are managed together in a general ICU, while developed countries have now moved towards a more decentralised system with dedicated intensive care units for managing specific categories of patients i.e. coronary ICUs for cardiac conditions and neurosurgical ICUs for neurological cases. The intensive care units described in this study catered to both medical and surgical cases; however the most common reasons for admission were post-surgical treatment and neurological trauma (head injury). Several studies from Nigeria have also reported a similar pattern in ICU admission diagnosis; one study conducted in Jos University teaching hospital stated that over 50% of its ICU admissions were post-surgical and trauma cases(233);another reported that over 75% of ICU patients in Calabar teaching hospital in the south eastern part of Nigeria were post-surgical patients(214); Onyekwulu et al.(2015) also reported that almost half (49.3%) of the ICU admissions in large teaching hospital in the eastern part of Nigeria were post-surgical cases(213). According to Ugwu et al. (2000) a significant proportion of surgeries in Nigeria are complicated by cardiac arrest and this might be a reason for the high post-op ICU admission rates (241). In addition to surgical complications, the lack of designated high dependency units in most of these tertiary centres could also contribute to the high ICU admission rates for post-op patients(213). The high prevalence of neurological trauma cases in Nigerian ICU can be attributed to the poor conditions of the motorways as most of the cases are due to road traffic accidents (213, 229, 233). Despite the fact that ALRTIs are a major contributor to morbidity and mortality figures in Nigeria, the interviewees' responses suggested that surgical cases were given priority over

patients with ALRTI who required ventilatory support. These findings are an indication that there is a need for a shift from the generalised ICU organisation in these hospitals towards specialised ICUs, as this will free up bed spaces for other cases and also provide more specialised management for the surgical and trauma patients.

Participants mentioned that admission and discharge decisions were a shared responsibility between the consultant anaesthetist in charge of the ICU and the consultant of the admitting unit, and only one of the tertiary care study centres had an intensivist. Over the years intensive care medicine has become more specialised and in most high income countries, specially trained critical care physicians known as intensivists now manage the ICUs. There are several benefits associated with having a full time intensivist in charge of an ICU such as shorter ICU length of stay, lower ICU mortality rates and more efficient bed utilisation (242-244). However, in Nigeria and most developing countries where ICUs were initially set up to provide ventilatory support for post-surgical cases, anaesthesiologists and the admitting consultants were responsible for ICU patients and despite current changes in critical care practice anaesthesiologists have maintained control of the intensive care units(245).

8.5 Generalisability

Generalisability which is defined as 'the extent, to which a study's findings can be applied in different settings', is considered to be quite problematic by some qualitative researchers as they believe that it is more suited to quantitative research(246, 247). In the case of this present study, respondents were recruited from hospitals in only one state in Nigeria, and questions might arise about how true these findings are in relation to hospitals in other parts of Nigeria. There are

certain points to consider in relation to the generalisability of this study's findings, the first being the distinction in the meaning of 'study sample' when used in quantitative studies vs. a qualitative studies. In quantitative studies a study sample is expected to be representative of a general population, while in a qualitative study the participants are chosen to be key informants who have a potentially important perspective on the issue being studied. Therefore, the study population in quantitative research is chosen based on size, while in qualitative research the population is chosen based on relevance and how much information they can provide(248). The participants in this study were considered to be key informants being medical staff who are involved in the care of critically ill patients in Nigerian hospitals, and as such possess specialised knowledge and capable of providing a full spectrum of views on critical care practice.

It has also been suggested that a possible approach to judging the generalisability of qualitative study findings to another setting is by taking into consideration any similarities in terms of location, participants and social contexts (249). Therefore it is important to note that all public hospitals irrespective of state/ location in Nigeria are ranked according to the three tiers of government such that all hospitals at the same level of the healthcare system are similar in terms of structure and organisation. Based on this, it is conceivable that the findings on structure and delivery of critical care from this study are generalisable to government owned hospitals in other parts of Nigeria. Findings relating to the participants' understanding or definition of critical care are however not generalisable as these were personal views based on acquired knowledge or training.

Though it is important to note that broadly speaking, all the definitions provided by the participants expressed a common understanding of critical care and need

for critical care, this is the one aspect of this study's findings which may not be generalisable as these definitions were based on personal knowledge or training. This is an inescapable dilemma as studies have also reported variations in the definition of critical care across settings.

8.6 Strengths

One of the key strengths of this study is its philosophical underpinning; this study was conducted within a pragmatic framework which allowed the researcher to achieve the study aims without the paradigmatic restrictions typical of qualitative research. Despite the flexibility provided by adopting a pragmatic approach, this study showed theoretical rigor by providing justification for the research question, obtaining ethical clearance for data collection and also by choosing appropriate methods for conducting the study. Secondly, procedural rigor was demonstrated by using a recognised framework to perform the thematic analysis and as such, transcripts were analysed in a consistent and transparent manner (109). Interpretative rigor was also achieved through independent coding by the researcher and supervisor as this ensured that researcher bias was minimised.

By using semi structured interviews both the researcher and the respondents were allowed some degree of flexibility which encouraged an open in-depth discussion of certain topics which might not have been possible with structured interviews or other methods of data collection. Furthermore data saturation was reached for each of the study's objectives, meaning that the responses provided by the participants were adequate for answering the research question and further interviewing of new participants was unlikely to have provided additional findings. This assumption is based on the fact that no new information was obtained with

subsequent interviews as the thematic analysis processes was exhaustive and no further codes were generated.

8.7 Limitations

Although the study participants were recruited to represent the major clinical specialities, none of the doctors were from the department of obstetrics and gynaecology and this study's findings did not include gynaecological patients as typical ICU patients. Although other studies from Nigeria have reported that gynaecological patients were commonly found in the ICU following surgery, the contribution of such patients to the ICU patient population in these studies was minimal.

8.8 Reflexivity

The initial idea to conduct a qualitative study came about as a result of a tripartite meeting between the researcher(RA) and her PhD supervisors (PM and JVT). The protocol for the external validation study was being developed, when it became apparent that dissimilarities in critical care practices between Nigeria and the UK might pose certain issues in terms of assessing the CATs ability to predict outcome variables associated with critical illness. The researcher (RA) had certain professional connections to the study centres, being a former employee and clinician in one of the tertiary care hospitals and it is likely that this might have had an impact on the respondents who agreed to participate as they might have been more comfortable with discussing issues and working practices with the researcher than they might have been with an unfamiliar interviewer.

The researcher (RA) also considered the possibility that her professional background might bias data collection and analysis, and in a bid to minimise researcher influence and maintain a "degree of distance of self", the interview questions were initially drafted by the study supervisor. Furthermore, coding of the interview scripts was independently done by both the supervisor (PM) and researcher (RA) to ensure transparency.

Also, as the researcher (RA) did not begin to conduct interviews until several months after data collection for the quantitative aspect of this thesis had begun, it is possible that the researcher's observations of the critical care system during this time may have changed and influenced some of the interview questions.

The researcher(RA) was concerned about not having undergone any formal training in qualitative research methods as this might have an effect on the interview process; as such a pilot interview was initially conducted and shared with the thesis supervisor (PM) for review before going ahead with the rest of the interviews. The researcher (RA) was able to enrol in a qualitative research methods course which informed the post-interview stages of this study (transcribing, coding and analysing) and this also helped the researcher(RA) become more reflexive in the interpretation of the study findings.

8.9 Summary

The purpose of this study was to gain insight into critical care practice in Nigeria using qualitative methods. Fourteen health care workers in four secondary and tertiary care hospitals were interviewed using a semi-structured approach; thematic analysis of the interview transcripts was conducted and 3 key themes relating to the study objectives were identified.

Participants generally had a good understanding of critical care, as their definitions were comparable with those used in more developed settings, encompassing the salient aspects of critical care such as severity of illness, level of monitoring and need for organ support.

Inequalities in terms of financial resources and economic development were the main factor responsible for disparities in critical care practice across settings; intensive care units were only found in tertiary hospitals in Nigeria and these hospitals did not have high dependency units. Availability of bed space and financial resources were the two major determinants of admission and discharge in the tertiary centres offering ICU services; as a result bed rationing was the norm for deciding on admission and discharge of patients. There were no explicit protocols or guidelines being followed in these hospitals for admitting and discharging critically ill patients, as most decisions were made on first come first serve basis and clinicians' assessments of patients' suitability for ICU care.

The hospitals in this study operated an open unit ICU managing all types and ages of patients together; post-surgical and trauma patients were prioritised over most medical cases, and even within medical cases neurological patients were more commonly found in the ICU compared to patients with severe respiratory illnesses. As there were no high dependency units, patients who had made good recovery

but still required some level of monitoring were either kept back in the ICU or discharged to the general wards.

Finally, intensive care medicine training was lacking in the tertiary hospitals; only one of the two centres had an intensivist and critically ill patients were being co-managed by the admitting physicians and anaesthetists.

The implications of these study findings with regard to the contextualisation and interpretation of the CATs validation study will be discussed in the next chapter.

9: Thesis conclusions and recommendations.

9.1 Recap of the thesis objectives and key findings

Data prospectively collected between December 2013 and July 2014 on the clinical assessment and management of 1016 children and adults presenting with symptoms suggestive of an acute lower respiratory tract infection to the emergency departments of four large hospitals (two tertiary and two secondary care facilities) in Lagos Nigeria, were used to assess the external validity of the Community Assessment Tools (CATs) for the severity assessment of patients with ALRTIs. In addition, an exploratory study involving doctors and nurses in the participating hospitals was conducted using qualitative methods of inquiry to gain an understanding of critical care practice in Nigeria in comparison to more developed healthcare settings.

This chapter recaps the key findings of the five objectives of this thesis which were treated as standalone chapters, as well as implications for clinical practice and potential areas for future work.

Objective 1: **assessing the external validity of the individual CATs criterion for predicting the need for hospital based care and the risk of in-hospital mortality in adults and children with ALRTI**

Unadjusted and adjusted logistic regression analyses were used to assess the relationship between each CATs criterion and five predefined clinical outcomes; need for hospital admission, intravenous therapy (antibiotics or fluids), supplemental oxygen therapy and the risk of in-hospital mortality. Each CATs criterion was significantly associated with outcomes which were indicative of

the need for hospital care or the risk of death in the unadjusted analysis, and none of these criteria were redundant in the adjusted analyses. A comparison of two multivariable modelling approaches for the adjusted analyses did not reveal much difference in the performance of the CATs criteria. Study outcomes were independently predicted by similar CATs criteria irrespective of whether all seven criteria were included based on pre-determined clinical significance or if only criteria found to be statistically significant at $p < 0.05$ were included using a forward stepwise regression approach.

After adjusting for the effect of other CATs criteria, only criterion G (causing other clinical concern) predicted all five study outcomes in children < 16 years. Except for the outcome 'need for intravenous fluids'; criterion G was seen to perform similarly in the adult analyses, predicting the other four study outcomes irrespective of modelling approach used. Criteria C (oxygen saturation levels of $\leq$92%), G (causing other clinical concern), and D (respiratory exhaustion) were seen to be the only independent predictors of in-patient mortality in both adults and children.

This study also examined the predictive accuracy of each CATs criterion using AUROC, specificity, sensitivity, negative and positive predictive values. Criteria A (severe respiratory distress) and B (increased respiratory rate) had better sensitivity for predicting all study outcomes in both adults and children, while the other four criteria C to G were seen to be more specific predictors. In terms of overall predictive accuracy, AUROC values of the criteria were seen to fall within the range of 0.53-0.86 in adults and 0.49-0.77 in children, indicating low to moderate levels of accuracy at predicting the study outcomes. On the whole, CATs criteria A (severe respiratory distress) and C (oxygen saturation levels $\leq$92%) had the highest AUROC values in adults and children for

predicting the need for hospital admission, supplemental oxygen and intravenous therapy (antibiotics and fluids); while criterion C (oxygen saturation levels of $\leq$92%) and the subjective criterion G (causing other clinical concern) had the highest accuracy for predicting in-hospital mortality.

Objective 2: assessing the predictive performance of CATs as a combined un-weighted or weighted scoring index for severity assessment of ALRTI in adults and children

This thesis assessed the performance of the CATs criteria as combined un-weighted and weighted scoring systems on the premise that scoring systems or indexes which are able to assign patients into specific treatment or risk groups will aid better decision making in practice than single standalone parameters. The performance of the newly created CATs scoring systems reaffirmed this theory, as in comparison to the standalone criteria these scoring systems were shown to have better predictive accuracy in terms of AUROC values for identifying high risk adults and children who would benefit from hospital based care. Of the two scoring systems, the weighted CATs score had a higher range of AUROC values; however, this may not necessarily translate into better performance when used in practice as the differences in AUROC values between the un-weighted and weighted scores were clearly marginal. It was also difficult to identify a suitable weighted score threshold for predicting any of the outcomes, as the use of statistically derived weights did not allow for a sensible consensus on appropriate scoring scales. On the other hand, an un-weighted CATs score of $\geq$3 was identified as a suitable threshold for making management decisions regarding the need for hospital care and interventions

in both adults and children. It is worth mentioning that one thing which was common to both CATs scoring systems, was that the calibration of their predictions as assessed by the Hosmer-Lemeshow goodness-of-fit test were better in adults than in children, although this was likely as a result of the effect of sample size on the performance of this goodness-of-fit test.

Objective 3: assessing the predictive performance of the combined un-weighted CATs score in comparison to the CRB65 score in adults with ALRTI

As part of the external validation, this thesis also comparatively validated the performances of CATs and CRB65: the simplified variant of CURB65 which is a well-established and widely used pneumonia severity scoring system. These comparative analyses were performed using data from adult participants, as one of the variables included in the CRB65 score – 'blood pressure' is not routinely recorded in children presenting to Nigerian hospitals without suspicion of a cardiac illness.

The unweighted CATs was seen to consistently predict the need for hospital based care and risk of in-hospital mortality with higher accuracy and calibration than the CRB65 in adults with ALRTI. A further comparison of these scores in a sensitivity analysis involving adults with a clinical diagnosis of pneumonia revealed that although the un-weighted CATs score was better at identifying the need for hospital based care and life saving interventions such as intravenous antibiotic and supplemental oxygen therapy, the difference in AUROC values between CATs and CRB65 was marginal and statistically

insignificant for predicting the risk of in-hospital mortality and on the whole the CRB65 score was better calibrated for this purpose.

Objective 4: assessing the value of the subjective criterion G (causing clinical concern) in the performance of CATs when used as a combined un-weighted score

The fourth study in this thesis explored the value of subjective judgement in clinical prediction by comparing the performance of a six-criteria un-weighted CATs score (without the subjective criterion G) against a seven criteria un-weighted CATs score inclusive of criterion G. This comparison revealed that the seven-point un-weighted CATs scoring system was superior to the six-point score in terms of identifying the risk of in-hospital mortality in both adults and children. In addition to improving mortality prediction, the inclusion of the subjective criterion G was also seen to significantly improve the performance of the seven-point un-weighted CATs at predicting the need for hospital admission and intravenous antibiotics in adults, and intravenous fluids in children. Although the reported differences in AUROC values of the two versions of the un-weighted CATs score were marginal, these findings provide additional evidence to support the usefulness of intuitive clinical judgement in practice.
Most of the reasons for meeting criterion G were based on abnormal findings recorded by the doctors during the clinical assessment of patients on initial presentation to the study hospitals and no differences were reported in the validity of such intuitions with regards to the level of seniority or experience of the doctors.

Objective 5: qualitative exploration of critical care practice and utilization in Nigerian health facilities

The qualitative aspect of this thesis was aimed at providing contextual insights into critical care practices in the participating hospitals, which was premised on the fact that the use of a clinical prediction tool to inform decisions regarding ICU admissions of critically ill patients will most likely be influenced by the heterogeneity of critical care practice across settings.

Three key themes were identified from the thematic analysis of the responses of the fourteen doctors and nurses who were interviewed at the four study hospitals.

The first theme which focused on the definition and structure of critical care revealed that the participants' definitions were mostly comparable and consistent with definitions used in other published writings from developed countries; which meant that, broadly speaking there was a common understanding of critical care and the need for critical care across these settings. Irrespective of this, the structure of critical care in the Nigerian hospitals was quite distinguishable from the structure in more developed high income countries. The hospitals in this thesis lacked clearly defined levels of care for managing critically ill patients as only tertiary care hospitals provided specialised critical care services and there were no designated Step Down or High Dependency units which in turn resulted in an overburdened ICU with potential negative impacts on patient care. The second theme identified cost of critical care services and ICU-bed space availability as the major factors determining admission and discharge; healthcare in Nigeria is financed by out of pocket payments and as a result most patients are unable to afford the high

cost of ICU care, also the limited number of ICU beds and lack of High Dependency Units results in very strict and sometimes inappropriate rationing of critical care in these hospitals.

The third and final theme which described ICU organisation and patient profile revealed that the ICUs in Nigerian hospitals were run by anaesthesiologists and operated as open or mixed units catering to both medical and surgical cases. This is unlike ICUs in developed countries where the intensive care units are more decentralised, and operate as specialised units managed by intensivists. Furthermore, ICUs in Nigerian hospitals tend to admit more surgical, particularly neurosurgical and trauma cases. Although several explanations have been offered for why surgical cases were given more priority over medical cases in these hospitals (see discussion in chapter 8), one key reason was the lack of designated High dependency or Step Down units as most of these surgical cases were post-surgery patients who required monitoring in a more specialised unit than a hospital ward.

.

9.2 Discussion of the thesis findings

Based on the findings from this thesis it can be said that the CATs has shown geographical generalisability, seeing as it performed reasonably well and replicated most of the findings reported in the other CATs studies conducted in the UK despite case-mix and geographical differences. These findings are particularly relevant in low and middle income countries, as majority of the existing clinical prediction tools for ALRTI are developed in high income countries and are rarely validated for use in resource poor geographical settings where the reported burden of such diseases is greater. The evidence provided in this thesis will potentially improve the management of acute lower respiratory tract infections in low and middle income countries; as clinicians working in such settings will be more confident in using a prediction tool which has been validated in the context of medical practice in a developing country. A further advantage which makes the CATs even more applicable for use in resource-constrained settings, is the simplicity of its criteria which does not require extensive laboratory investigations and can be used by clinical staff at different levels of expertise. CATs only requires the use of a finger-tip pulse oximeter which is a low-cost, non-invasive point-of-care equipment used in recording oxygen saturation levels, and based on the reported performance of the CATs criterion C (low oxygen saturation levels $\leq$92%) in this thesis, it can be said that the CATs provides a rapid and relatively simple means of identifying illness severity in patients of all ages in resource-poor, 'under-pressure' and 'high through-put' LMIC settings.

At present, more than 100 developing countries have adopted the use of the WHO *Integrated Management of Childhood Illnesses* (IMCI) strategy in the care

of sick children under the age of five presenting to first-level outpatient healthcare facilities(250). The IMCI guideline covers the most common conditions responsible for morbidity and mortality in this age group such as ALRTI, malaria, measles, diarrhoea and malnutrition. This strategy was recommended to improve the accurate assessment, identification, and treatment of sick children, as well as ensure the prompt referral of the severely ill for further care in hospital. However, this strategy does not cover all childhood conditions, and local adaptations of the IMCI have been proposed to include the major causes of childhood morbidity and mortality specific to each country. As the CATs criteria are not specific to ALRTI alone, but are able to identify the most likely presentations of severe illness and deterioration in children (irrespective of the condition), the genericness of the CATs provides a unique solution for the local adaptations of the IMCI guidelines. Also, CATs only requires the use of few clinical signs which will not require the need for extensive training of health workers without a formal clinical background. If adapted into the WHO IMCI, CATs could ensure consistency in prioritising the most critically ill patients for ED referral, and also standardise the recognition of severe illnesses across countries which will make comparison and multi-country evaluations of the effectiveness of care possible. While the IMCI applies to outpatient management of childhood illnesses, the WHO *Emergency Triage Assessment and Treatment* (ETAT) guidelines were developed for use in developing countries to ensure the timely triage of critically ill children and the initiation of urgent assessment and emergency interventions in order to reduce mortality in children presenting to ED(251). Like the IMCI, the ETAT guidelines also focus on the emergency management of children with severe ALRTI/pnemonia, malaria, diarrhoea, malnutrition and dehydration. ETAT

prioritises patients into three groups: children having emergency signs and in need of urgent immediate treatment; those having priority signs and need full assessment and treatment; and thirdly non-urgent patients who are able to wait for treatment and assessment. Using simple clinical signs, both ETAT and CATs are both able to provide guidance on the need for administering intravenous fluids and airway management with supplemental oxygen to severely ill children, however the CATs also has the potential for use as a mortality predictor, while the WHO ETAT does not. In times of extreme surge, triage decisions will prioritise patients who are more likely to survive from treatment; while the use of ETAT may be limited in such scenarios, the CATs and its threshold scores will allow for such prioritisation of care.

An important finding from this study which corroborates the reports of the previous CATs validation studies is that the CATs criteria perform better at discriminating between patients who will or will not benefit from hospital based care when combined into weighted and un-weighted scoring indices than when used as individual standalone triggers. Using the CATs as a combined scoring system will provide a means of defining suitable score thresholds or cut-off points which would reflect the severity of a patient's condition and thus guide clinicians in making more objective therapeutic decisions. Despite the weighted combined CATs score being superior to its un-weighted counterpart in terms of a higher range of AUROC values, the statistical differences between the two scoring systems were marginal (see results section in chapter 5). It is also quite likely that the complexities of calculating the weighted scores would limit its clinical utility in practice, as clinicians are more likely to favour a simpler and more user-friendly tool over one which requires complicated mathematical imputations. The un-weighted CATs score also had better specificity and

positive predictive values which makes it a more useful tool for triaging, as it will limit excessive and liberal hospital admission of patients who would otherwise cope well outside a hospital without compromising the overall safety of patients as its good positive predictive values will ensure that the actual patients who would benefit more from admissions are correctly identified. This is not to say that the there is no role for the use of weighted CAT scores; these findings only indicate that the weighted index may require refinement possibly through expert clinical consensus scoring using the Delphi technique or by creating a nomogram to help clinicians calculate scores in practice.

Calibration which is an equally important aspect of describing the validity of severity scoring systems appeared to better when the combined CATs scores were used in adults than in children; particularly for predicting outcomes 'need for supplemental oxygen, intravenous antibiotics and fluids'. This does not necessarily mean that the predictions of the combined CATs scores in children are not accurate, giving that a major drawback of the Hosmer-Lemeshow goodness-of-fit test which was used to assess calibration in this study is its reliance on sample size. (141). Significant p-values suggesting miscalibration are more likely to be discovered in larger samples and as this study's data was heavily populated by children, this possibly explains why the calibration of the combined CATs score was poorer in this patient group. It has been suggested by Vergouwee et al. and Peek et al. (252, 253) that the use of Hosmer-Lemeshow test should be avoided in external validation studies, instead they have recommended that other methods such as calibration slopes and intercept proposed in a paper by Cox (254), might be more useful in this regard.

Giving the correlation between influenza and pneumonia such that more cases of pneumonia are reported during influenza seasons or pandemics, a sensitivity

analysis was conducted to compare the CATs with a well validated and established pneumonia mortality prediction tool in adult patients with a diagnosis of pneumonia. Although the discriminatory ability of the CATs was comparable with that of the CRB65, the CRB65 score appeared to be a better calibrated tool for predicting in-patient mortality in adults with pneumonia; which is not surprising as this was the purpose for which it was developed. However as pneumonia was the most prevalent diagnosis and major cause of the mortality amongst adults in the study population, it will be worth conducting a head to head randomised trial comparing the CATs and CRB65 scores in a less homogenous sample of adult patients as this may eliminate any bias which a high number of pneumonia cases might have introduced.

It is also worth mentioning that this thesis adopted a pragmatic approach aimed at reflecting existing conditions in Nigerian hospitals and as such pneumonia diagnosis was not based on radiological evidence. Although the concept of radiologically confirming a diagnosis of pneumonia is not foreign to low income settings, it has been reported that pneumonia in Nigeria is usually diagnosed based on clinical grounds particularly as a result of infrastructural and economic constraints. With chest radiographs being the accepted gold standard for confirming a diagnosis of pneumonia, it is possible that the findings from this thesis may not be comparable with other studies where chest X-rays have been used to confirm the presence of pneumonia. On the other hand, it is unlikely that all clinical decisions made during a pandemic event will be supported by extensive clinical investigations as demand may outstrip clinical capacity both in terms of staff and resources. Chest x-rays may not be readily available to confirm a diagnosis of pneumonia and clinicians will thus have to rely on clinical signs and examinations alone. By assessing the CATs in

a low-resource setting, this thesis therefore provides evidence to support the suitability of the CATs for use in a somewhat similar under-pressure scenario such as a pandemic event.

As mentioned earlier, time and resource constraints may limit the extent to which patients will be investigated in the event of a pandemic or a similar high-pressure scenario, as such it is likely that there will be more scenarios where clinicians may have to rely on gut feelings or intuition to make decisions regarding admission. The findings from this thesis complement reports from other studies which are in support of the use of intuitive decision making among clinicians, which will be quite useful in difficult scenarios where a clinician may have strong suspicions regarding the condition of a patient in spite of seemingly normal physiological parameters. As earlier mentioned, these reported differences between the AUROC values of the un-weighted CATs score with only objective criteria and the un-weighted CATs score with the subjective criterion G (causing other clinical concern) were quite marginal, but there is reason to believe that this might be as a result of instances where clinical intuitions may have been used tacitly and therefore unrecorded. Although based on the findings of this thesis it can be hypothesised that intuition is independent of a clinician's level of expertise, exploring the relationship between level of expertise and subjective medical decision making was beyond the scope of this thesis objectives. Caution should be exercised in applying this finding in practice as it contradicts the opinions of other authors, and further research will be required to truly establish this assumption.

One limitation of this thesis was the low numbers of patients identified as needing critical care admission, which made it impossible to conduct the required statistical analyses for assessing the performance of the CATs score

in relation to predicting the need for mechanical ventilation or ICU admission. Despite findings from the qualitative study not indicating much differences between Nigerian clinicians and those in more developed settings in terms of their definition and understanding of the need for critical care, it was evident from the responses of the participants that surgical patients were prioritised over medically unwell patients in regard to ICU admissions in the study hospitals. Given that several other studies conducted in Nigeria have reported similar patterns in critical care admissions, it may not be appropriate for validation studies to consider outcomes such as 'admission to critical care' or 'need for mechanical validation' as suitable outcomes in LMIC without the independent assessment of these outcomes by researchers using standardised criteria in order to avoid misclassification.

Regardless of this limitation, it is reasonable to conclude that CATs presents a simple rule by which clinicians in low-resource clinical settings can identify children and adults with ALRTI who are more likely to benefit from hospital based acute care interventions. There is no explicit statement in the literature on predictive research on how many external validation studies are required for a tool to be considered acceptable for use in practice, it is only logical to expect that the more external validation studies conducted in a variety of clinical settings and populations the higher the possibility of clinicians gaining confidence in such a tool.

9.2.1 Implications for clinical practice and future research

This section outlines the potential clinical implications of the thesis findings as well as recommendations which could direct further research on the CATs.

9.2.1.1 Implications for practice

- CATs is well suited to low-resource settings based on its simplicity and the use of criteria which do not require extensive diagnostic investigations.
- The CATs criteria have more clinical validity when used as a combined scoring system than as stand-alone predictors of severity; however the reliance on statistically derived weights may limit the use of the weighted score in practice.
- Although both the Un-weighted CATs and CRB65 scores demonstrated good predictive performance in adults with pneumonia in Nigeria; the CATs is better for predicting hospital interventions, while the CRB65 is a better mortality predictor.
- Gut-feelings and clinical intuition are able to improve clinical decision making when used alongside objective clinical signs of illness severity. As such the inclusion of criterion G in the CATs is particularly useful in the context of resource poor settings, where doctors do not have access to laboratory or other diagnostic investigation results before making management decisions.
- ICU admissions in Nigeria decisions are not solely dependent on clinical signs, as other factors such as bed availability and cost of admission are also considered. It is therefore recommended that careful consideration of such factors is required before using the CATs in new settings.
- Unweighted CATs score thresholds could be potentially adapted to suit the different tiers of the Nigerian health system, as this will ensure a more efficient use of resources by preventing patients bypassing the referral system and overburdening the tertiary hospitals.

- In relation to the above point, the CATs could serve as a link between the WHO IMCI and ETAT, by ensuring consistency in the identification and referral of patients by health workers in first-level facilities in Nigeria.

9.2.1.2 Recommendations for future research

- As this was the first geographical validation of the CATs criteria, it will be worth conducting further validation studies in a variety of settings to provide further evidence of its generalisability.
- Further research is needed in terms of a head to head randomised trial of CATS vs CRB-65 as this may eliminate any bias which may have been present in this study's data.
- It will be interesting to know if the level of experience of a doctor affects the validity of their clinical intuitions; this is quite important as medical staff illness or absence during a surge may result in less experienced staff replacing the regular doctors.
- In terms of using the CATs as a weighted scoring index, further research is needed to explore different approaches to assigning weights to the CATs criteria in order to ensure an unbiased comparison of the scores.
- Further research involving a comparison of the unweighted CATs score and WHO ETAT would be useful in validating the CATs for use in Nigeria
- Further validation studies using multiple external data to test the potential for using different thresholds of the unweighted CATs score in outpatient and hospital settings is required before piloting the CATs in practice.

References

1. Laupacis A, Sekar N, Stiell l G. Clinical prediction rules: A review and suggested modifications of methodological standards. JAMA. 1997;277(6):488-94.
2. Burke W. Genetic Tests:Clinical Validity and Clinical Utility. Current protocols in human genetics / editorial board, Jonathan L Haines [et al]. 2014;81:9.15.1-9..8.
3. Ewout WS. Clinical Prediction Models: a practical approach to development, validation and updating. Gail M KK, Samet J, Tsiatis A, Wong W, editor. New York: Springer; 2008.
4. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. European Heart Journal. 2014;35(29):1925-31.
5. Moons KGM, Kengne AP, Woodward M, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. Heart. 2012;98(9):683-90.
6. Moons KGM, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? BMJ. 2009;338.
7. Royston P, Moons KGM, Altman DG, Vergouwe Y. Prognosis and prognostic research: Developing a prognostic model. BMJ. 2009;338.
8. Altman DG, Vergouwe Y, Royston P, Moons KGM. Prognosis and prognostic research: validating a prognostic model. BMJ. 2009;338.
9. Moons KGM, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. BMJ. 2009;338.
10. Hutchinson R. Prognosis. Lancet. 1934;226:697-8.
11. Adams F. The Genuine Works of Hippocrates. Birmingham, AL: The Classics of Medicine Library; 1985.

12. Christakis NA. The ellipsis of prognosis in modern medical thought. Soc Sci Med. 1997;44(3):301-15.
13. Thomas J. Foretelling the Future Of Prognostication: A Historically

Inspired Domain-Based Approach For The Elderly. Yale Medicine Thesis Digital Library: Yale University; 2012.
14. Schnur S. MORTALITY RATES IN ACUTE MYOCARDIAL INFARCTION. II. A PROPOSED METHOD FOR MEASURING QUANTITATIVELY SEVERITY OF ILLNESS ON ADMISSION TO THE HOSPITAL*. 1953;39(5):1018-25.
15. Hendriksen JMT, Geersing GJ, Moons KGM, de Groot JAH. Diagnostic and prognostic prediction models. Journal of Thrombosis and Haemostasis. 2013;11:129-41.
16. Toll DB, Janssen KJM, Vergouwe Y, Moons KGM. Validation, updating and impact of clinical prediction rules: A review. 2008;61(11):1085-94.
17. Moons KGM, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: II. External validation, model updating, and impact assessment. Heart. 2012;98(9):691-8.
18. Steyerberg EW, Borsboom GJ, van Houwelingen HC, Eijkemans MJ, Habbema JD. Validation and updating of predictive logistic regression models: a study on sample size and shrinkage. Stat Med. 2004;23(16):2567-86.
19. Francis NA, Cals JW, Butler CC, Hood K, Verheij T, Little P, et al. Severity assessment for lower respiratory tract infections: potential use and validity of the CRB-65 in primary care. Primary Care Respiratory Journal. 2012;21(1):65-70.
20. Lim WS, Macfarlane JT. Importance of severity of illness assessment in management of lower respiratory infections. Curr Opin Infect Dis. 2004;17(2):121-5.

21. Lim WS, van der Eerden MM, Laing R, Boersma WG, Karalus N, Town GI, et al. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. Thorax. 2003;58(5):377-82.
22. Fine MJ, Auble TE, Yealy DM, Hanusa BH, Weissfeld LA, Singer DE, et al. A Prediction Rule to Identify Low-Risk Patients with Community-Acquired Pneumonia. New England Journal of Medicine. 1997;336(4):243-50.
23. Moura FE, Perdigao AC, Siqueira MM. Seasonality of influenza in the tropics: a distinct pattern in northeastern Brazil. Am J Trop Med Hyg. 2009;81(1):180-3.
24. Shek LP, Lee BW. Epidemiology and seasonality of respiratory tract virus infections in the tropics. Paediatr Respir Rev. 2003;4(2):105-11.
25. Health Do. Pandemic Flu: Managing Demand and Capacity in Healthcare Organisations (Surge). In: Health Do, editor. London: Her Majesty's Stationery Office; 2009.
26. Chen K-F, Hsieh Y-H, Gaydos CA, Valsamakis A, Rothman RE. Derivation of a clinical prediction rule to predict hospitalization for influenza in EDs. The American Journal of Emergency Medicine.31(3):529-34.
27. Liao Q, Ip DKM, Tsang TK, Cao B, Jiang H, Liu F, et al. A clinical prediction rule for diagnosing human infections with avian influenza A(H7N9) in a hospital emergency department setting. 2014;12(1):1-9.
28. Health. Do. Pandemic influenza: surge capacity and prioritisation in health services - provisional UK guidance. In: Health Do, editor.: (Department of Health, Richmond House, 79 Whitehall, London SW1A 2NJ, UK,; 2007.
29. Lim WS. Pandemic flu: clinical management of patients with an influenza-like illness during an influenza pandemic. Thorax. 2007;62(suppl 1):1-46.
30. Challen K, Bright J, Bentley A, Walter D. Physiological-social score (PMEWS) vs. CURB-65 to triage pandemic influenza: a comparative validation study using community-acquired pneumonia as a proxy. BMC Health Services Research. 2007;7(1):1-8.
31. Team DPIP. UK Influenza Pandemic Preparedness Strategy 2011. In: Health Do, editor. DH website, in electronic PDF format only.2011.
32. NHS DoH. Swine Flu clinical package for use when there are exceptional demands on healthcare services. COI for the Department of Health; 2009.
33. Ewig S, Ruiz M, Mensa J, Marcos MA, Martinez JA, Arancibia F, et al. Severe community-acquired pneumonia. Assessment of severity criteria. Am J Respir Crit Care Med. 1998;158(4):1102-8.
34. &amp DoH, NHS. Swine flu clinical package for use when there are exceptional demands on healthcare services. In: &amp DoH, NHS, editors.: COI for Department of Health Gateway ref 12368; 2009.
35. Semple MG, Myles PR, Nicholson KG, Lim WS, Read RC, Taylor BL, et al. An Evaluation of Community Assessment Tools (CATs) in Predicting Use of Clinical Interventions and Severe Outcomes during the A(H1N1)pdm09 Pandemic. PLoS ONE. 2013;8(9):e75384.
36. Myles PR, Nguyen-Van-Tam JS, Lim WS, Nicholson KG, Brett SJ, Enstone JE, et al. Comparison of CATs, CURB-65 and PMEWS as Triage Tools in Pandemic Influenza Admissions to UK Hospitals: Case Control Analysis Using Retrospective Data. PLoS ONE. 2012;7(4):e34428.
37. Nguyen-Van-Tam JS, Openshaw PJM, Hashim A, Gadd EM, Lim WS, Semple MG, et al. Risk factors for hospitalisation and poor outcome with pandemic A/H1N1 influenza: United Kingdom first wave (May–September 2009). Thorax. 2010;65(7):645-51.
38. Goodacre S, Challen K, Wilson R, Campbell M. Evaluation of triage methods used to select patients with suspected pandemic influenza for hospital admission: cohort study. Health Technol Assess. 2010;14(46):173-236.
39. Elimian, O K. Community Assessment Tools (CATs): A comparison of weighted and un-weighted scores for predicting use of critical care facilities and death in a UK pandemic influenza hospital cohort: University of Nottingham; 2013.

40. Venkatesan S, Myles PR, McCann G, Kousoulis AA, Hashmi M, Belatri R, et al. Development of processes allowing near real-time refinement and validation of triage tools during the early stage of an outbreak in readiness for surge: the FLU-CATs Study. Health Technology Assessment (Winchester, England). 2015;19(89):1-132.
41. Organisation. WH. Top 10 leading causes of death in the world 2000 and 2012: http://www.who.int/mediacentre/factsheets/fs310/en/; [
42. Akanbi MO, Ukoli, C.O., Erhabor, G.E., et al. The burden of respiratory disease in Nigeria. Afr J Resp Med. 2009 March;4(2):10-7.
43. Cooper RS, Osotimehin B, Kaufman JS, Forrester T. Disease burden in sub-Saharan Africa: what should we conclude in the absence of data? The Lancet. 1998;351(9097):208-10.
44. Lanata CF, Rudan I, Boschi-Pinto C, Tomaskovic L, Cherian T, Weber M, et al. Methodological and quality issues in epidemiological studies of acute lower respiratory infections in children in developing countries. International Journal of Epidemiology. 2004;33(6):1362-72.
45. Rudan I, Tomaskovic L, Boschi-Pinto C, Campbell H, Group WHOCHER. Global estimate of the incidence of clinical pneumonia among children under five years of age. Bulletin of the World Health Organization.82(12):895-903.
46. WHO. Global Health Observatory data/country statistics/Nigeria: http://who.int/gho/mortality_burden_disease/en/; 2012 [
47. WHO. Nigeria-summary statistics (2002-present) http://apps.who.int/gho/data/node.country.country-NGA2013 [
48. Liapikou A, Torres A. The clinical management of lower respiratory tract infections. Expert Review of Respiratory Medicine. 2016;10(4):441-52.
49. Örtqvist Å. Treatment of community-acquired lower respiratory tract infections in adults. European Respiratory Journal. 2002;20(36 suppl):40s.
50. Hitoshi O, Taro K, Akira S. Major Issues and Challenges of Influenza Pandemic Preparedness in Developing Countries. Emerging Infectious Disease journal. 2008;14(6):875.
51. Lim WS, Baudouin SV, George RC, Hill AT, Jamieson C, Le Jeune I, et al. BTS guidelines for the management of community acquired pneumonia in adults: update 2009. Thorax. 2009;64(Suppl 3):iii1-iii55.
52. Cashat-Cruz M, Morales-Aguirre JJ, Mendoza-Azpiri M. Respiratory tract infections in children in developing countries. Seminars in Pediatric Infectious Diseases. 2005;16(2):84-92.
53. Steyerberg EW MK, van der Windt DA, Hayden JA, Perel P, et al. Prognosis Research Strategy (PROGRESS) 3: Prognostic Model Research. PLoS Med. 2013;10(2):e1001381.
54. Moons KGM, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and ElaborationThe TRIPOD Statement: Explanation and Elaboration. Annals of Internal Medicine. 2015;162(1):W1-W73.
55. Bouwmeester W, Zuithoff NPA, Mallett S, Geerlings MI, Vergouwe Y, Steyerberg EW, et al. Reporting and Methods in Clinical Prediction Research: A Systematic Review. PLoS Medicine. 2012;9(5):e1001221.
56. Collins GS, de Groot JA, Dutton S, Omar O, Shanyinde M, Tajar A, et al. External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. BMC Medical Research Methodology. 2014;14:40-.
57. Riley RD, Sauerbrei W, Altman DG. Prognostic markers in cancer: the evolution of evidence from single studies to meta-analysis, and beyond. Br J Cancer. 0000;100(8):1219-29.
58. Laupacis A, Wells G, Richardson W, et al. Users' guides to the medical literature: V. how to use an article about prognosis. JAMA. 1994;272(3):234-7.
59. Hayden JA, van der Windt DA, Cartwright JL, Côté P, Bombardier C. Assessing Bias in Studies of Prognostic Factors. Annals of Internal Medicine. 2013;158(4):280-6.

60. Lisa M McShane DGA, Willi Sauerbrei, Sheila E Taube, Massimo Gion and Gary M Clark. REporting recommendations for tumor MARKer prognostic studies (REMARK). Nat Clin Prac Oncol. 2005;2(8):416-22.
61. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD StatementThe TRIPOD Statement. Annals of Internal Medicine. 2015;162(1):55-63.
62. Lijmer JG, Mol B, Heisterkamp S, et al. EMpirical evidence of design-related bias in studies of diagnostic tests. JAMA. 1999;282(11):1061-6.
63. Simon R, Altman DG. Statistical aspects of prognostic factor studies in oncology. British Journal of Cancer. 1994;69(6):979-85.
64. Altman DG. Systematic reviews of evaluations of prognostic variables. BMJ : British Medical Journal. 2001;323(7306):224-8.
65. World Population Review 2015 [Available from: http://worldpopulationreview.com/world-cities/lagos-population/.
66. Ogunbekun I, Ogunbekun A, Orobaton N. Private health care in Nigeria: walking the tightrope. Health Policy Plan. 1999;14(2):174-81.
67. Olakunde B. Public health care financing in Nigeria: Which way forward? Annals of Nigerian Medicine. 2012;6(1):4-10.
68. FMOH. Revised National Health Policy. In: Nigeria FMoH, editor. Abuja, Nigeria2004.
69. Bank W. Improving Primary Health Care Delivery in Nigeria. World Bank Working Papers: The World Bank; 2010. p. 90.
70. Abodunrin O, Akande T, Osagbemi G. Awareness and perception toward referral in health care: A study of adult residents in Ilorin, Nigeria. Annals of African Medicine. 2010;9(3):176-80.
71. Akande TM. REFERRAL SYSTEM IN NIGERIA: STUDY OF A TERTIARY HEALTH FACILITY. Annals of African Medicine. 2004;3(3):130-3.
72. Woodhead M, Blasi F, Ewig S, Garau J, Huchon G, Ieven M, et al. Guidelines for the management of adult lower respiratory tract infections - Full version. Clinical Microbiology and Infection. 2011;17:E1-E59.
73. Holmes WF, Macfarlane JT, Macfarlane RM, Hubbard R. Symptoms, signs, and prescribing for acute lower respiratory tract illness. The British Journal of General Practice. 2001;51(464):177-81.
74. Scott JAG, Wonodi C, Moïsi JC, Deloria-Knoll M, DeLuca AN, Karron RA, et al. The Definition of Pneumonia, the Assessment of Severity, and Clinical Standardization in the Pneumonia Etiology Research for Child Health Study. Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America. 2012;54(Suppl 2):S109-S16.
75. Macfarlane J, Lewis SA, Macfarlane R, Holmes W. Contemporary use of antibiotics in 1089 adults presenting with acute lower respiratory tract illness in general practice in the U.K.: implications for developing management guidelines. Respiratory Medicine.91(7):427-34.
76. Macfarlane JT, Colville A, Guion A, Macfarlane RM, Rose DH. Prospective study of aetiology and outcome of adult lower-respiratory-tract infections in the community. Lancet. 341. England1993. p. 511-4.
77. Obuchowski NA, Lieber ML, Wians FH. ROC Curves in Clinical Chemistry: Uses, Misuses, and Possible Solutions. Clinical Chemistry. 2004;50(7):1118-25.
78. Obuchowski. Sample Size Tables For Receiver Operating Characteristic Studies. American Journal of Roentgenology. 2000;175(3):603-8.
79. Streiner DL, Norman GR. Correction for multiple testing: Is there a resolution? 2011;140(1):16-8.
80. Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. BMJ. 1995;310(6973):170.
81. Rothman KJ. No adjustments are needed for multiple comparisons. Epidemiology. 1990;1(1):43-6.

82. Perneger TV. What's wrong with Bonferroni adjustments. Bmj. 1998;316(7139):1236-8.
83. Savitz DA, Olshan AF. Multiple comparisons and related issues in the interpretation of epidemiologic data. Am J Epidemiol. 1995;142(9):904-8.
84. Gardner MJ, Altman DG. Confidence intervals rather than P values: estimation rather than hypothesis testing. BMJ. 1986;292(6522):746-50.
85. A.S G. A Course in Econometrics. Cambridge MA: Harvard University Press; 1991.
86. Slinker BK, Glantz SA. Multiple regression for physiological data analysis: the problem of multicollinearity. Am J Physiol. 1985;249(1 Pt 2):R1-12.
87. Läuter H. Weisberg, S.: Applied Linear Regression, 2nd Ed., J. Wiley & Sons, Inc., New York 1985, XIV, 324 pp., £35.75. Biometrical Journal. 1988;30(7):877-.
88. Neter J, Wasserman, W. & Kutner, M. H. Applied Linear Regression Models. Homewood, IL: McGraw-Hill/Irwin; 2004.
89. Mason R. L. GRF, Hess J. L. Statistical Design and Analysis of Experiments: Applications to Engineering and Science. 2nd Edition ed: Wiley; 2003.
90. Moons KGM, Harrell FE, Steyerberg EW. Should scoring rules be based on odds ratios or regression coefficients? Journal of Clinical Epidemiology. 2002;55(10):1054-5.
91. Tripepi G, Jager KJ, Dekker FW, Zoccali C. Statistical methods for the assessment of prognostic biomarkers (Part I): discrimination. Nephrol Dial Transplant. 25. England2010. p. 1399-401.
92. Tripepi G, Jager KJ, Dekker FW, Zoccali C. Statistical methods for the assessment of prognostic biomarkers(part II): calibration and re-classification. Nephrology Dialysis Transplantation. 2010;25(5):1402-5.
93. Hanley JA, McNeil BJ. THE MEANING AND USE OF THE AREA UNDER A RECEIVER OPERATING CHARACTERISTIC (ROC) CURVE. Radiology. 1982;143(1):29-36.
94. Hosmer DW, Lemeshow S. Applied Logistic Regressson. New York: Wiley; 2000.
95. WHO WHO. The process of obtaining informed consent.
96. Greene J, Caracelli V. Making Paradigmatic Sense of Mixed Methods Inquiry. In: Tashakkori A, Teddlie C, editors. Handbook of Mixed Methods in Social & Behavioral Research. California: Sage; 2003.

97. Maxcy S. “Pragmatic threads in mixed methods research in the social sciences: The search for multiple modes of inquiry and the end of the philosophy of formalism”. In: Tashakkori AT, C, editor. Handbook of Mixed Methods in Social & Behavioral Research,. Thousand Oaks, CA: Sage; 2003.

98. Johnson RB, Onwuegbuzie AJ. Mixed Methods Research: A Research Paradigm Whose Time Has Come. Educational Researcher. 2004;33(7):14-26.
99. Freshwater D, Avis M. Analysing interpretation and reinterpreting analysis: exploring the logic of critical reflection. Nursing Philosophy. 2004;5(1):4-11.
100. Morgan DL. Paradigms Lost and Pragmatism Regained: Methodological Implications of Combining Qualitative and Quantitative Methods. Journal of Mixed Methods Research. 2007;1(1):48-76.
101. Gill P, Stewart K, Treasure E, Chadwick B. Methods of data collection in qualitative research: interviews and focus groups. 2008;204(6):291-5.
102. Allmark P, Boote J, Chambers E, Clarke A, McDonnell A, Thompson A, et al. Ethical Issues in the Use of In-Depth Interviews: Literature Review and Discussion. Research Ethics Review. 2009;5(2):48-54.
103. Polit D, Tatano BC. Essentials of Nursing Research: Methods, Appraisal, and Utilization. sixth ed. Philadelphia PA: Lippincott, Williams and Wilkins; 2006.

104. Morse JM. The Significance of Saturation. Qualitative Health Research. 1995;5(2):147-9.

105. Fossey E, Harvey C, McDermott F, Davidson L. Understanding and Evaluating Qualitative Research. Australian and New Zealand Journal of Psychiatry. 2002;36(6):717-32.
106. Byrne MM. Evaluating the findings of qualitative research. AORN J. 2001;73(3):703-6.
107. Guest G, Bunce A, Johnson L. How Many Interviews Are Enough?: An Experiment with Data Saturation and Variability. Field Methods. 2006;18(1):59-82.
108. Newton N. The use of semi-structured interviews in qualitative research: strengths and weaknesses.: University of Bristol.; 2010.
109. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology. 2006;3(2):77-101.
110. Aronson J. {A Pragmatic View of Thematic Analysis}. The Quality Report. 1994;2(1):1989-91.
111. Boyatzis RE. Transforming qualitative information: Thematic analysis and code development. London: Thousand Oaks: Sage; 1998.
112. Hayes N. Theory-led thematic analysis: social identification in small companies. In: Hayes N, editor. Doing Qualitative Analysis in Psychology. United Kingdom: Psychology Press; 1997.
113. Challen K, Goodacre SW, Wilson R, Bentley A, Campbell M, Fitzsimmons C, et al. Evaluation of triage methods used to select patients with suspected pandemic influenza for hospital admission. Emergency Medicine Journal. 2012;29(5):383-8.
114. O'Driscoll BR, Howard LS, Davison AG. BTS guideline for emergency oxygen use in adult patients. Thorax. 2008;63(Suppl 6):vi1-vi68.
115. Mandell LA, Wunderink RG, Anzueto A, Bartlett JG, Campbell GD, Dean NC, et al. Infectious Diseases Society of America/American Thoracic Society Consensus Guidelines on the Management of Community-Acquired Pneumonia in Adults. Clinical Infectious Diseases. 2007;44(Supplement 2):S27-S72.
116. Kuti BP, Adegoke SA, Ebruke BE, Howie S, Oyelami OA, Ota M. Determinants of Oxygen Therapy in Childhood Pneumonia in a Resource-Constrained Region. ISRN Pediatrics. 2013;2013:6.
117. Nizami S, Bhutta Z, Hasan R, Husen Y. Role of chest X-Ray in diagnosis of lower respiratory tract infections in children less than five years of age in community. Pakistan Journal of Medical Sciences. 2005;21(4).
118. Myles PR N-V-TJ, Lim WS, Nicholson KG, Brett SJ,, et al. Comparison of CATs, CURB-65 and PMEWS as triage tools in pandemic influenza admissions to UK hospitals: Case Control Analysis Using Retrospective Data. PLoS ONE. 2012;7(4:):e34428.
119. Myles PR N-V-TJ, Lim WS, Nicholson KG, Brett SJ, et al. Comparison of CATs, CURB-65 and PMEWS as triage tools in pandemic influenza admissions to UK hospitals: Case Control Analysis Using Retrospective Data. PLoS ONE. 2012;7(4:):e34428.
120. Lazzerini M, Sonego M, Pellegrin MC. Hypoxaemia as a Mortality Risk Factor in Acute Lower Respiratory Infections in Children in Low and Middle-Income Countries: Systematic Review and Meta-Analysis. PLoS ONE. 2015;10(9):e0136166.
121. Plüddemann A, Thompson M, Heneghan C, Price C. Pulse oximetry in primary care: primary care diagnostic technology update. The British Journal of General Practice. 2011;61(586):358-9.
122. Grassino A, Macklem PT. Respiratory muscle fatigue and ventilatory failure. Annu Rev Med. 1984;35:625-47.
123. Steven M. Selbst M, FAAP, FACEP. Pediatric Emergency Medicine Secrets,. Selbst SM, editor. Philadephia: Elsevier Saunders; 2000. 664 p.
124. Shann F, Hart K, Thomas D. Acute lower respiratory tract infections in children: possible criteria for selection of patients for antibiotic therapy and hospital admission. Bulletin of the World Health Organization. 1984;62(5):749-53.
125. Mulholland EK, Simoes EA, Costales MO, McGrath EJ, Manalac EM, Gove S. Standardized diagnosis of pneumonia in developing countries. Pediatr Infect Dis J. 1992;11(2):77-81.

126. Singanayagam A, Chalmers JD, Hill AT. Severity assessment in community-acquired pneumonia: a review. QJM. 2009;102(6):379-88.
127. Steyerberg EW. Clinical Prediction Models: a practical approach to development, validation and updating. Gail M KK, Samet J, Tsiatis A, Wong W, editor. New York: Springer; 2008.
128. Daskivich TJ, Fan K-H, Koyama T, Albertsen PC, Goodman M, Hamilton AS, et al. Effect of Age, Tumor Risk, and Comorbidity on Competing Risks for Survival in a U.S. Population–Based Cohort of Men With Prostate Cancer. Annals of internal medicine. 2013;158(10):709-17.
129. Tu JV, Mazer CD, Levinton C, Armstrong PW, Naylor CD. A predictive index for length of stay in the intensive care unit following cardiac surgery. Cmaj. 1994;151(2):177-85.
130. Haybittle JL, Blamey RW, Elston CW, Johnson J, Doyle PJ, Campbell FC, et al. A prognostic index in primary breast cancer. British Journal of Cancer. 1982;45(3):361-6.
131. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. Journal of Chronic Diseases. 1987;40(5):373-83.
132. Daskivich TJ, Kwan L, Dash A, Greenfield S, Litwin MS. Weighted Versus Unweighted Charlson Score to Predict Long-term Other-cause Mortality in Men with Early-stage Prostate Cancer. 2014;66(6):1002-9.
133. Tu YK, Kellett M, Clerehugh V, Gilthorpe MS. Problems of correlations between explanatory variables in multiple regression analyses in the dental literature. 2005;199(7):457-61.
134. Stoltzfus JC. Logistic Regression: A Brief Primer. Academic Emergency Medicine. 2011;18(10):1099-104.
135. Cleveland W. {LOWESS}: {A} {P}rogram for {S}moothing {S}catterplots by {R}obust {L}ocally {W}eighted {R}egression. The American Statistician. 1981;35(1).
136. Simon R. Clinical trial designs for evaluating the medical utility of prognostic and predictive biomarkers in oncology. Per Med. 2010;7(1):33-47.
137. Nemes S, Jonasson J, Genell A, Steineck G. Bias in odds ratios by logistic regression modelling and sample size. BMC Medical Research Methodology. 2009;9(1):56.
138. Gupta S, QasemMamun AN. Effect of Sample Size on Correlation and Regression Coefficients. International Journal of Research in Economics and Social sciences. 2015;5(7).
139. &amp DoH, amp, amp, NHS. Swine flu clinical package for use when there are exceptional demands on healthcare services. In: &amp DoH, amp, amp, NHS, editors.: COI for Department of Health Gateway ref 12368; 2009.
140. Vergouwe Y, Steyerberg EW, Eijkemans MJ, Habbema JD. Validity of prognostic models: when is a model clinically useful? Semin Urol Oncol. 2002;20(2):96-107.
141. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for some traditional and novel measures. Epidemiology (Cambridge, Mass). 2010;21(1):128-38.
142. Kramer AA, Zimmerman JE. Assessing the calibration of mortality benchmarks in critical care: The Hosmer-Lemeshow test revisited. Crit Care Med. 2007;35(9):2052-6.
143. Feudtner C, Hexem KR, Shabbout M, Feinstein JA, Sochalski J, Silber JH. Prediction of Pediatric Death in the Year after Hospitalization: A Population-Level Retrospective Cohort Study. Journal of Palliative Medicine. 2009;12(2):160-9.
144. Hosmer DW, Hosmer T, Le Cessie S, Lemeshow S. A comparison of goodness-of-fit tests for the logistic regression model. Stat Med. 1997;16(9):965-80.
145. Akobeng AK. Understanding diagnostic tests 3: receiver operating characteristic curves. Acta Pædiatrica. 2007;96(5):644-7.
146. Macfarlane J, Boldy D. 2004 update of BTS pneumonia guidelines: what's new? Thorax. 2004;59(5):364-6.
147. Neill AM, Martin IR, Weir R, Anderson R, Chereshsky A, Epton MJ, et al. Community acquired pneumonia: aetiology and usefulness of severity criteria on admission. Thorax. 1996;51(10):1010-6.

148. Fine Mj HLJMAR, et al. The hospital admission decision for patients with community-acquired pneumonia: Results from the pneumonia patient outcomes research team cohort study. Archives of Internal Medicine. 1997;157(1):36-44.
149. Hoare Z, Lim WS. Pneumonia: update on diagnosis and management. BMJ : British Medical Journal. 2006;332(7549):1077-9.
150. Niederman MS, Mandell LA, Anzueto A, Bass JB, Broughton WA, Campbell GD, et al. Guidelines for the management of adults with community-acquired pneumonia. Diagnosis, assessment of severity, antimicrobial therapy, and prevention. Am J Respir Crit Care Med. 2001;163(7):1730-54.
151. Eccles S, Pincus C, Higgins B, Woodhead M. Diagnosis and management of community and hospital acquired pneumonia in adults: summary of NICE guidance2014 2014-12-03 12:41:22.
152. Charles PGP, Wolfe R, Whitby M, Fine MJ, Fuller AJ, Stirling R, et al. SMART-COP: A Tool for Predicting the Need for Intensive Respiratory or Vasopressor Support in Community-Acquired Pneumonia. Clinical Infectious Diseases. 2008;47(3):375-84.
153. Bauer TT, Ewig S, Marre R, Suttorp N, Welte T, The Capnetz Study G. CRB-65 predicts death from community-acquired pneumonia*. Journal of Internal Medicine. 2006;260(1):93-101.
154. Yan Man S, Lee N, Ip M, Antonio GE, Chau SSL, Mak P, et al. Prospective comparison of three predictive rules for assessing severity of community-acquired pneumonia in Hong Kong. Thorax. 2007;62(4):348-53.
155. Ewig S, Welte T. CRB-65 for the assessment of pneumonia severity: who could ask for more? Thorax. 2008;63(8):665-6.
156. Ewig S, Torres A, Woodhead M. Assessment of pneumonia severity: a European perspective. European Respiratory Journal. 2006;27(1):6-8.
157. Capelastegui A, España PP, Quintana JM, Areitio I, Gorordo I, Egurrola M, et al. Validation of a predictive rule for the management of community-acquired pneumonia. European Respiratory Journal. 2006;27(1):151-7.
158. Altman DG, Bland JM. Diagnostic tests 3: receiver operating characteristic plots. BMJ. 1994;309(6948):188.
159. Guidelines NGa. Pneumonia in adults: diagnosis and management. 2014.
160. Yandiola PPE, Capelastegui A, Quintana J, Diez R, Gorordo I, Bilbao A, et al. Prospective Comparison of Severity Scores for Predicting Clinically Relevant Outcomes for Patients Hospitalized With Community-Acquired Pneumonia. CHEST Journal. 2009;135(6):1572-9.
161. Marrie TJ, Huang JQ. Low-risk patients admitted with community-acquired pneumonia. Am J Med. 2005;118(12):1357-63.
162. Bont J, Hak E, Hoes AW, Schipper M, Schellevis FG, Verheij TJM. A prediction rule for elderly primary-care patients with lower respiratory tract infections. European Respiratory Journal. 2007;29(5):969-75.
163. Moskop JC, Iserson KV. Triage in Medicine, Part II: Underlying Values and Principles. 2007;49(3):282-7.
164. O'Laughlin DT, Hick JL. Ethical issues in resource triage. Respir Care. 2008;53(2):190-7; discussion 7-200.
165. Aacharya RP, Gastmans C, Denier Y. Emergency department triage: an ethical analysis. 2011;11(1):1-13.
166. WHO. The Nigerian Health System [Available from: http://www.who.int/pmnch/countries/nigeria-plan-chapter-3.pdf.
167. Soyibo A OO, Lawanson AO. National Health Accounts of Nigeria 2003-2005: Incorporating Sub-National Health Accounts of States. Main report submitted to Federal Ministry of Health. University of Ibadan;Ibadan2009.
168. WHO PoARI. Acute respiratory infections in children : case management in small hospitals in developing countries, a manual for doctors and other senior health workers 1990.

169. Olowu A, Elusiyan J, Esangbedo D, Ekure E, Esezobor C, Falade A, et al. Management of community acquired pneumonia (CAP) in children: Clinical practice guidelines by the Paediatrics Association of Nigeria (PAN). Nigerian Journal of Paed. 2015;42(4):283-92.

170. Onyedum CC, Chukwuka JC. Admission profile and management of community acquired pneumonia in Nigeria-5 year experience in a tertiary hospital. Respiratory Medicine. 2011;105(2):298-302.
171. Angus DC, Marrie TJ, Obrosky DS, Clermont G, Dremsizov TT, Coley C, et al. Severe Community-acquired Pneumonia: Use of Intensive Care Services and Evaluation of American and British Thoracic Society Diagnostic Criteria. American Journal of Respiratory and Critical Care Medicine. 2002;166(5):717-23.
172. Mbata GC, Chukwuka CJ, Onyedum CC, Onwubere BJC. The CURB-65 scoring system in severity assessment of Eastern Nigerian patients with community-acquired pneumonia: a prospective observational study. 2013;22:175.
173. M.O T. Mortality predictors in community-acquired pnuemonia. Nigerian journal of clinical practice. 2009 September;12(3):298-301.
174. Mbata GC, Chukwuka CJ, Onyedum CC, Onwubere BJC, Aguwa CN. Comparison of two predictive rules for assessing severity of communityacquired pneumonia. African Journal of Respiratory Medicine. 2014;10(1).

175. Mbata GC, Chukwuka CJ, Onyedum CC, Onwubere BJC, Aguwa EN. The Role of Complications of Community Acquired Pneumonia on the Outcome of the Illness: A Prospective Observational Study in a Tertiary Institution in Eastern Nigeria. Annals of Medical and Health Sciences Research. 2013;3(3):365-9.
176. Desalu OO, Ojo OO, Busari OA, Fadeyi A. Pattern of respiratory diseases seen among adults in an emergency room in a resource-poor nation health facility. The Pan African medical journal. 2011;9:24.
177. Chalmers JD, Singanayagam A, Akram AR, Mandal P, Short PM, Choudhury G, et al. Severity assessment tools for predicting mortality in hospitalised patients with community-acquired pneumonia. Systematic review and meta-analysis. Thorax. 2010;65(10):878-83.
178. Loke YK, Kwok CS, Niruban A, Myint PK. Value of severity scales in predicting mortality from community-acquired pneumonia: systematic review and meta-analysis. Thorax. 2010;65(10):884-90.
179. Siegel RE. Clinical opinion prevails over the pneumonia severity index. Am J Med. 2005;118(12):1312-3.
180. Matheny ME, Ohno-Machado L, Resnic FS. Discrimination and calibration of mortality risk prediction models in interventional cardiology. Journal of Biomedical Informatics.38(5):367-75.
181. Kienle GS, Kiene H. Clinical judgement and the medical profession. Journal of Evaluation in Clinical Practice. 2011;17(4):621-7.
182. Croskerry P. A universal model of diagnostic reasoning. Acad Med. 84. United States2009. p. 1022-8.
183. Hammond KR. Intuitive and analytic cognition: Information models. In: Sage A, editor. Concise Encyclopedia of Information Processing in Systems and Organizations. Oxford, UK: Pergamon Press; 1990. p. 306–12.
184. Evans JS. Dual-processing accounts of reasoning, judgment, and social cognition. Annu Rev Psychol. 2008;59:255-78.
185. Epstein S. Integration of the cognitive and the psychodynamic unconscious. Am Psychol. 1994;49(8):709-24.
186. Stolper E, Van de Wiel M, Van Royen P, Van Bokhoven M, Van der Weijden T, Dinant GJ. Gut Feelings as a Third Track in General Practitioners' Diagnostic Reasoning. Journal of General Internal Medicine. 2011;26(2):197-203.

187. Barrows HS, Norman GR, Neufeld VR, Feightner JW. The clinical reasoning of randomly selected physicians in general medical practice. Clin Invest Med. 1982;5(1):49-55.
188. Norman G. Dual processing and diagnostic errors. Adv Health Sci Educ Theory Pract. 2009;14 Suppl 1:37-49.
189. Regehr G, Cline J, Norman GR, Brooks L. Effect of processing strategy on diagnostic skill in dermatology. Acad Med. 1994;69(10 Suppl):S34-6.
190. de Vries M, Witteman CL, Holland RW, Dijksterhuis A. The unconscious thought effect in clinical decision making: an example in diagnosis. Med Decis Making. 30. United States2010. p. 578-81.
191. Dhaliwal G. Going with Your Gut. Journal of General Internal Medicine. 2011;26(2):107-9.
192. Cooper N. Models of clinical reasoning. In: Cooper N, Frain J, editors. ABC of Clinical Reasoning. Oxford: Wiley; 2016.
193. Klein G. The Power of Intuition: How to Use Your Gut Feelings to Make Better Decisions at Work. New York: Currency-Doubleday; 2003.
194. Smith A, Goodwin D, Mort M, Pope C. Expertise in practice: an ethnographic study exploring acquisition and use of knowledge in anaesthesia. Br J Anaesth. 2003;91(3):319-28.
195. Henry SG. Recognizing tacit knowledge in medical epistemology. Theor Med Bioeth. 2006;27(3):187-213.
196. Hamm R. Clinical intuition and clinical analysis: Expertise and the Cognitive Continuum. . In: Dowie J AE, editor. Professional judgement A reader in clinical decision making Cambridge: Cambridge University Press; 1988. p. 78-104.
197. Custers EJ. Medical education and cognitive continuum theory: an alternative perspective on medical problem solving and clinical reasoning. Acad Med. 2013;88(8):1074-80.
198. Norman G, Monteiro S, Sherbino J. Is clinical cognition binary or continuous? Acad Med. 88. United States2013. p. 1058-60.
199. Mickleborough T. Intuition in medical practice: A reflection on Donald Schön's reflective practitioner. Medical Teacher. 2015;37(10):889-91.
200. Van den Bruel A, Thompson M, Buntinx F, Mant D. Clinicians' gut feeling about serious infections in children: observational study. BMJ. 2012;345.
201. Lodha R, Bhadauria PS, Kuttikat AV, Puranik M, Gupta S, Pandey RM, et al. Can clinical symptoms or signs accurately predict hypoxemia in children with acute lower respiratory tract infections? Indian Pediatr. 2004;41(2):129-35.
202. Weber M, Usen S, Palmer A, Jaffar S, Mulholland E. Predictors of hypoxaemia in hospital admissions with acute lower. Arch Dis Child. 1997;76(4):310-4.
203. Zhang L, Mendoza-Sassi R, Santos JCH, Lau J. Accuracy of symptoms and signs in predicting hypoxaemia among young children with acute respiratory infection: a meta-analysis [Review article]. The International Journal of Tuberculosis and Lung Disease. 2011;15(3):317-25.
204. Djulbegovic B, Hozo I, Beckstead J, Tsalatsanis A, Pauker SG. Dual processing model of medical decision-making. 2012;12(1):1-13.
205. Benner P. From Novice To Expert. AJN The American Journal of Nursing. 1982;82(3).
206. Woolley A, Kostopoulou O. Clinical Intuition in Family Medicine: More Than First Impressions. The Annals of Family Medicine. 2013;11(1):60-6.
207. Dreyfus H, Dreyfus S, Athanasiou T. Mind over machine:the power of human intuition and expertise in the era of the computer. Oxford: Basil Blackwell; 1986.
208. Spadaccini J, Esteves JE. Intuition, analysis and reflection: An experimental study into the decision-making processes and thinking dispositions of osteopathy students. International Journal of Osteopathic Medicine.17(4):263-71.
209. Murthy S, Wunsch H. Clinical review: International comparisons in critical care - lessons learned. 2012;16(2):1-7.
210. Murthy S, Leligdowicz A, Adhikari NKJ. Intensive Care Unit Capacity in Low-Income Countries: A Systematic Review. PLoS ONE. 2015;10(1):e0116949.

211. Wunsch H, Angus DC, Harrison DA, Collange O, Fowler R, Hoste EA, et al. Variation in critical care services across North America and Western Europe. Crit Care Med. 2008;36(10):2787-93, e1-9.
212. Adudu OP, Adudu OG. *Working practices and patient outcome in the intensive care unit of the University of Benin Teaching Hospital*Working practices and patient outcome in the intensive care unit of the University of Benin Teaching Hospital. *Journal of Medicine and Biomedical Research*Journal of Medicine and Biomedical Research. 2004;3(1):67-72.
213. Onyekwulu FA, Anya SU. Pattern of admission and outcome of patients admitted into the Intensive Care Unit of University of Nigeria Teaching Hospital Enugu: A 5-year review. Niger J Clin Pract. 18. India2015. p. 775-9.
214. Ilori IU, Kalu QN. Intensive care admissions and outcome at the University of Calabar Teaching Hospital, Nigeria. J Crit Care. 27. United States: 2012 Elsevier Inc; 2012. p. 105 e1-4.
215. Abubakar A, Ojo E, EL Nafaty A, Edomwonyi N. An Audit Of One Year Intensive Care Practice In a Developing Country. The Internet Journal of Anesthesiology. 2007;18(2).
216. Baker T. Critical care in low-income countries. Trop Med Int Health. 14. England2009. p. 143-8.
217. Okafor UV. Challenges in critical care services in Sub-Saharan Africa: Perspectives from Nigeria. Indian Journal of Critical Care Medicine : Peer-reviewed, Official Publication of Indian Society of Critical Care Medicine. 2009;13(1):25-7.
218. Bolaji BO, Kolawole KL. The Intensive Care Unit of the University Teaching Hospital, Ilorin, Nigeria: A Ten Year Review (1991–2001). Southern African Journal of Anaesthesia and Analgesia. 2005;11(4).
219. *Adudu OP*, *Adudu OG*. *Working practices and patient outcome in the intensive care unit of the University of Benin Teaching Hospital*. *Journal of Medicine and Biomedical Research*. 2004;3(1):67-72.
220. Department o, Health. Comprehensive critical care: a review of adult critical care services. London: Department of Health; 2000.

221. Tobi K, Efobi C. Characteristics and outcome of critically ill patients admitted to a general intensive care unit: Time for specialized units in Nigeria? African Journal of Medical and Health Sciences. 2015;14(2):153-4.
222. Nguyen Y-L, Kahn JM, Angus DC. Reorganizing Adult Critical Care Delivery. American Journal of Respiratory and Critical Care Medicine. 2010;181(11):1164-9.
223. Thompson DR, Clemmer TP, Applefeld JJ, Crippen DW, Jastremski MS, Lucas CE, et al. Regionalization of critical care medicine: task force report of the American College of Critical Care Medicine. Crit Care Med. 1994;22(8):1306-13.
224. Barnato AE, Kahn JM, Rubenfeld GD, McCauley K, Fontaine D, Frassica JJ, et al. Prioritizing the organization and management of intensive care services in the United States: the PrOMIS Conference. Crit Care Med. 2007;35(4):1003-11.
225. Singh JM, MacDonald RD. Pro/con debate: Do the benefits of regionalized critical care delivery outweigh the risks of interfacility patient transport? 2009;13(4):1-7.
226. Kahn JM, Linde-Zwirble WT, Wunsch H, Barnato AE, Iwashyna TJ, Roberts MS, et al. Potential value of regionalized intensive care for mechanically ventilated medical patients. Am J Respir Crit Care Med. 2008;177(3):285-91.
227. Glance LG, Li Y, Osler TM, Dick A, Mukamel DB. Impact of patient volume on the mortality rate of adult intensive care unit patients. Crit Care Med. 2006;34(7):1925-34.
228. Newgard CD, McConnell KJ, Hedges JR, Mullins RJ. The benefit of higher level of care transfer of injured patients from nontertiary hospital emergency departments. J Trauma. 63. United States2007. p. 965-71.
229. Dunser MW, Baelani I, Ganbold L. A review and analysis of intensive care medicine in the least developed countries. Crit Care Med. 2006;34(4):1234-42.

230. Buor D. Analysing the primacy of distance in the utilization of health services in the Ahafo-Ano South district, Ghana. Int J Health Plann Manage. 2003;18(4):293-311.
231. Metcalfe MA, Sloggett A, McPherson K. Mortality among appropriately referred patients refused admission to intensive-care units. 1997;350(9070):7-11.
232. Goldhill DR, McNarry AF, Hadjianastassiou VG, Tekkis PP. The longer patients are in hospital before Intensive Care admission the higher their mortality. Intensive Care Med. 2004;30(10):1908-13.
233. Isamade ES, Yiltok SJ, Uba AF, Isamade EI, Daru PH. Intensive care unit admissions in the Jos University Teaching Hospital. Niger J Clin Pract. 2007;10(2):156-61.
234. Yiltok SJ, Isamade ES, Uba AF. Outcome of Burn patients managed in a general intensive care unit. Nigerian Journal Of Surgery. 2005;11:1-4.
235. Evans TW, Nava S, Mata GV, Guidet B, Estenssoro E, Fowler R, et al. Critical Care Rationing: International Comparisons. 2011;140(6):1618-24.
236. Ward NS, Teno JM, Curtis JR, Rubenfeld GD, Levy MM. Perceptions of cost constraints, resource limitations, and rationing in United States intensive care units: results of a national survey. Crit Care Med. 36. United States2008. p. 471-6.
237. Danis M, Linde-Zwirble WT, Astor A, Lidicker JR, Angus DC. How does lack of insurance affect use of intensive care? A population-based study. Crit Care Med. 2006;34(8):2043-8.
238. Fowler RA, Noyahr LA, Thornton JD, Pinto R, Kahn JM, Adhikari NK, et al. An official American Thoracic Society systematic review: the association between health insurance status and access, care delivery, and outcomes for patients who are critically ill. Am J Respir Crit Care Med. 2010;181(9):1003-11.
239. Giannini A, Consonni D. Physicians' perceptions and attitudes regarding inappropriate admissions and resource allocation in the intensive care setting. British Journal of Anaesthesia. 2006;96(1):57-62.
240. Vincent JL. European attitudes towards ethical problems in intensive care medicine: results of an ethical questionnaire. Intensive Care Med. 1990;16(4):256-64.
241. Ugwu BT, Isamade ES, Isamade EI. Intraoperative cardiac arrest-A tropical experience. West African Journal of Medicine. 2000;99:277-80.

242. Pollack MM, Katz RW, Ruttimann UE, Getson PR. Improving the outcome and efficiency of intensive care: the impact of an intensivist. Crit Care Med. 1988;16(1):11-7.
243. Fuchs RJ, Berenholtz SM, Dorman T. Do intensivists in ICU improve outcome? Best Pract Res Clin Anaesthesiol. 2005;19(1):125-35.
244. Pronovost PJ, Angus DC, Dorman T, Robinson KA, Dremsizov TT, Young TL. Physician staffing patterns and clinical outcomes in critically ill patients: a systematic review. JAMA. 288. United States2002. p. 2151-62.
245. Obaseki D, Adeniyi B, Kolawole T, Onyedum C, Erhabor G. Gaps in capacity for respiratory care in developing countries. Nigeria as a case study. Ann Am Thorac Soc. 2015;12(4):591-8.
246. Michael Huberman, Miles MB. The Qualitative Researcher's Companion: SAGE Publications, Inc; 2002.
247. Leung L. Validity, reliability, and generalizability in qualitative research. Journal of Family Medicine and Primary Care. 2015;4(3):324-7.
248. Morse JM. Qualitative generalizability. *Qualitative Health Research*. 1999;9(1):5-6.
249. WM T. Research Methods: The Concise Knowledge Base. Cincinnati, Ohio: Atomic Dog Publishing; 2005.
250. WHO/UNICEF. IMCI adaptation guide A guide to identifying necessary adaptations of clinical policies and guidelines, and to adapting the charts and modules for the WHO/UNICEF course. 2002.
251. Tamburlini G, Di M, Maggi RS, Vilarim JN, Gove S. Evaluation of guidelines for emergency triage assessment and treatment in developing countries. Archives of Disease in Childhood. 1999;81(6):478-82.

252. Peek N, Arts DGT, Bosman RJ, van der Voort PHJ, de Keizer NF. External validation of prognostic models for critically ill patients required substantial sample sizes. Journal of Clinical Epidemiology. 2007;60(5):491.e1-.e13.
253. Vergouwe YS, E W.Eijkemans, R.Habbema D. Sample size considerations for performance assessment of prediction models: a simulation study. Control Clin Trials. 2003;24(2 Suppl)::S43-S4.
254. Cox DR. Two further applications of a model for binary regression. Biometrika. 1958;45(3-4):562-5.

Appendices

APPENDIX I: TRIPOD checklist -methods section

Section/ Topic	Study type	Checklist Item
Source of data	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.
	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.
Participants	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.
	D;V	Describe eligibility criteria for participants.
	D;V	Give details of treatments received, if relevant.
Outcome	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.
	D;V	Report any actions to blind assessment of the outcome to be predicted.
Predictors	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.
	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.

Sample size	D;V	Explain how the study size was arrived at.
Missing data	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.
Statistical analysis methods	D	Describe how predictors were handled in the analyses.
	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.
	V	For validation, describe how the predictions were calculated.
	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.
	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.
Risk groups	D;V	Provide details on how risk groups were created, if done.
Development vs. validation	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.

*Items relevant only to the development of a prediction model are denoted by D;

*Items relating solely to the validation of a prediction model are denoted by V;

*Items relating to both are denoted D; V.

APPENDIX II: Data collection Forms

Hospital :
Hospital number:

Data Collection Form for Patients under 16 years

PART 1: INITIAL ASSESSMENT IN EMERGENCY DEPARTMENT

1. PATIENT DEMOGRAPHICS

Patient details (complete all fields)

Forename : Surname :

D.O.B.: ☐ ☐/ ☐ ☐/ ☐ ☐ ☐ ☐ or age (years): or age if < 12months:

Gender : M ☐ F ☐

Ethnicity : Yoruba ☐ Igbo ☐ Hausa ☐ Other ☐

Was patient referred: Y ☐ N ☐

Route of referral : Private clinic: ☐ Primary Health Centre : ☐ General Hospital: ☐

Date of presentation to referral hospital: ☐ ☐/ ☐ ☐/ ☐ ☐ ☐ ☐

Date of first presentation to this hospital : ☐ ☐/ ☐ ☐/ ☐ ☐ ☐ ☐

2. PRESENTING SYMPTOMS

2.1 Inclusion Criteria (tick all boxes that apply)		
☐	Cough	Date of onset: ☐ ☐/ ☐ ☐/ ☐ ☐ ☐ ☐ or Duration (in days):
☐	Difficulty in breathing	Date of onset: ☐ ☐/ ☐ ☐/ ☐ ☐ ☐ ☐ Or Duration (in days):
☐	Difficulty in feeding/drinking	Date of onset: ☐ ☐/ ☐ ☐/ ☐ ☐ ☐ ☐ Or Duration (in days):
☐	Fast breathing	Date of onset: ☐ ☐/ ☐ ☐/ ☐ ☐ ☐ ☐ Or Duration (in days):
☐	Lower chest in-drawing	Date of onset: ☐ ☐/ ☐ ☐/ ☐ ☐ ☐ ☐ Or Duration (in days):

2.2 other symptoms (tick all boxes that apply)					
Fever		Diarrhoea		Nasal flaring	
Vomiting		Seizures		Persistent crying	

3. EXAMINTION

Respiratory rate (breaths per minute)	
Oxygen saturation on pulse oximetry	
Blood Pressure (mmHg) Systolic Diastolic	
Heart rate (beats per minute)	
Temperature (^{o}C)	

4. TRIAGE CRITERIA

CATs Criteria for patients under 16 years (tick all boxes that apply)	
Criteria met	Presenting symptoms
☐ **A**	**Severe respiratory distress** Lower chest wall indrawing, sternal recession , grunting or noisy breathing when calm
☐ **B**	**Increased respiratory rate** Measured over at least 30 seconds, ≥50 breaths per minute if under 1 year or ≥40 breaths per minute if over 1year
☐ **C**	**Oxygen saturation ≤92% on pulse oximetry breathing air or on oxygen** Absence of cyanosis is a poor discriminator for severe illness
☐ **D**	**Respiratory exhaustion or apnoeic episode** Apnoea defined as a ≥20 second pause in breathing
☐ **E**	**Evidence of severe clinical dehydration or clinical shock** Sternal capillary refill time > 2seconds, reduced skin turgor, sunken eyes or fontanelle
☐ **F**	**Altered conscious level** Strikingly agitated or irritable, seizures or floppy infant
☐ **G**	**Causing clinical concern to the clinical team or specialist doctor** e.g. rapidly progressive or unusually prolonged illness
☐ **None identified**	

5. PROVISIONAL DIAGNOSIS

Pneumonia	
Acute bronchitis	
Bronchiolitis	

6. ADMITTED TO HOSPITAL OR DISCHARGED FROM EMERGENCY DEPARTMENT

Admitted Y ☐ N ☐	Date(☐ ☐/ ☐ ☐/ ☐ ☐ ☐ ☐)
Admit for observation in ED	
Admit to ward	
Admit to ICU	
Discharge from ED	
Reason for discharge:	Not severe enough to admit to ward/ICU Y ☐ N ☐
	Financial reasons Y ☐ N ☐

7. INTERVENTIONS AT TIME OF DECISION

Intravenous fluids		Intravenous antibiotics	
Supplemental oxygen		Mechanical ventilation	
Antipyretics		Others (please state):	

Hospital :
Hospital number:

PART 2: WARD/ ICU ASSESSMENT AND OBSERVATIONS

8. PAST MEDICAL HISTORY

8.1 Comorbidities (tick all that apply)			
Asthma		Endocrine disease	
Sickle cell anaemia		Malnutrition	
Malaria		HIV +VE	
Neurological disease		Tuberculosis	
Malignancy : haematological Solid tumour		Oral steroids or immune suppressing medication/treatment	
Renal disease		Congenital heart disease	

8.2 Risk factors
Does any member of family smoke: Y ☐ N ☐

9. PRE-ADMISION MEDICATIONS

Self-medication for this illness prior to presentation : Y ☐ N ☐			
Prescribed medication for this illness from referral hospital : Y ☐ N ☐			
Name of medication	Dose	Date started	Days completed

10. BIRTH HISTORY & IMMUNISATION HISTORY (CHILDREN < 5YEARS)

10.1 Mode of delivery
Spontaneous vaginal delivery ☐ Assisted vaginal delivery ☐ Caesarean section ☐
Gestational age at birth :
Birth weight :
Neonatal problems : Y ☐ N ☐
Details if yes:

10.2 Immunisation	
At Birth	BCG ☐ OPV ☐ HBV 1 ☐
6 weeks	DTP1 ☐ OPV1 ☐ HBV 2 ☐ Hib 1 ☐
10 weeks	DPT 2 ☐ OPV2 ☐ Hib 2 ☐
14 weeks	DPT 3 ☐ OPV 3 ☐ HBV 3 ☐ Hib 3 ☐
9 months	Measles + Yellow fever + Vitamin A ☐

11. OBSERVATIONS

Day 1= (first day on admission on wards) ☐☐/☐☐/☐☐☐☐													
indicate oxygen administration with a ✓ **indicate oxygen free day on the ward with an **X **indicate discharge with **D** **indicate death with **O**													
1		5		9		13		17		21		25	

2		6		10		14		18		22		26	
3		7		11		15		19		23		27	
4		8		12		16		20		24		28	

Day 1= (first day on admission in ICU) ☐☐/☐☐/☐☐☐☐												
indicate mechanical ventilation with a ✓ **indicate ventilation free day with an **X **indicate discharge **D** **indicate death with **O**												
1		5		9		13		17		21		25	
2		6		10		14		18		22		26	
3		7		11		15		19		23		27	
4		8		12		16		20		24		28	

12. OUTCOMES

12.1 Length of stay	
Length of stay on Oxygen	
Length of stay on mechanical ventilator	
Length of stay on the wards	
Length of stay in ICU	
Moved from ICU to ward Y ☐ N ☐ Moved from ward to ICU Y ☐ N ☐	

12.2 Discharge from ward or Death		
Discharge date:	☐☐/☐☐/☐☐☐☐	
Was patient ready for discharge On this date? (tick as appropriate)	Ready Y ☐ N ☐	
	Not ready – but discharged self, due to financial reasons Y ☐ N ☐	
	Other reasons:	
Death	Ward ☐	Emergency department ☐ ICU ☐
Cause of death :		

Hospital :
Hospital number:

Data Collection Form for Patients 16 years and older

PART 1: INITIAL ASSESSMENT IN EMERGENCY DEPARTMENT

1. PATIENT DEMOGRAPHICS

Patient details (complete all fields)
Forename : Surname :
D.O.B.: ☐ ☐/ ☐ ☐/ ☐ ☐ ☐ ☐ or age (years): Gender : M ☐ F ☐
Occupation:
Ethnicity : Yoruba ☐ Igbo ☐ Hausa ☐ Other ☐
Was patient referred: Y ☐ N ☐
Route of referral : Private clinic: ☐ Primary Health Centre : ☐ General Hospital: ☐
Date of presentation to referral hospital: ☐ ☐/ ☐ ☐/ ☐ ☐ ☐ ☐
Date of first presentation to this hospital : ☐ ☐/ ☐ ☐/ ☐ ☐ ☐ ☐

2. PRESENTING SYMPTOMS

2.1 Inclusion criteria (tick all boxes that apply)		
☐	Cough Productive of sputum: Y ☐ N ☐	Date of onset: ☐ ☐/ ☐ ☐/ ☐ ☐ ☐ ☐ or Duration (in days):
☐	Difficulty in breathing	Date of onset: ☐ ☐/ ☐ ☐/ ☐ ☐ ☐ ☐ Or Duration (in days):

☐	Fast breathing	Date of onset: ☐ ☐/ ☐ ☐/ ☐ ☐ ☐ ☐ Or Duration (in days):
☐	Chest Pain	Date of onset: ☐ ☐/ ☐ ☐/ ☐ ☐ ☐ ☐ Or Duration (in days):

2.2 other symptoms (tick all boxes that apply)					
Fever		Headache		Myalgia	
Sore throat		Nausea & vomiting		General unwellness	

3. EXAMINTION

Respiratory rate (breaths per minute)	
Oxygen saturation on pulse oximetry	
Blood Pressure (mmHg) Systolic Diastolic	
Heart rate (beats per minute)	
Temperature (^{o}C)	

4. TRIAGE CRITERIA

4.2 CATs Criteria for patients 16 years and over (tick all boxes that apply)	
Criteria met	Presenting symptoms
☐ **A**	**Severe respiratory distress** Severe breathlessness: e.g. unable to complete sentences in one breath, use of accessory muscles, supraclavicular recession, tracheal tug or feeling suffocation.
☐ **B**	**Increased respiratory rate** Measured over at least 30 seconds. Over 30 breaths per minute.
☐ **C**	**Oxygen saturation ≤92% on pulse oximetry breathing air or on oxygen** Absence of cyanosis is a poor discriminator for severe illness
☐ **D**	**Respiratory exhaustion or apnoeic episode** New abnormal breathing pattern e.g. alternating fast and slow rate or long pauses between regular breaths
☐ **E**	**Evidence of severe clinical dehydration or clinical shock** Systolic blood pressure <90mmHg and/or diastolic blood pressure <60mmHg. Sternal capillary refill time > 2seconds, reduced skin turgor
☐ **F**	**Altered conscious level** New confusion, strikingly agitated or seizures
☐ **G**	**Causing clinical concern to the clinical team or specialist doctor** e.g. rapidly progressive or unusually prolonged illness
☐ **None identified**	

5. PROVISIONAL DIAGNOSIS

Pneumonia	
Acute Bronchitis	
Exacerbation of COPD	

6. ADMITTED OR DISCHARGED FROM EMERGENCY DEPARTMENT

Admitted Y ☐ N ☐	Date(☐ ☐/ ☐ ☐/ ☐ ☐ ☐ ☐)
Admit for observation in ED	
Admit to ward	
Admit to ICU	
Discharge from ED	
Reason for discharge	Not severe enough to admit to ward/ICU Y ☐ N ☐
	Financial reasons Y ☐ N ☐

7. INTERVENTIONS AT TIME OF DECISION

Intravenous fluids		Intravenous antibiotics	
Supplemental oxygen		Mechanical ventilation	
Antipyretics		Others (please state):	

Hospital :
Hospital number:

PART 2: WARD/ ICU ASSESSMENT AND OBSERVATIONS

8. PAST MEDICAL HISTORY

8.1 Comorbidities Tick all that apply			
Asthma		Other chronic lung disease	
Sickle cell anaemia		COPD	
Malaria		HIV positive	
Cardiovascular disease		Tuberculosis	
Pregnant Y ☐ N ☐ Number of weeks?		Oral steroids or immune suppressing medication/treatment	
Diabetes		Malignancy : haematological Solid tumour	

8.2 Risk factors								
Does any house hold member smoke: Y ☐ N ☐								
Smoking	Daily		Occasional		Ex -smoker		Never	
Alcohol	Quantity :							
	Frequency :							

9. PRE-ADMISSION MEDICATIONS

Self-medication for this illness prior to presentation : Y ☐ N ☐
Prescribed medication for this illness from referral hospital : Y ☐ N ☐

Name of medication	Dose	Date started	Days completed

10. OBSERVATIONS

Day 1= (first day on admission on wards) ☐☐/☐☐/☐☐☐☐												
indicate oxygen administration with a ✓ **indicate oxygen free day on the ward with an **X **indicate discharge with **D** **indicate death with **O**												
1		5		9		13		17		21		25	
2		6		10		14		18		22		26	
3		7		11		15		19		23		27	
4		8		12		16		20		24		28	

Day 1= (first day on admission in ICU) ☐☐/☐☐/☐☐☐☐												
indicate mechanical ventilation with a ✓ **indicate ventilation free day with an **X **indicate discharge with **D** **indicate death with **O**												
1		5		9		13		17		21		25	
2		6		10		14		18		22		26	

3		7		11		15		19		23		27	
4		8		12		16		20		24		28	

11. OUTCOMES

11.1 Length of stay	
Length of stay on Oxygen	
Length of stay on mechanical ventilator	
Length of stay on the ward	
Length of stay in ICU	
Moved from ICU to ward	☐ ☐/ ☐ ☐/ ☐ ☐ ☐ ☐
Moved from ward to ICU	☐ ☐/ ☐ ☐/ ☐ ☐ ☐ ☐

11.2 Discharge from ward or Death		
Discharge date :	☐ ☐/ ☐ ☐/ ☐ ☐ ☐ ☐	
Was patient ready for discharge On this date? (tick as appropriate)	Ready Y ☐ N ☐	
	Not ready – but discharged self, due to financial reasons Y ☐ N ☐	
	Other reasons:	
Death	Ward ☐ Emergency department ☐ ICU ☐	
Cause of death :		

APPENDIX III: Qualitative interview questionnaire

1. How is 'critical care' defined?

2. Are there different levels of critical care? Is this the same for adults and children?

3. What treatment or care is provided in critical care facilities (that is different from treatment/care provided on regular wards)? Is this the same for adults and children?

4. What factors influence admission to critical care in general (irrespective of a given presenting diagnosis)?

5. What factors influence admission to critical care specifically in patients presenting with acute respiratory tract illness? Are any predefined criteria used? If yes, what are these? Is this the same for adults and children?

6. What factors influence discharge from critical care facilities?

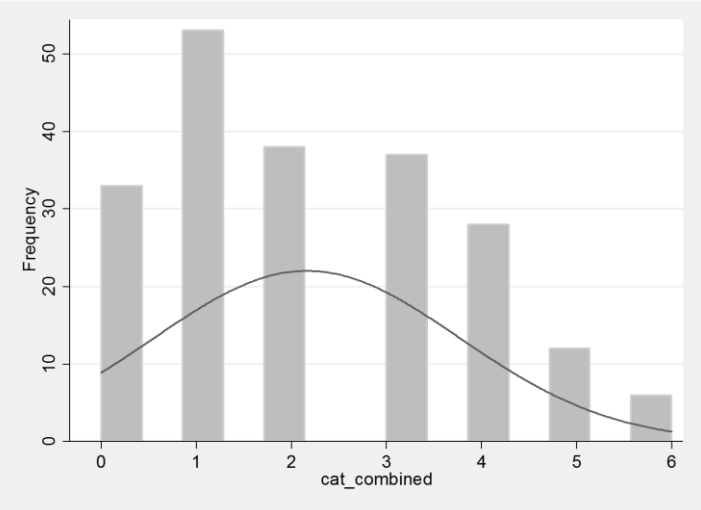
7. How is the condition of patients monitored in critical care (for e.g. SOFA scores /APACHE scores or similar)?

8. What is done if a patient is assessed as needing critical care but cannot be admitted due to shortage of beds?

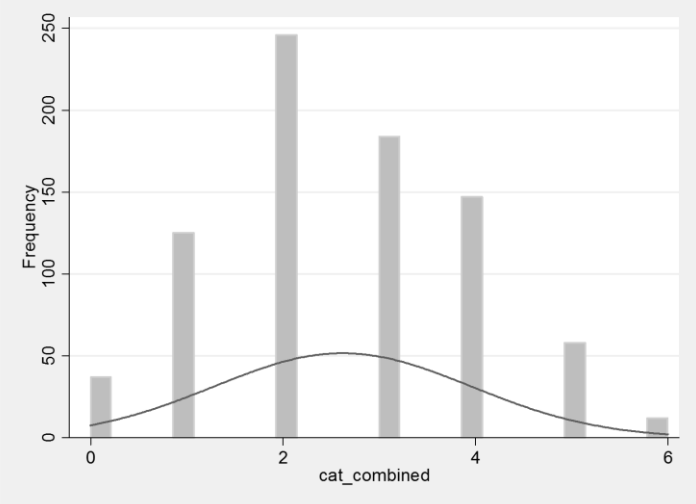
9. Is there an explicit or implicit 'maximum acceptable length of stay' in critical care facilities (for e.g. if a patient has been admitted to critical care for a month but no improvement is seen)? Is this the same for adults and children?

APPENDIX IV: Distribution of un-weighted CATs scores by patient age group

Distribution of un-weighted combined CATs scores for adults

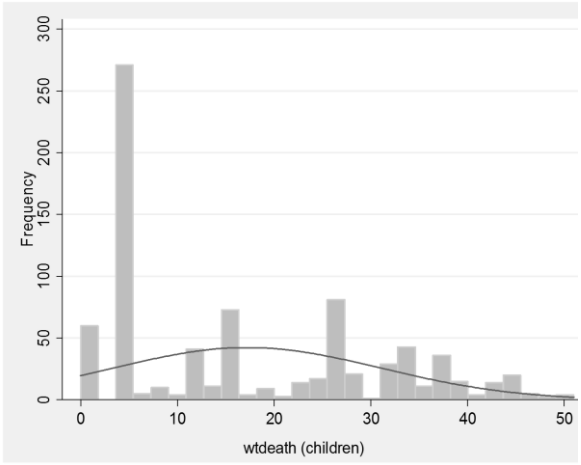


Distribution of un-weighted combined CATs scores for children

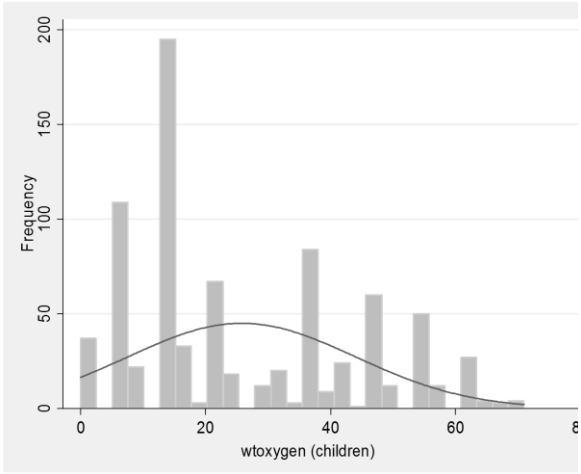


APPENDIX V: Distribution of weighted CATs combined scores in children

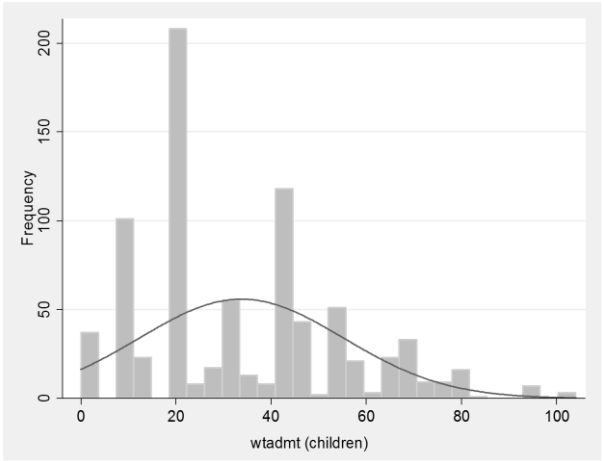
1. In-hospital mortality [Median (IQR) = 15 (5-29)]



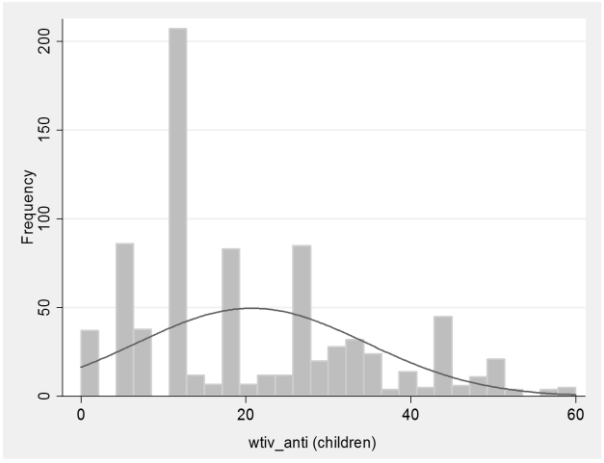
2. Supplemental oxygen [Median (IQR) = 21 (13-40)]



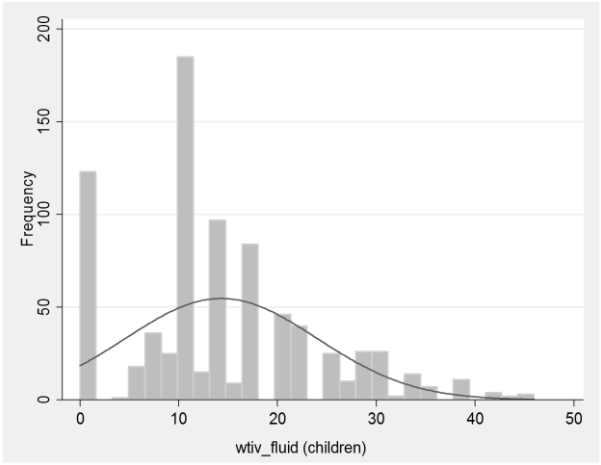
3. Admission [Median (IQR) = 32 (21-46)]



4. IV antibiotic [Median (IQR) = 18 (12-31)]

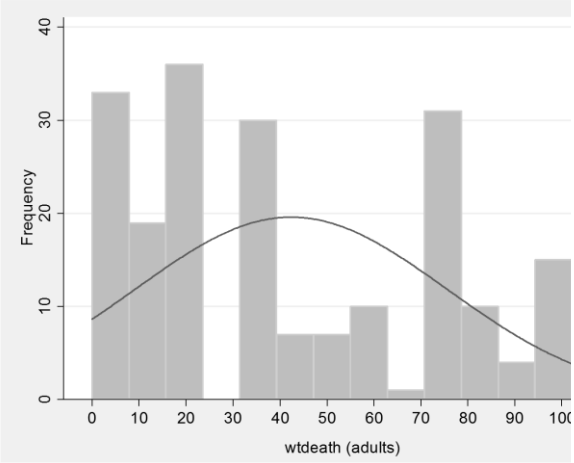


5. IV fluid [Median (IQR) = 14(9-21)]

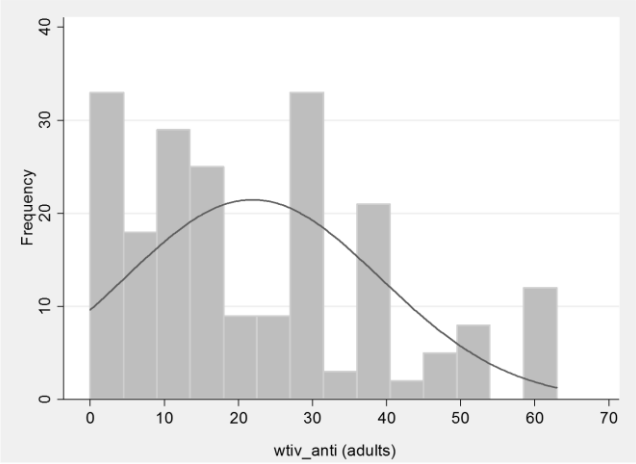


APPENDIX VI: Distribution of weighted CATs combined scores in adults

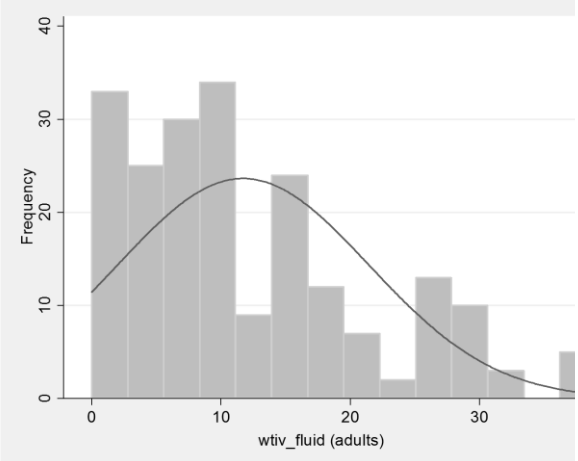
1. In-hospital mortality [Median (IQR) = 33 (15-75)]



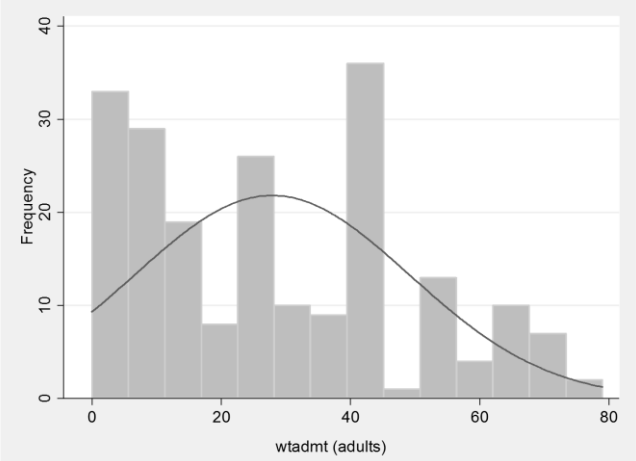
3. Admission [Median (IQR)] =23 (10-44)



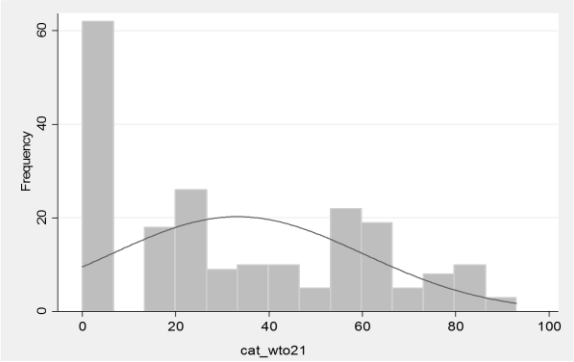
4. IV fluid [Median (IQR) = 10 (4-17)



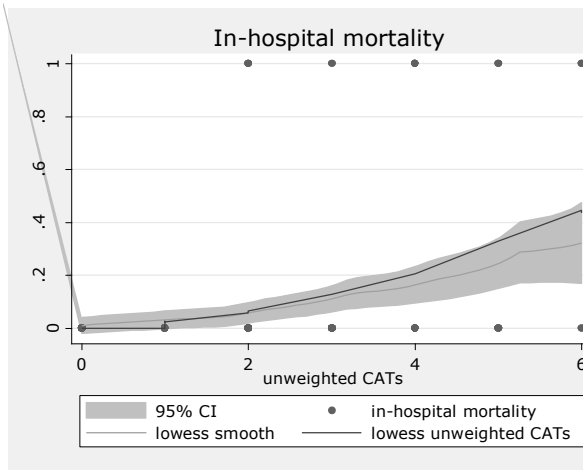
5. IV antibiotic [Median (IQR)] = 16 (10-31)

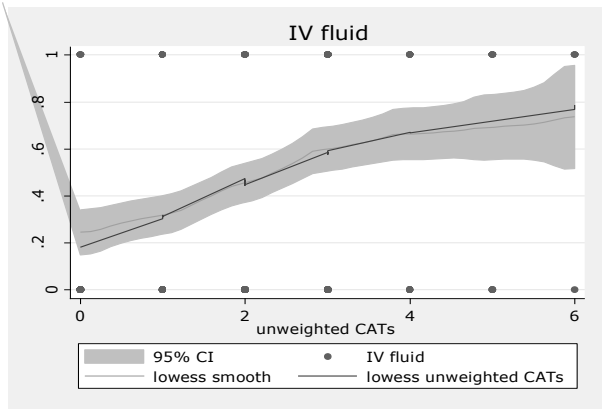


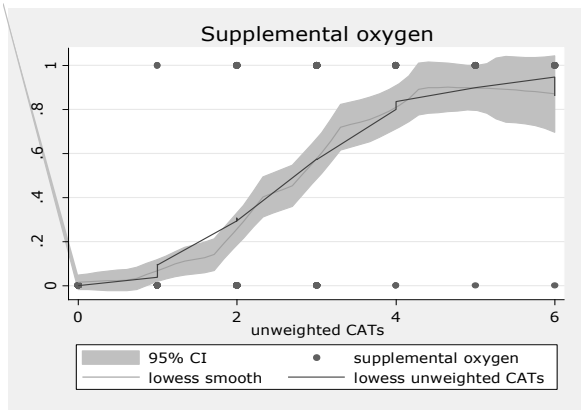
2. Supplemental oxygen [Median (IQR) = 24 (6-54)]

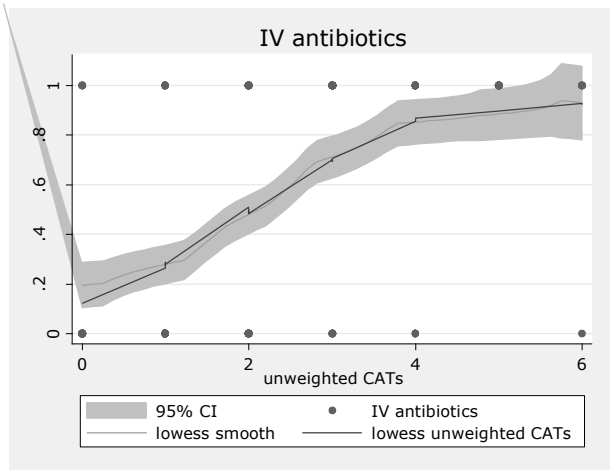


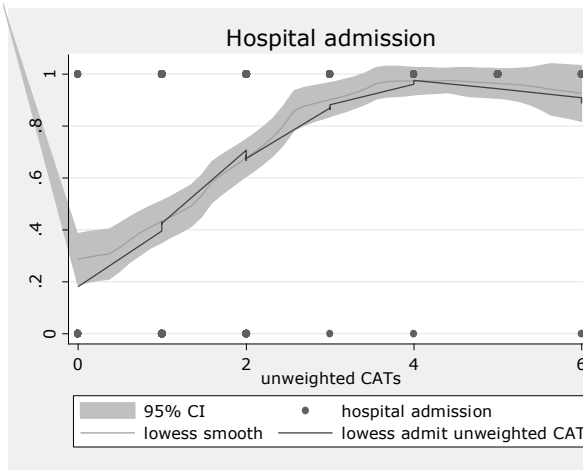
APPENDIX VII: Lowess plots showing linearity of logit assumption for un-weighted CATs scores in adults



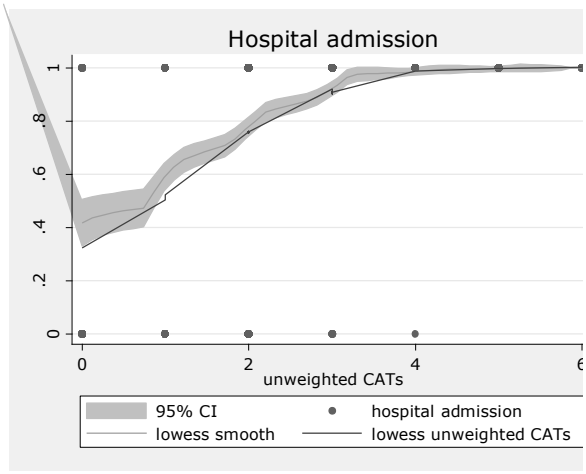


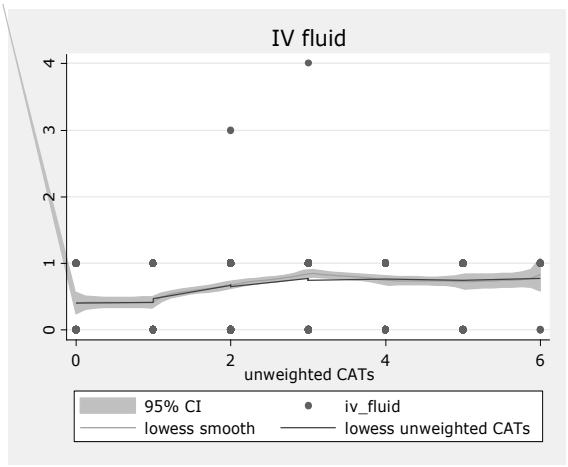


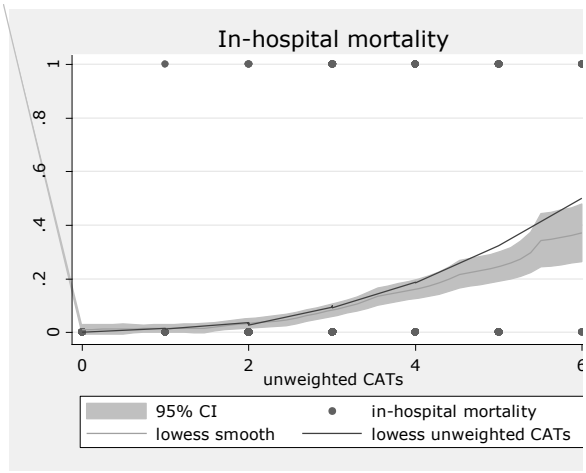


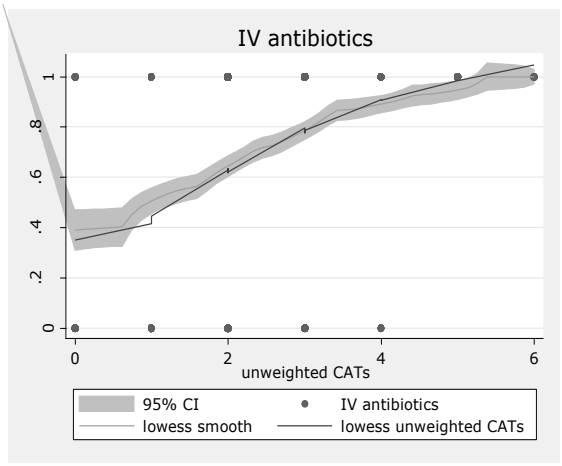


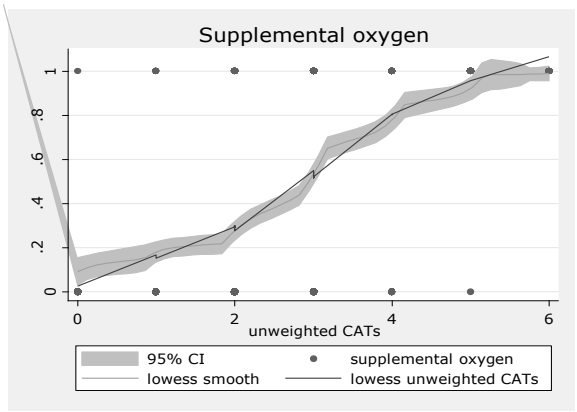
APPENDIX VIII: Lowess plots showing linearity of logit assumption for un-weighted CATs scores in children



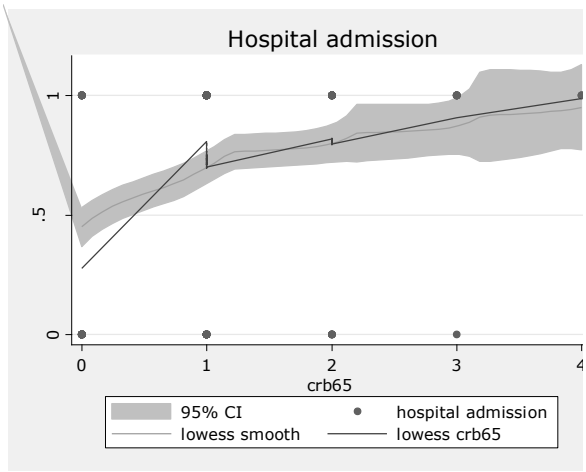


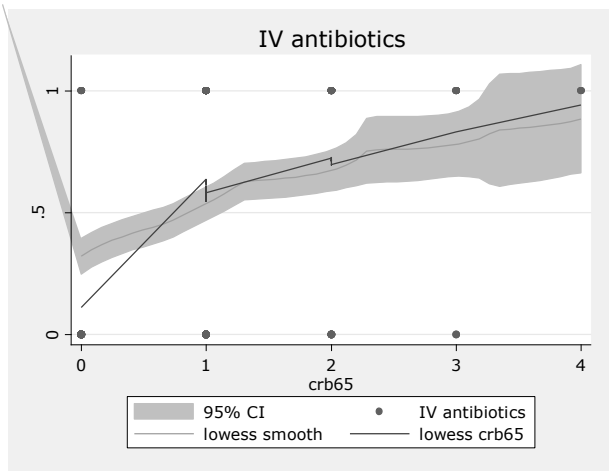


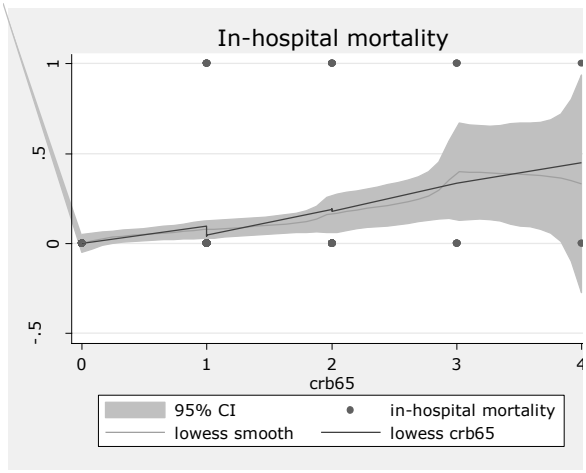


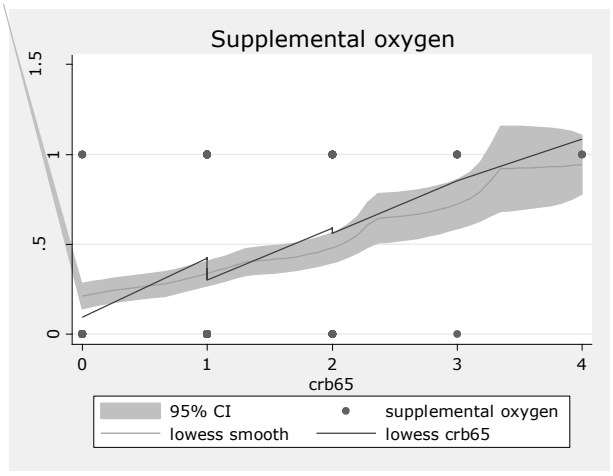


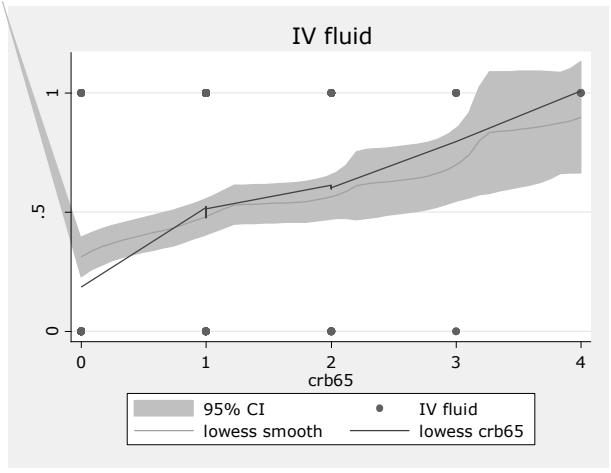
APPENDIX IX: Lowess plots assessing the linearity of logit assumption for CRB-65 scores in adults with ALRTI





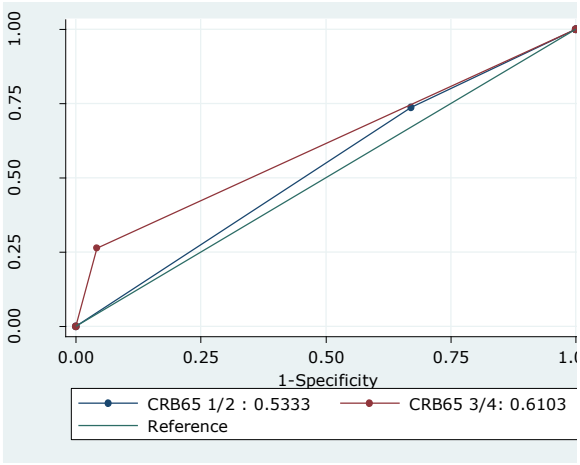




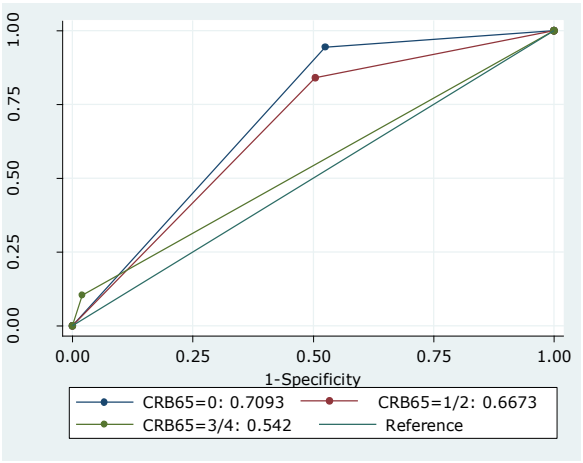


APPENDIX X: Comparison of ROC curves for CRB65 score thresholds in adults with ALRTI

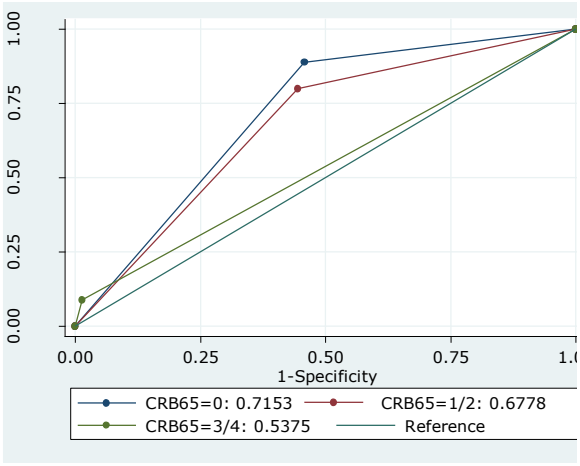
In-hospital mortality



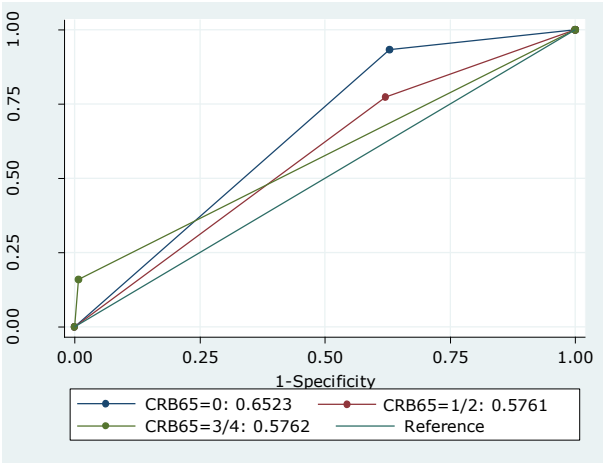
IV antibiotics



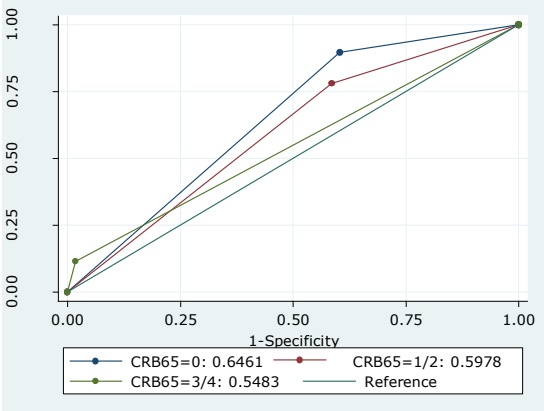
Admission



Supplemental oxygen



IV fluid



CRB65=0	Blue line
CRB65=1/2	Red line
CRB65=3/4	Light green line

APPENDIX XI: Data extracts from interview transcripts

Themes	Sub-themes	Data extracts
Definition and structure of critical care	Care involving organ support	**"**CC is an increased level of care provided by doctors, nurses, health workers to patients who need some form of organ support" (SSR-LUTH) **"**critical care, is the care given to patients who require higher organ support, such as cardiac and respiratory , renal support, more than two organ support" (AJR-LASUTH) "it is basically taking care of critically ill patients who need multi organ support" (ASR-LASUTH) "is the care given to a patient who is critically ill and needs organ support" (SNO-LUTH) "means intensive care provided in specialized units for patients that are critically ill or have one or more organ system failure, and is essentially the same for adults and children" (SJR-LUTH) "CC means the care of patients requiring organ support in the ICU" (SSR-LASUTH)
	Care involving close monitoring	"it equally means intensive care, where patients are being cared for or given a high level of care in terms of patient care giver proportion whereby ,there is a nurse to one patient, the monitors ..the level of monitoring is more than what is found in an ordinary ward" (AJR-LUTH) "we try to sort out patients that are critically ill , which means haemodynamically unstable patients or with severe organ dysfunction and would benefit from closer monitoring than other patients, based on that recommendation we can get them into critical care" (MJR-LASUTH)

		"what i think it is, is the care of the sick patient that are not able to express needs or that require close monitoring basically" (SR-RGH) "Well on the general ward the support or care usually is about 2 nurses to 24 patients while the HDU and ICU operate on a 1:1 (a nurse for a patient)" (SSR-LUTH).
Critical care delivery : structure and services (Levels, capacity, cost)	Critical care delivery in tertiary care hospitals	"in the theatre we have a recovery room which performs the functions of an HDU and we have a 4 bedded ICU" (SSR-LASUTH) ", no but we try: What we have is an ICU, we don't have an HDU, but, we are building a new ICU :and the old ICU which has four beds will be used as the HDU" (AJR-LASUTH) "In LASUTH we do not have different levels of CC; we have just the intensive care unit. we have a 4 bedded ICU, but even though it is 4 bedded we have rapid turnovers (ASR-LASUTH) "Yes there are different levels, even in the main emergency room we have some patients we keep in the resuscitation room where they have dedicated oxygen administration with hourly monitoring of their vitals. They also have access to suctioning and are typically kept closer than other patients to the nurses' station. Then there is the ICU unit where patients require MV and will have a dedicated anaesthetist on call and also specially trained ICU nurses in addition to the primary team handling that patient." (MJR-LASUTH) "We do have an ICU there are six beds and LUTh is a 700 bed hospital and we have a two bed HDU; well on the general

		ward the support or care usually are about 2 nurses to 24 patients while the HDU and ICU operate on a 1:1 (a nurse for a patient). also in the ICU you have anaesthetist but not in the HDU , where the anaesthetists are invited if needed" (SSR-LUTH) "we have the ICU which is `shared by all departments in the hospital, however the surgical specialties have access to a portion of the recovery room which functions as a HDU, it is just beside the ICU but not within the ICU and sometimes patients who are critically ill but have not met the criteria for being in the ICU might stay there and may benefit from some level of intensive care from anaesthetists, and nurses. Sometimes they might have access to ventilators. if they recover they might be taken to the open ward but if they deteriorate then they are taken into the ICU" (SJR-LUTH) "We are trying to create some kind of HDU right beside the recovery room, where if the patient is doing well we transfer to the ICU. In the HDU here we keep post-op patients or patients who need close monitoring. If the patient is not doing well then we move the patient to the ICU. when patients recover we take them to the wards" (SNO-LUTH) "in LUTH presently we have HDU, that is the high dependency unit, so that one the level is a little bit lower than that of the ICU; well in the ICU , we have five beds, but it is not limited to adults alone, beds are allocated as the patients come, so it's first come first serve" (ASR-LUTH)

		"the levels are the ICU, the HDU or the recovery room, and the operating room/theatre" (AJR-LUTH) "the practice in LUTH has been modified in such a way that will be appropriate for the level of skill and resources available to us and in terms of the different levels you would expect internationally , we clearly do not fall within that ambience yet but we have been able to modify as much as possible the practice of CC to suit our local limitations in terms of manpower skills and available resources o there is an ICU and a recovery room / post resuscitation room which serves as HDU " (CO-LUTH)
	Critical care delivery in Secondary care hospitals	"unfortunately we do not have an ICU in this hospital so the best thing we can do here is that we stabilise any critically ill child and refer to a tertiary institution for further management" (SMO-MASSEY) "Presently there is no formal ICU care here, although we do get critically ill patients that are meant for ICU care" I am aware of the different levels of critical care, I believe we offer critical care at the first level here. if we have a patient who needs a chest tube we refer but if it is respiratory failure secondary to an infection or metabolic changes we can manage such a patient here" (MO-RGH) "no we don't have an ICU, .we refer such patients to LUTH, LASUTH and general hospital Lagos" (SR-RGH) "we have an oxygen concentrator and tanks, suction but when you talk about ICU care you are referring to nurse patient ratio and doctor patient ratio which we can't provide" (SR-RGH).

Factors affecting admission to critical care	Decision making	"The head of the team on call headed by the consultant followed by the senior registrar, take the decision on who is to be admitted".(AJR-LUTH) **"**a lot is based on clinical judgement, because in LUTh admission into the ICU is ultimately a decision made by the head of the managing unit and the anaesthesiologist" (SJR-LUTH) **"**The decision is made by the consultant anaesthetist" (SNO-LUTH) "it's a joint decision between the consultant anaesthetist and the consultant of the managing team" (SSR-LUTH) "Both the managing team and the anaesthetist. so the managing team decides who needs critical care based on assessment and the anaesthetist also review prior to accepting the patients" (MJR-LASUTH) "the ICU has a protocol for admitting patients, the person requesting ICU admission will get in touch with the anaesthetist who will then review and ascertain whether or not the patient will benefit from ICU admission/care" (SSR-LASUTH)
	Use of triage criteria	"We have an initial triage of patients who come into the ER using the emergency severity index which is an internationally accepted triage rule which classifies patients based on their clinical presentation into levels 1-5 and this helps us immediately identify the patients we need to give priority attention. in the adult emergency once we are through the triage stage and realize the patient falls into the ESR1 and 2 which we consider as a critically ill

		patient who needs resuscitation" (CO-LUTH) "For now at least from my experience as I just recently joined the unit. There is really no guideline used, the head of the team on call headed by the consultant followed by the senior registrar, take the decision on who is to be admitted. However where there are classical cases that qualify for admission, for example a young fit adult in a RTA (that is age is considered), then the GCS, who has sustained severe traumatic brain injury; such cases will not require a consultant to make such a decision. but the guideline is not defined" (AJR-LUTH) ", I don't know whether, that rule does exist for patients , but presently with experience and time we just happen to know those that are more urgent or can improve on icu admission; That notwithstanding anyone that we intubate outside the ICU area, depending on the case most of them have a reduced GCS some of them may need to be supported with ventilators, we usually admit such straight into the ICU" (ASR-LUTH) "A lot is based on clinical judgement, because in LUTh admission into the ICU is ultimately a decision made by the head of the managing unit and the anaesthesiologist. But other than the GCS there isn't really a scoring system used by the department of anaesthesia" (SJR-LUTH) "yes we do, you know we have intensivists here "Well Dr Odom, who is the intensivist in the ICU, he has a protocol for admission, but it is not like we have to look at this all the time we already know the protocol, so once a

		patient is not doing well, especially on SP02, GCS <9, (ASR-LASUTH)
	Factors for admission	"Basically finance is a huge factor. it cost about 250 thousand naira to admit a patient for a week in the ICU and they pay per day after a week" (SSR-LUTH); "Prognosis of the diagnosis of the patient to be admitted, if the benefit to the patient is not justifiable compared to the high level of cost then the patient will rather be managed on the ward".(SSR-LUTH) "there are certain things we consider, we look at patient's age, the illness if it is possible the patient will recover, we do a quick triage" (SNO-LUTH) "We only have 6 beds here and one is for isolation. Most times we have the issue of lack of bed space and we try to use our HDU from time to time" (SNO-LUTH). "Well, one thing is that at times we have more people trying to come into the ICU than the number of bed spaces available, so what we usually do is first of all see cases that are reversible, so not terminal stage cases. So the younger people that we know can benefit from the ICU care are given preference before the patients with terminal illnesses or with bad prognosis. Then again if there is going to be any operation that is going on and we know that such surgery might need post-op ICU care we usually reserve a bed for such cases." (ASR-LUTH) "Yes definitely, you find patients admitted into the ICU where the patient is not able to afford the fee, also bed space is a factor, here we have a five bed

		ICU, and if the beds are filled that becomes an issue as well" (AJR-LUTH) "well two tings clearly determine this: 1 is the availability of bed space in the ICU , to the best of my knowledge the last time i checked , the ICU had about 8/10 bed spaces available, and if this number is situated against a back drop of a seven hundred bed hospital and tertiary centre in a city with such a large population ranging between 4 to 9 million, this is grossly inadequate .the second which sadly is a major determinant is the ability of the patient to afford the admission that is the cost implication"(CO-LUTH) **"**The ICU has a protocol for admitting patients, the person requesting ICU admission will get in touch with the anaesthetist who will then review and ascertain whether or not the patient will benefit from ICU admission/care. often times the review will be that the patient will indeed benefit from ICU care but there maybe or may not be bed space" (SSR-LASUTH) "Then sometimes the ICU is full so we can't cope or admit new patients. also sometimes patients cannot even afford ICU fees" (AJR-LASUTH) "Okay, well firstly, the commonest factor is availability of bed space. there are limited beds in the ICU" so if there is no space available it becomes difficult to admit patients, secondly the financial factor as most of the patients provide health care cost out of pocket and paying for ICU admission is more expensive than being on the wards".(MJR-LASUTH) "Yes you see finance and bed space are indeed major factors in LASUTH, as i said

		we have only just 4 beds. And we may have 6 people requiring admission into the ICU at the same time so we need to triage and find out who will benefit from ICU and who will not. Not all the patients who want to come into the ICU will benefit from the admission so we have to triage. Also if they do not have enough money how do you manage them? This is a general hospital and patients buy all they need for care, it's out of pocket payment they pay for everything from syringes to drugs. So if they cannot procure all that is required it will be impossible to manage such patients. So the monetary aspect is very important it is a major factor for admission into ICU in LASUTH, not even LASUTH alone; everywhere any ICU" (ASR-LASUTH).
Factors affecting discharge	Length of stay	"Like i said no, even right now we have a patient that has been in the ICU for 127 days so that patient can stay and there is no maximum length of stay, it comes down to what the patient or the family can afford" (AJR-LUTH) "We have to ensure the patient is fit to be discharged before we release them to the ward; we even had a case of a patient who was coming in and out of the ICU for six months. This patient stayed for 3 months was discharged to the ward, and in less than 3 days she was brought back. so you can't determine the length of stay.(ASR-LASUTH) "The patient can stay as long as they can afford the bills or as long as is needed for care" (MJR-LASUTH) "We have patients staying longer than 4/5 months." (SNO-LUTH) ", i have seen patients stay from as little as a few days and others up to 120 days.

		But the decision to move a patient out rests on the admitting consultant, the anaesthetists and the patient's family" (SJR-LUTH)
	Factors affecting discharge	"there are three possible exits, one is if the patient improves and gets better based on established physiological parameters i.e. Oxygen saturation. the second determinant is that patient deteriorates and dies, and the third is if the relatives are not able to afford the continued care in the ICU and may need to be taken off the MV for care elsewhere" (CO-LUTH) "obvious clinical improvement, or in the case where the bed is needed for a more critical patient, we can step down the intensity of care of one patient to allow for another patient" (SJR-LUTH) "improvement in the condition of the patient, and if the level of care can be stepped down" (SSR-LUTH)
Critical care Patient profile and pathway	Patient profile	**"**The spectrum includes both medical and surgical patients… the common denominator for these patients will be that their cardiopulmonary function cannot be sustained unsupported " (CO-LUTH) **"**Well mostly patients with head injury, post-op patients are common too, well respiratory tract infection wow as a cause of ICU admission, well once in a while but it's not that common. We do have burns patients, tetanus mainly from paediatric patients. we rarely see adult tetanus cases" (AJR-LUTH) " well there is no preference for them, we only admit them when they have gotten to the stage of respiratory failure" (ASR-LUTH)

		"We see a variety of cases, and majority of our patients are neurology cases, who have haemorrhagic stroke, tumours, we also see medicine patients who are in coma, who have renal problems, we see sepsis, ARDS: most of our pneumonia cases do not need ICU admissions, most of them are managed on the wards with oxygen, but those that develop ARDS or actually require ventilatory support who aren't doing well on oxygen via face mask or nasal cannula, they have to come to the ICU. an HDU would have been appropriate but since we don't have that we use the ICU"(AJR-LASUTH) "Yes we do get ALRTI patients, but not often. in a situation where we have a more critical case such as post op patient from an open heart surgery, the surgical patient takes precedence over the ALRTI "(ASR-LASUTH) "well do see occasionally , but not too common we do see cases of diphtheria" (SNO-LUTH) "Head injury patients, post-op especially those who had extensive surgery, and critically ill pre-op patients who need stabilization before surgery. but not a lot of paediatric patients get into the ICU except burns victims" (SSR-LUTH) "if the ALRTI patient is deemed to require ventilatory support, yes; but we don't not get them often" (SSR-LUTH) "All kinds of patients needing organ support, most of the time we get more surgical patients than medical patients especially neuro surgical cardio thoracic and post-op. occasionally we see general surgery patients in renal failure. sometimes we take paediatric patients to

		the ICU especially patients with tracheoesophageal fistula" (SSR-LASUTH) "Yes i am aware patients with ALRTI may be in need of respiratory support, but the ICU here is pressed for space so i am not sure that it will admit such patients. ideally every patient who needs admission should go to the ICU , however they will be admitted based on priority because of space constraints" (SSR-LASUTH)
	Route of admission	"Most are post-op, but we still get people from outside the hospital for ICU care. they could be post -op or medically ill patients" (ASR-LASUTH) "Patients come in either from the ward, theatre or the ER, for patients from the ER these are patients with severe head injury people who cannot maintain their airway, people with respiratory failure and burns or patients who had a stroke and are unconscious. From the ward, patients who have deteriorated are taken into the ICU" (SJR-LUTH)
	Use of ICU severity scores	"Well we just started using the APACHE score, like i said we have an intensivist that takes care of all these. Most of the time we weigh the risk and benefit of the patient on the ventilator on a daily basis, if they are improving or deteriorating, in view of weaning them off the ventilator. but now we have begun using the APACHE2" (ASR-LASUTH) "i am aware of the APACHE but i am not sure if it is followed strictly in the ICU" (MJR-LASUTH) "We usually look at the indication for admission, that's the number one thing then we look at the organs involved, whether it is the central nervous system

		which is involve then we use the Glasgow coma scale..” (ASR-LUTH) “OK. we are familiar with the APACHE score, so we use that” (AJR-LASUTH) “Yes i am aware of the APACHE, but i am not sure as most of my patients don't go to the ICU simply because there is no neonatal ICU as i am a paediatric surgeon`. but i believe they do severity risk assessment in the ICU” (SSR-LASUTH) “mainly the GCS is used, for trauma patients” (SSR-LUTH)
	Route of discharge	**“**most of our patients stay for long periods, because we do not have the HDU unit to keep them in when their condition is less critical , so we make sure they are fine and fit in the ICU before we send them to the ward”(ASR-LASUTH) “they go straight to the ward most times” (SSR-LUTH) “if we had a functioning HDU we might be able to step down care for patients in the ICU. But i am afraid most of the patients here go straight from the ICU to the regular wards or from wards to the ICU where there might not be any other form of monitoring but i am aware that some wards now have electronic monitors” (SSR-LASUTH)
	Outcome of patients not admitted	“well what then happens is that the continued care of the patient is undertaken on the ward or the designated space in the ED as the HDU, but after a while we need to move these patients and most end up on the wards or they continue staying in the ED” (CO-LUTH) “well we have a privately run ICU in the hospital premises , if the patient is able

		to afford the cost of the private ICU , then we advise them to use that service" (ASR-LASUTH) "Ok we explain the situation and give them the option of going elsewhere i mean there are private ICUs but they are quite expensive and might not have as much staff as compared to the ICUs in the government hospitals. we might also sometimes manage the patient in the ED and try to care for them as closely as possible until there is an available bed in the ICU" (MJR-LASUTH) "it depends on the kind of patient it is , if it is a surgical case, requiring post up ICU care, we actually keep them in the recovery room, or use our ventilators "..actually in the theatre, that's the best we can do when the bed spaces are full, but they are usually managed on the wards and as soon as we get a free bed then we can quickly bring the patient in.(AJR-LASUTH) "most times we carry out conservative management and intubate the patients wherever they are" (SNO-LUTH) "there might be a joint review with the anaesthetist to determine if the patient currently in the ICU can manage in the HDU or the open ward so that someone else might be able to move into that space. Other times we can postpone surgeries in the event of an elective decision to admit in the ICU. At other times we refer to other tertiary centres and on a few occasions we send the patients to LASUTH" (SJR-LUTH). "If this was to happen, we have a 4 bedded section in the A&E for patients who require some form of intensive care, known as the resuscitation room, there

		are monitors and we also provide endotracheal intubation, but there are no nurses here, monitoring is done by the doctors, usually there is only one doctor and the anaesthetist come to review" (SSR-LUTH) "well we do give them the option of referral but most of these patients are managed conservatively and i am afraid I cannot say what the outcome of such cases will be" (SSR-LASUTH)

APPENDIX XII: Conference presentations based on this thesis

Oral Presentations

- Institute of Lung Health Joint Respiratory Research day (Leicester-2016): comparative validation of CATs vs. CRB65 for severity assessment of ALRTI in adults: an external validation.
- School of medicine divisional seminar-EPH (Nottingham-2015): Validation and comparison of CATs for prognosis of ALRTI.

Poster presentation

- Midlands Academy of Medical Sciences Research Festival (Leicester-2016): comparative validation of CATs vs. CRB65 for severity assessment of ALRTI in adults: an external validation.

LAGOS STATE UNIVERSITY TEACHING HOSPITAL, IKEJA

1-5, OBA AKINJOBI ROAD, IKEJA, LAGOS. P.M.B. 21005, TEL: 01-4710670

www.lasuth.org. E-mail: dsct@lasuth.org

DIRECTORATE OF CLINICAL SERVICES AND TRAINING

HEALTH RESEARCH AND ETHICS COMMITTEE
REG.NO. NHREC 04/04/2008
(www.nhrec.net)

PROJECT TITLE: A PROSPECTIVE EVALUATION OF THE USE OF COMMUNITY ASSESSMENT TOOLS (CATs) IN PREDICTING THE NEED FOR ACUTE CARE INTERVENTIONS, CRITICAL CARE AND MORTALITY IN PATIENTS PRESENTING WITH ACUTE LOWER RESPIRATORY TRACT INFECTIONS IN A NIGERIAN UNIVERSITY HOSPITAL: A VALIDATION STUDY.

REF. NO.: LREC/10/06/354

PRINCIPAL INVESTIGATOR: DR. RUBY E ARMON

ADDRESS: DEPT. OF EPIDEMIOLOGY AND PUBLIC HEALTH, UNIVERSITY OF NOTTINGHAM, NOTTINGHAM, UNITED KINGDOM.

DATE OF RECEIPT OF VALID APPLICATION: 02/09/13

DATE OF APPROVAL: 10/09/2013

This is to inform you that the study described here in the submitted protocol, the consent forms, advertisements and other participant information materials have been reviewed and given full approval by the Health Research and Ethics Committee of LASUTH. (LREC)

This approval dates from 10/09/2013 to 10/04/2014. If there is any delay in starting the Research, Please inform the LREC so that the dates of approval can be adjusted accordingly. Note that no participant accrual or activity related to this research may be conducted outside of these dates. All informed consent forms used in this study must carry the LREC assigned number and duration of LREC approval. In a multiyear research, endeavor to submit your annual report to the LREC early in order to obtain renewal of your approval and avoid disruption of your research.

THE NATIONAL CODE FOR HEALTH RESEARCH AND ETHICS(www.nhrec.net) REQUIRES YOU TO COMPLY WITH ALL INSTITUTIONAL GUIDELINES, RULES AND REGULATIONS AND WITH THE TENETS OF THE CODE INCLUDING ENSURING THAT ALL ADVERSE EVENTS ARE REPORTED PROMPTLY TO THE LREC. NO CHANGES ARE PERMITTED IN THE RESEARCH WITHOUT PRIOR APPROVAL BY LREC EXCEPT IN CIRCUMSTANCES OUTLINED IN THE CODE. THE LREC RESERVES THE RIGHT TO CONDUCT COMPLIANCE VISIT TO YOUR RESEARCH SITE WITHOUT PREVIOUS NOTIFICATION.

DR. (MRS.) F.O AJOSE
CHAIRMAN
HREC-LASUTH

DR. A. ADEDOKUN
MD, FMCGP
CMAC/Director of Clinical Services and Training
08086748003, 08033277108

PROF. D. A.A. OKE
B.Sc. (Hon.) MBBS, FMCP
Chief Medical Director
0802313752, 08058995250
+2347032959145

N

N RUBY
nent of Epidemiology & Public Health,
iversity of Nottingham,
of Pharmacist,
Kingdom.

PERMISSION TO CARRY OUT RESEARCH

have been directed to convey the approval of the Permanent Secretary, Lagos State Service Commission for the above named student to carry out her research titled "A **ctive evaluation of the use of Community Assessment Tools *(CATs) in predicting d for acute care interventions, critical care and mortality in patients presenting with wer respiratory tract infections in Nigerian hospitals: a validation study."**

py of your findings should be forwarded to the Health Service Commission.

Dr. (Mrs.) A. M. Onayiga
For: Permanent Secretary

dical Director,
General Hospital,
d Hospital, Yaba,
Hospital, Lagos,
Street Children's Hospital.

H

LAGOS UNIVERSITY TEACHING HOSPITAL

HEALTH RESEARCH AND ETHICS COMMITTEE

PRIVATE MAIL BAG 12003, LAGOS, NIGERIA
e-mail address: luthethics@yahoo.com

Chairman
ASSOC. PROF. N. U. OKUBADEJO
MB. ChB, FMCP

Administrative Secretary
MR. D. J. AKPAN
B.Sc. BUS. ADMIN, MIHSAN

Chief Medical Director:
PROF. AKIN. OSIBOGUN
MBBS (Lagos), MPH (Columbia), FMCPH FWACP

Chairman, Medical Advisory Committee
DR. M. O. OGUNLEWE
BDS, FWACS.

LUTH HREC REGISTRATION NUMBER: NHREC: 19/12/2008a
Office Address: Room 107, 1st floor, LUTH Administrative Block
Telephone: 234-1-5850737, 5852187, 5852209, 5852158, 5852111

2nd September, 2013

NOTICE OF EXPEDITED REVIEW AND APPROVAL

PROJECT TITLE: "A PROSPECTIVE EVALUATION OF THE USE OF COMMUNITY ASSESSMENT TOOLS (CATs) IN PREDICTING THE NEED FOR ACUTE CARE INTERVENTIONS, CRITICAL CARE AND MORTALITY IN PATIENTS PRESENTING WITH ACUTE LOWER RESPIRATORY TRACT INFECTIONS IN A NIGERIAN UNIVERSITY HOSPITAL: A VALIDATION STUDY".
HEALTH RESEARCH COMMITTEE ASSIGNED NO.: ADM/DCST/HREC/1023
NAME OF PRINCIPAL INVESTIGATOR: DR. RUBY E. ARMON
ADDRESS OF PRINCIPAL INVESTIGATOR: DEPT. OF EPIDEMIOLOGY AND PUBLIC HEALTH, UNIVERSITY OF NOTTINGHAM, UNITED KINGDOM.
DATE OF RECEIPT OF VALID APPLICATION: 23-05-13

This is to inform you that the research described in the submitted protocol, the consent forms, and all other related materials where relevant have been reviewed and given full approval by the Lagos University Teaching Hospital Health Research Ethics Committee (LUTHHREC).

This approval dates from 02-09-2013 to 02-09-2014. If there is delay in starting the research, please inform the HREC so that the dates of approval can be adjusted accordingly. Note that no participant accrual or activity related to this research may be conducted outside of this dates. All informed consent forms used in this study must carry the HREC assigned number and duration of HREC approval of the study. In multiyear research, endeavor to submit your annual report to the HREC early in order to obtain renewal of your approval and avoid disruption of your research.

The National code for Health Research Ethics requires you to comply with all institutional guidelines, rules and regulations and with the tenets of the code including ensuring that all adverse events are reported promptly to the HREC. No changes are permitted in the research without prior approval by the HREC except in circumstances outlined in the code. The HREC reserves the right to conduct compliance visits to your research site without previous notification.

DR. N. U. OKUBADEJO
CHAIRMAN, UTH HEALTH RESEARCH ETHICS COMMITTEE