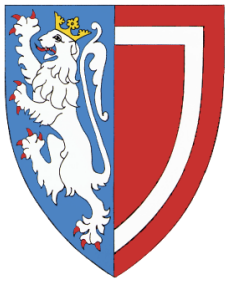
EARLY IDENTIFICATION AND PREDICTION OF MULTIPLE ORGAN FAILURE FOLLOWING MAJOR TRAUMA



LYNN HUTCHINGS

BALLIOL COLLEGE

UNIVERSITY OF OXFORD

SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN MUSCULOSKELETAL SCIENCES

MICHAELMAS TERM 2013

ABSTRACT

INTRODUCTION

Trauma is the main cause of death in working-age adults in the UK. Multiple organ failure (MOF) is associated with a high proportion of late trauma deaths, and MOF survivors have poor long-term outcomes. Early prediction of patients at risk of MOF would assist treatment decisions and allow targeted interventions.

METHODS

A cohort of major trauma patients requiring intensive care unit (ICU) treatment at the John Radcliffe Hospital was identified. Data were obtained from the two national databases of the Trauma Audit Research Network and the Intensive Care National Audit and Research Centre, and from a local ICU database with hourly data recording.

Literature review and questionnaire analysis of trauma clinicians identified candidate predictors of MOF, grouped into patient, injury, physiological, laboratory and management variables. MOF scoring systems were reviewed to determine the most appropriate for use in trauma patients.

Prediction models of post-trauma MOF were developed using logistic regression at a range of times from 0 to 48 hours after injury. Models were internally validated using bootstrapping.

RESULTS

517 adult trauma patients were identified from 2003-2011. Overall mortality was 14.9%, with 491 patients surviving more than 48 hours, and therefore being at risk of MOF development. For these 491 patients, MOF incidence depended on the definition, and ranged from 23% (Denver score) to 58% (SOFA score). MOF was associated with mortality, time to ICU admission, and length of ICU and hospital stay. MOF could be predicted with an accuracy of up to 81.3% at 2 hours post-injury, and 84.2% at 12 hours post-injury using small numbers of clinical variables. Age, head injury, abdominal injury, maximum heart rate and the need for vasopressors were strong predictors of all definitions of MOF.

CONCLUSIONS

Post-trauma MOF can be predicted early after injury using combinations of clinical variables. Further validation of the identified variables on external populations would allow development of a clinical score to assist clinicians in trauma management.

ACKNOWLEDGEMENTS

My thanks extend far beyond those I can mention on this page.

First and foremost I should like to thank my supervisors. Extensive thanks go to Keith Willett, for believing in me in the first place, allowing me the great freedom of creating a study from scratch, and supporting me with advice and funds to get the project up and running. Your support throughout my career to date cannot be emphasised enough, and I am extremely grateful for all you have done for me. To Duncan Young, for helping me to formulate and develop the idea until it became a viable research project, for the pieces of paper that magically transformed into coffee, and for the endless sessions of advice and support that extended far beyond what was required from you. To Peter Watkinson, for assisting me through the numerous sticky patches, for the helpful statistical support, and for the sound advice on developing a clinical research career.

A huge amount of thanks must also go to the NIHR Oxford Biomedical Research Centre for their generous funding, and their support of my wider development as a clinical researcher. In particular, I would like to thank Denise Best for always being on hand for advice and help.

The staff of the Kadoorie Centre made it an excellent place to work, study, and learn, and the endless supply of cake definitely enhanced my productivity. I would like to thank both the trauma and ICU sides for including me in all your activities, and being so supportive of a project that spanned both parts of the department.

I would also like to thank the staff of NDORMS, and specifically Afsie Sabokbar for being endlessly supportive - from before I even applied to the time that I submitted. For the project logistics, extensive thanks are due to TARN and ICNARC for their support from the outset, and for allowing me to use their datasets. In addition, grateful thanks are due to Severn Deanery, especially those responsible for training in T&O, for allowing me time out of programme to pursue my research interests. Thanks to the T&O staff of Frenchay, particularly Tim Chesser and Mike Kelly, for putting up with my distractions as I pieced the project idea together, and the T&O department of GWH, particularly Mike Rigby, for turning me back into a clinician, and absorbing the stresses of submission.

Outside of the office, the beauty of rowing on the River Thames in the early morning provided the best possible opportunity for clearing my head, and testing my body rather than my mind. To the clubs of OUWBC, OUWLRC and BCWBC, I'll be forever grateful for the time I spent playing at being an athlete, and for the friends I made in the process. Specific mentions to Kasia, for the care pathway and the frequent sanity checks, and to Jill, for being the twin I never knew I had. And finally, thanks I can't put into words to Ross, for supporting me throughout this venture, and for endlessly bringing supplies to my desk.

ABSTRACT I

ACKNOWLEDGEMENTS III

LIST OF FIGURES X

LIST OF TABLES XII

LIST OF ABBREVIATIONS XV

CHAPTER 1 BACKGROUND 1

1.1. Trauma 1
1.1.1. Trauma: The Global Problem 1
1.1.2. Causes of Traumatic Injury 2
1.1.3. Impact of Trauma 3
1.1.4. Trauma in the UK 4

1.2. Outcomes After Major Trauma 7
1.2.1. Trauma Deaths and the Trimodal Distribution 7
1.2.2. Non-Fatal Outcomes 10

1.3. Defining And Quantifying Major Trauma 11
1.3.1. Major Trauma Definitions 11
1.3.2. Anatomical Scores 12
1.3.3. Physiological Scores 14
1.3.4. Combined Anatomical and Physiological Scores 16

1.4. Multiple Organ Failure 18
1.4.1. Defining Multiple Organ Failure 18
1.4.2. Post-Trauma MOF 19
1.4.3. Outcomes Following Post-Trauma MOF 21
1.4.4. Aetiology and Pathophysiology of Post-Trauma MOF 21
1.4.5. Variables Associated with Post-Trauma MOF 28
1.4.6. Prediction of Post-Trauma MOF 31

1.5. The Research Project 33
1.5.1. Creating a Predictive Model for Post-Trauma MOF 33
1.5.2. The Research Question 34
1.5.3. Aims and Objectives 34
1.5.4. Outline of Methods 35
1.5.5. Structure of the Thesis 36

CHAPTER 2 METHODOLOGY OVERVIEW 38

2.1. Outline of Methods 38

2.2. Input Variable Selection 42
2.2.1. Literature Review 42
2.2.2. Questionnaire Analysis 42
2.2.3. Creating the Input Variable List 42

2.3. Outcome Measure Selection 43
2.3.1. Primary Outcome – MOF 43
2.3.2. Secondary Outcomes 44

2.4. Data Sources 44
2.4.1. TARN Database 45
2.4.2. ICNARC Database 46
2.4.3. CareVue 46
2.4.4. Additional Data Sources 47

2.5. Data Considerations 47
2.5.1. Data Permissions 47
2.5.2. Data Availability 49

2.6. Study Population 49
2.6.1. The John Radcliffe Hospital Trauma Population 49
2.6.2. Inclusion Criteria 51
2.6.3. Exclusion Criteria 51
2.6.4. Population Size 52

2.7. Data Linkage 53

2.8. Data Extraction 54

2.9. Data Cleaning and Analysis 55
2.9.1. Data Cleaning 55
2.9.2. Timing of Data 56
2.9.3. Statistical Analysis 57

CHAPTER 3 INPUT VARIABLES 60

3.1. Selecting Candidate Input Variables 60

3.2. Literature Review 61
3.2.1. Introduction 61
3.2.2. Methods 62
3.2.3. Results 65
3.2.4. Discussion of Review Findings 71
3.2.5. Candidate Input Variables from Literature Review 73

3.3. Questionnaire Analysis 75

3.3.1. Methods 75
3.3.2. Results 75
3.3.3. Candidate Input Variables - Questionnaire Analysis 81

3.4. Candidate Input Variable List 84

CHAPTER 4 MULTIPLE ORGAN FAILURE 88

4.1. Defining Multiple Organ Failure 88

4.2. Review of MOF Scoring Systems 89
4.2.1. Review Methods 89
4.2.2. Results of Review 91

4.3. Discussion of MOF Scoring Systems 99
4.3.1. Overview 99
4.3.2. Developing the Ideal MOF Score 101
4.3.3. Use of Scoring Systems 106
4.3.4. Choosing a MOF Scoring System for Trauma Research 108
4.3.5. MOF – Study Definition 113

CHAPTER 5 DATA LINKAGE 115

5.1. Data Linkage Overview 115

5.2. Identifying Patients From TARN 116
5.2.1. Accessing TARN Data 116
5.2.2. Initial TARN Query 116
5.2.3. Patient Identifiers 116

5.3. Identifying Patients From ICNARC 118
5.3.1. Recreating the ICNARC Database 118
5.3.2. The ICNARC Coding Method 121
5.3.3. Initial ICNARC Query 122
5.3.4. Patient Identifiers 124

5.4. Identifying Patients From ICU Admin 124
5.4.1. Accessing ICU Admin 124
5.4.2. Identifying Trauma Patients 124
5.4.3. Patient Identifiers 125

5.5. Process of Data Linkage 125
5.5.1. Overall Process Summary 125
5.5.2. Matching TARN, ICNARC and ICU Admin 126
5.5.3. Matching ICNARC and ICU Admin 127
5.5.4. Resolving Patients From TARN Query 129
5.5.5. Resolving Patients from ICU Queries 130
5.5.6. Comparing Results From TARN and ICU Matching 134
5.5.7. Matching to CareVue 135

5.5.8. Finalising the Study Population 136

CHAPTER 6 DATA EXTRACTION AND ANALYSIS 138

6.1. Overview of Data Extraction 138
6.1.1. Extracting Appropriate Data 138
6.1.2. Method of Data Extraction 138
6.1.3. Data Transformation and Cleaning 139

6.2. Data Extraction From CareVue 139
6.2.1. Overall CareVue Structure 139
6.2.2. CIS Structure 140
6.2.3. CIS Extraction Process 149
6.2.4. ISM Structure 150
6.2.5. ISM Extraction Process 157

6.3. Data Extraction From TARN 158

6.4. Data Extraction From ICNARC 159

6.5. Data Extraction From Other Sources 160

6.6. Data Transformation and Cleaning 160
6.6.1. Anonymising Data 160
6.6.2. Data Dictionary 160
6.6.3. Data Cleaning 161
6.6.4. Missing Data 161

6.7. Data Analysis 162
6.7.1. Overview of Analysis 162
6.7.2. Variable Selection for Multivariate Analysis 162
6.7.3. Model-Building Strategy 163
6.7.4. Internal Validation 165

6.8. Summary of Methods 167

CHAPTER 7 RESULTS 169

7.1. Trauma Population Description 169
7.1.1. Patient Demographics 169
7.1.2. Injury Characteristics 169
7.1.3. Survival 171
7.1.4. Timings 172
7.1.5. Interventions 175

7.2. Population at Risk of MOF 176
7.2.1. Population Numbers 176
7.2.2. Incidence of MOF 177
7.2.3. MOF and Mortality 178

7.2.4. Onset and Duration of MOF 181
7.2.5. Organ System Dysfunction 184
7.2.6. Comparison of MOF and No MOF Groups 191

7.3. Univariate Analysis of MOF Predictors 195
7.3.1. Candidate Input Variables 195
7.3.2. Results of Univariate Analysis 198

7.4. Variables for Multivariate Analysis 198

7.5. Multivariate Analysis 199
7.5.1. Model Development 199
7.5.2. Denver MOF Multivariate Models 200
7.5.3. SOFA MOF Multivariate Models 204
7.5.4. Marshall MOF Multivariate Models 208
7.5.5. Denver Combined Model 212
7.5.6. SOFA Combined Model 219
7.5.7. Marshall Combined Model 225
7.5.8. Comparison of Included and Excluded Patients from Combined and Extended Models 231

7.6. Internal Validation 232
7.6.1. Denver Combined and Extended Models 232
7.6.2. SOFA Combined and Extended Models 233
7.6.3. Marshall Combined and Extended Models 235

7.7. Summary of Results 237

CHAPTER 8 DISCUSSION 239

8.1. Prediction of MOF Following Major Trauma 239

8.2. Describing the Current Trauma Population 240

8.3. Defining MOF After Major Trauma 243

8.4. Predicting MOF 245

8.5. Study Limitations 248

8.6. Recommendations for Future Work 253

8.7. Conclusions 256

BIBLIOGRAPHY 258

APPENDICES 276

LIST OF FIGURES

Figure 1.1: Ranking of RTCs in leading global causes of death by age, from Global Burden of Disease: 2004 Update (WHO, 2008) 4
Figure 1.2: Representation of the trimodal distribution of deaths after major trauma 8
Figure 1.3: Immune response phases after trauma 24
Figure 3.1: Flow diagram of study selection (as per PRISMA guidelines) 65
Figure 3.2: Median number of questionnaire responses for each category of input variable 76
Figure 3.3: Questionnaire responses for patient variables 77
Figure 3.4: Questionnaire responses for injury variables 78
Figure 3.5: Questionnaire responses for physiology variables 79
Figure 3.6: Questionnaire responses for laboratory variables 80
Figure 3.7: Questionnaire responses for management variables 81
Figure 4.1: Summary of the different organ systems used in the 42 MOF scoring systems identified 92
Figure 5.1: Diagram of ICNARC database table relationships 121
Figure 5.2: Results of matching of patients from TARN query 130
Figure 5.3: Results of matching ICU queries to local TARN files 132
Figure 5.4: Results of matching of patients from ICU queries 133
Figure 5.5: Summary of final resolution of patient matching across TARN and ICU databases 135
Figure 6.1: Relationships between Patient group of tables in CIS 143
Figure 6.2: Relationships between Observation group of tables in CIS 144
Figure 6.3: Relationships between the Intervention group of tables in CIS 146
Figure 6.4: Relationships between Orders group of tables in CIS 147
Figure 6.5: Linkage between Orders and Intervention Groups in CIS 148
Figure 6.6: Schematic diagram of relationships between main data tables in CIS 149
Figure 6.7: Relationships between Chart group of tables in ISM 153
Figure 6.8: Relationships between Med group of tables in ISM 154
Figure 6.9: Relationships between IO group of tables in ISM 156
Figure 7.1: Age distribution for a) males and b) females. 169
Figure 7.2: Kaplan-Meier survival curve for trauma population 171
Figure 7.3: Distribution of deaths by post-injury day 172
Figure 7.4: Box and whisker plot of time to arrival at the John Radcliffe for direct admissions and transfers. 173
Figure 7.5: Length of AICU stay in days 174
Figure 7.6: Day of hospital discharge for trauma survivors, given as post-injury day. 175
Figure 7.7: Total number of units of blood products used for patients requiring transfusion within 24 hours of injury 176
Figure 7.8: Flow diagram of patient numbers for population at risk of MOF 177
Figure 7.9: Classification of 289 patients developing MOF according to scoring system 178
Figure 7.10: Kaplan-Meier survival curves categorised into MOF and no MOF groups according to the Denver score 179

Figure 7.11: Kaplan-Meier survival curves categorised into MOF and no MOF groups according to the SOFA score 180
Figure 7.12: Kaplan-Meier survival curves categorised into MOF and no MOF groups according to the Marshall score 181
Figure 7.13 (a-c): Duration of MOF for patients defined with MOF for each scoring system. 184
Figure 7.14: Organ dysfunction by individual systems for Denver, SOFA and Marshall scores 185
Figure 7.15: Denver combined test model output 213
Figure 7.16: ROC curves for Denver combined test model 215
Figure 7.17: Denver extended test model output 217
Figure 7.18: ROC curves for Denver extended test model at 0 – 12 hours, 0 – 24 hours and 0 – 48 hours post-injury 218
Figure 7.19: SOFA combined test model output 220
Figure 7.20: ROC curves for SOFA combined test model 221
Figure 7.21: SOFA extended test model output 223
Figure 7.22: ROC curves for SOFA extended test model 224
Figure 7.23: Model output for Marshall combined test model 226
Figure 7.24: ROC curves for Marshall combined test model 227
Figure 7.25: Model output for Marshall extended test model 229
Figure 7.26: ROC curves for Marshall extended test mode 230
Figure 7.27: Bootstrapping results for Denver combined test model 232
Figure 7.28: Bootstrapping results for Denver extended model 233
Figure 7.29: Bootstrapping results for SOFA combined test model 234
Figure 7.30: Bootstrapping results for SOFA extended model 234
Figure 7.31: Bootstrapping results for Marshall combined test model 235
Figure 7.32: Bootstrapping results for Marshall extended model 236

LIST OF TABLES

Table 1.1: Systemic Inflammatory Response Syndrome (SIRS) Criteria..... 23
Table 2.1: Availability of main data sources 49
Table 3.1: Input variable categories..... 61
Table 3.2: Inclusion/exclusion criteria for literature review..... 63
Table 3.3: Candidate input variables identified from literature review 74
Table 3.4: Candidate input variables identified from questionnaire analysis 83
Table 3.5: Candidate input variables final list..... 86
Table 4.1: Summary of organ systems used in 42 MOF scoring systems (Papers in chronological order)..... 94
Table 4.2: Measures of respiratory and cardiovascular failure 96
Table 4.3: Measures of renal and hepatic failure 97
Table 4.4: Measures of haematological and GI dysfunction..... 98
Table 4.5: Measures of CNS failure..... 99
Table 4.6: Characteristics of the ideal descriptor of organ dysfunction..... 103
Table 4.7: The Marshall Multiple Organ Dysfunction Syndrome Score..... 104
Table 4.8: The Sequential Organ Failure Assessment Score 106
Table 4.9: The Denver Postinjury Multiple Organ Failure Score..... 109
Table 4.10: Comparison of the components of the Denver, SOFA and Marshall MOF scoring systems..... 112
Table 5.1: Variables collected in the Opening Section of the Core Dataset of TARN 117
Table 5.2: Summary of ICNARC data files 119
Table 5.3: Data periods for matching 126
Table 5.4: Results of matching initial queries from TARN, ICNARC and ICU Admin 127
Table 5.5: Results of matching trauma patients across ICNARC and ICU Admin..... 128
Table 7.1: Injury severity details for trauma population 170
Table 7.2: Injury details for trauma population..... 170
Table 7.3: Timing of onset of MOF..... 182
Table 7.4: Duration of MOF (unadjusted for death)..... 182
Table 7.5: Breakdown of organ dysfunction by combinations of organ systems affected for Denver, SOFA and Marshall scores. 188
Table 7.6 (a-c): Duration of organ dysfunction by Denver, SOFA and Marshall scores (unadjusted for death) 189
Table 7.7 (a-c): Onset of organ dysfunction by Denver, SOFA and Marshall scores. 191
Table 7.8: Population comparison of Denver MOF and no MOF groups..... 192
Table 7.9: Population comparison of SOFA MOF and no MOF groups..... 193
Table 7.10: Population comparison of Marshall MOF and no MOF groups. 194
Table 7.11: Categorisation of variables for univariate analysis – Patient, Injury and Physiological variables 196
Table 7.12: Categorisation of variables for univariate analysis – Laboratory and Management variables 197

Table 7.13: Model selection for Denver MOF models 200
Table 7.14: Goodness-of-fit summary measures for selected Denver MOF models 200
Table 7.15: Category predictors for Denver MOF models 201
Table 7.16: AUC and 95% CI for ROC curves for Denver MOF models 201
Table 7.17: Denver 0 – 2 hours – model output 201
Table 7.18: Denver 0 – 4 hours – model output 201
Table 7.19: Denver 0 – 6 hours – model output 202
Table 7.20: Denver 0 – 12 hours – model output 202
Table 7.21: Denver 0 – 24 hours – model output 203
Table 7.22: Denver 0 – 48 hours 203
Table 7.23: Model selection for SOFA MOF models 204
Table 7.24: Goodness-of-fit summary measures for selected SOFA MOF models 204
Table 7.25: Category predictors for SOFA MOF models 204
Table 7.26: AUC and 95% CI for ROC curves for SOFA MOF models 205
Table 7.27: SOFA 0 – 2 hours – model output 205
Table 7.28: SOFA 0 – 4 hours – model output 205
Table 7.29: SOFA 0 – 6 hours – model output 206
Table 7.30: SOFA 0 – 12 hours model output 206
Table 7.31: SOFA 0 – 24 hours model output 207
Table 7.32: SOFA 0 – 48 hours model output 207
Table 7.33: Model selection for Marshall MOF models 208
Table 7.34: Goodness-of-fit summary measures for selected Marshall MOF models 208
Table 7.35: Category predictors for Marshall MOF models 208
Table 7.36: AUC and 95% CI for ROC curves for Marshall MOF models 209
Table 7.37: Marshall 0 – 2 hours – model output 209
Table 7.38: Marshall 0 – 4 hours – model output 209
Table 7.39: Marshall 0 – 6 hours – model output 210
Table 7.40: Marshall 0 – 12 hours – model output 210
Table 7.41: Marshall 0 – 24 hours - model output 211
Table 7.42: Marshall 0 – 48 hours – model output 211
Table 7.43: Goodness-of-fit measures for Denver combined test model 213
Table 7.44: Category predictors for Denver combined test model 214
Table 7.45: AUC for ROC curves for Denver combined test model 216
Table 7.46: Goodness-of-fit measures for Denver extended test model 217
Table 7.47: Category predictors for Denver extended test model 217
Table 7.48: AUC for ROC curves for Denver extended test model 218
Table 7.49: Goodness-of-fit measures for SOFA combined test model 220
Table 7.50: Category predictors for SOFA combined test model 220
Table 7.51: AUC for ROC curves for SOFA combined test model 222
Table 7.52: Goodness-of-fit measures for SOFA extended test model 223
Table 7.53: Category predictors for SOFA extended test model 223
Table 7.54: AUC for ROC curves for SOFA extended test model 224
Table 7.55: Goodness-of-fit for Marshall combined test model 226
Table 7.56: Category predictors for Marshall combined test model 226

Table 7.57: AUC for ROC curves for Marshall combined test model 228
Table 7.58: Goodness-of-fit for Marshall extended test model 229
Table 7.59: Category predictors for Marshall extended test model 229
Table 7.60: AUC for ROC curves for Marshall extended test model 230

LIST OF ABBREVIATIONS

ACCP	American College of Chest Physicians
ACoTS	Acute Coagulopathy of Trauma Shock
ACP	Augmented Care Period
ACS	Acute Coronary Syndrome
AICU	Adult Intensive Care Unit
AIS	Abbreviated Injury Scale
ALI	Acute Lung Injury
ALP	Alkaline phosphatase
ALT	Alanine transaminase
AP	Anatomic Profile
APACHE	Acute Physiology and Chronic Health Evaluation
APC	Activated Protein C
APS	Acute Physiological Score
APTT	Activated Partial Thromboplastin Time
ARDS	Adult Respiratory Distress Syndrome
ASCOT	A Severity Classification of Trauma
AST	Aspartate transaminase
AUC	Area under the curve
BD	Base Deficit
BMI	Body Mass Index
CARS	Compensatory Anti-Inflammatory Reaction Syndrome
CASP	Critical Appraisal Skills Programme
CCU	Coronary Care Unit
CI	Cardiac Index
CK	Creatine kinase
CNS	Central Nervous System
CritPal	Critical Care Patient Liaison Association
CRP	C-Reactive Protein
CV	Cardiovascular
CVP	Central venous pressure
DALYs	Disability-Adjusted Life Years
DAMPs	Damage-associated molecular patterns
DCO	Damage Control Orthopaedics
DCS	Damage Control Surgery
Denver TOF	Denver Trauma Organ Failure Score
DIC	Disseminated Intravascular Coagulation
DOB	Date of Birth
ECC	Ethics and Confidentiality Committee
ED	Emergency Department
ENT	Ear, Nose and Throat
ETC	Early Total Care
FDPs	Fibrin Degradation Products
FFP	Fresh Frozen Plasma
GGT	Gamma glutanyl transferase
GI	Gastrointestinal
HR	Heart rate

HRT	Hormone Replacement Therapy
ICAM	Intercellular adhesion molecule
ICD	International Classification of Disease
ICD-9	International Classification of Disease 9th Edition
ICD-10	International Classification of Disease 10th Edition
ICISS	International Classification of Disease Injury Severity Score
ICNARC	Intensive Care National Audit and Research Centre
ICU	Intensive Care Unit
IEDs	Improvised Explosive Devices
IL	Interleukin
IQ	Interquartile
ISS	Injury Severity Score
IV	Intravenous
FiO2	Fraction of inspired oxygen
GCS	Glasgow Coma Scale
GNP	Gross National Product
LDH	Lactate dehydrogenase
LODS	Logistic Organ Dysfunction System
LRTI	Lower Respiratory Tract Infection
MAP	Mean arterial pressure
MODS	Multiple Organ Dysfunction Syndrome
MOF	Multiple Organ Failure
MPM	Mortality Prediction Model
MSOF	Multi-Stage Organ Failure
MTCs	Major Trauma Centres
MTOS	Major Trauma Outcomes Study
NAO	National Audit Office
NCEPOD	National Confidential Enquiry into Peri-Operative Deaths
NHS	National Health Service
NIGB	National Information Governance Board
NISS	New Injury Severity Score
NO	Nitric Oxide
NPV	Negative Predictive Value
NSAIDs	Non-steroidal anti-inflammatory drugs
ONS	Office of National Statistics
OR	Odds Ratio
$PaCO_2$	Partial arterial pressure of carbon dioxide
PaO_2	Partial arterial pressure of oxygen
PAC	Percentage accuracy in classification
PACS	Picture archiving and communication system
PAR	Pressure-adjusted heart rate
PAS	Patient Administration System
PAWP	Pulmonary Artery Wedge Pressure
PbR	Payment by Results
PCT	Procalcitonin
PEEP	Positive End-Expiratory Pressure
PHI	Pre-hospital Index
PICU	Paediatric Intensive Care Unit
PLA	Phospholipase
PMH	Previous medical history
PPV	Positive Predictive Value

PRBCs	Packed red blood cells
PRISMA	Preferred Reporting Items for Systematic Review and Meta-Analyses
Ps	Probability of Survival
PTT	Prothrombin Time
PTS	Physiologic Trauma Score
RCTs	Randomised Controlled Trials
REC	Research Ethics Committee
RISC	Revised Injury Severity Classification
ROC	Receiver Operating Characteristic
RR	Respiratory Rate
RTCs	Road Traffic Collisions
RTS	Revised Trauma Score
SAPS	Simplified Acute Physiology Score
SBP	Systolic blood pressure
SCCM	Society of Critical Care Medicine
SD	Standard deviation
SEM	Standard error of the mean
SGOT	Serum gluatmic oxaloacetic transaminase
SIRS	Systemic Inflammatory Response Syndrome
SOF	Single Organ Failure
SOFA	Sequential Organ Failure Assessment
SQL	Structured Query Language
SRR	Survival Risk Ratio
TARN	Trauma Audit and Research Network
TASH	Trauma-associated severe haemorrhage score
TBI	Traumatic Brain Injury
TNF	Tumour necrosis factor
TS	Trauma Score
T-SQL	Transact - Structured Query Language
WCC	White Cell Count
WHO	World Health Organisation

CHAPTER 1 BACKGROUND

1.1. TRAUMA

1.1.1. TRAUMA: THE GLOBAL PROBLEM

Trauma is a physical wounding of the body causing injury. Trauma can occur by a wide range of mechanisms, including both intentional and accidental causes. It is a massive global disease problem. Nearly 6 million people die from traumatic injuries each year, accounting for 10% of total global mortality.[1] Over 100 million people are affected by trauma annually, with an estimated 50 million people currently living with a significant disability resulting from a traumatic injury.

Of the 6 million global trauma deaths, approximately 4 million result from unintentional injuries, such as road traffic collisions (RTCs) and falls. The remaining 2 million deaths result from intentional causes, including self-inflicted injuries, violence and war. Men are impacted heavily, with trauma accounting for 1 in 8 male deaths, in comparison to 1 in 14 female deaths.[1]

In addition to the death toll, trauma is a major cause of long-term health problems. Over 90 million individuals sustain traumatic injuries severe enough to require medical attention each year, resulting in 190 million lost Disability-Adjusted Life Years (DALYs) or 13% of the total global DALY loss.[2] The overall worldwide burden of disease from trauma is increasing, with rises in both the proportion of global deaths, and the disability burden in survivors.

1.1.2. CAUSES OF TRAUMATIC INJURY

Geographical variation is seen in the causes of traumatic injury. In countries with increased use of firearms, such as the USA and South Africa, higher levels of penetrating trauma are seen. Globally, RTCs are the leading cause of accidental injury, accounting for 1.3 million deaths worldwide.[2] Even in countries with high levels of penetrating trauma, RTCs remain a leading cause of mortality, accounting for 33,687 deaths in the USA in 2010, as compared to 30,470 deaths attributed to firearm injury.[3] In the UK, strict gun control laws make penetrating trauma uncommon, and blunt trauma from RTCs is the predominant cause of injury.

Death rates from RTCs have been stabilising in many high-income countries, but continue to increase globally. RTCs are predicted to become the 5th leading cause of global deaths by 2030, overtaking diarrhoeal diseases, Human Immunodeficiency Virus (HIV), tuberculosis and lung malignancies. By 2030, it is estimated RTCs will be responsible for 2.4 million global fatalities, or 3.6% of all deaths each year. Nearly half of all RTC fatalities can be attributed to "vulnerable road users" – pedestrians, cyclists and motorcyclists. Exact figures on the impact of RTCs remain unclear, since under-reporting is commonplace, and non-fatal injuries are frequently unrecorded.[4] However, the World Health Organisation (WHO) estimated between 20 and 50 million non-fatal injuries result from RTCs each year. Accounting for under-reporting, it is believed that approximately 20 severely injured patients are seen for every fatality.[1]

Military trauma creates different patterns of injury, which have altered over time, depending on the nature of the conflict. Historically, high levels of penetrating trauma have been seen. Recent military engagements have seen an increase in blast injuries from

improvised explosive devices (IEDs), which create combination patterns of blunt and penetrating trauma.[5] Whilst injury causes and patterns may be different in military trauma, it has long been a driver of changes in civilian trauma care. Examples stretch from the widespread use of the Thomas splint for femoral fractures at the turn of the 20th century[6] to the recent use of tourniquets and haemostatic dressings in the field.[7]

1.1.3. IMPACT OF TRAUMA

The impact of trauma is greatest in the young and those of lower socioeconomic status. Analysis of RTC statistics show similar patterns to those from trauma of all-causes, and shows that for those aged 15-29 years of age, RTCs are the leading global cause of death. (Figure 1.1) This has major social and economic consequences, far beyond the cost to health care systems.

The Global Status Report on Road Safety[4] emphasised the disproportionate impact of RTCs on lower socioeconomic classes, even within higher-income countries.[8] Individuals left with physical disabilities resulting from trauma can face social stigma and discrimination, leading to educational, occupational and financial deprivation. Poorer individuals are less likely to have the resources to manage the direct and indirect costs that follow from traumatic injury, with costs for medical care compounded by loss of income.

Financially, trauma places a massive strain on health care services, impacting on bed occupancy, in-hospital and community resources, and demanding input from a multi-disciplinary array of health professionals. The cost of RTCs has been estimated at a global average of 1-3% of Gross National Product (GNP).[4] The financial impact is severe even in

high-income countries, with the estimated lifetime costs of RTCs in the USA greater than $70 billion.[9]

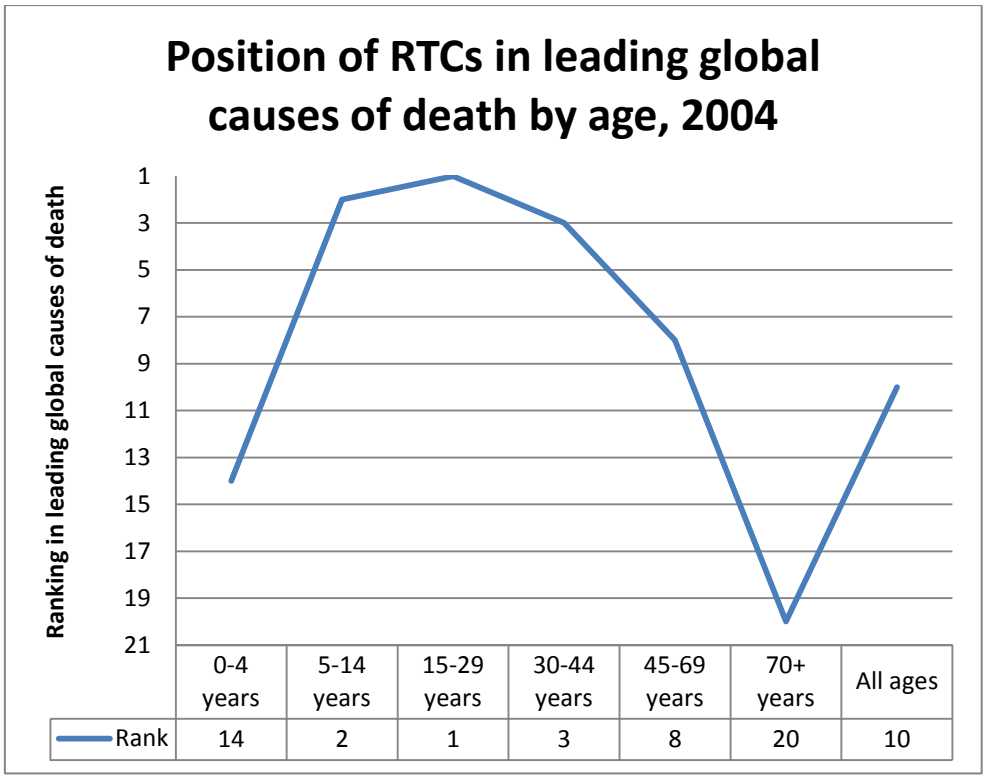


Figure 1.1: Ranking of RTCs in leading global causes of death by age, from Global Burden of Disease: 2004 Update (WHO, 2008)

1.1.4. TRAUMA IN THE UK

In the UK, blunt trauma from RTCs predominates, with penetrating trauma seen rarely outside of a few specific urban areas. Additional common causes of blunt trauma include falls, industrial accidents and sporting injuries. Estimates of trauma deaths vary depending on source. In 2004, the Office for National Statistics (ONS) attributed 16,000 deaths to trauma.[10] Police data reported 3298 deaths from RTCs in 2006,[4] but police statistics have been criticised for underestimating the incidence of major trauma.[11] Sharp decreases in road deaths were seen in the 1970s and 1980s, linked to significant changes

in vehicle design and legislation, but subsequent improvements have been minimal, with some indications that mortality in now increasing.[12]

The annual cost of trauma treatment to the National Health Service (NHS) was estimated in 1999 to be £1.6 billion, or 7% of the total NHS budget,[13] and is likely to have increased further. However, medical costs form only a small part of the total cost of trauma, with human factors related to loss of economic activity and social costs having a far higher impact. An attempt to value the costs of traumatic injury in the mid-1990s estimated medical expenditure to be only 5% of the total societal cost of a trauma death, and 9% of the cost of non-fatal injuries[14].

Care of the trauma patient in hospital frequently involves personnel from a wide range of specialities, including emergency physicians, intensivists, anaesthetists, and an array of surgical specialities. The way in which trauma care in the UK is delivered has changed considerably over the past 30 years.

The traditional model of care in the UK involved patients being taken to their nearest hospital following injury. Patients who required specialist care were then transferred to larger centres at a later date. This differed from the model developed in countries such as the USA where specialised trauma centres were established with concentration of skills and personnel required for trauma care. Patients were admitted to such centres directly, even if other hospitals were geographically closer to the scene of injury.

A seminal study by Anderson in 1988 retrospectively examined 1000 consecutive deaths from injury.[15] Of the patients who reached hospital alive, 20% of the subsequent deaths were deemed preventable. Case-by-case analysis of these deaths indicated that they

were predominantly due to uncontrolled haemorrhage, hypoxia or delays in surgical treatment. The study described *"serious deficiencies in the services for managing severe injury in England and Wales"* and advised consideration of the formation of trauma centres.

Studies in the USA showed increasing evidence of the benefits of trauma centres.[16–18] Direct comparison of two counties with different systems of trauma care in California, showed far better outcomes in the county where patients were taken directly to a single major trauma centre.[16] Fears persisted about the effect of longer transit times to centralised units, despite evidence that centralisation actually reduced delays in treatment.[19]

The Stoke trauma unit experiment in the UK trialled centralisation of services, but the concept was rejected[20] despite controversies over data interpretation.[21,22] Few changes had occurred by the time of the publication of a Royal College of Surgeons report in 2000, which reiterated many of Anderson et al's findings.[23] Data from the Trauma Audit and Research Network (TARN) showed limited change in outcomes following major trauma from 1994-2000.[24]

Evidence from the USA continued to increase on the benefits of centralisation of services.[25,26] In 1994, the John Radcliffe Hospital in Oxford became unique in the UK by developing a system of 24-hour consultant cover provided by trauma specialists, and effectively began to function as a major trauma centre. Whilst this improved local outcomes, nationwide outcomes from trauma remained poor, as further emphasised by the National Confidential Enquiry into Peri-Operative Deaths (NCEPOD) report in 2007, and the National Audit Office (NAO) report into trauma in 2010.[27,28]

Following the appointment of the first National Clinical Director for Trauma in 2009, the move to create major trauma centres and trauma networks accelerated. In April 2012, a series of regional trauma networks went live across England, centred around 26 Major Trauma Centres (MTCs). The John Radcliffe Hospital in Oxford was formally named as an MTC, reflecting the way in which it had been effectively operating as such for many years prior to the official nationwide designations. From data obtained from the John Radcliffe and other units, and extrapolated from trauma centre development in other countries, it has been estimated that an additional 450-600 lives could be saved following this reconfiguration of trauma services.[29]

1.2. Outcomes After Major Trauma

1.2.1. Trauma Deaths and the Trimodal Distribution

The epidemiological work of Baker in the 1970s in San Francisco led to the first description of the trimodal distribution of trauma deaths.[30] Baker studied 437 consecutive trauma deaths, and found that approximately 60% of deaths occurred at scene or in the emergency department (ED), 20% occurred in hospital but within 2 days of injury, and 20% occurred after 2 days. A simplified representation of the trimodal distribution of trauma deaths is given in Figure 1.2.

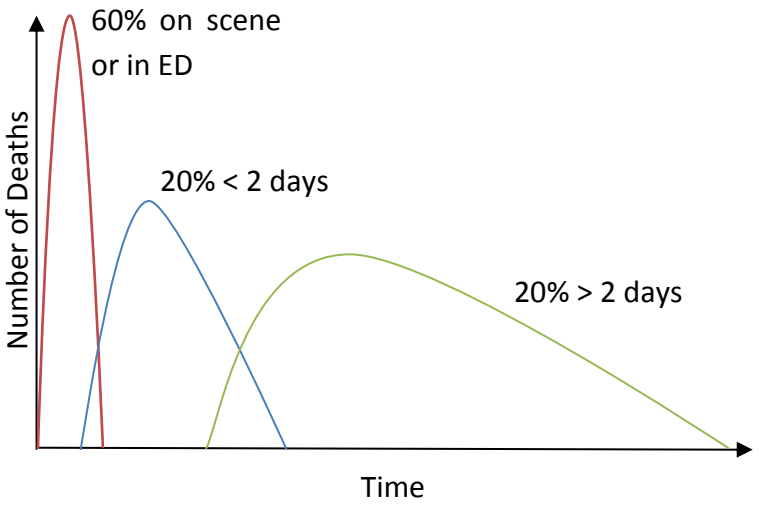


Figure 1.2: Representation of the trimodal distribution of deaths after major trauma

Case analysis showed each of the three peaks resulted from different causes of death. Following the widespread adoption of the trimodal distribution as a model for trauma deaths, different strategies were employed to reduce deaths in each peak. The first peak occurred within the first few hours of injury, and resulted from catastrophic and non-survivable injuries. Focus on reducing deaths in the first peak has come not from medical care, but from accident prevention strategies, including improvements in road and vehicle design, and in health and safety legislation.

The second peak of deaths occurred in the hours immediately following injury, and was largely due to massive neurological damage to the head or spine, or to uncontrolled haemorrhage. The emergency management of trauma at this early stage after injury has been a focus for research into different resuscitation strategies and control of haemorrhage. Examples of this include the move towards permissive hypotension[31,32] and the development of aggressive early massive transfusion protocols.[33,34] The second

peak of deaths is also affected by logistical influences. The *Trauma:Who Cares?* report showed improved outcomes in patients who received early input from senior staff.[27]

The third peak of deaths occurred more than 2 days after injury, and Baker attributed 78% of these to sepsis or multiple organ failure (MOF). Subsequent epidemiological studies have demonstrated widely varying estimates of MOF prevalence, but have consistently shown its importance as the major cause of late post-trauma mortality.[35–38]

Despite its widespread integration into trauma teaching, the relevance of the trimodal distribution has been widely debated. Demetriades et al analysed 4,151 deaths from a Californian trauma registry, and concluded that no evidence of a trimodal distribution was seen.[39] However, when the data on penetrating and blunt trauma were analysed separately, very different distributions were seen, with the proportions of penetrating trauma deaths continually decreasing with elapsed time from injury, but blunt trauma deaths showing small increases at 1-3 and > 7 days.[39] These findings echoed those by Sauaia et al in 1995, where analysis of deaths in Denver, Colorado showed high levels of pre-hospital and ED mortality in penetrating trauma.[40]

Analysis of predominantly blunt trauma deaths in the Netherlands showed no trimodal distribution.[41] A small-scale study in New Zealand found over 80% of deaths occurred at scene, with a subsequent gradual reduction over time.[42] A literature review of 22 publications examining causes of trauma deaths found some evidence of a shift towards a bimodal distribution of early versus late deaths with a decrease in deaths secondary to haemorrhage. However, the authors concluded *"no considerable changes in the… pattern of death were found"*.[43]

1.2.2. NON-FATAL OUTCOMES

Evidence suggests that mortality from trauma has been declining over time. Regel et al found a reduction in trauma-related mortality from 37 to 22% in the two decades from the early 1970s to the early 1990s.[44] A mortality rate of around 20% has been recorded in a number of other studies in the past 15 years.[45] Nast-Kolb et al showed a significant decrease in mortality from blunt trauma from 28.7% in 1975-1980 to 13.9% in 1998-1999, despite increases in mean age, and no change in injury severity.[38]

Mortality has been the major outcome measure in trauma research, but with evidence of declining mortality, other outcomes have been considered. The Trauma Recovery Project represented an early attempt to quantify quality of life and functional outcomes, and found trauma patients suffered from severe limitations at 12 and 18 months post-injury.[46] More recently, an Australian team argued that future gains in trauma survival were likely to be small, and that quality of life and functional outcomes were key. However, many potential problems in quantifying these areas were identified, including the range of trauma measurement instruments available, the importance of both function and quality of life, feasibility of use, extent and timing of follow-up and issues with extrapolation back to pre-injury status.[47] The Victoria State Trauma Registry in Australia has now integrated the collection of data on functional outcomes with clinical and more traditional outcome measures. However, no international consensus exists on which measures of function are deemed important.

1.3. Defining And Quantifying Major Trauma

1.3.1. Major Trauma Definitions

There is no universally applied definition of major trauma.[48,49] In simple terms, major trauma or "polytrauma" involves a massive physical insult, resulting in severe injury to one or more regions of the body. This is often quantified by use of a scoring system to assess trauma severity.

A number of scoring systems have been developed to quantify the severity of trauma. The majority of these are graded scales, with increasing scores representing increasing severity, but often no definitive cut-off is given to denote major trauma.

The most widely used means of denoting trauma severity has been via anatomical scores, particularly the Injury Severity Score (ISS) or the New Injury Severity Score (NISS). Physiological and combined anatomical and physiological scores have also been developed.

Scores to classify trauma severity have been developed for a wide range of uses, including early patient triage, resource allocation, comparison of patient populations, measurement of changes to an individual patient's condition, and evaluation of patient treatment. Most scores have used mortality as the outcome measure. The following provides a brief overview of the main scores for measuring severity and predicting mortality in trauma.

1.3.2. ANATOMICAL SCORES

ABBREVIATED INJURY SCALE

The Abbreviated Injury Scale (AIS) was first developed in 1969 by a joint committee of the American Medical Association, the Society of Automotive Engineers and the Association of the Advancement of Automotive Medicine. It was designed to indicate the severity of injuries from RTCs. The AIS classifies each injury according to its potential "threat to life" on a scale of 1-6, where 1 indicates a minor injury and 6 an unsurvivable one.

The AIS has developed into a widely-used means of anatomical scoring for medical purposes, and as the basis of other scoring systems including the ISS and NISS. The coding manual is regularly revised and updated, and from 1985 penetrating injuries were also included. The AIS manual has increased in complexity to account for all potential injuries and now runs to multiple volumes.[50]

INJURY SEVERITY SCORE

The ISS was first developed in 1974 primarily as a description of blunt trauma resulting from RTCs.[51] The body is divided into six anatomical regions – Head, Face, Chest, Abdomen, Extremities and External. Each individual injury is classified according to the AIS. To calculate the ISS, the highest AIS for each anatomical region is taken and squared. The scores from the three highest-scoring anatomical regions are then added together to give a total from 0 to 75. If any injury scored an AIS of 6, then the ISS is automatically 75. The cut-off value to denote major trauma varies between research studies but is in the range of 15-20.

ISS is in widespread international use. It has been validated, found to be reproducible[52,53] and shown to correlate with mortality,[54] morbidity and hospital stay.[55] It has been validated for use in comparing retrospective population data[54] and for cost analyses.[56] However, it has a number of limitations. These include its lack of suitability as a triage tool since not all injuries are apparent at presentation, the lack of weighting to body region injured, and the formation of the same score from markedly different injury patterns.[55]

NEW INJURY SEVERITY SCORE

The NISS was developed to overcome one of the limitations of the ISS – its inability to account for multiple severe injuries in the same body region. To calculate the NISS, the three highest AIS scores are taken regardless of body region, squared and summed. This again gives a score from 0-75, and an AIS of 6 automatically scores an NISS of 75. This modification was suggested to improve mortality prediction,[57] and has been independently shown to be more accurate than ISS in prediction of in-hospital mortality,[58] hospital stay and intensive care admissions.[59]

ANATOMIC PROFILE

The Anatomic Profile (AP) was an earlier scale developed in response to criticisms of the ISS. It attempted to assign a weighting to injuries from different body regions. The AP consisted of four components, each relating to a specific array of body areas – A (Head, brain and spinal cord), B (Thoracic), C (Front of neck, abdomen, pelvis, spine without cord) and D (Femoral artery, above knee crush or amputation, face and all others). Each component was assigned a weighting to create an overall score linked to the probability of survival.[60] The AP has not been widely used either for research or clinical purposes.

ICD-9 INJURY SEVERITY SCORE

Codes from the International Classification of Disease 9th Edition (ICD-9) were used to create the ICD-9 Injury Severity Score (ICISS). A dataset of over 300,000 patients sustaining over 500,000 traumatic injuries was used to calculate a Survival Risk Ratio (SRR) for each ICD-9 code. The ICISS was then the product of the SRRs for all of the patient's recorded injuries. The original developers demonstrated ICISS had increased predictive power of mortality then ISS with a lower misclassification rate.[61] Subsequent studies showed ICISS to have improved prediction of hospital length of stay in comparison to ISS and other trauma scores.[62] An Australian study studying the impact of using International Classification of Disease 10th Edition (ICD-10) rather than ICD-9 codes concluded that the performance of ICISS is affected by the data and methods used to develop the SRRs, the ICD version used and the number and type of covariates considered.[63]

1.3.3. PHYSIOLOGICAL SCORES

REVISED TRAUMA SCORE

Champion et al first proposed a physiology-based trauma severity score called the Triage Index in 1980.[64] This was later updated to the Trauma Score (TS)[65] and the Revised Trauma Score (RTS).[66]

The original Trauma Score was based on measurement of respiratory rate (RR), respiratory effort, systolic blood pressure (SBP), capillary refill and the Glasgow Coma Scale (GCS). The RTS simplified this to RR, SBP and groupings of GCS values representing mild, moderate and severe head injury. Data was taken from the first available dataset following injury. Values of RR, SBP and GCS were given designated scores based on

deviation from normal and inputted into a formula, giving a single output ranging from 0 to 7.8408. The GCS component was heavily weighted to compensate for the poor accuracy of the TS for serious head injuries. The RTS has been used to correlate with probability of survival (Ps) and as a triage tool for determining the need for trauma centre treatment.[67]

PREHOSPITAL INDEX

The Prehospital Index (PHI) was developed in 1986 for triage use. Computerised analysis of data from over 300 patients was used to determine variables easily obtained in the prehospital setting that correlated with mortality within 72 hours of injury or the need for emergency surgery within 24 hours of injury. SBP, heart rate, RR and level of consciousness were found to be the best predictors, and each was given a score from 0-5. From the total of 20, 0-3 was classified as minor trauma and 4-20 as major trauma. The original validation study found a misclassification rate of 5.4% in relation to prediction of mortality.[68]

24 HOUR ICU POINT SYSTEM

Vassar et al developed a scoring system to be used in the first 24 hours of ICU admission.[69] PaO2/FiO2 ratio, GCS and fluid balance were scored to total 0-12 points, and were the combination of variables most predictive of mortality.

PHYSIOLOGIC TRAUMA SCORE

The Physiologic Trauma Score (PTS) was developed to predict mortality by measurement of variables that could be collected on admission to hospital The score combined the

Systemic Inflammatory Response Syndrome (SIRS) criteria, GCS and age[70] and was shown to have improved predictive ability to anatomic scores.[71]

1.3.4. COMBINED ANATOMICAL AND PHYSIOLOGICAL SCORES

TRISS

The TRISS methodology was designed to improve the predictive power for mortality by combining anatomical and physiological variables. TRISS is an acronym created from the component parts that include Trauma Score, ISS and age.[72] The TRISS methodology allows the estimation of Ps for an individual patient based on coefficients for each required variable. These coefficients were devised from Walker-Duncan regression analysis on the data from over 80,000 patients enrolled in the Major Trauma Outcomes Study (MTOS) of the early 1980s in the USA.[73]

The TRISS methodology has been widely used and accepted, and also widely criticised. Its use on a sample of 5445 patients in California showed a misclassification rate for death of 4.3%. However, misclassification rates over 20% were seen in older and more severely injured patients, patients with polytrauma, or those who needed ICU admission.[74] TRISS-like models have been created for a number of trauma subsets, such as penetrating trauma,[72] intubated patients[75] and children.[76] Modifications of the anatomical component to NISS, AP or ICISS have been proposed to improve prediction.[77,78] TRISS has been criticised for not being sufficiently accurate for use at an individual patient level.[79]

A SEVERITY CHARACTERISATION OF TRAUMA (ASCOT)

ASCOT was developed by the same team as TRISS, as a means to update measurement of its components. ASCOT was developed on the MTOS data, and was developed to have

different weightings for blunt and penetrating trauma.[80] On prospective validation by the original developers, neither TRISS nor ASCOT met the criteria for model calibration for penetrating trauma and sensitivity in blunt trauma was low.[81] An independent research team found high misclassification rates for both, and concluded that the small gain in accuracy with ASCOT was offset by its increased complexity.[82]

CRAMS

CRAMS is an early scoring system developed as a triage tool for determining need for transfer to a major trauma centre. The name is based on the constituent measures of circulation, respiration, abdominal injury, motor capacity and speech. Scores range from 0-10 with a score of 8 or less indicating major trauma.[83] The composition of the original test population has been criticised and high misclassification rates were found in other populations.[84]

TARN MODEL

TARN, the UK-wide body collecting national clinical audit data on trauma patients, adopted the TRISS model for mortality prediction in 1989. However, problems were noted with the high levels of missing data for RR, and the high number of excluded patients. In 2006, a new TARN model involving modification of the TRISS methodology was proposed for mortality prediction. After initial development on blunt trauma patients, a number of models were trialled on trauma population subsets, with the final model including age, a transformation of ISS, sex and a sex by age interaction. This was shown to be appropriate for more patient groups, result in less missing data and have improved predictive power compared to TRISS.[85]

REVISED INJURY SEVERITY CLASSIFICATION

The Revised Injury Severity Classification (RISC) was developed from data on over 2000 patients from the Trauma Registry of the German Society for Trauma Surgery. Variables included are age, NISS, head injury, extremity injury, GCS, prothrombin time (PT), base deficit (BD), signs of bleeding and early cardiac arrest. These are differentially weighted to create a probability of survival. The RISC score showed improved specificity compared to other trauma scores.[86]

MGAP

MGAP was designed to be used for prehospital triage of trauma patients. The name amalgamates the constituent parts of mechanism, GCS, age and SBP. Patients are given points based on the severity of each measurement, and divided into three groups – low risk (23-29 points), intermediate (18-22 points) and high risk (< 18 points), which were related to increasing mortality. MGAP was shown to be more specific and accurate for triage than the RTS.[87] Criticism of the scoring of the mechanism component has led to the development of the GAP score, based on the remaining variables.[88]

1.4. MULTIPLE ORGAN FAILURE

1.4.1. DEFINING MULTIPLE ORGAN FAILURE

Late deaths after trauma often occur due to failure of multiple organ systems. The term "multiple organ failure" was first used by Eiseman in 1977,[89] although descriptions of sequential organ failure were present prior to this date. At a basic level, MOF results from a progressive and systematic failure of body systems to meet physiological demands.

Failing organ systems require medical support to maintain physiological parameters. The nature of this support depends on the nature of the failing system.

The terminology of MOF has been complex and confusing, with significant overlap to terms also used to describe inflammation and infection. This has led to a lack of clarity as to the exact definition of MOF.[90,91] A large number of scoring systems have been proposed which attempt to quantify the extent of organ failure, largely based on physiological parameters or laboratory measurements. In general terms, most of these consist of individual definitions of failure for a number of different organ systems, which are combined to give an overall score. However, these scoring systems differ in the measures used, the definitions of what constitutes a failing organ, and even in the body systems that can be considered to have failed. There is no gold-standard definition of MOF, and many different scoring systems remain in widespread use. The wide array of MOF scoring systems will be discussed in detail in Chapter 4.

1.4.2. POST-TRAUMA MOF

Issues in the definition of MOF have complicated accurate assessment of its incidence following trauma. The reported incidence has ranged from less than 10% to over 40%, depending both on the definition of MOF used, and the population studied.[38,92,93]

Ciesla et al prospectively studied the incidence of MOF over a 12-year period in more than 1300 severely injured patients.[93] The unadjusted incidence of MOF and mortality was unchanged throughout this period, but the injury severity of the patients developing MOF was greater. Nast-Kolb et al showed an increase in MOF incidence from 25.6% in 1975-1980 to 33.6% in 1998-1999, with a marked increase in the age of the patients developing MOF.[38] Evidence of improved survival over time for the same level of injury

severity had been suggested in other studies,[94] leading to the proposal that increased survival of the severely injured in the early phase post-injury has resulted in more patients being at risk of MOF development. However, other studies have suggested a decline in overall MOF incidence,[93] although levels remain high in the most severely injured.[95,96] MOF has been consistently demonstrated as a significant cause of trauma mortality, although estimates vary from approximately 30-80% of late deaths.[30,38,44,92,93,97]

Moore et al suggested that MOF following trauma could be divided into early and late MOF.[98] Early MOF was defined as that present on Day 3 following trauma and accounted for 40%, and late MOF was defined as having an onset occurring after Day 3 and accounted for 60% of the total. Debate has persisted about whether early and late MOF are clinically useful entities, and whether they represent different pathophysiological processes. Some differences in inflammatory marker patterns have been seen between patients with early or late MOF[99] which may mean they result from different processes[100] but these findings have not been widespread.

A major distinction is the difference between ongoing organ dysfunction and acute derangements secondary to inadequate resuscitation in the early phase following injury, which can often be hard to distinguish in clinical practice. For this reason, in trauma, MOF is not defined until after 48 hours following injury. Evidence suggests a decrease in early MOF[101] may reflect improvements in early resuscitation strategies.[100]

Early respiratory dysfunction has been suggested as a key component of post-trauma MOF. A study by Ciesla et al showed that respiratory dysfunction preceded cardiovascular dysfunction by a mean of 0.6 days, and renal dysfunction by a mean of 5.5 days.[102] Post-trauma patients developing respiratory failure predominantly require support due to the

original injury, not from the potential secondary insults of Adult Respiratory Distress Syndrome (ARDS) or pneumonia.[103]

Children have lower rates of MOF and MOF-related mortality following trauma than adults.[104] A study of 579 paediatric trauma patients with an ISS > 15, showed a MOF incidence less than 1%.[105] The rare cases of MOF in children are seen early after injury,[106] and may represent a different physiological response to injury, but the mechanisms and processes behind this remain unclear.[107]

1.4.3. Outcomes Following Post-Trauma MOF

MOF not only confers an increased short-term mortality risk for the year following injury, but also a longer-term increased risk of significant morbidity.[108] A study of 322 trauma patients by Ulvik et al demonstrated increased morbidity and mortality up to seven years post-injury in patients who had suffered from MOF in their initial inpatient stay.[109] Survivors of MOF were four times more likely to require assistance with daily activities than patients who had never developed MOF, and had increased community care costs.

Livingston et al studied 241 trauma patients who had required intensive care treatment, and showed only 50% had returned to work or education at three years post-injury.[110] In the general ICU population, high levels of psychological problems including anxiety and depression are reported[111] and physical issues may persist for long periods after discharge.[112]

1.4.4. Aetiology and Pathophysiology of Post-Trauma MOF

The aetiology of MOF is complex and not clearly understood. MOF was first described in trauma patients by Faist et al in 1983,[36] and was subsequently termed “generalised

autodestructive inflammation" by Goris et al.[92] Many attempts have been made to explain the inflammatory responses that result in significant organ dysfunction following trauma. Meakins first introduced the concept of MOF being the end-point of a series of physiological "hits".[113] The one-hit model postulates that the initial trauma creates a pro-inflammatory state, which will lead to subsequent progression to MOF in patients with insufficient physiological reserve. The multiple-hit model postulates that whilst the initial trauma creates a pro-inflammatory state, it takes further insults to trigger an uncontrolled inflammatory response resulting in end-organ dysfunction. These insults may result from medical interventions such as transfusions or operative procedures, or from the potential sequelae of injury such as infection.[98,114,115]

MOF was originally viewed as a manifestation of uncontrolled infection. In the 1980s, the gastrointestinal tract was considered to be the major driving force for MOF.[116,117] It was proposed that splanchnic hypoperfusion caused ischaemia of the gut mucosa, which increased gut permeability and allowed translocation of bacteria into the systemic circulation. However, selective gut decontamination was not shown to reduce MOF rates in surgical patients,[118,119] and increasing evidence built that MOF could be seen in the absence of infection, and instead reflected a systematic inflammatory response.[120–122]

The Systemic Inflammatory Response Syndrome (SIRS) criteria were defined by a consensus conference of the American College of Chest Physicians and the Society of Critical Care Medicine (ACCP/SCCM) in 1992 to describe the clinical findings resulting from an excessive inflammatory response to an insult of infectious or non-infectious origin.[123] SIRS is defined when two or more of the four criteria given in Table 1.1 are

present. The criteria were designed to be non-specific, with the main purpose being to allow standardisation of illness severity for patient and population comparisons.

SIRS criteria
Temperature < 36^0C or > 38^0C
Heart rate > 90 beats per minute
Respiratory rate > 20/min or $PaCO_2$ < 4.3 kPa (32 mmHg)
White Cell Count < 4 of > 12 x 10^9 cells or > 10% immature neutrophils

Table 1.1: Systemic Inflammatory Response Syndrome (SIRS) Criteria
($PaCO_2$ = Partial arterial pressure of carbon dioxide)

The SIRS criteria have been used for trauma severity scoring, both by amalgamation into the PTS (as previously discussed) and by use as an independent measure. A SIRS score of 2 on admission was shown to be an independent predictor of increased mortality and length of stay in trauma patients in a study of nearly 5,000 patients at a single US centre.[124] Breakdown of the predictive capacity of the SIRS components on a dataset of nearly 10,000 patients found temperature to be the most significant predictor of mortality.[125] A small study of 75 patients in Serbia found SIRS criteria on Day 4 post-injury as highly predictive of MOF development.[126] However, due to the non-specific nature of the SIRS criteria and a lack of consensus on timing of use, they have not been widely incorporated into clinical trauma assessment.

It has been increasingly recognised that post-trauma MOF results from severe immunological dysfunction, but the mechanisms behind this have been widely debated. The overall pattern appears to be one of an early phase of hyperinflammation, followed by a period of immunosuppression.[98] The immunosuppressed period is often termed the Compensatory Anti-Inflammatory Reaction Syndrome (CARS). It has been postulated that

post-trauma MOF can result from either phase. A representation of this process is shown in Figure 1.3.

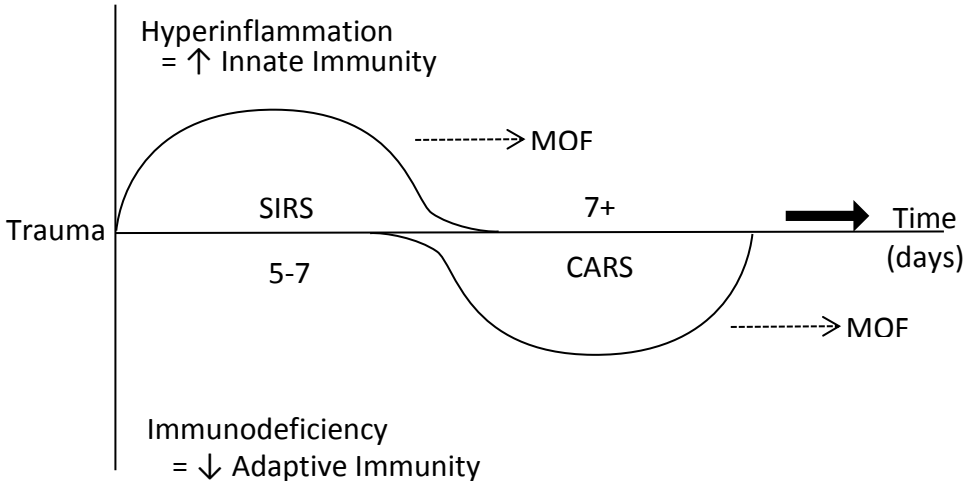


Figure 1.3: Immune response phases after trauma

The hyperinflammation phase occurs due to an exaggerated release of proinflammatory mediators following direct cell injury and ischaemia/reperfusion resulting from the original trauma. Mediators identified in this phase include interleukins, prostaglandins, growth factors and tumour necrosis factor (TNF)-α.[127] Increased priming of neutrophils is accompanied by a macrophage oxidative burst. Activation of the complement system stimulates multiple pathways to induce recruitment of active immune cells to the site of injury.[128,129] The level of complement C3 has been demonstrated to be elevated in the serum of trauma patients and correlated with injury severity.[130]

Severe blood loss from trauma has been shown to result in decreased endothelial-derived nitric oxide (NO) release, which results in increased platelet aggregation and

neutrophil infiltration. Increased vessel permeability is further enhanced by a rise in adhesion molecules and selectins.[131] The overall inflammatory response is thought to result from both local and systemic processes, depending on tissue damage, haemodynamics, oxygen delivery, lactic acidosis and the release of inflammatory mediators.[127]

Significant research efforts have been made to establish specific inflammatory markers that can be measured after trauma, and used to monitor the immune response. A wide range of markers has been studied, including interleukins (IL), C-reactive protein (CRP), procalcitonin (PCT), phospholipase A_2 (PLA_2) and TNF-α. Studies have attempted both to describe the levels and patterns of inflammatory markers over time after injury, and to find biomarkers that can be correlated with subsequent MOF.

A graphical representation of the course of cytokine release after trauma is given in Figure 1.4. Levels of TNF-α and IL-1β increase within the first few hours to maximal levels around 4-6 hours, followed by IL-6 and IL-8 which hit maximal levels at 6-24 hours. The anti-inflammatory cytokine IL-10 is also released in the early hours following injury, but reaches a peak at more than 24 hours post-injury.[132,133]

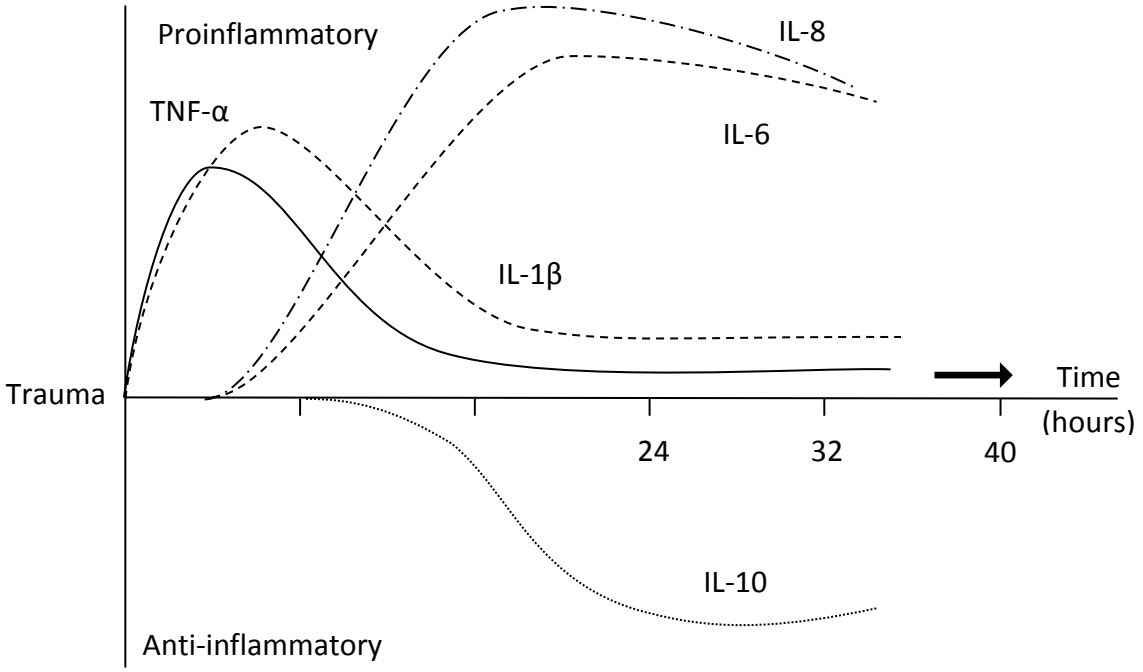


Figure 1.4: Graphic representation of cytokine release after trauma

Results of clinical studies to correlate biomarkers with MOF have shown differing results. Nast-Kolb et al found significantly higher levels of neutrophil elastase, lactate, antithrombin III, IL-6 and IL-8 in patients who died from MOF as early as 24 hours post-injury, although the number of patients studied was small.[99] Partrick et al studied IL-6 and IL-8 levels in 27 patients and found differences between MOF and non-MOF groups.[134] Jastrow et al studied a large number of inflammatory markers at multiple time points from 2 to 24 hours after injury in 48 patients, and identified inducible protein 10, IL-10, IL-6, IL-1Ra and eotaxin as potential MOF predictors.[135] Lausevic et al found IL-6 on the day of admission to be the greatest predictor of MOF development from a range of biomarkers.[126] Cohen et al found increased activated protein C (APC) levels resulted in a 1.6 x increased risk of MOF.[136] Waydhas et al used raised neutrophil elastase, CRP and a

low platelet count to predict MOF with an accuracy of 79%.[114] Other studies have looked at a diverse range of markers, including intercellular adhesion molecule-1 (ICAM-1),[137,138] PLA_2,[139] TNF-α,[140,141] selectins,[138] and neutrophil priming.[142,143]

Transfer of the results of these studies into clinical practice has been problematic. Whilst levels of inflammatory markers can be correlated with severity of injury, and with outcome, sensitivity and specificity is often low. The exact timing at which differences between groups become significant has varied between studies. The range of values for many inflammatory markers appear to vary widely between individuals, and lower levels have been seen in females,[144] making it hard to extrapolate to individual outcomes.[132] Many of the potential biomarkers, such as IL-6, IL-8 and TNF-α have short half-lives complicating their measurement in clinical practice. A recent review showed whilst many associations could be seen between general inflammatory markers and post-trauma MOF, none had sufficient accuracy for use as a clinical biomarker.[132]

A systematic review of randomised controlled trials (RCTs) involving immunomodulative agents showed little evidence for improved outcomes.[145] It seems clear that the mechanisms behind post-trauma MOF are not understood to the extent of translating into clinical improvement. More recent work has focussed on the roles of neutrophil priming, damage-associated molecular patterns (DAMPs) and neurohormonal regulation of cell-mediated immunity.[127,146] Increased understanding of the complex interactions between these pathways is required to affect clinical interventions.

1.4.5. VARIABLES ASSOCIATED WITH POST-TRAUMA MOF

Aside from biomarkers, significant attempts have been made to identify variables that are associated with the development of post-trauma MOF. These may be used to describe and classify patient groups at risk of MOF, to predict MOF development and to assess the effects of interventions.

Several variables pertaining to the patient have been suggested as having an association with post-trauma MOF. Increasing age has been associated with MOF in a number of studies.[98,147–149] Elderly patients are liable to sustain more severe injuries from minor mechanisms of trauma[150] and have less physiological reserve, resulting in an increased risk of MOF.[151] In addition, pre-injury co-morbidities and pharmacy may alter the physiological response to injury, complicating assessment and resuscitation.[152]

The impact of sex on MOF development remains unclear. Animal models have demonstrated that sex hormones affect the synthesis and release of cytokines, and may impact on outcome.[153] Oestrogen is thought to be protective, and male sex has been demonstrated as a risk factor for post-operative infection.[154] Sperry et al found an association between male sex, IL-6 expression and MOF in trauma patients.[155] However other studies have shown no link between sex and MOF development.[147]

Injury severity is associated with both mortality and MOF.[147,149] Keel et al found an association between severe head and abdominal injuries and MOF, and between extremity and pelvic injuries and sepsis.[156] Sauaia et al found an ISS > 25 to be a strong and independent predictor of MOF development.[147]

The so-called "terrible triad" of hypothermia, coagulopathy and acidosis in trauma has long been associated with poor outcomes. Hypothermia and acidosis strongly predict mortality, and are associated with increased transfusion requirements, ARDS and MOF.[157] Elevated BD and lactate are predictive of transfusion requirements and MOF.[147,158] Persistent elevation of BD, even after resolution of lactate to normal predicts mortality[159,160] and post-operative complications.[161]

In 2003, Brohi et al introduced the concept of the Acute Coagulopathy of Trauma Shock (ACoTS).[162] Coagulopathy in trauma was previously thought to be due to fluid administration. In a study of over 1000 patients, it was determined that an acute coagulopathy is seen in trauma that is unrelated to fluid administration, but is linked to injury severity, and confers an increased risk of mortality.[162] This association was reinforced in a US study of over 7000 trauma patients which showed an increased adjusted odds of death of 35% with an abnormal prothrombin time (PT) on arrival and 326% with an abnormal activated partial thromboplastin time (APTT).

Recognition of ACoTS has led to changes in early management, with more aggressive transfusion protocols and correction of haematological parameters. Massive transfusion protocols have been instituted with a marked change in the ratio of blood products administered, with Packed Red Blood Cells (PRBCs) and Fresh Frozen Plasma (FFP) given in 1:1 ratios. Transfusion of more than 6 units of PRBCs in the first 12 hours after injury has been shown to be an early risk factor for MOF.[148,149,163] A study by Watson et al found an increased risk of MOF of 2.1% for each unit of FFP administered, but has been criticised for use of secondary data and failing to control for confounders.[164]

The contribution of surgical intervention to post-trauma MOF remains unclear. A major historical model for fracture surgery in the critically ill trauma patient was termed Early Total Care (ETC). This involved operating on the trauma patient as soon as possible, with long operating times and multiple procedures carried out in the early phase after injury. Anecdotally the outcomes from this were poor for some severely injured patients, and in the 1990s a move was made to Damage Control Surgery (DCS). DCS involves minimising the early surgical burden by carrying out only life or limb saving interventions to stop haemorrhage, reduce contamination or prevent further injury.[165,166] Further definitive surgery is then performed at a later date. Debate has persisted about the timing of these definitive procedures.

Waydhas et al demonstrated that patients with higher pre-operative levels of inflammation had an increased risk of post-operative MOF,[114] and higher rates of MOF have been seen in patients undergoing orthopaedic fixation on Day 2-4 post-injury.[167] Pape et al showed a higher release of proinflammatory cytokines in pelvic surgery than spinal surgery, associated with the magnitude of intervention and blood loss, but not with operative timing.[168] The association of surgical type and timing with MOF remains unclear. A recent review found only 5 retrospective studies including MOF as an outcome measure, and no overall associations were found with surgical type and timing and MOF development.[169]

A full systematic review was performed to identify variables associated with post-trauma MOF and is discussed in detail in Chapter 3.

1.4.6. PREDICTION OF POST-TRAUMA MOF

A limited amount of work has been performed on predictive modelling of post-trauma MOF. Sauaia et al studied data from 411 patients at a single institution in Denver, Colorado, and used logistic regression to identify which combinations of variables best predicted MOF development.[97] Models were built on admission and at 12, 24 and 48 hours following injury, using variables indicating tissue injury, patient factors, shock and measures from the Acute Physiological Score – Acute Physiology and Chronic Health Evaluation III (APS-APACHE III). MOF was defined according to the Denver Postinjury Multiple Organ Failure Score. Injury severity was assessed by ISS, AP and GCS. The resultant models were internally validated on the same population.

Models at all time-points showed higher specificity than sensitivity, using a 20% probability cut-off for MOF development. Using Receiver Operating Characteristic (ROC) curves, the area under the curve (AUC) was significantly increased in the 12-hour model compared to the admission model, with small further increases at 24 and 48 hours. The authors argued that use of age, ISS, transfusion amount and platelet count at 12 hours offered reasonable predictive power for MOF development.

A number of issues complicate the application of these findings to the UK trauma population. 30% of the patients in the study suffered from penetrating trauma, which has been suggested in other studies to be associated with a decreased incidence of post-trauma MOF.[40,93] Whilst a sample size of 411 patients was used, data on all variables used in the models were available for only 121 of them, and whilst the findings have been routinely commented on in trauma literature, they have not been externally validated.

More recently, the Denver team attempted to derive an instrument to predict MOF from data available in the ED, called the Denver Trauma Organ Failure Score (Denver TOF). This defined MOF by the Sequential Organ Failure Score (SOFA) and used data from over 4000 patients. Points were given for age over 65 years, emergency intubation, low haematocrit, low SBP, low urea and a raised WCC. An increasing score correlated with an increased probability of MOF development.[170] This was found to be more accurate than use of SOFA, lactate and BD in the ED.[171] However, this cohort again contained nearly 20% of patients with penetrating trauma.

Two further studies were published following the commencement of this project. Minei et al used data on over 900 blunt trauma patients from the Inflammation and the Host Response to Injury Collaborative Research Program database to develop a multivariate model of predictors of organ failure.[172] MOF was defined using the Marshall scale and an incidence of 29% was recorded. Male sex, pre-existing heart or liver disease, chest or extremity injury, ISS, APACHE II score, volume of infused crystalloid, vasopressor use, FFP transfusion and worst BD and lactate. However, the significance of these variables was recorded at different time-points. The entry criteria required hypotension, an elevated BD, a need for early transfusion and a high injury severity, and whilst these entry criteria were based on information from previous studies indicating them as markers of high severity and poor outcome, they were not correlated with MOF development.

A recent Australian study by Dewar et al studied 140 patients, 94% with a blunt mechanism of trauma.[173] MOF was defined using the Denver score, and an incidence of 15% was recorded. Multivariate logistic regression was used to create models on admission and at 24 hours post-injury. The admission model contained age and the

admission platelet count, and gave an AUC of 0.71. The 24-hour model contained age, minimum bilirubin, maximum creatinine and admission platelet count and gave an AUC of 0.84. Patients with MOF were found to be 8 times more likely to die than patients with similar injuries and shock parameters who did not develop MOF. The authors noted the differences in the variables found to predict MOF in comparison to the 1998 Sauaia et al study.

1.5. THE RESEARCH PROJECT

1.5.1. CREATING A PREDICTIVE MODEL FOR POST-TRAUMA MOF

Modelling combinations of clinical variables to predict post-trauma MOF has not been previously performed in the UK. Knowledge of the relationship between clinical variables and post-trauma MOF would be useful to create risk-adjustment strategies to allow case-mix adjustment, and hence audit outcomes from different hospitals, teams and groupings with different case mixes. Current decision-making on the patients at greatest risk of MOF is largely dependent on the knowledge and opinions of individual clinicians. Such decisions can be complex, and robust research would help to guide the clinician in making decisions on resource allocation, treatment plans and patient management.

This project therefore aims to create a prediction model for post-trauma MOF using routine clinical variables that are available and can be collected in the first 48 hours following injury.

For any model to be clinically useful it needs to give an indication of the likelihood of MOF within a time-scale that would allow current and future treatments aimed at MOF prevention to be delivered. Post-trauma MOF is defined as dysfunction present from 48

hours after injury.[37] Clinically useful prediction therefore needs to occur in this 48-hour window. The earlier that MOF can be predicted, the greater the opportunity for changes in patient management to modify the outcome.

One of the problems in identifying a post-trauma MOF biomarker has been the lack of feasibility in clinical practice. The use of routinely collected data focuses on variables that are frequently collected, easily accessible and considered important to clinicians. This also allows a predictive model to be quickly assimilated into clinical practice in the future. In addition, knowledge of the association of clinical findings and outcome, even if not quantified numerically, aids clinical decision-making and prognostication.

In addition, the project was designed to address some of the problems noted in previous studies. The use of retrospective data sources allows access to a large amount of data, which would take a considerable time to collect prospectively. A larger sample size allows an increased number of potential MOF predictor variables to be considered, and reduces problems with missing data.

1.5.2. THE RESEARCH QUESTION

Using routine clinical data collected in the first 48 hours after injury, is it possible to predict those patients at highest risk of post-trauma MOF?

1.5.3. AIMS AND OBJECTIVES

Aim: To develop a predictive model using clinical variables to predict post-trauma MOF within the first 48 hours following injury

Objectives

- To examine the use of retrospective and routinely collected data in trauma patients in providing viable research data
- To accurately link patient data across multiple data sources
- To use predictive modelling techniques to apply to post-trauma MOF prediction
- To attempt to quantify patient risk of MOF after trauma
- To describe the outcomes from the current package of care for severely injured trauma patients in the UK
- To identify potentially modifiable risk factors to focus future research
- To create a predictive model for post-trauma MOF in the UK trauma population to allow future patient stratification into treatment groups, case-mix adjustment for observational studies and case selection and enrichment for RCTs

1.5.4. OUTLINE OF METHODS

A full overview of the methodology employed is given in Chapter 2.

In brief, a retrospective study was carried out using routinely collected clinical data, stored in electronic formats. The patient population was adult trauma patients admitted to the Adult ICU (AICU) of the John Radcliffe Hospital in Oxford from 2003-2011.

Relevant input variables that may be associated with post-trauma MOF were identified from a systematic literature review, questionnaire analysis of clinicians involved in trauma care, and clinical review. The most appropriate measure to define the outcome measure of MOF was chosen following literature review and analysis of previous studies.

Data on the input and outcome variables, and for population descriptors were collected from a range of electronic sources. These included two national databases – TARN and the Intensive Care National Audit and Research Centre (ICNARC) database – and the local AICU database, called CareVue. Additional local data sources were also used to minimise missing data. Patients were matched across the three main data sources prior to inclusion. Data were extracted using a range of techniques depending on data source, and details for each variable were recorded in a data dictionary.

All data were then cleaned and collated into a single data system. Statistical analysis was then performed to describe the study population and the distribution of post-trauma MOF. Univariate analysis was performed on individual input variables to assess association with MOF development. Significant variables meeting set criteria were then inputted into multivariate logistic regression analysis to determine combinations of variables that could predict post-trauma MOF at 2, 4, 6, 12, 24 and 48 hours following injury. The most significant models were then internally validated on the same sample.

1.5.5. STRUCTURE OF THE THESIS

The following chapters describe in detail the methodology and analysis employed to answer the research question.

Chapter 2 provides an overview of the methodology. Chapters 3-6 provide details on different aspects of the methods used. Chapter 3 describes the process of input variable selection. Chapter 4 describes the use of MOF scoring systems, and the selection of specific scoring systems for outcome measures. Chapter 5 describes the process of data matching across data sources to obtain the study population. Chapter 6 describes the

process of data extraction from each of the data sources and the process of cleaning for analysis.

The results of this process are given in Chapter 7, with discussion and future recommendations given in Chapters 8.

Details of data extraction, the data dictionaries for the collected variables and statistical analyses are given in the appendices.

CHAPTER 2 METHODOLOGY OVERVIEW

2.1. OUTLINE OF METHODS

A retrospective observational study was performed to collate routinely collected clinical data on major trauma patients admitted to the AICU of the John Radcliffe Hospital in Oxford.

The research question was examined by following a multi-stage process:

1. Determining the input variables
2. Determining the outcome variables
3. Data matching of the patients across data sources
4. Extraction of data from multiple sources
5. Data cleaning
6. Data analysis

The exact data collected were determined by a systematic process prior to data extraction. Input variables consisted of those variables that had the potential to be predictive of post-trauma MOF. These were determined by a systematic review of the current literature, and analysis of the responses to questionnaires by clinicians involved in trauma care. Results from both of these were considered in the light of knowledge of current clinical practice in the UK to create a definitive list of input variables.

MOF was the primary outcome measure. Due to the lack of a gold-standard measure of MOF, a review was performed of the different MOF scoring systems available. These

were evaluated for use in the study, based on the specificity of the scoring systems for post-trauma MOF and their use in other studies. As a result, three MOF scoring systems were chosen (the Denver, SOFA and Marshall scores) and data collected to allow calculation of all three. A breakdown of the variables required for each of the scoring systems was created. Secondary outcomes collected included mortality, single organ failure, and hospital and AICU lengths of stay.

In addition to the data required for the input and output variables, data to allow a full description of the trauma population studied were also collected.

The trauma population was defined as adults over 16 years of age who had suffered a traumatic mechanism of injury of sufficient severity to require admission to the AICU of the John Radcliffe Hospital from 2003-2011. Patients had to be matched to the TARN database, and a minimum of one of the ICU data sources. A level of trauma severity is required for inclusion in TARN, details of which are further discussed later in this chapter. Patients were excluded if they died prior to AICU admission, or were admitted to a specialist ICU (i.e. cardiac or neurosciences) without admission to AICU.

Data were collated from a range of electronic sources. The two national databases of TARN and ICNARC both gave permission for use of their data in this study. The third major data source was CareVue – a local system used in AICU to essentially create detailed electronic charts of the patient stay by recording hourly physiological and management data in conjunction with the results of laboratory testing and other investigations. Patients were matched across the TARN, ICNARC and CareVue databases. Additional local data sources were used to resolve matching problems and to address issues with missing data.

Data were extracted from each of the data sources using a range of techniques. This was a two-stage process – first of identifying potential fields representing a particular variable across the data sources, and secondly of extracting the data in the relevant fields in accordance with the definitions for each variable set out in the data dictionary.

After extraction, all data were cleaned and collated. Variables were either considered fixed (i.e. they did not alter with time following injury, such as patient age or injury severity) or time-dependent (i.e. variation over time was possible, such as physiological variables or laboratory results). The time of injury was taken as time zero. Each of the time-dependent variables was then considered for each of the following time periods after injury:

- 0 – 2 hours
- 0 – 4 hours
- 0 – 6 hours
- 0 – 12 hours
- 0 – 24 hours
- 0 – 48 hours

Time-dependent variables were considered sequentially for each of the above time periods. For example, should the minimum value of a variable be required, all data from 0-2 hours would be assessed to obtain the minimum value. For the 0-4 hours period, the data from 2-4 hours would be additionally considered, and the minimum value from all measures available was taken for that time period.

Variables representing the components of the MOF scoring systems were allocated scores according the Denver, SOFA and Marshall criteria. An overall MOF score was created for each patient who had data available from days 3 – 100 post-injury for each of the Denver, SOFA and Marshall scores. If data were unavailable for any day within the first 100 days (i.e. due to ICU discharge), the patient was assumed not to have MOF on that day. This was translated into binary fields indicating the presence or absence of MOF at any time in the first 100 days post-injury for each of the MOF scoring systems.

Statistical analysis was carried out following a strict analysis plan written in advance. Initial analyses were performed to describe the study population to give an indication of the current composition of major trauma patients seen at a UK trauma centre. The incidence and distribution of post-trauma MOF was described in detail, and the outcomes according to the three MOF scoring systems of Denver, SOFA and Marshall were compared.

Input variables were further assessed on data completeness and suitability for further analysis. All input variables with sufficient data underwent univariate analysis to assess association with post-trauma MOF, as defined by each of the Denver, SOFA and Marshall scores in turn. Variables meeting set criteria following univariate analysis were included in multivariate analysis.

Multivariate logistic regression analysis was used to determine which combinations of variables most accurately predicted MOF. Models were created at 2, 4, 6, 12, 24 and 48 hours post-injury for MOF as defined by each of the Denver, SOFA and Marshall scores, making 18 models in total. The fit and accuracy of these models were assessed, and the most clinically significant models were internally validated on the same sample.

2.2. INPUT VARIABLE SELECTION

2.2.1. LITERATURE REVIEW

A full systematic review was performed according to a pre-defined protocol. The purpose of the review was to assess the current literature and evaluate the available evidence on variables that may be associated with post-trauma MOF. The review was focussed on variables routinely used in clinical practice that are available to the clinician within the first 48 hours after injury. Experimental and non-standard investigations or procedures were excluded. Both positive and negative associations found on univariate and multivariate analyses were considered in the review. From this review, a list of potential input variables was created.

2.2.2. QUESTIONNAIRE ANALYSIS

A simple questionnaire was given to 60 clinicians involved in the management of major trauma patients. Respondents were asked select variables from a list which they considered to be associated with post-trauma MOF. Responses were pooled to assess which variables are currently used by doctors in clinical practice, and a second list of potential input variables was created.

2.2.3. CREATING THE INPUT VARIABLE LIST

The definitive list was taken by combining variables obtained from the review and questionnaire, and assessing the utility of each variable in current clinical practice. Variables that could not be clearly defined or are not routinely used in the UK were excluded to create a definitive list. Component variables of scoring systems (i.e APACHE

II, SIRS) were included only if separately identified as variables with an independent association with MOF.

Input variables were divided into five groups for the selection process - patient, injury, physiological, laboratory and management variables. Patient variables were considered to be intrinsic to the individual and present prior to the traumatic event i.e. demographics and pre-existing health conditions. Injury variables pertained to the traumatic event itself and the resultant injuries. Physiological variables were measureable clinical parameters which did not require laboratory-based testing, whereas laboratory variables were predominantly blood test results. Management variables were those that related to the clinical management of the patient, and included a range of surgical and non-surgical interventions.

2.3. Outcome Measure Selection

2.3.1. Primary Outcome – MOF

MOF was the primary outcome measure.

The definition of MOF remains problematic with no accepted gold-standard. Since the concept of MOF was first introduced in the 1970s more than 30 scoring systems have been developed to attempt to quantify and define MOF. Different scoring systems are used in different research areas, with trauma-specific scales not widely used in intensive care research. Selection of a single scoring system to define MOF for the study was therefore problematic.

A review of the literature was performed to assess the range and extent of MOF scoring systems and their use in related studies. These were then assessed against the available data to determine the most appropriate scoring system for this study.

No single scoring system was found to be appropriate for comparison to pre-existing literature. In order to increase the applicability of the study findings, it was decided that data would be collected for three MOF scoring systems – the Denver, SOFA and Marshall scores.[37,174,175]

2.3.2. SECONDARY OUTCOMES

Data were collected on a defined list of secondary outcomes, including mortality and length of AICU and hospital stays. All variables required for the composition of the three MOF scoring systems, secondary outcomes, population description and input variables were combined to a single list for data extraction.

2.4. DATA SOURCES

Data were obtained from three main sources and a number of additional sources. All sources were pre-existing systems used to collect and record clinical data. All sources, bar one, were held in electronic format. Data sources were selected to provide as wide a range of data as possible, given this was a retrospective study. All data sources contained data routinely accessible to members of the clinical team involved in trauma patient care.

The three main data sources were the two national databases of TARN and ICNARC, and CareVue, a local system for recording clinical data in the AICU.

2.4.1. TARN DATABASE

The TARN database is a nationwide project covering England and Wales. Details on trauma patients fulfilling specific criteria are submitted from individual hospitals to a central database via a web-based portal. Submission to TARN used to be optional, but since the 1st April 2012 has been compulsory for all MTCs in England and Wales. The John Radcliffe has submitted data in some form since the 1990s, but routine complete submissions have been performed since 1st January 2003. The criteria for entry into TARN are given in Appendix 1. Minor changes in the criteria have occurred over time, but these tend to impact more on defining minor trauma at the bottom end of eligibility for inclusion, and would not impact on the severe trauma patients included in this study.

Information in TARN is focussed on the trauma itself and its subsequent treatment and management. Data collected include patient demographics, injury details, management, interventions, operative procedures and outcomes. A single TARN entry is created for each traumatic episode – if a patient is admitted to more than one hospital following injury, data from both are combined into a single TARN entry. Data could therefore be obtained on both direct admissions and patients who were transferred in from surrounding centres.

TARN data on individual patients is entered into a web-based portal by administration staff at the end of the patient's hospital stay. The data required is predominantly obtained from written hospital notes. Summary TARN data on local patients can be obtained from the central TARN office. Permission was granted from TARN for the use of local data for this study.

2.4.2. ICNARC DATABASE

ICNARC receive data submissions for all patients admitted to intensive care in the UK. A single entry is created for each ICU admission. The John Radcliffe has submitted data to ICNARC since 1st June 2006. Submissions are entered onto a local system by administration staff using data collected prospectively during the patient's ICU stay, before being uploaded to the central ICNARC database. Permission was obtained from ICNARC for the use of the local submission data in this study.

Data collected on the ICNARC database includes patient demographics, physiological details in the first 24 hours of ICU admission and measures of outcome.

2.4.3. CAREVUE

CareVue is an electronic records system designed and operated by Philips Medical Systems. It is used in the John Radcliffe AICU to provide an electronic record of the patient's entire ICU stay. Information contained includes patient demographics, physiological monitoring, laboratory results, prescriptions and drugs administered, interventions and nursing and clinical notes. Standard physiological variables are recorded hourly. Entries of clinical data into the system are not automatic, and must be verified by nursing staff. Individual data entries are checked by local nursing staff before uploading to the system, and these staff are trained not to allow input of erroneous or unrepresentative data. Data inputted are therefore meant to reflect typical data for a patient for each time period, as extreme values should not be verified.

Two different systems of CareVue have been used at the John Radcliffe Hospital. An older system has collected data since 1999. The original system was superseded by a new system on 1st June 2008. Whilst similar data are held in both systems, the newer system is

far more extensive. The differences between systems will be further discussed in Chapter 6.

2.4.4. ADDITIONAL DATA SOURCES

The AICU at the John Radcliffe Hospital maintains its own internal administration database, known as ICU Admin. This contains patient identifiers, and details pertaining to patient coding for the Payment by Results (PbR) system.

The John Radcliffe Hospital Patient Administration System (PAS) collects data on patients and their movements throughout the hospital It contains basic patient demographics and identifiers. Each patient should have a single entry on PAS which is subdivided into each medical contact at the hospital, including both inpatient and outpatient contacts. This system was moved to a new Electronic Patient Record in 2011-2012. The old PAS system was left available for research and audit purposes.

Other electronic data sources available included the radiology picture archiving and communication system (PACS), and the internal trauma administration record - Bluespiers (by Bluespiers International Ltd). Written hospital notes were also used in the absence of electronic alternatives to obtain specific data.

2.5. DATA CONSIDERATIONS

2.5.1. DATA PERMISSIONS

Independent permission was granted by both TARN and ICNARC for use of their data at the study design phase.

Obtaining individual informed consent for the use of personal data was not practical in this patient group for a number of reasons. Firstly, the study population is retrospective and to obtain a suitable sample size, dates back to over a decade from the present day. Secondly, the mortality in this population is high, even following ICU discharge. Contacting relatives for retrospective consent would be difficult in such an emotionally sensitive context, as well as time consuming and complex. Thirdly, contacting only patients who had survived would introduce significant bias to the sample.

Since obtaining informed consent was not practical, patients user groups representing both trauma (The Oxford Trauma Users group) and intensive care (CritPaL – the Critical Care Patient Liaison Association) were contacted to investigate the acceptability of the use of routine clinical data for research purposes in this study. A positive and supportive response was received from both groups. The study design was then presented to the local Caldicott Guardian, who supported the use of patient data without consent in this context.

Ethics approval was given by the Berkshire Research Ethics Committee (REC) in March 2011 (REC ref 11/SC/0015). Approval from the National Information Governance Board (NIGB) was sought under Section 251 of the NHS Act 2006, which allows the common law duty of confidentiality to be set aside in specific circumstances. The Ethics and Confidentiality Committee (ECC) of the NIGB granted approval in April 2011 under study reference 3-04(X)/2011. Both REC and NIGB approval have been reconfirmed each year since the study commencement. Research and Development approval was given by the Oxford University Hospitals NHS Trust (previously the Oxford Radcliffe Hospitals NHS Trust) with reference 6407.

2.5.2. DATA AVAILABILITY

The availability of the main data sources of TARN, ICNARC and CareVue varied. Complete TARN submissions were present from 1/1/2003, which was therefore set as the start date for data collection. ICNARC data was available from 1/6/2006. CareVue data was present from 1999, but switched from an old to a new system on 1/6/2008. The differences in data availability impacted on the data extraction and management process, and are summarised in Table 2.1.

Year	2003	2004	2005	2006	2007	2008	2009	2010	2011
TARN	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ICNARC	No	No	No	From 1/2006	Yes	Yes	Yes	Yes	Yes
CareVue	Old	Old	Old	Old	Old	Switch 6/2008	New	New	New

Table 2.1: Availability of main data sources

TARN and ICNARC data are both entered retrospectively, often after the patient's discharge, creating a time lag in the availability of data from these sources. The end date was set to maximise the number of included patients, without introducing bias from the time lag in submissions. The total time period used for the study was therefore 1/1/2003 to 15/5/2011.

Since ICNARC was not available for the entire time period, the use of ICNARC data was minimised. To assist in linking patients across database, ICNARC submissions and the ICU Admin database were compared and an exact 1:1 match of patients was found.

2.6. STUDY POPULATION

2.6.1. THE JOHN RADCLIFFE HOSPITAL TRAUMA POPULATION

During the period of this study, MTCs had not yet been formally founded across England. However, the John Radcliffe Hospital in Oxford was one of the few units that had

effectively been functioning as an MTC before they were formally established. The John Radcliffe took trauma patients from across a wide geographical area, covering a large section of the south-central part of the UK, and the whole Thames Valley region. Now that MTCs have been established, the next nearest units are Bristol to the west, Birmingham to the north-west, Southampton to the south, and the London centres of St Mary's, St George's, King's and The Royal London to the south-east. The area between these approximates the area from which referrals and transfers were taken prior to the establishment of MTCs nationwide.

Trauma patients were both admitted directly to the John Radcliffe, and transferred from surrounding units for specialist care. Many of these transfers took place in the early phase after trauma, following initial stabilisation of the patient. Both direct admissions and transfers were included in the study population, providing transfer occurred within 48 hours of injury.

Prior to 2007, neurosurgery, plastic surgery and Ear, Nose and Throat (ENT) services were housed on a separate site at the Radcliffe Infirmary, around 2 miles from the site of the John Radcliffe. The ED has always been housed at the John Radcliffe, but when services were on split sites, patients with isolated head injuries were transferred from ED to neurosciences ICU at the Radcliffe Infirmary. From 2007, a new extension to the John Radcliffe brought all surgical services to one site. The pattern of admission of trauma patients between AICU and neurosciences ICU remained unchanged. Isolated head injuries were admitted to neurosciences ICU and polytrauma patients with associated head injuries were admitted to AICU.

Patients admitted to neurosciences ICU (and other specialist ICUs) without an admission to AICU were excluded from the study for a number of reasons. Firstly, the neurosciences ICU does not use the CareVue system, reducing the information available for these patients. Secondly, isolated head injuries are directly responsible for a high number of trauma deaths in the absence of MOF and measures of head injury severity correlate strongly with mortality, complicating analysis of the impact of MOF.[148]

For the purposes of the study, all patients admitted to neurosciences ICU without an AICU stay were taken to have isolated head injuries, and all patients admitted to AICU were taken to have polytrauma, reflecting the divide that occurred in clinical practice.

2.6.2. INCLUSION CRITERIA

The following inclusion criteria were applied to the study population:

- Aged 16 years or over
- Traumatic mechanism of injury
- Admitted to AICU of the John Radcliffe Hospital
- Arrival at the John Radcliffe Hospital within 48 hours of injury
- Admission to the John Radcliffe Hospital between 1/1/2003 and 15/5/2011
- Present in the TARN database
- Present in the CareVue database
- Present in either ICNARC or the ICU Admin databases

2.6.3. EXCLUSION CRITERIA

The following exclusion criteria were applied to the study population:

- Death prior to AICU admission i.e. pre-hospital, en route or ED

- Admission to other specialist ICU (neurosciences ICU, cardiac ICU) without admission to AICU
- Did not meet criteria for inclusion in the TARN database
- No available data on time of injury
- Unable to match patient between CareVue, TARN and one of ICNARC or ICU Admin

2.6.4. POPULATION SIZE

The study aimed to include as many patients meeting the inclusion/exclusion criteria within the relevant time period. Initial analyses based on the frequency of AICU admission amongst TARN submissions from the John Radcliffe, suggested a sample size of 500-600 patients was possible.

The number of variables that could be inputted into a predictive model was determined by the number of patients in the study population and the extent of missing data. Since the exact size of the population could not be established until data matching was complete, an estimate of the required sample size was made to determine if sufficient data were likely to be available for a robust predictive model. This estimated sample size was based upon advice from the Department of Statistics at the University of Oxford and published information on model development.

A key consideration in sample size for predictive models is the relationship between the number of predictive variables and the number of outcome events.[176] An insufficient number of events per variable has been linked with over-estimation of model performance, and with an increased risk of bias in the observed regression coefficients.[176,177] A minimum of 10 - 20 events per variable has been accepted as

reducing risk of problems with model development.[176,178] Vittinghoff et al argued that fewer than 10 events per variable could be acceptable in specific circumstances,[179] but Peduzzi et al noted significant problems with bias if fewer than 10 events per variable were used.[177]

As a conservative estimate, a sample size of 20 patients x (number of variables +1) was used. This assumed the use of a generalised linear model, with a significance level of 5% and power of 80%. From this calculation, a sample size of 420 patients was estimated for a 20 variable model, with the recognition that smaller numbers could be possible without a significant adverse effect on model stability. Potential interactions between variables could result in increased random effects, which could be compensated for by an increase in patient numbers or a decrease in the number of input variables.

2.7. DATA LINKAGE

Data linkage involved the use of patient identifiers to match patients across the TARN, ICNARC and CareVue databases. Further data sources, including ICU Admin, PAS and hospital notes were used to resolve any linkage problems.

Linkage was performed from two separate starting points. The TARN database was queried for eligible patients. The ICNARC and ICU Admin databases were also queried for eligible patients and the results of these queries combined. This resulted in three groups of patients – those present in TARN and ICNARC/ICU Admin, those present in TARN but not ICNARC/ICU Admin and those present in ICNARC/ICU Admin but not TARN. A secondary process then occurred to resolve the matching of patients in the last two groups.

Once complete linkage between TARN and ICNARC/ICU Admin had been performed, the resultant patients were matched to CareVue. If matching was successful, patients were labelled as major trauma patients within CareVue to aid future data extraction.

The primary identifier for linkage was hospital number. Secondary identifiers used included name, postcode, address, date of birth, NHS number, date of admission and date of injury.

A single linkage file was created of all patients in the study population to allow linkage across all data sources. Data retained in this file included patient identifiers, and database identifiers such as TARN number and ICNARC admission number. This file was stored on a secure NHS research server according to the plans submitted to and agreed with the NIGB ECC, and access was limited to the primary researcher and three named supervisors only.

Each patient in the linkage file was also given a specific "TraumaDBID" number. This identification number was then used to anonymise all further data files. The only patient identifier kept in the final data for analysis was sex, due to its potential relevance to the research question.

The plan for data management in the original study protocol is given in Appendix 2.

2.8. DATA EXTRACTION

The data extraction process was complex, and is fully explained in Chapter 6. In essence, a staged process was used to identify potentially relevant data for each required variable from TARN, ICNARC and CareVue. Potential data fields were assessed for the presence

and utility of the data contained in them. Data were then extracted if the fields were relevant and met the criteria for each variable as outlined in the data dictionary.

The process of data extraction varied depending on the format of the original data source, and the way it could be accessed. Data from TARN were obtained by direct queries sent to the national TARN office, who returned results in Microsoft Excel spreadsheets. ICNARC data were stored in comma separate variable (.csv) files, which were converted into a Microsoft Access database. CareVue data were stored in Microsoft SQL Server, and required queries to be written in Structured Query Language (SQL) for data extraction.

Data from all sources were converted into Microsoft Excel files to allow combining of data from different sources, and data cleaning. The exact details of data extraction for each variable and for each data source were recorded in the data dictionary to allow for future replication.

2.9. DATA CLEANING AND ANALYSIS

2.9.1. DATA CLEANING

Data cleaning was performed to remove impossible and implausible values. Assessment of data suitability depended on the variable in question and hard and soft limits for data exclusion were set for each variable. Full details of these limits are included in the data dictionary. Data was standardised for units, timing and data type, and duplicate entries were removed.

2.9.2. TIMING OF DATA

The study aimed to create predictive models for post-trauma MOF at 2, 4, 6, 12, 24 and 48 hours following injury. The time of injury was taken as time zero for the models, and was chosen as modelling needs to occur from a fixed time-point that is:

- Consistently measured for each patient
- Available from the data sources
- Relevant to the research question
- Not subject to uncontrolled variability
- Not affected by a process that may potentially alter the results

Whilst it is recognised that the exact timing of injury may not always be accurately recorded, it was felt that other timing points such as hospital admission were already susceptible to uncontrolled variation, and that interventions in this period could impact on the patient's outcome. To address potential problems in data accuracy, hospital notes were used to ensure data completeness. The earliest recorded time in any format was taken as the time of injury i.e from ambulance sheets, ED clerking or other clinical notes. A small number of patients were excluded due to there being no accurate data on the timing of their injury.

For each variable that was time-dependent, values for that variable were obtained for the time periods 0 – 2 hours, 0 – 4 hours, 0 – 6 hours, 0 – 12 hours, 0 – 24 hours and 0 – 48 hours after injury, or classified as missing data. Data present within 2 hours of injury would therefore be used in all models, whereas data present at 25 hours would only be used in the 0 – 48 hours model.

2.9.3. STATISTICAL ANALYSIS

Data was analysed according to a pre-determined statistical analysis plan. This was divided into several sections:

1. Description of the whole trauma population
2. Analysis of MOF and the population at risk of MOF
3. Univariate analysis of potential MOF predictors
4. Selection of variables for multivariate analysis
5. Multivariate analysis of potential MOF predictors and model building
6. Internal validation of relevant multivariate models

MOF scoring was performed according the Denver, SOFA and Marshall scores for each organ system for days 3 – 100 post-injury. This was used for descriptions of organ failure in the population, and also converted into a binary field to indicate the presence or absence of MOF at any stage in the first 100 days after trauma.

Input variables were assessed for their relationship with MOF, to determine the best categorisation for further analysis. Categorisation is dependent on graphical and statistical relationships, published recommendations and clinical appropriateness, in keeping with accepted model-building strategy. Univariate analysis was performed to assess the association of each input variable with MOF, as defined by each of the Denver, SOFA and Marshall scores.

Assessment of the suitability of a variable for use in multivariate analysis was based on the following data rules:

- $P < 0.25$ on univariate analysis

- Missing data ≤ 50%
- For variables with binary outcomes, > 5% patients have least frequent outcome (to exclude rare events)
- If both minimum and maximum values for the same laboratory variable or physiological parameter were collected and examined in univariate analysis, and for both $p < 0.250$, then the variable definition with the lowest p value was used
- For categorical variables with ≥ 3 categories, consider collapsing categories if any category has ≤ 5% of data
- Remove clinically inter-related variables, retaining the one with greatest clinical utility and relevance
- If a treatment variable is specific to an injury, remove the treatment and consider the injury data only
- For composite scores, remove composite and only use significant individual component variables
- Exclude variables with data that remain clinically implausible or unreliable despite data cleaning

Using these rules, a list of variables was made for multivariate analysis at each time-point and for each outcome scale, making 18 models in total (2, 4, 6, 12, 24 and 48 hours post-injury for MOF defined by each of the Denver, SOFA and Marshall scales). Multivariate logistic regression was run for each model using a manual process based on backward elimination.

Each model was assessed for accuracy, goodness-of-fit, and predictive value. Models were considered to have the greatest potential for future clinical use if they showed high

predictive value relative to the number of input variables, and could predict MOF early after injury. Internal validation of models was performed using bootstrapping. Further details on the process of model formation, assessment and validation are given in Chapters 6 and 7.

CHAPTER 3 INPUT VARIABLES

3.1. SELECTING CANDIDATE INPUT VARIABLES

To create a predictive model of post-trauma MOF, plausible input variables are required as the basis for the model building. These input variables should represent measures or processes that are considered likely to be associated with post-trauma MOF, and may therefore be useful in predicting its development.

Sample size calculations indicated that the estimated number of patients in the study population would be adequate for consideration of a 20 variable model. Therefore, an estimate of up to 100 potential candidate variables was made for consideration for univariate analysis, to allow for a 20 variable final model.

The selection of these candidate input variables was performed using two main processes:

1. A systematic review of the current trauma literature on variables associated with post-trauma MOF
2. Questionnaire analysis of trauma clinicians to establish variables considered to be associated with post-trauma MOF in clinical practice

Lists of candidate input variables were created from each of these processes, and were then analysed to create a single final list. For inclusion on the final list, variables needed to be potentially measurable given the available data sources, clinically relevant and realistic for measurement and use in a clinical setting.

Input variables were grouped into five main categories, which were further subdivided into sub-categories as given in Table 3.1. This categorisation was used throughout the study.

Input variable category	Sub-categories
Patient	Basic patient data Chronic health conditions
Injury	Injury severity Injury details
Physiology	*No sub-categories*
Laboratory	Haematology Biochemistry
Management	Transfusion Emergency Surgery Emergency Intervention

Table 3.1: Input variable categories

3.2. LITERATURE REVIEW

3.2.1. INTRODUCTION

A wide range of studies have attempted to find variables associated with post-trauma MOF. However, many of these have used experimental or novel tests or interventions that are not available to the clinician in daily practice. Studies focussing on clinical variables have often studied individual variables in isolation, and therefore a synthesis of the available evidence was required. This review aimed to systematically assess the current literature to identify clinical variables that may be associated with an increased risk of post-trauma MOF, to determine potential input variables for predictive modelling.

In practice, an assessment of a patient's risk of post-trauma MOF is predominantly based on the clinical judgement of the treating doctor. To mirror the real-life situation, this

review was focussed on variables available to the medical team within the first 48 hours of injury that are routinely collected in clinical practice.

3.2.2. METHODS

REVIEW PROTOCOL

A full systematic review protocol is given in Appendix 3. Reporting of the review in this chapter follows the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) guidelines.[180,181]

SEARCH STRATEGY

The bibliographic databases PubMed MEDLINE and EMBASE were searched for papers covering a 20-year period from 1992-2011 inclusive. Hand searching was also performed of major trauma journals, with additional snowballing from identified citations and known reviews. Identified papers were English language studies covering traumatic injury in human patients.

The following search term was used for MEDLINE:

(((Multiple organ failure [MeSH Terms]) OR exp. organ failure OR exp. organ dysfunction) AND ((Multiple trauma [MeSH Terms]) OR exp. polytrauma OR exp. severe injury) AND ((Humans [MeSH] AND English [lang] AND ("1992" [PDat] : "2011" [PDat]))

This was adapted for the EMBASE search facility as follows:

1. Exp *Multiple organ failure/
2. Exp *trauma/
3. 1 and 2
4. Limit 3 to Human, English and 1992-2011

Electronic citation alerts were set-up, and a final search was conducted on 15th May 2012.

INCLUSION/EXCLUSION CRITERIA

Only primary quantitative research was included, to avoid the potential for bias from lower levels of evidence. For inclusion, studies were required to report hospital data from adult trauma patients within 48 hours of injury, and have MOF as a stated outcome measure. Major trauma needed to be defined by ISS, NISS, defined injury criteria to multiple body regions or specific clinical criteria. MOF needed to be defined by a recognised scoring system or physiological variables.

Studies needed to assess an association between MOF and clinical variables available in routine practice that were present or measured within 48 hours of injury. Variables were excluded if they were obtained using non-standard laboratory tests or procedures. Studies involving pre-hospital data, non-trauma-specific ICU populations or specific trauma subsets such as isolated head injury, burns or the elderly were excluded. A summary of the inclusion/exclusion criteria for the review is given in Table 3.2.

	Inclusion criteria	**Exclusion criteria**
Research type	Quantitative primary research	Case reports, case series, expert opinion, discussion, anecdotal evidence, conference abstracts
Timing	Hospital data < 48 hours post-injury	Pre-hospital data > 48 hours post-injury
Population	Major trauma Adults	Not exclusively major trauma Extremes of age Specific trauma subsets
Variables studied	Clinical factors available in routine practice	Laboratory-based experimental parameters Novel technologies
Outcome measure	MOF defined by scoring system or physiological parameters	Single organ failure MOF without a stated definition

Table 3.2: Inclusion/exclusion criteria for literature review

STUDY SELECTION AND DATA EXTRACTION

Citations obtained from the search strategy were screened for duplicates and relevance. Abstracts were then screened for eligibility against the inclusion/exclusion criteria. The full texts of remaining papers were assessed for both eligibility and a basic quality assessment.

Data extraction was performed onto a standardised spreadsheet recording basic study details, population definitions, methodology, population statistics and findings, with a full list of headings given in the review protocol in Appendix 3. Variables associated with post-trauma MOF were grouped into patient, injury, physiological, laboratory and management variables. The timing of variable measurement was noted. Positive and negative findings on both univariate and multivariate analyses were recorded.

Due to heterogeneity in study type, focus and data reporting, quantitative synthesis was not possible. A qualitative review with descriptive summary measures was therefore performed.

QUALITY ANALYSIS

Quality analysis was performed using the tools from the Critical Appraisal Skills Programme (CASP). An initial assessment was made using the CASP screening questions, followed by an extended secondary process using the full CASP tool appropriate for the study type. This included appraisal of potential sources of bias, and of precision and clarity of results. The final quality assessment was scored on a scale of 1 – 3 for each main category, with 3 indicating the highest quality. Scores were not summed due to the potential problems in use of aggregate scores for quality analysis.[182]

3.2.3. RESULTS

All results tables are given in Appendix 4.

PAPER SELECTION

The search strategy identified 479 citations, of which 54 were duplicates. 363 of the 425 abstracts screened did not meet the inclusion/exclusion criteria, leaving 62 full papers for assessment. 34 of these were not eligible on inclusion/exclusion criteria following assessment of the full paper. No papers were considered ineligible due to the screening quality assessment. This resulted in 28 studies for inclusion in the review.[97,98,103,144,147,149,155,160,164,183–201] The flow diagram of study selection is given in Figure 3.1.

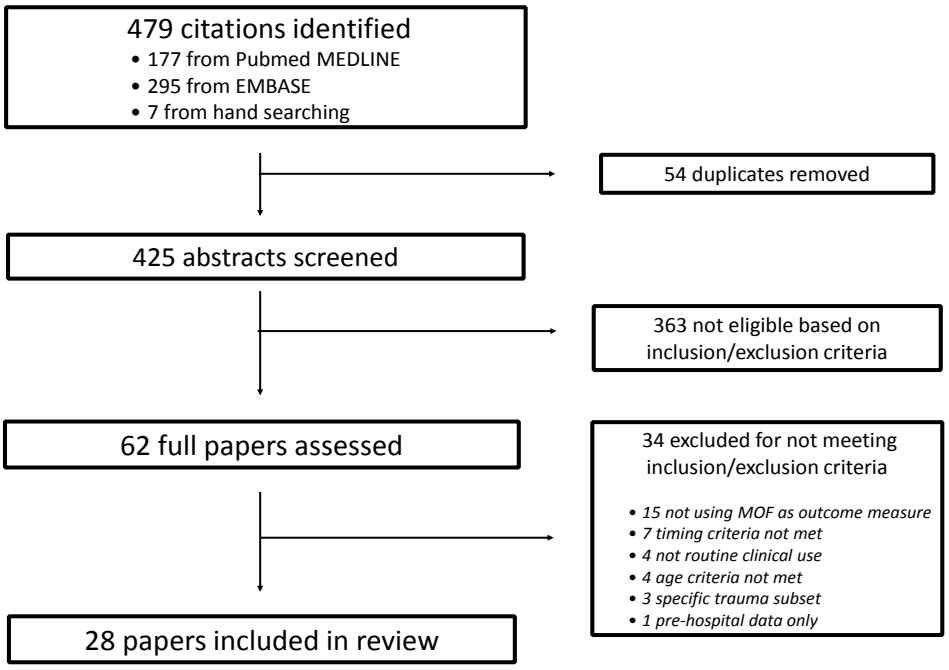


Figure 3.1: **Flow diagram of study selection (as per PRISMA guidelines)**

POPULATION DEFINITIONS

Specific age criteria were defined in 22 of the studies, with the remainder stated as adult,[160] skeletally mature[198] or with no named criteria. Seven studies specified an upper age limit, which ranged from 65 years[201] to 90 years.[196,202] ISS or the AIS components were used to define major trauma in 18 studies. The lower limit for ISS ranged from 9[186,195] to 25.[198] An ISS of 15 was the most frequently used cut-off value, seen in 9 of the studies.[97,147,149,163,184,192,200,203,204] Additional injury definitions included the need for ICU admission,[163,187,191,192,204,205] activation of a transfusion protocol[189] or specific injury criteria.[183,185,197,198]

A minimum patient survival time was specified in 17 studies, the majority specifying 48 hours.[97,149,160,163,184,192,199,203–206] Primary admission[201,206] or early admission after trauma[97,147,185,192,195,197] were widely specified. Other inclusion criteria seen in multiple papers were hypotension[185,190,196,198–200,202,207] and early blood transfusion.[185,186,189,190,196,199,200,202,207] Head injury was a common exclusion criterion, specified in seven papers.[149,185,192,196,199,202,203] Other population exclusions were dependent on the study focus.

MOF DEFINITIONS

The most common MOF scoring system was the Denver Postinjury MOF score, used in 12 studies.[97,147,149,160,163,184,189,192,198,200,203,204] Other scores used were the Marshall Multiple Organ Dysfunction Score (Nine studies)[185,190,196,199,201,202,205–207] the SOFA score (Three studies),[186,187,193] and the Goris score (Two studies).[183,195] Two studies used defined individual criteria for each organ system.[191,197] A single study used the Knaus score with

Fry's additional definitions for liver and gastrointestinal systems in addition to the Marshall score.[205]

Common additional outcome measures included mortality in 17 studies (61%),[160,183–187,189,190,193,195–198,202,203,205,207] sepsis or infection in ten studies (36%),[186,187,189,190,195,196,201,202,206,207] and ICU stay in four studies (14%).[186,198,203,206]

POPULATION STATISTICS

Overall patient numbers ranged from 63[200] to 8724.[193] Combined analysis of age and injury data was not possible due to differences in result presentation and a lack of access to primary data. All study populations showed male sex predominance, ranging from 64%[190,207] to 90%.[189] Blunt trauma was a stated inclusion criterion in five studies,[190,196,199,202,207] and accounted for a range of 55%[189] to 96%[193] of patients in the remaining 15 studies reporting this data. The incidence of MOF varied widely from 3%[201] to 56%,[195,199] with a median value of 25%. Mortality was reported in all but one study, and ranged from 2%[149] to 53%.[189]

VARIABLES ASSOCIATED WITH MOF

- **PATIENT VARIABLES**

Age was associated with MOF in ten studies on univariate analysis[97,149,163,183,184,187,192,200,204,205] and eight studies on multivariate analysis.[97,147,163,184,187,192,203,204] A cut-off value of 55 years was used to create a binary variable in six of the studies.[147,163,184,187,203,204] Frink et al[206] found no association between age and MOF development on univariate analysis, whilst Sauaia et al[147] found no

association on univariate analysis, but did demonstrate age as an independent predictor of MOF on multivariate analysis.

Male sex was identified as an independent risk factor for MOF in two multivariate analyses.[196,202] However, no association was found between male sex and MOF in four univariate analyses[97,163,187,204] and one multivariate analysis.[205] Both Oberholzer et al[195] and Frink et al[206] demonstrated an association with MOF only in males with a high ISS.

Ciesla et al[203] showed obesity to be an independent risk factor for MOF after adjustment for age, injury severity and blood transfusion, with obese patients slower to resolve metabolic acidosis than non-obese patients despite similar volumes of IV fluid used in resuscitation.[199] Two studies found increasing body mass index (BMI) to increase risk of MOF even if within the normal range.[190,196] Moore et al[204] showed an association between MOF development and pre-existing heart disease, while co-morbidities were not associated in three studies.[97,147,163]

- **INJURY VARIABLES**

ISS was identified as a risk factor for MOF in ten studies on univariate analysis[97,147,149,163,184,187,192,195,204,205] and ten on multivariate analysis,[97,147,163,184,187,192,196,202–204] with an ISS of 25 being commonly used as a cut-off value for creation of a binary variable. Direct comparison of ISS and NISS showed similar[187] or improved[184] prediction with NISS. Whilst one study showed an association of MOF with severe head injury (AIS ≥ 3) on multivariate analysis,[187] six further studies showed no association between MOF and admission or early GCS.[97,147,163,202,204,206] Two studies showed an association on univariate analysis that was not repeated on multivariate analyses.[187,204]

The RISC score was shown to be a consistent independent predictor of MOF.[186] Other injury variables, including mechanism of injury,[97,147,149,163,184,204] time from injury to admission[97] and a prolonged extrication time[187] were not associated with subsequent MOF development in any studies.

- **PHYSIOLOGY VARIABLES**

MOF was associated with systolic hypotension on admission in three studies on univariate analysis,[163,187,204] but no association was found on multivariate analyses.[185,187,196] Beilman et al[185] showed hypothermia in the first 6 hours following injury to be associated with increased risk of MOF, even after controlling for fluid administration, transfusion of blood products and an increased BD. A single study showed cardiac index on admission to be predictive of MOF, but only univariate analysis was conducted.[160] Ciesla et al[149] found an association between organ dysfunction in the first 48 hours and subsequent MOF.

The SIRS criteria were linked with subsequent MOF in a single study.[191] The APACHE II and III scoring systems were found to be predictive of MOF on multivariate analyses,[97,187,196,205] APACHE II on admission[187] and APACHE III at 24[205] and 48 hours. [97]

- **LABORATORY VARIABLES**

Lactic acidosis, as indicated by a raised lactate or negative BD, was associated with MOF in multiple studies.[97,147,160,163,204,205] Sauaia et al[97] showed an elevated lactate to be an important independent predictor of subsequent MOF in models derived at 12, 24 and 48 hours post-injury. The combination of a BD > 8mEq/L and a lactate > 2.5 mmol/L in the first 12 hours was demonstrated to strongly predict MOF development in three studies, with persistence of the lactic acidosis from 13-24 hours conferring additional risk.[147,163,204]

A number of studies did not find BD a significant predictor when the contribution of other variables such as age, injury severity, hypothermia and transfusion were taken into consideration.[184,185,196,202]

Associations between markers of coagulopathy and MOF were found on univariate analyses only.[191–193] Platelet count was consistently associated with MOF at time points from admission to 48 hours in a single modelling study.[97] Hyperglycaemia in the first 24 hours was associated with MOF development.[202] Brattstrom et al[187] found an association between MOF and raised blood ethanol on admission, while no association was found by Zeckey et al[201]

- **MANAGEMENT VARIABLES**

The need for blood product transfusion was associated with MOF in multiple studies on both univariate[147,149,163,184,187,189,192,204–207] and multivariate analyses.[97,147,163,184,187,189,192,196,202–205,207] Transfusion of Packed Red Blood Cells (PRBCs) of > 6 units in the first 12 hours was predictive of MOF in several studies,[147,163,187,202–204] with the need for massive blood product transfusion predictive up to 24 hours.[97,189,205] Zallen et al[200] was the only study not to find an association, and the small population numbers increase the risk of a Type 2 error.

A high ratio of FFP to PRBCs was predictive of MOF on multivariate analysis in two studies[192,207] and on univariate analysis in a single study in a lower-risk trauma group.[186] A single study showed association between MOF and the age of transfused blood, with increasing risk with units aged > 14 or > 21 days.[200] Watson et al[207] found an increased risk of MOF with increasing FFP use, but no relationship with platelet or cryoprecipitate transfusion.

The use of vasopressors or inotropes was predictive even after accounting for intravenous (IV) fluid and blood product administration[186,196] with predictive value up to 48 hours.[97] Borgman et al[186] found an association between the need for emergency operative intervention and subsequent MOF, whereas Tuttle et al[198] found no differentiation in risk with the two surgical strategies of ETC and Damage Control Orthopaedics (DCO). One study showed no relationship between the timing of intubation and MOF development.[197] Total IV fluid administration was associated with MOF on univariate but not multivariate analysis in one study,[187] but was associated on multivariate analysis in a different study.[186]

QUALITY ANALYSIS

Wide variation in study quality was seen. Adjustment for potential confounders varied significantly, and descriptions of the extent and duration of follow-up were predominantly unclear. Reporting of results varied in both clarity and precision, with only four studies obtaining high scores on quality analysis .[97,147,192,202]

A number of the studies reported only positive associations, indicating the potential for bias from selective reporting.

3.2.4. DISCUSSION OF REVIEW FINDINGS

The 28 studies included in the review showed a wide range of variables have been associated with post-trauma MOF. Comparing data from these studies has highlighted both areas of broad consensus and areas of debate. A number of variables related to the patient and the injury could be used to assess patient risk on or soon after arrival. Age, obesity and injury severity showed a consistent association with MOF. Evidence for an

association between MOF and sex was less conclusive, although two studies demonstrated a decreased incidence of MOF in pre-menopausal females, associated with lower plasma cytokine levels.[196,202] These studies postulated a potential immunomodulatory effect of oestrogens, but levels of sex steroid hormones in trauma remain unmeasured, with no in vitro evidence to support this assumption.

The "terrible triad" of hypothermia, coagulopathy and acidosis were consistently demonstrated to be risk factors for MOF. Global hypoperfusion is indicated by increasing blood lactate and a metabolic acidosis, but the cause-and-effect of these parameters has been debated, with early derangements often attributed to inadequate resuscitation.[149] However, Blow et al, demonstrated an increased risk of MOF in patients with persistent hypoperfusion despite aggressive treatment protocols to normalise their parameters, suggesting an underlying physiological derangement.[160]

No significant evidence was found to support many variables routinely considered markers of poor outcome, such as pre-existing chronic disease, injury mechanism, prolonged extrication times or a need for emergency intubation. The lack of evidence on surgical decision making was marked, with few studies examining its impact, and with both the duration of initial operative intervention and the differing strategies of ETC and DCO having no impact on MOF incidence.[198,206]

A number of significant limitations exist with this review. Comparison between studies is complicated by the lack of a standard definition for a trauma population, and for MOF itself. Whilst cross-study consensus indicates the importance of increasing ISS in relation to MOF, the cut-off value used to denote severe trauma varied widely. The definition of MOF varied widely, and while the Denver and Marshall scores were the most commonly

used, comparison of the two has shown a marked increase in MOF incidence with use of the Marshall score.[208,209] Standardisation of population definitions and outcome measurement would vastly improve the ability to compare study results.

Interpretation of the results must also take into consideration the quality of the studies suitable for review inclusion. The problems of trauma research are well documented, leading to a reliance on cohort studies. The majority of the studies suitable for review inclusion were Level III cohort studies, and there was little formal reporting of extent and duration of follow-up and of allocation for missing data, leading to low scores on quality analysis. Additionally, a number of papers came from the same research groups, with four first authors being responsible for more than one study. It is unclear without access to primary data whether any of these involved overlapping cohorts of patients.

Limitations must also be considered at review level. Efforts were made to identify all relevant studies fitting the criteria, but individual author contact for unpublished data was not performed. Criteria for quality assessment were based on the recognised CASP system, but these remain subjective.

3.2.5. CANDIDATE INPUT VARIABLES FROM LITERATURE REVIEW

Candidate input variables in each of the categories of patient, injury, physiology, laboratory and management variables were identified from the literature review (Table 3.3). A pragmatic decision was taken to include all variables that had a positive association with MOF from the review, with no attempt to assign increased importance to variables identified in multiple studies.

Variable category	Variable sub-category	Variables
Patient	Basic patient data	Age BMI Sex
	Chronic health conditions	Heart disease
Injury	Injury severity	ISS NISS Head injury
Physiology	*No sub-categories*	Temperature Cardiac index SBP APACHE II APACHE III SIRS RISC score
Laboratory	Haematology	APTT FDPs Fibrinogen Platelet count PT
	Biochemistry	BD Ethanol Glucose Lactate
Management	Transfusion	PRBCs FFPs Platelets Cryoprecipitate Age of transfused blood
	Emergency surgery	Initial operation
	Emergency intervention	Vasopressors/inotropes IV fluid volume

Table 3.3: Candidate input variables identified from literature review

3.3. QUESTIONNAIRE ANALYSIS

3.3.1. METHODS

A simple checklist questionnaire was created with over 100 potential candidate variables, based on a wide range of clinical data. These were selected from an overview of trauma literature, and via discussion with a small panel of trauma professionals. The list was designed to cover both a range of variables and a range of timings. A copy of the questionnaire is given in Appendix 5.

Questionnaires were distributed to 60 trauma and ICU professionals at a targeted regional academic meeting, with a verbal explanation given in addition to the written instructions on the questionnaire. Respondents were asked to choose around 15 variables that they felt to be most associated with post-trauma MOF. No upper or lower limits were imposed on the number of variables that could be chosen, with 15 variables stated as a guidance figure only. An input variable was considered significant for inclusion in MOF prediction models if it was selected in around 20% of responses.

3.3.2. RESULTS

48 questionnaires were returned, giving an 80% response rate. 33 respondents returned data on their grade, with 88% (29) being consultants or senior trainees. The number of variables chosen ranged from 11-33, with a median of 15. Breakdown of the median number of responses for each category of variable showed an emphasis from respondents on patient and injury variables (Figure 3.2). Only one free-text response was received, which suggested the use of experimental parameters. Nine responses (19%) was considered the threshold to meet for inclusion in the candidate variable list.

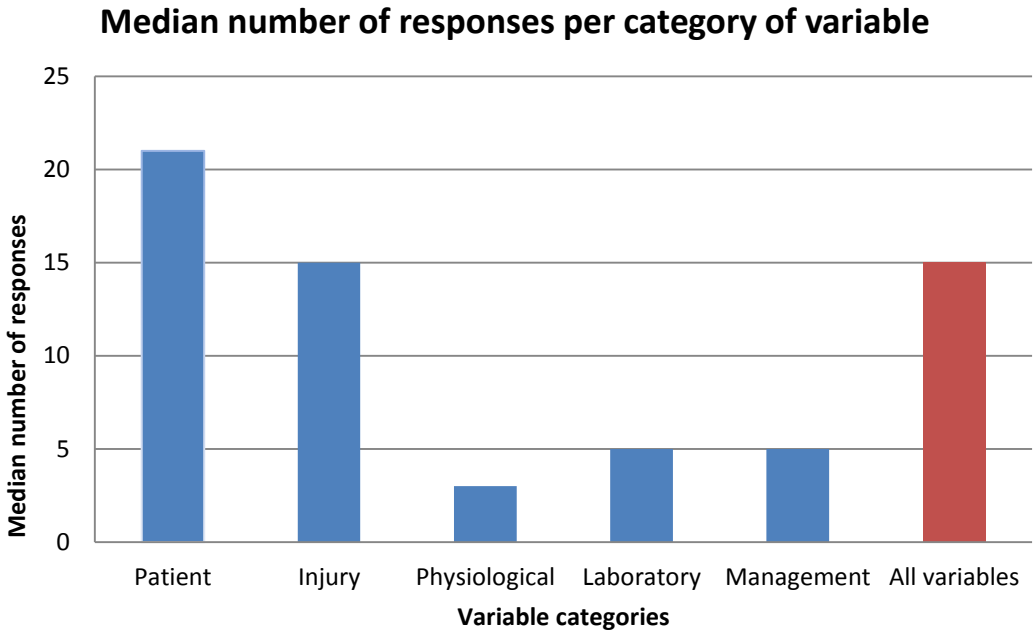


Figure 3.2: Median number of questionnaire responses for each category of input variable

PATIENT VARIABLES

Age and pre-existing cardiovascular (CV) disease both received 34 responses (71%). Obesity and all listed chronic diseases received high numbers of responses. Sex and ethnicity received very few responses (Figure 3.3).

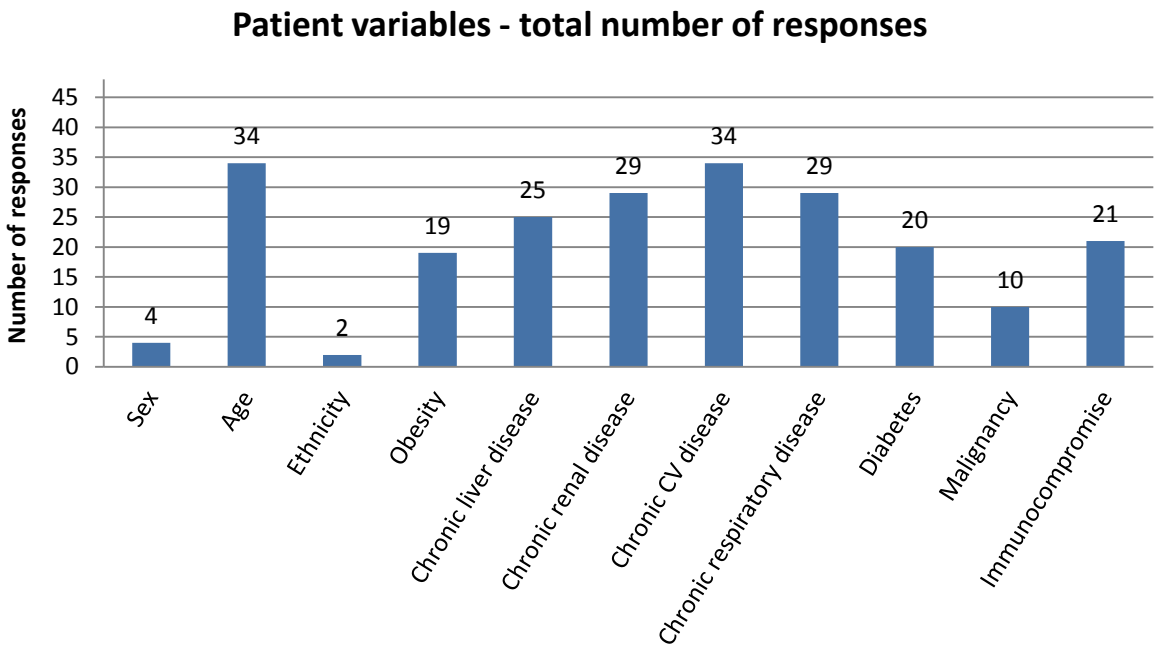


Figure 3.3: Questionnaire responses for patient variables

INJURY VARIABLES

Overall injury severity received 34 responses (71%). Other variables obtaining more than 20% response rate were mechanism of injury (42%), time from injury to hospital arrival (35.4%) and extrication time (19%). Types of injury, including head trauma (27%), chest trauma (40%), and abdominal trauma (33%) also received a significant number of responses. GCS was considered important both on admission (33%) and as the lowest GCS recorded (29%) (Figure 3.4).

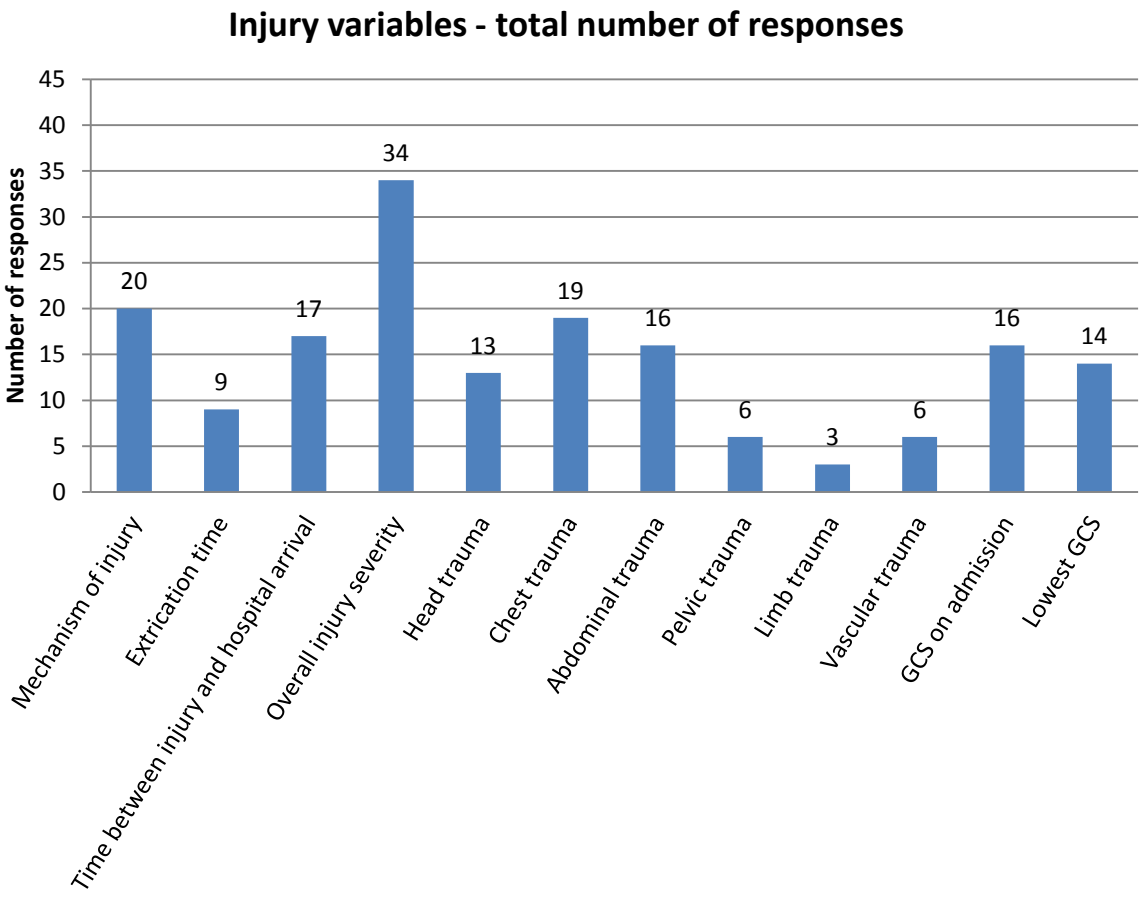


Figure 3.4: Questionnaire responses for injury variables

PHYSIOLOGY VARIABLES

The APACHE II score received a high number of responses (46%) but few other physiological variables received more than a small number of responses. Other variables over the threshold for inclusion as a candidate variable were urine output in the first 24 hours (27%), temperature on arrival (23%), lowest recorded temperature (19%), SBP on arrival (19%) and the SIRS score (21%) (Figure 3.5).

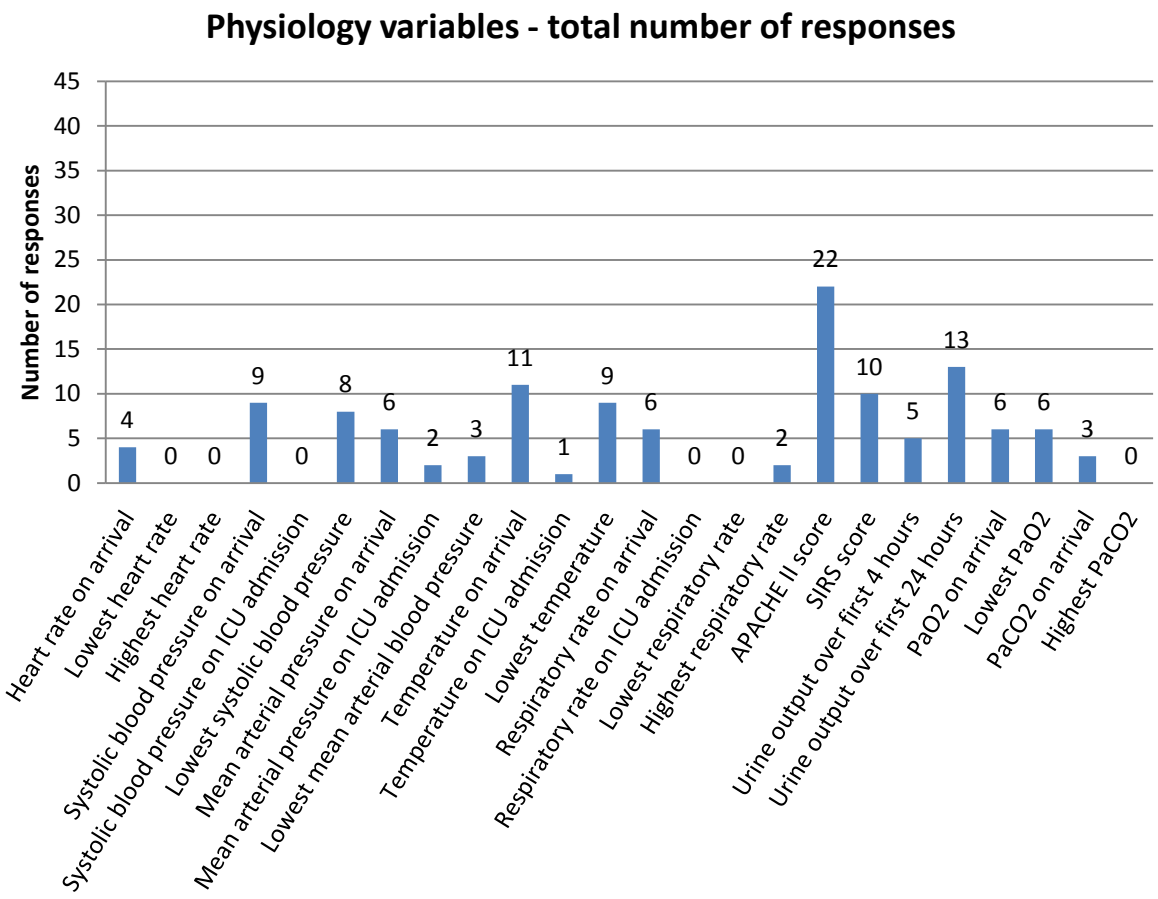


Figure 3.5: Questionnaire responses for physiology variables

Where more than one timing option was given for a variable, results were combined to see if the significance was lost by respondents selecting different timing options. Mean arterial pressure (MAP) and PaO2 received sufficient responses to include in the candidate list when this was done. RR was just below, with 8 (17%) responses.

LABORATORY VARIABLES

Limited consensus was seen for laboratory variables (Figure 3.6). Signs of lactic acidosis were emphasised with higher numbers of responses seen for BD (29%), lactate level (33%) and pH (35%) on hospital arrival, as well as highest lactate (19%) and most acidotic pH (19%) recorded. Creatine kinase (CK) levels received 29% responses.

Combining responses on variables at different time points, haemoglobin levels and creatinine also received sufficient responses for inclusion.

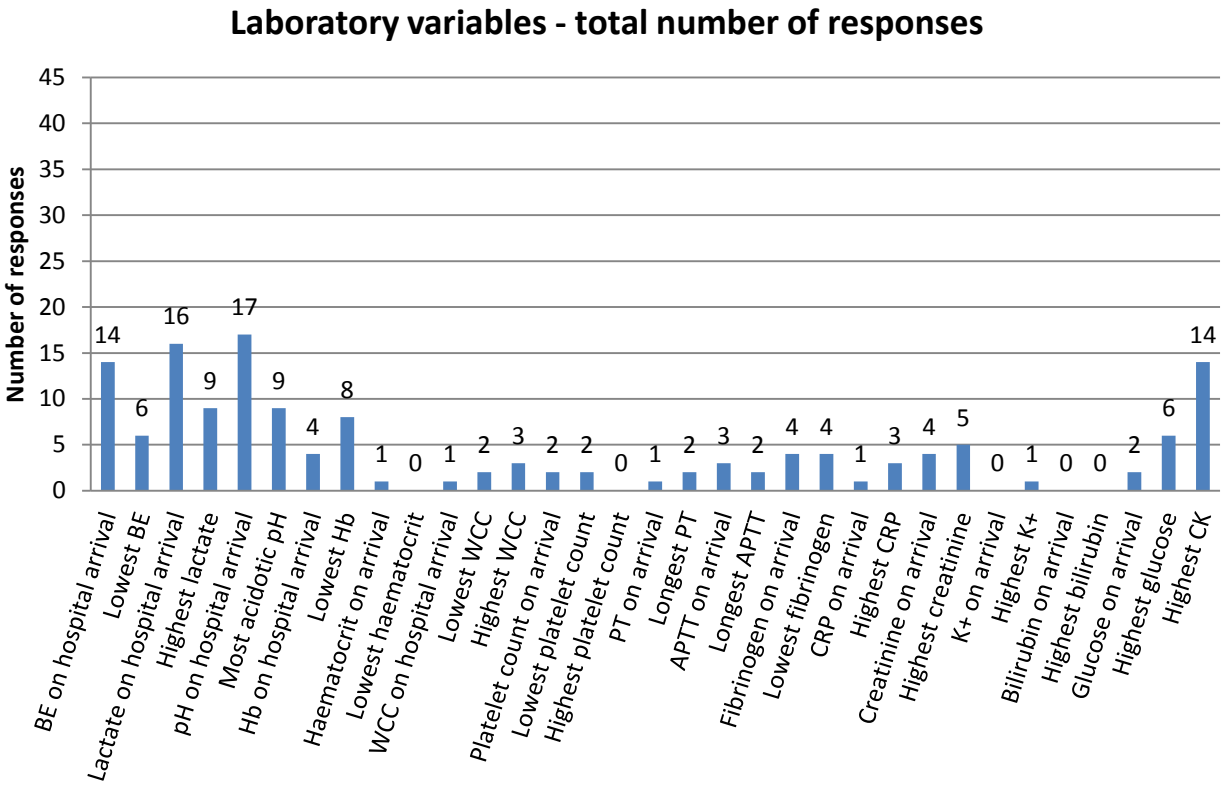


Figure 3.6: Questionnaire responses for laboratory variables

MANAGEMENT VARIABLES

Overall transfusion volume received a high number of responses (48%), with the need for > 4 units of PRBCs (25%) also past the response threshold (Figure 3.7). Other important variables were an initial operative time of > 4 hours (23%), intubation at scene (19%) and ventilation with an Fi02 > 0.4 (19%).

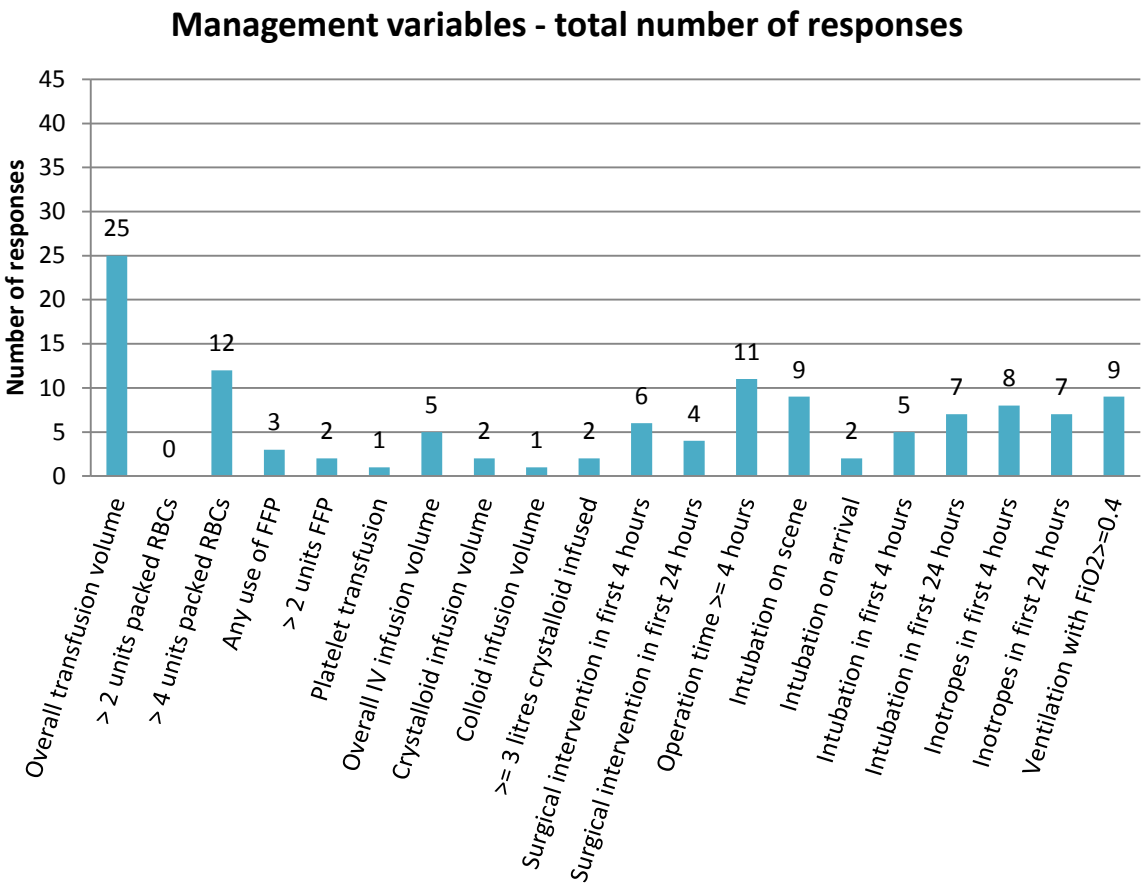


Figure 3.7: Questionnaire responses for management variables

3.3.3. CANDIDATE INPUT VARIABLES - QUESTIONNAIRE ANALYSIS

The trauma and ICU clinicians surveyed in the questionnaire emphasised their use of patient and injury variables in assessment of the risk of post-trauma MOF. Other variables received lower emphasis, particularly laboratory parameters. Across all variables, age, chronic CV disease and injury severity received the highest number of responses. Blood markers of coagulopathy were not emphasised but the clinical results of trauma coagulopathy were, with the need for transfusion receiving a high number of responses.

The questionnaire responses showed limited areas of consensus between clinicians, with a large number of variables receiving a few responses each. This suggests that the variables used by doctors in clinical practice vary widely from clinician to clinician.

The questionnaire process was designed to obtain an indication of the variables used by clinicians to assess patient risk of MOF, but was not designed to be a definitive exploration of variable combinations used. The questionnaire was given to professionals at a regional targeted academic meeting, recognising that this would create limitations in interpretation, as it would reflect practice in one region only. However, this approach allowed a high response rate to be obtained, and involved a large number of clinicians involved in the care of the trauma patients that formed the study cohort. Candidate input variables were identified for each of the five variable categories and are shown in Table 3.4.

Variable category	Variable sub-category	Variables
Patient	Basic patient data	Age Obesity
	Chronic health conditions	CV disease Respiratory disease Renal disease Liver disease Immunocompromise Diabetes Malignancy
Injury	Injury severity	ISS NISS Head injury Chest injury Abdominal injury GCS – admission, lowest
	Injury details	Mechanism Extrication time Time to hospital arrival
Physiology	*No sub-categories*	SBP MAP Temperature APACHE II SIRS Urine output PaO_2
Laboratory	Haematology	Haemoglobin – arrival, lowest Fibrinogen – arrival, lowest
	Biochemistry	BD – arrival, lowest Lactate – arrival, highest pH – arrival, acidotic Creatinine – arrival, highest CK – highest
Management	Transfusion	Transfusion volume PRBCs
	Emergency surgery	Early operation Operation time
	Emergency intervention	Early intubation Inotropes Ventilation $FiO_2 > 0.4$

Table 3.4: Candidate input variables identified from questionnaire analysis

3.4. CANDIDATE INPUT VARIABLE LIST

The final candidate input variable list was compiled using the lists created from the literature review and questionnaire analysis. Each variable was discussed with a panel of trauma specialists, and the overall list reviewed to assess if any potentially important variables had not been identified.

Several scoring systems were highlighted as having an association with MOF, such as APACHE II, SIRS and RISC. However, the association between the individual components of these scores and MOF was not always clear from the available evidence. After significant discussion, it was decided not to use scoring systems but to consider their constituent variables. The reasons for not using overall scoring systems as input variables were as follows:

1. Correlation with MOF is likely since the majority of scoring systems are designed to be predictive of mortality, and MOF is known to be a strong determinant of post-trauma mortality. This would cause issues with co-linearity.
2. The contribution of individual components of a score could not be determined if the overall score was used. This would impose a high data collection burden for any resultant model that included a scoring system, with unproven benefit for many of the required fields. This could also affect the utility of the end product.
3. There would be a need to abide by the data collection rules of the scoring system in question, which may cause conflict and duplication in data extraction.

4. Some scoring systems e.g. APACHE III, are based on proprietary algorithms, and are therefore not freely available.

Scoring systems that had been highlighted as associated with post-trauma MOF were therefore broken down to their individual components, and the candidate input variable list amended accordingly. To avoid confusion over the term “emergency surgery”, a distinction was made between operations and emergency invasive surgical procedures, which may or may not be performed in theatre. After discussion, emergency invasive procedures were defined as those involving intrusion into a body cavity and were then further reduced to chest drain or intracranial pressure (ICP) bolt insertion.

The final list is given in Table 3.5. Whilst timings were investigated in the formation of the candidate list, timings were left open, with an assessment of data availability required to determine at which time-points the input variables could be considered. This will be further discussed in later chapters.

Variable category	Variable sub-category	Variables
Patient	Basic patient data	Age BMI Sex
	Chronic health conditions	CV disease Diabetes Immunocompromise Liver disease Malignancy Renal disease
Injury	Injury severity	AIS regions ISS NISS GCS
	Injury details	Extrication time Injury mechanism Time to hospital
Physiology	*No sub-categories*	BP – SBP and MAP Cardiac Index FiO_2 Heart Rate (HR) PCO_2 PO_2 RR Temperature Urine output
Laboratory	Haematology	APTT FDPs Fibrinogen Haematocrit Haemoglobin Platelet count PT WCC
	Biochemistry	Albumin BD Bilirubin CK Creatinine Ethanol Glucose Lactate pH Potassium (K+) Sodium (Na+) Urea

Table 3.5: Candidate input variables final list

Variable category	Variable sub-category	Variables
Management	Transfusion	Age of transfused blood Any transfusion Cryoprecipitate FFPs Platelets PRBCs Total blood products
	Emergency surgery	Chest drain insertion ICP bolt insertion Intra-operative time Operative intervention Time to first operation
	Emergency intervention	CPR Intubation IV fluids Time to intubation Time to vasopressors/inotropes Vasopressors/inotropes Ventilation

Table 3.5 continued: Candidate input variables final list

CHAPTER 4 MULTIPLE ORGAN FAILURE

4.1. DEFINING MULTIPLE ORGAN FAILURE

The term "multiple organ failure" was first used by Eiseman et al in 1977.[89] Since this time, many attempts have been made to define, categorise and formalise a description of MOF. Despite this, the language used to define MOF has remained inconsistent. Baue reviewed the history of MOF descriptions and found 46 different expressions used to describe the syndrome.[91] These ranged from "whole body inflammation", "malignant intravascular inflammation" and "autocannibalism", to the more commonly used terms of MOF, Multiple Organ Dysfunction Syndrome (MODS) and Multi-Stage Organ Failure (MSOF).

Baue also described the array of systems that have been used to score and define MOF. Whilst stressing that his review was not intended to be definitive, he recorded 20 different MOF scoring systems developed between 1980 and 1996. These differed in the definitions used for failure of each organ system, the measurement of severity, and even in the organ systems that could be considered to be failing. Baue called for a consistent system and language to be used in future research, but none has yet been accepted into widespread use. Previous reviews had noted that inconsistencies in MOF definition and scoring made comparisons between studies difficult, and therefore made it hard to determine if changes in patient care were impacting on MOF incidence.[90]

As was seen in the literature review of input variables in Chapter 2, the use of MOF scoring systems in trauma research has been inconsistent. This greatly reduces the ability

to compare results between studies. Studies in the review demonstrated inconsistencies not just in the MOF scoring system used, but in the application of systems, with different thresholds used to define MOF, and the use of entities derived from the MOF scores as outcome measures. Such derived measures included maximum MOF score, MOF score on a specific post-injury day and summed MOF scores.

In order to address these issues, a review of MOF scoring systems was undertaken, to establish the most suitable outcome measure for this study.

4.2. Review of MOF Scoring Systems

4.2.1. Review Methods

Searching bibliographic databases for information on MOF scoring systems was complicated by the lack of standardised terminology to describe them. At the time of the review, there was no grouping or MesH term in PubMed to account for MOF scoring systems. After the completion of this review, in 2013, the MeSH term "Organ Dysfunction Scores" was added, but has not yet been retrospectively applied. A search of this MeSH term on 10th February 2013 revealed only 7 studies, none of which were descriptions of scoring systems.

A number of attempts were made to formulate a search term on PubMed to allow accurate searching for MOF scoring systems. However, search strategies were either too non-specific, resulting in a high number of citations, and including all papers that used a MOF scoring system, or were too focussed and missed known papers.

Since previous non-systematic reviews of MOF scoring systems existed, an alternative approach was taken. Starting with known reviews, particularly Baue's 2006 paper, an

initial listing was made of MOF scoring systems. Snowball searching from references was then used to identify further scales to add to the listing. Papers were also identified from scoring systems used in papers in the input variables review in Chapter 2, and from generalised reading across literature on outcome measurement in trauma and ICU populations.

Using the list of scoring systems, the definitive and original descriptive paper was obtained for each score. This was then assessed to determine if the score met the following inclusion criteria:

- Stated as designed to diagnose, measure or assess severity of MOF
- Used criteria from more than one organ system
- Had been applied to a human population, and was not just a theoretical construct
- Contained a listing of the organ systems considered
- Contained a description of how organ dysfunction was assessed for each organ system considered
- Contained a description of how measures of organ dysfunction related to numerical or descriptive scoring
- Contained sufficient information for replication of the score in a novel patient population
- Original paper described the score in English

No restrictions were placed on the age of the original study, nor on the patient group in which the original score was developed, providing MOF could be reasonably expected as an important outcome in that population. The review therefore aimed to list all MOF

scoring systems available in the present day, before assessing their relevance to research in post-trauma MOF.

Details were taken from the original study describing the score, the definitions of organ dysfunction used for each system, the relationship of these descriptions to scoring, the overall scoring system, and any details on how the score was intended to be used. Significant modifications of scoring systems were considered as separate scores.

4.2.2. RESULTS OF REVIEW

In total, 42 different MOF scoring systems were identified.[35,36,92,98,114,116,149,174,175,188,210–241] These were developed between 1969 and 2006, with two providing descriptions of the physiological process of organ failure prior to Eiseman's first use of the term "multiple organ failure".[210,211] A large proliferation of new scoring systems was seen in the 1980s and 1990s, with few scores developed in the past 15 years. The two most recent scores[149,188] were modifications of previous scoring systems, and no new scores were found after 2006. A complete listing of all MOF scoring systems identified in order of the date of first reference, with breakdown of the individual organ systems criteria is given in Appendix 6.

Differences were noted in the number and composition of organ systems used to define MOF. For each organ system, wide variations in the variables used to measure dysfunction were seen.

ORGAN SYSTEMS

Marked differences were seen in the organ systems used to compose the MOF scores. Most scores used combinations from respiratory, cardiovascular (CV), renal, hepatic,

haematological, gastrointestinal (GI), central nervous system (CNS) and endocrine systems. Other systems seen were measures of infection or immune status,[214,222,238,239] or "shock".[229] Tilney et al divided the GI system into two components – upper and lower.[211]

The total number of systems used varied between three[114,210] and eight,[222,229,231] with a median of six. A graphical summary of the organ systems used across all 42 scoring systems is given in Figure 4.1.

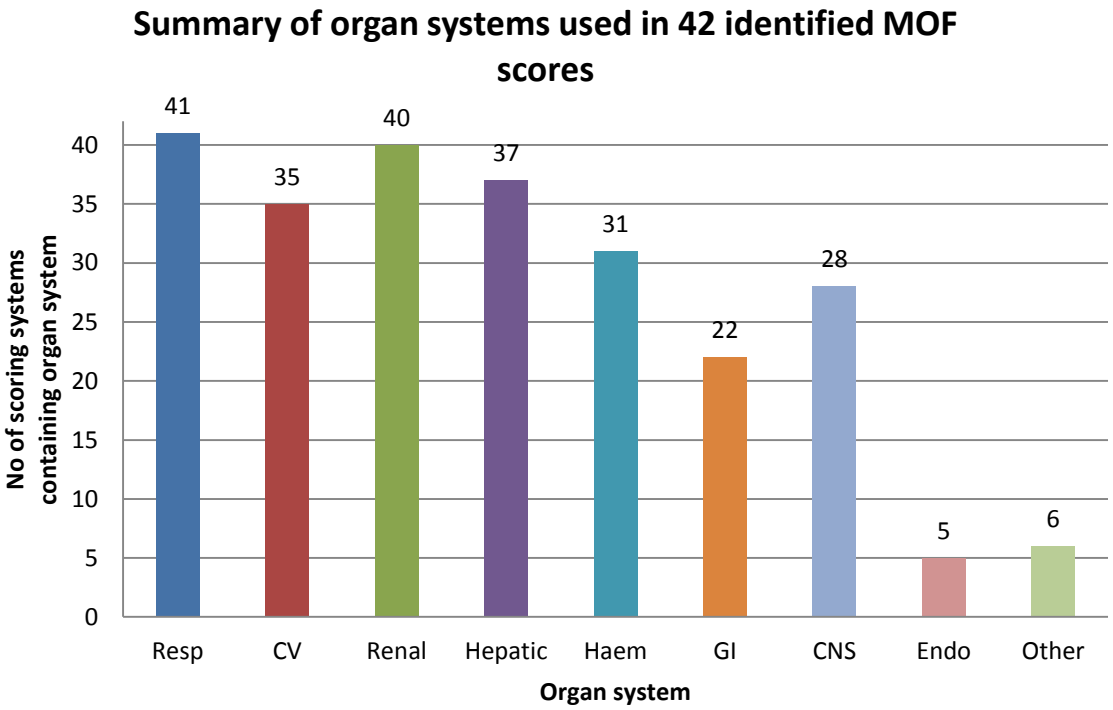


Figure 4.1: Summary of the different organ systems used in the 42 MOF scoring systems identified

Some consensus could be seen in the organ systems considered important in defining MOF. Use of measures of respiratory dysfunction were seen in 41 of the 42 scoring systems, with Bell et al (1983) being the only score not to include respiratory failure. Renal dysfunction was measured in 40 of the 42 scales, and in all scales from 1980

onwards. Cardiovascular and hepatic dysfunction were commonly measured, being seen in 35/42 (83%) and 37/42 (88%) of scores respectively. CNS dysfunction was measured in 28 scoring systems (67%) and GI dysfunction in 22 (52%). Endocrine and other organ systems were used far less frequently. A detailed summary of the distribution of the organ systems used across different scores is given in Table 4.1.

Some evidence of convergence over time was seen, with the majority of recent scores containing measures of respiratory, cardiovascular, renal and hepatic dysfunction. GI dysfunction was common in earlier scoring systems, but not present in the eight most recent scores (Table 4.1).

Scoring system	Resp	CV	Renal	Hepatic	Haem	GI	CNS	Endo	Other	Total
Skillman et al, 1969	x	x		x						3
Tilney et al, 1973	x	x		x		x	x	x	Upper and lower GI	7
Fry et al, 1980	x		x	x		x				4
Faist et al, 1983	x	x	x	x	x	x				6
Pine at al, 1983	x	x	x	x			x			5
Bell et al, 1983		x	x	x	x	x	x	x		7
Marshall WG and Dimick, 1983	x	x	x						Sepsis	4
Stevens, 1983	x	x	x	x	x	x	x			7
Manship et al, 1984	x	x	x	x	x					5
Goris et al, 1985	x	x	x	x	x	x	x			7
Knaus et al, 1985	x	x	x		x		x			5
Norton, 1985	x		x	x	x	x	x			6
Maetani et al, 1986	x		x	x	x	x				5
Carrico et al, 1986	x	x	x	x	x		x	x		7
Shen & Zhang, 1987	x	x	x	x		x	x			6
Bihari et al, 1987	x		x	x	x		x			5
Marshall et al, 1988	x	x	x	x	x		x	x	Immunology	8
Darling et al, 1988	x	x	x	x	x	x	x			7
Bumaschny et al, 1988	x	x	x		x	x				5
Rubin et al, 1990	x		x	x	x		x			5
Dorinsky and Gadek, 1990	x	x	x	x	x	x	x			7
Tran et al, 1990	x	x	x	x	x	x	x			7
Bone, 1991	x		x	x	x		x			5
Villar et al, 1991	x	x	x	x	x	x	x		Shock	8
Ruokonen et al, 1991	x	x	x	x	x	x	x			7
Moore et al, 1991	x	x	x	x	x	x	x	x		8
Henao et al, 1991	x	x	x	x	x	x				6
Shoemaker et al, 1992	x	x	x		x		x			5
Martin et al, 1992	x	x	x	x	x	x				6
Dobb, 1992 (Royal Perth Hospital)	x	x	x	x	x	x	x			7
Hebert et al, 1993	x	x	x	x	x	x	x			7
Kollef M, 1993	x	x	x	x	x	x	x			7
Fagon et al, 1993 (ODIN)	x	x	x	x	x		x		Infection	7
Saffle et al, 1993	x	x	x			x			Infection	5
Marshall et al, 1995 = MODS score	x	x	x	x	x		x			6
Moore et al, 1996	x	x	x	x						4
SOFA - Vincent et al, 1996	x	x	x	x	x		x			6
LOD System, Le Gall et al, 1996	x	x	x	x	x		x			6
Waydhas et al, 1996	x		x	x						3
Brussels score, 1997	x	x	x	x	x		x			6
Denver score modification, 2004	x	x	x	x						4
Denver score current modification, 2006	x	x	x	x						4
Total number of scores using organ system	41	35	40	37	31	22	28	5	6	**Median = 6**

Table 4.1: Summary of organ systems used in 42 MOF scoring systems (Papers in chronological order)

MEASURING ORGAN FAILURE

Even for scores using the same organ systems, considerable variation was seen in the parameters used to measure organ failure. Often several measures were used in combination within the same score. For the commonest seven organ systems, all the measures of dysfunction seen were collated, and are given in Tables 4.2 to 4.5.

Three main categories of organ dysfunction measures were seen – physiological, laboratory and clinical parameters. Respiratory dysfunction was predominantly measured by physiological parameters, although the presence of ARDS and the need for ventilation were commonly used. Measures of cardiovascular dysfunction included both physiological, laboratory and clinical parameters (Table 4.2). The use of physiological parameters was not seen in any other organ systems.

	Respiratory	**Cardiovascular**
Physiological	Vital capacity Alveolar-arterial gradient Dead space: Tidal Volume Alveolar hypoventilation Fraction of inspired oxygen (FiO2) Positive End-Expiratory Pressure (PEEP) Respiratory rate PCO_2 PO_2 Respiratory alkalosis Pulmonary Artery Wedge Pressure (PAWP) PaO_2/FiO_2 ratio	Systolic blood pressure Mean arterial pressure Right atrial pressure PAWP Cardiac output Cardiac index Heart rate
Laboratory	N/A	Acidosis
Clinical	Mechanical ventilation ARDS Signs of barotrauma Signs of pulmonary oedema	Arrhythmias ECG changes Signs of circulatory inadequacy Cardiac damage at post mortem Signs of congestive heart failure Use of inotropes, diuretics or catecholamines Myocardial infarction Signs of cardiogenic shock Enterocarditis Need for volume loading Oedema Cardiac arrest Intra-aortic balloon pump use

Table 4.2: Measures of respiratory and cardiovascular failure

The measurement of renal and hepatic dysfunction used both laboratory and clinical parameters (Table 4.3). Over time, the use of clinical parameters decreased, and convergence was seen on the use of creatinine to measure renal dysfunction and bilirubin to measure hepatic dysfunction.

	Renal	Hepatic
Laboratory	Creatinine Creatinine clearance Potassium Urea Nitrogen	Bilirubin Aspartate transaminase or serum glutamic oxaloacetic transaminase (AST or SGOT) Alkaline phosphatase (ALP) Alanine transaminase (ALT) Lactate dehydrogenase (LDH) Prothrombin time (PTT) Icteric index Albumin Gamma glutanyl transferase (GGT)
Clinical	Haemodialysis Peritoneal dialysis Oliguria/urine output Azotaemia Signs of chronic renal failure Plasma apheresis	Jaundice Liver histology Hepatic encephalopathy Need for blood product transfusion Hepatitis Biliary obstruction

Table 4.3: Measures of renal and hepatic failure

Some clinical parameters were used in the measurement of haematological failures, but most scores used laboratory measures. However, there was little consensus seen in the early scores with a wide range of laboratory measures used. In the more recent scores, platelet count was frequently used as a measure of haematological failure (Table 4.4).

A large number of measures of GI failure were seen, with little consensus observed across the 22 scores which included a GI component. The majority of the measures used were clinical, and were often poorly defined (Table 4.5). The absence of GI dysfunction from later scoring systems may reflect the lack of consensus on the best measure of GI failure.

CNS failure was consistently determined by clinical parameters. Whilst a large number of clinical parameters were seen, particularly in the early scores, all of the later scores used GCS as the sole measure of CNS dysfunction.

	Haematological	GI
Laboratory	Platelets Prothrombin time Activated partial thromboplastin time Fibrin degradation products Fibrinogen Euglobulin Disseminated Intravascular Coagulation White cell count Thrombocyte count Haematocrit Immature blood cell forms Positive protamine sulphate test D-dimers Thrombin clotting time	Bilirubin Amylase ALP Haemoglobin
Clinical	Need for clotting factors Need for blood product transfusion Ecchymoses Active bleeding Chronic haematological disease	Bleeding in nasogastric tube Rectal bleeding Need for blood product transfusion Bleeding seen on endoscopy Perforation Pancreatitis Bowel ischaemia Ileus Acalculous cholecystitis Mesenteric venous thrombosis Necrotising enterocolitis Need for parenteral nutrition Diarrhoea

Table 4.4: Measures of haematological and GI dysfunction

	CNS
Clinical	Coma Impaired cognition GCS Confusion Agitation Meningitis Intracranial haemorrhage Infarction Encephalopathy Psychosis Seizures Neuropathy Focal neurological deficits

Table 4.5: **Measures of CNS failure**

4.3. DISCUSSION OF MOF SCORING SYSTEMS

4.3.1. OVERVIEW

The majority of work to develop severity-of-illness scoring systems has originated from research on ICU populations. The major models used in ICU outcome prediction are designed to measure risk of death, and include APACHE II and III, the Simplified Acute Physiology Score (SAPS) and the Mortality Prediction Model (MPM) II. In general, these models are based on logistic regression models using variables collected on the first day of ICU admission and are used to predict in-hospital mortality.[242,243] They have been used to measure ICU performance, including appropriateness and effectiveness of care.[244,245] Use of mortality models allows for case-mix adjustment, to effect comparison between units.[246]

Scoring systems can measure severity-of-illness by two main approaches – the use of measurement of treatment, or of measurement of patient characteristics and physiology. These measures can be combined into single systems, such as with APACHE II and III.

Selection of a good scoring system depends on its intended use, and will depend on the accuracy and goodness of fit of the system. Accuracy is based on score validity and reliability, and goodness of fit on calibration, and ability to discriminate between outcomes.[246]

Organ failure scoring systems provide a different type of ICU scoring to the mortality prediction of the severity-of-illness scores. MOF scoring systems provide a description of the state of a patient at a particular time. Whilst the majority of severity-of-illness scores take one set of data on a patient early in their ICU stay, and use this for prediction, MOF scores can be measured repeatedly and can therefore reflect changes in the patient's condition after the first 24 hours in ICU. MOF scores can be used on a daily basis to determine changes in the patient condition. The variety of purposes for which MOF scores have been used is reflected in the differences seen in their development and structure.

Analysis of all 42 MOF scoring systems indicates both areas of consensus and areas of ongoing debate. A clear distinction exists in the use of physiological or laboratory variables to measure organ dysfunction, and the use of variables representing clinical features or measures of therapeutic intervention. In clinical practice, many decisions are made based on therapeutic variables, but established practice may vary considerably between units, and therefore impact on the value of a score assigned to a patient with the same underlying physiology. For example, the criteria for the threshold for transfusion of blood products may vary, resulting in different levels of severity assigned to a patient, depending on whether laboratory criteria or the act of transfusion itself is used as the determinant of organ system dysfunction.

The cardiovascular system has perhaps provided the most debate on measurement of dysfunction. The use of inotropes is frequently used as a descriptor of dysfunction, but thresholds for inotrope use are largely decided by individual clinicians. A range of physiological measures of cardiovascular dysfunction have been proposed, but no single measure has been accepted. Over time, increasing consensus has been seen in measures of organ dysfunction for other organ systems.

Other issues noted in the development of MOF scoring systems have been the calibration of scores against mortality, the weighting of different organ system dysfunction, the timing and nature of variable measurement, and the use of scores to create secondary outcome measures, such as organ-failure free days.[247] The majority of MOF scores have been calibrated against mortality, but since mortality can occur without MOF, and MOF can occur without mortality, it remains a surrogate measure of the accuracy of a MOF score.

4.3.2. DEVELOPING THE IDEAL MOF SCORE

Early MOF scores were based on the opinions and descriptions of individual clinicians. More recent scores have developed using highly structured processes, in an attempt to ameliorate some of the issues outlined above. Score development has varied from clinical consensus (e.g. SOFA) to literature and database review (e.g. Marshall) to mathematical analysis (e.g. Logistic Organ Dysfunction Score – LODS). The development of the Marshall and SOFA scores will be considered in more detail.

MARSHALL MULTIPLE ORGAN DYSFUNCTION SCORE (MODS)

The Marshall MODS was developed in 1995 by following a two-stage process.[175] Firstly, a literature review was performed to evaluate previous measures of organ dysfunction, and create a list of the optimal descriptors. In the second part of the process, the optimal variables identified in the review were tested in a prospective cohort study.

The systematic review identified 30 papers from 1976 and 1993 that documented diagnostic criteria to define MOF in at least 20 patients. Descriptors of organ dysfunction were assessed for validity using three principles:

1. Construct validity – the extent to which the descriptor was meaningful measure of the characteristic of interest
2. Content validity – the extent to which the descriptor represents the range of possible processes
3. Criterion validity – the extent to which the descriptor predicts outcome

Construct validity was maximised by examining the full range of available literature. For evaluation of content validity, a list of characteristics of the "ideal" descriptor of organ dysfunction was made, (Table 4.6) and identified descriptors were measured against this. Criterion validity was assessed by evaluation against ICU mortality.

Characteristics of the ideal descriptor of organ dysfunction
1. Simple 2. Readily, routinely and reliably measured 3. Objective and independent of clinical judgement 4. An accurate reflection of function in a given organ system 5. Specific for the function of that organ system 6. Unaffected by transient abnormalities 7. Minimally altered by specific patterns of therapeutic intervention 8. Maximally abnormal after resuscitation and well before time of death 9. A reflection of an acute homeostatic insult, not of chronic primary disease 10. Reproducible in a heterogeneous group of critically ill patients 11. Abnormal in one direction only 12. Continuous rather than dichotomous

Table 4.6: Characteristics of the ideal descriptor of organ dysfunction[248]

Following this process, Marshall determined that seven organ systems should be used to define MOF – respiratory, renal, hepatic, haematological, cardiovascular, GI and CNS. For five of these systems, individual physiological variables were identified that met validity criteria. However, descriptors for the GI and cardiovascular systems were described as suboptimal.

For the second part of the process, a database of nearly 700 surgical ICU admissions was used to evaluate candidate variables. This was split into a development set and a validation set. Individual variables were calibrated against survival in the development set, and each variable categorised into groups based on the values associated with increasing levels of ICU mortality. These groups were scores on a scale of 0 to 4, with 0 correlating with an ICU mortality risk of < 5% and 4 correlating with an ICU mortality risk > 50%. These were then tested against the validation set.

The final score excluded the GI system as no variable that could be accurately measured, and met the criteria in Table 4.6, was found to be correlated with mortality. No single

variable for the cardiovascular system correlated consistently with mortality, and therefore a composite measure of cardiovascular function was developed. It was recognised that the use of a composite measure risked improving specificity at the potential cost of sensitivity, and increased the data collection burden. However, the development of the pressure-adjusted heart rate (PAR) measure, derived from heart rate, central venous pressure (CVP) and MAP was deemed by the authors to be the measure of cardiovascular function that best reflected organ dysfunction (Table 4.7).

The overall MOF scoring system was shown to have 93.1% accuracy, with a positive predictive value of 97.8% and negative predictive value of 47.7%. Both the score on ICU admission and the change in score over time were significant predictors of mortality.[175]

MARSHALL						
Organ system	**Variable**	**Score allocation**				
		0	**1**	**2**	**3**	**4**
Respiratory	*PaO_2/FiO_2 (mmHg)*	>300	226-300	151-225	76-150	≤75
Renal	*Creatinine (umol/L)*	≤100	101-200	201-350	351-500	>500
Hepatic	*Bilirubin (umol/L)*	≤20	21-60	61-120	121-240	>240
Cardiac	*PAR =(HR/MAP)xCVP*	≤10.0	10.1-15.0	15.1-20.0	20.1-30.0	>30.0
Coagulation	*Platelets (x $10^3/mm^3$)*	>120	81-120	51-80	21-50	≤20
CNS	*GCS (Scale)*	15	13-14	10-12	7-9	≤6

Table 4.7: The Marshall Multiple Organ Dysfunction Syndrome Score[175]

SOFA SCORE

The SOFA score was developed in 1994 via a consensus group of clinicians on behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine.[174] Several principles were agreed in advance of variable selection. Firstly, that

organ dysfunction should be seen as a continuum. Secondly, that organ dysfunction could evolve over time and so scoring should be able to be repeated. Thirdly, that measurement of MOF should be based on a small number of simple and objective variables that could be easily and routinely measured in all ICUs. Ideally, variables should also be independent of therapeutic interventions.

Six organ systems were chosen for consideration – respiratory, coagulation (haematological), hepatic, cardiovascular, CNS and renal. The chosen variable for each was categorised according to degrees of severity, and scored as 1-4 (Table 4.8). Of note, the SOFA score was designed to describe MOF rather than predict mortality. However, retrospective database studies have shown a correlation between the scores for each organ system and mortality.[174]

The principle to exclude therapeutic interventions was broken for the cardiovascular system, since no single descriptor was considered to be suitable by the consensus group. Measurement of cardiovascular dysfunction was determined by a mixture of MAP and inotrope use.

Due to the descriptive design of the SOFA score, the original authors did not evaluate its predictive accuracy. This has been shown by later studies to be similar to the Marshall score.[249]

SOFA					
Organ system	**Variable**	**Score allocation**			
		1	**2**	**3**	**4**
Respiratory	*PaO_2/FiO_2 (mmHg)*	<400	<300	<200 AND respiratory support	<100 AND respiratory support
Renal	*Creatinine (umol/L)*	110-170	171-299	300-440	>440
Hepatic	*Bilirubin (umol/L)*	20-32	33-101	102-204	>204
Cardiac	*Inotropes (ug/kg/min)*	(MAP < 70mmHg)	Dopamine ≤ 5 OR Dobutamine any dose	Dopamine > 5 OR adrenaline ≤ 0.1 OR noradrenaline ≤ 0.1	Dopamine > 15 OR adrenaline > 0.1 OR noradrenaline > 0.1
Coagulation	*Platelets ($x\ 10^3/mm^3$)*	<150	<100	<50	<20
CNS	*GCS (Scale)*	13-14	10-12	6-9	<6

Table 4.8: The Sequential Organ Failure Assessment Score[174]

4.3.3. USE OF SCORING SYSTEMS

Despite the different methods of development, the Marshall and SOFA scoring systems involve the same six organ systems, and the same variables to measure dysfunction, with the exception of the cardiovascular system. There are differences in the cut-off values for each variable and the method of allocation to a score. Examination of all of the available MOF scores showed marked convergence over time. Other recent scores, including the Brussels score, Logistic Organ Dysfunction Score (LODS) and the newest version of the Denver score also use very similar variables. However, a consensus on a single MOF scoring system has yet to be reached.

Part of the issues in arriving at a consensus on a single score has been in the range of measures that can be derived from a scoring system, and the ways in which that score is

used. The developers of the SOFA score aimed for it to be used as a descriptive parameter, and therefore did not initially set a cut-off value to denote organ failure. Subsequent studies by the original authors suggest a value of 3 for a single system to indicate single organ failure (SOF). In addition, the maximum SOFA score was found to correlate with mortality.[250] The authors stopped short of suggesting a score to indicate MOF, but later advocated use of the ICU admission SOFA score, and a delta score representing the difference between maximum and ICU admission SOFA scores.[251] Other studies have found predictive value from the mean SOFA score, the SOFA score at 48 and 96 hours after admission and the change in SOFA score at 48 hours.[252]

SOFA has been used both as a severity-of-illness score to predict mortality on admission, and as an organ failure score. A review of uses of the SOFA score showed similar predictive ability to severity-of-illness scores with admission SOFA, but improved prediction if the two were used in combination. However, the review authors concluded that differences in study design precluded recommendations on which of the SOFA derivatives (maximum SOFA, change in SOFA, mean SOFA etc.) were most appropriate.[253]

The Marshall score was initially described as taking the total of the worst scores for each organ system over the patient stay, to create a maximum score that was correlated with mortality. However, it has subsequently been used as a baseline measure on admission, as serial values, and as a binary field with a cut-off of >5 taken to indicate MOF.[164,254,255] There are also differences in the data collected, with the original data collection rules for SOFA specifying the worst values for each variable, and Marshall a representative value. These original data collection rules are not always followed in subsequent studies.

4.3.4. CHOOSING A MOF SCORING SYSTEM FOR TRAUMA RESEARCH

Given the problems outlined above, which MOF scoring system should be used to define MOF in trauma research? Choosing a score must be based on the utility of the scoring system itself, its use in other studies, its applicability to trauma patients and the way in which it would be used in this study.

From the noted convergence over time, and in the absence of a long-accepted scale, the use of a recent scoring system seemed most appropriate. Formal citation analysis was complicated by the vastly different uses of MOF scoring systems and the time range over which they were developed, but informal analysis showed the SOFA and Marshall scores were in common use in ICU research, with the Denver score seen frequently in trauma research. This was reflected in the distribution of use of MOF scoring systems in the review in Chapter 3, where at least one of the Denver, SOFA or Marshall scores was used in 86% of the included papers. It was therefore decided to focus on these three scores and assess their use in trauma.

Few attempts have been made to compare the use of different MOF scoring systems in trauma patients. The Denver score was developed by a panel of trauma experts in Denver, Colorado, USA. The first version was proposed in 1991,[231] and underwent significant modification in 1996,[98] before more recent amendments to the respiratory and cardiovascular components.[149,188,208] The score was designed for use in trauma patients with an ISS > 15 who survived more than 48 hours after injury, and were aged 16 years or over. The initial score used data on organ dysfunction from eight systems, but this was reduced to four in the 1996 version. The worst values obtained for the required variables are scored 0 to 3 for each of the respiratory, cardiovascular, renal and hepatic

systems. The scores are summed, with a score of > 3 taken to denote post-trauma MOF (Table 4.9).

DENVER					
Organ system	**Variable (Units)**	**Score allocation**			
		0	**1**	**2**	**3**
Respiratory	*PaO_2/FiO_2 (mmHg)*	> 250	250-175	175-100	<100
Renal	*Creatinine (umol/L)*	<159	160-210	211-420	>420
Hepatic	*Bilirubin (umol/L)*	<34	34-68	69-137	>137
Cardiac	*Inotropes (ug/kg/min)*	None	1 inotrope Small dose	1 inotrope Moderate dose OR > 1 inotrope Small dose	1 inotrope Large dose OR > 2 inotropes Moderate dose

Table 4.9: The Denver Postinjury Multiple Organ Failure Score

The Denver group have compared their own MOF score with the Marshall score to assess population incidence and congruency of classification of patients as MOF or non-MOF. Both scores were compared against a number of adverse outcomes occurring within 28 days of injury, designed to cover resource utilisation (length of ICU stay, > 7 days of mechanical ventilation) and patient-specific outcomes (ventilator-free days, and death). A cut-off of >5 for MOF as defined by the Marshall scale was used. Whilst the original Marshall study did not specify a cut-off value, > 5 has been used in multiple studies.[164,185,199]

For the 1389 patients studied, 694 (50%) were defined as not having MOF, and 304 (22%) defined as having MOF by both the Denver and Marshall scores. However, 391 (28%) patients were variably defined as either having MOF or not, depending on which score was used. The incidence of MOF was far higher using the Marshall score (49.7%) than the

Denver score (22.2%), although a higher case-fatality rate was noted with the Denver score (30% versus 15%). This indicated the Denver score as being more specific but less sensitive for post-trauma MOF.[208]

Analysis of ROC curves showed a cut-off value of > 8 for the Marshall score was more specific and sensitive for mortality than the accepted cut-off of 5. Both scores performed better on goodness-of-fit measures for patient-specific outcomes rather than measures of resource allocation. Cardiac dysfunction was the best predictor of death for both systems, with haematological and CNS dysfunction poorly predictive.

The study authors concluded that both scores *"perform[ed] reasonably well"* but suggested elements of both might be combined to improve prediction. Whilst recommending their own score (Denver) due to its increased simplicity, they accepted the improved accuracy of aspects of the Marshall score, such as renal dysfunction. This comparative study used worst daily values to calculate both scores, not the representative scores suggested by the original Marshall score developers.[208]

The original developers of the SOFA score investigated its application specifically in trauma patients, and found it able to reliably describe organ dysfunction after injury. SOFA score on admission was significantly higher in non-survivors, and associated with an increased ICU stay in survivors. The SOFA score was most predictive in the early phase of ICU care, with only the respiratory component being related to outcome after day 4.[256] The SOFA score has become increasingly seen in trauma research, with a cut-off of > 5 or two failed systems (>3) taken to denote post-trauma MOF.[187,257] No research was found to compare the SOFA score with the Marshall or Denver scores in trauma patients.

No significant evidence was found to support the use of one of the Denver, SOFA or Marshall scores over the others. The contents of the scores were therefore compared to assess if there were any differences that would affect their use from the data available in this study.

All three scores were found to be very similar in content (Table 4.10). All scores contain respiratory dysfunction, as measured by PaO_2/FiO_2 ratio, renal dysfunction, as measured by creatinine levels, and hepatic dysfunction, as measured by bilirubin levels. The SOFA and Marshall scores both contain haematological dysfunction, as measured by platelet count, and CNS dysfunction, measured by GCS. The thresholds for allocating scores to values of these variables differ, but cover similar ranges. The main difference was in the definition of cardiovascular dysfunction – inotrope use for Denver, a mixture of inotrope use and MAP for SOFA, and the PAR composite measure for the Marshall score. All contained variables that would be freely available in routine clinical data.

Organ system	Variable	Units	MOF score	Score allocation				
			Denver	**0**	**1**	**2**	**3**	
			SOFA		**1**	**2**	**3**	**4**
			Marshall	**0**	**1**	**2**	**3**	**4**
Respiratory	*PaO_2/FiO_2*	*mmHg*	Denver	> 250	250-175	175-100	<100	
			SOFA		<400	<300	<200 AND resp support	<100 AND resp support
			Marshall	>300	226-300	151-225	76-150	≤75
Renal	*Creatinine*	*umol/L*	Denver	<159	160-210	211-420	>420	
			SOFA		110-170	171-299	300-440	>440
			Marshall	≤100	101-200	201-350	351-500	>500
Hepatic	*Bilirubin*	*umol/L*	Denver	<34	34-68	69-137	>137	
			SOFA		20-32	33-101	102-204	>204
			Marshall	≤20	21-60	61-120	121-240	>240
Cardiac	*Inotropes*	*ug/kg/min*	Denver	None	1 inotrope Small dose	1 inotrope Moderate dose OR > 1 inotrope Small dose	1 inotrope Large dose OR > 2 inotropes Moderate dose	
			SOFA		MAP < 70mmHg	Dopamine ≤ 5 OR Dobutamine any dose	Dopamine > 5 OR adrenaline ≤ 0.1 OR noradrenaline ≤ 0.1	Dopamine > 15 OR adrenaline > 0.1 OR noradrenaline > 0.1
	PAR = (HR/MAP)xCVP	*Beats/min*	Marshall	≤10.0	10.1-15.0	15.1-20.0	20.1-30.0	>30.0
Coagulation	*Platelets*	*$x\ 10^3/mm^3$*	Denver					
			SOFA		<150	<100	<50	<20
			Marshall	>120	81-120	51-80	21-50	≤20
CNS	*GCS*	*Scale*	Denver					
			SOFA		13-14	10-12	6-9	<6
			Marshall	15	13-14	10-12	7-9	≤6

Table 4.10: Comparison of the components of the Denver, SOFA and Marshall MOF scoring systems

4.3.5. MOF – STUDY DEFINITION

In the absence of significant evidence to merit the use of any single MOF scoring system, a pragmatic approach to MOF definition was taken. Due to the similarities in data required for collection, it was decided to calculate Denver, SOFA and Marshall scores for each trauma patient. Scores were calculated daily for each component organ system, and summed for a daily organ failure score. Assessment of neurological dysfunction can be complicated in patients who are intubated and sedated. Therefore, in keeping with the principles of the SOFA and Marshall scores, GCS was defined as being that observed in the absence of sedation, assessed via a sedation hold or similar.

Whilst recognising that MOF represents a continuum of processes, it has nonetheless been accepted in other studies that threshold values can be used to determine if a patient has developed MOF. A binary definition of MOF/no MOF has greater clinical utility than grading of illness severity. MOF was therefore defined if a threshold total score was reached on a day. The cut-off values for these MOF thresholds for each score were taken from published recommendations or values commonly used in other trauma studies. Despite suggestion that the higher cut-off of > 8 was more sensitive for mortality prediction with the Marshall score,[208] the lower cut-off of > 5 was used. An increase in the threshold value for MOF definition would inevitably increase the relationship with mortality, since it would represent a sicker cohort of patients. However, the purpose of this study was to assess post-trauma MOF, not post-trauma mortality, and the more widely used and accepted threshold of > 5 was felt to more appropriately represent a physiologically challenged post-trauma population.

- Denver – MOF if score > 3

- SOFA – MOF if score > 5
- Marshall – MOF if score > 5

These scores were designed to reflect a significant level of severity, and required organ failure in more than one organ system simultaneously. Each patient was therefore given a status of MOF/no MOF for each day post-injury, which allowed for consideration of the duration of MOF. However, with the main purpose being prediction of MOF, each patient was then coded according to whether they had ever developed MOF at any stage post-injury. Each patient therefore had three designations – MOF/no MOF by each of the Denver, SOFA and Marshall scores.

By using multiple scores, the incidence of MOF according to each score could be calculated and described for the first time in a trauma population. It was hoped that this might provide added information as to which is the most appropriate MOF scoring system for use in trauma practice and for further research.

CHAPTER 5 DATA LINKAGE

5.1. DATA LINKAGE OVERVIEW

To describe the study population, data for individual patients needed to be linked across the main data sources of TARN, ICNARC and CareVue. Patient identifiers were used to match patients, with hospital number used as the primary identifier. Secondary identifiers included first and last name, address, postcode, date of birth and NHS number. Date of injury and date of admission were also used to verify the correct patient was linked to the hospital attendance.

The linkage process was performed from two starting points. The TARN database was queried for trauma patients requiring ICU admission. Both the ICNARC and ICU admin databases were searched for trauma patients. Results from ICNARC and ICU admin were combined to cover the whole study period.

The results of TARN and ICNARC/ICU admin searches were combined. Patients who were matched to one but not the other underwent a secondary process of matching. Further data sources, including PAS and hospital notes were used to resolve matching problems.

Once linkage between TARN and ICNARC/ICU was complete, the patient list was matched to CareVue. Trauma patients were then labelled in CareVue to assist in future data extraction. Once patients were matched across all data sources, a single linkage file was created with patient identifiers and the matching SubmissionID from TARN and Admission number from ICNARC. Each patient was given a unique “TraumaDBID” number

to allow anonymisation in all further data files. The linkage file was stored on a secure NHS research server.

5.2. IDENTIFYING PATIENTS FROM TARN

5.2.1. ACCESSING TARN DATA

To identify patients for the study population, TARN data needed to be accessed in three ways. Data from TARN can be obtained directly from the national TARN office, but this is cleaned of all but a few limited patient identifiers. Additional identifiers could be obtained using the TARN web interface, but each patient record needed to be queried separately. The local TARN administrator also maintained files containing information on trauma patients whose details had been submitted to TARN, which contained additional identifiers and were ordered by date of admission. A combination of these three sources was used for data linkage.

5.2.2. INITIAL TARN QUERY

The national TARN database was searched using the following criteria:

- TARN submission from John Radcliffe AND
- Date of injury 1/1/2003 to 15/5/2011 AND
- Critical Care length of stay > 0 days OR
- Any entry in Critical Care section

5.2.3. PATIENT IDENTIFIERS

The following identifiers were available from the national TARN query:

- SubmissionID

- Age at date of injury
- Sex
- Date of arrival
- Postcode

Patient identifier data is contained within the Opening Section of the TARN submissions on the web interface. Variables collected in the Opening Section can be divided into three groups, as given in Table 5.1.

Numbers	**Patient data**	**Hospital details**
SubmissionID	First name	Date of arrival
ED patient number	Surname	Time of arrival
Hospital number	Sex	Transferred? In? Out?
NHS number	Date of Birth (DOB)	Previous hospital
	Age at date of injury	Reason for transfer in
	TARN eligible?	Next hospital
	Postcode	Reason for transfer out

Table 5.1: Variables collected in the Opening Section of the Core Dataset of TARN

Since identifiers from the Opening Section needed to be accessed via the web interface for individual patients in turn, this was very time-consuming. An initial assessment was made of the completeness of identifiers in the Opening Section to determine which would be most useful for linkage. Data on the first four patients from January, May and September each year were examined, up to a total of 100 patients. This showed hospital number, postcode and DOB were present in over 90% of patients and suitable for linkage use. Patient names had a high amount of missing data.

The files of the local TARN administrator did not contain the TARN SubmissionID, but did contain the date of admission, name, age and hospital number amongst other details on

TARN eligibility. Patient names could therefore be taken from the local files after matching to the hospital number and date of arrival in the Opening Section data.

5.3. IDENTIFYING PATIENTS FROM ICNARC

5.3.1. RECREATING THE ICNARC DATABASE

To identify patients in ICNARC, the structure of the database needed to be recreated using the data files retained for local patients. ICNARC submissions are retained locally on an NHS sever in text format.

Data for all AICU patients were drawn down from the local NHS server in July 2011. Data were extracted into tab-delimited text files, which could then be converted into other formats suitable for data processing. The national ICNARC database linkage structure was not retained in the local storage arrangements, and had to be re-established to essentially create a local version of the ICNARC database.

A total of 15 text files were obtained, each relating to a separate table in the national ICNARC database. Three of these files did not contain any data, and were excluded. The content of the remaining 12 files was examined to identify a potential primary key for database creation. To do this the files were exported into Microsoft Excel and cleaned for transcription errors. Following cleaning and identification of primary keys, the files were imported into Microsoft Access for database creation. A summary of the 15 text files is given in Table 5.2.

Table name	Data?	Summary of data fields	Primary key?
ACP Dataset Summary	Yes	Organ support, duration, outcome	None present
Admission	Yes	Admission details, location, transfer etc.	Admission number
Daily Assessment	Yes	Daily physiological, laboratory and clinical parameters for ICNARC model	None present
Departments	No	Referring doctor and department	N/A
Diagnosis	Yes	ICNARC Coding Model data	Admission number
Doctors	No	Details of speciality doctors	N/A
Hospital outcome	Yes	Destination, hospital stay, discharge status, outcome	Admission number
ICNARC_PMH	Yes	PMH	Admission number
MPM 0 and 24	Yes	Details for MPM	Admission number
Patients	Yes	Patient identifiers	Admission number
Scores	Yes	Multiple score details, including APACHE II, APACHE III, SOFA, ICNARC model	Scores information
Specialities	No	Referring doctor and speciality	N/A
Transfers	Yes	Transfer details	Admission number
Unit Consultant	Yes	AICU consultant	Admission number
Unit Outcome	Yes	Outcome data	Admission number

Table 5.2: Summary of ICNARC data files

(ACP = Augmented Care Period, PMH = Previous Medical History, MPM = Mortality Prediction Model)

The ICNARC admission number is designed to be unique to each admission, not to each patient. Therefore, if a patient required readmission to AICU, a new ICNARC admission number would be created. Admission number could therefore be used as a primary key in all tables that involved a single set of data. However, for tables that involved data over several days (Scores and Daily Assessment) or recorded readmission (ACP Dataset Summary), the admission number could not be used.

ACP DATASET SUMMARY

The ACP Dataset Summary table contained an Episode Number in addition to the Admission Number. The purpose of the Episode Number was unclear. Six Admission Numbers were associated with more than one Episode Number and these admissions were examined separately. All six were from the start of ICNARC submissions, and the additional Episode Numbers all seemed to result from imputational errors with no evidence of readmissions.

DAILY ASSESSMENT

The Daily Assessment table contained no unique fields, and a new primary key needed to be created. Data from the Admission Number and Day in Unit fields were concatenated to create a new field called Daily Assessment ID. 21 duplicate Daily Assessment IDs were found when the field was created. These were studied and universally found to be due to duplicate data entry.

SCORES

The Scores table contained a field called Scores Information that appeared to be intended to be a unique identifier, but contained a number of duplicates. The 14 duplicate rows related to 12 Admission Numbers and were all found to be due to errors in data entry.

Following resolution of these problems, and data cleaning, the database was recreated in Microsoft Access. The final database structure is given in Figure 5.1. The ICNARC database could then be queried using standard SQL queries. To effectively identify trauma patients, knowledge of the ICNARC Coding Method was required.

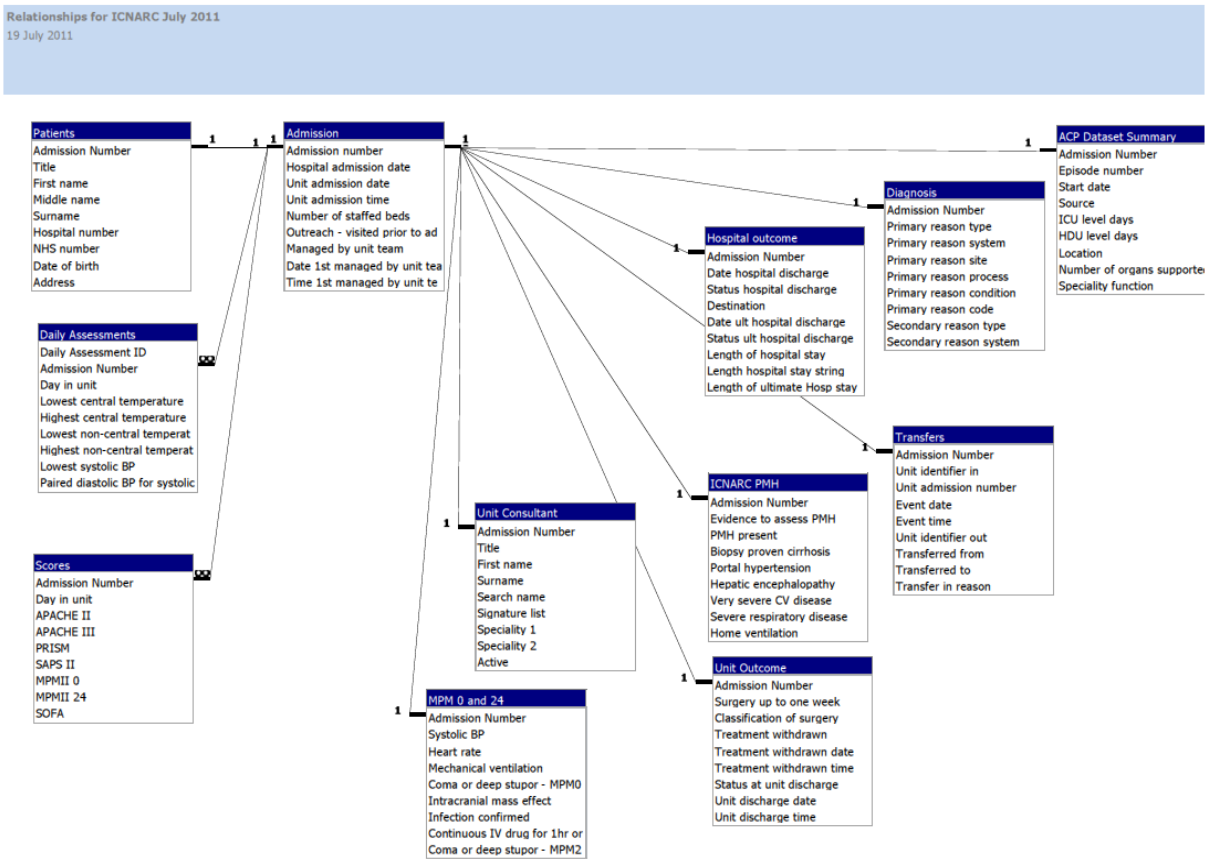


Figure 5.1: Diagram of ICNARC database table relationships

5.3.2. THE ICNARC CODING METHOD

The ICNARC Coding Method is a five-tier hierarchical system developed to code conditions leading to admission to ICU.[258] It was originally developed from the free-text reasons for admission given in the UK APACHE II study, before being prospectively tested on over 22,000 admissions to 62 ICUs in the UK. The method was developed to allow clear definition of the primary reason for ICU admission, to allow cross-mapping to other diagnostic coding methods, such as the ICD, and to improve precision to compare mortality rates for specific conditions.

The ICNARC Coding Method consists of five tiers – type, system, site, process or condition. These can be further defined as follows:

- Type – surgical or non-surgical
- System – organ or body system
- Site – anatomical site
- Process – physiological or pathological process
- Condition – named condition

The coding was developed with some redundancy to allow different conditions to be coded from more than one starting type or system. Therefore there are 1140 different codes relating to only 741 unique conditions. “Trauma” is included as a system, a process and a condition. If trauma is coded as a system, then no site definition is required.

The ICNARC Coding Method can be used for the primary reason for admission, the secondary reason for admission, and an ultimate reason for admission, used when the underlying reason becomes evident after the first 24 hours of ICU stay.

5.3.3. INITIAL ICNARC QUERY

The nature of the ICNARC Coding Method means multiple codes can be entered to represent trauma. The following fields were identified as containing trauma codes

- System: “Trauma”
- Process: “Trauma” OR “Trauma, perforation or rupture” OR “Trauma, perforation, fistula or rupture”
- Condition: “Trauma” OR a description of the precise nature of the injury

Many descriptions were permitted for trauma in the Condition tier. Preliminary searches showed that using these individually was inefficient and did not result in additional hits over and above using the generic trauma codes. Since the relevant codes in Process also contained non-trauma conditions, the results of the initial queries were hand-searched to verify all patients included in the study population were trauma patients.

The following SQL query was run using the Diagnosis, Admission and Patients tables to provide coding data and patient identifiers to match against the other data sources:

```
SELECT Diagnosis.[Primary reason system], Diagnosis.[Primary reason process], Diagnosis.[Primary reason condition], Diagnosis.[Admission Number], Admission.[Hospital admission date], Patients.[First name], Patients.[Surname], Patients.[Hospital number], Patients.[Date of Birth], Patients.[Sex], Patients.[Age], Patients.[Post code]
FROM Patients INNER JOIN (Admission INNER JOIN Diagnosis ON Admission.[Admission Number] = Diagnosis.[Admission Number]) ON Patients.[Admission Number] = Admission.[Admission Number]
WHERE (((Diagnosis.[Primary reason system])=”Trauma”)) OR (((Diagnosis.[Primary reason process])=”Trauma”)) OR (((Diagnosis.[Primary reason process])=”Trauma, perforation or rupture”)) OR (((Diagnosis.[Primary reason process])=”Trauma”)) OR (((Diagnosis.[Primary reason condition])=”Trauma”));
```

This search term was repeated for secondary and ultimate reasons, with direct replacement of the relevant search fields. Each search was carried out independently to ensure all trauma patients were captured, and to decrease the risk of errors in the hand-searching of the query results.

5.3.4. PATIENT IDENTIFIERS

A wide range of identifiers were available in the ICNARC database. The following identifiers were extracted to match to other data sources – first name, surname, hospital number, DOB, sex, age, and postcode.

5.4. IDENTIFYING PATIENTS FROM ICU ADMIN

5.4.1. ACCESSING ICU ADMIN

The ICU Admin database was used because ICNARC data was not available for the full time period of the study. It was used in data linkage as a means of identifying potential trauma patients in ICU. The ICU Admin database was stored on local NHS servers, and was accessible to AICU staff only. Many of the recorded details mirror those in the written AICU admission book.

Data were extracted in yearly segments, running from April to March. These were downloaded into password-protected Microsoft Excel files.

5.4.2. IDENTIFYING TRAUMA PATIENTS

Initial searches of the ICU Admin database focussed on searching for trauma patients by responsible consultant. However, many trauma patients were not listed under trauma consultants, and this method was found to be unreliable.

The second method involved searching the free-text field for the reason for admission. Specific terms were searched for, followed by hand-searching of remaining entries. To minimise the human error of hand-searching, the search was repeated twice.

5.4.3. PATIENT IDENTIFIERS

The following patient identifiers were contained in the ICU Admin database – first name, surname, hospital number, address and postcode.

5.5. PROCESS OF DATA LINKAGE

5.5.1. OVERALL PROCESS SUMMARY

Data linkage was carried out in a series of steps. Not all data were available at the start of the process, and several of the steps were repeated multiple times. The results presented here are amalgamated to simplify understanding of the numbers involved.

The steps were as follows:

1. Initial matching of lists of patients identified from TARN, ICNARC and ICU Admin
2. Matching of patients from the ICNARC and ICU Admin queries to make a single "ICU" list
3. Resolution of patients identified in the TARN query but not found in the "ICU" list
4. Resolution of patients in the "ICU" list not found in the TARN query
5. Comparison of results from steps 3 and 4 to identify any unresolved patients
6. Matching of the final list to CareVue to create the definitive patient list

To ease the matching process and reduce errors, the data was divided into five time periods. These periods were based on the data availability at the start of the matching process, and the division between the old and new systems of CareVue. A significant delay in TARN submissions was present at the start of the process, with a lag in access to the data from 2009-2010. This was therefore considered as a separate time period. A summary of the time periods and the data present in each is given in Table 5.3.

	Period 1	Period 2	Period 3	Period 4	Period 5
Dates	1/2003-6/2006	7/2006-5/2008	6/2008-12/2008	1/2009-12/2010	1/2011-5/2011
TARN	✓	✓	✓	✓	✓
ICNARC	x	✓	✓	✓	✓
ICU Admin	✓	✓	✓	✓	✓
CareVue	Old	Old	New	New	New

Table 5.3: Data periods for matching

5.5.2. MATCHING TARN, ICNARC AND ICU ADMIN

The initial TARN query was performed before data from 2009-2010 were available and identified 811 patients. From these, 215 matched to ICNARC and ICU Admin, and a further 262 matched to ICU Admin in the period before ICNARC submissions started. This totalled 477 (58.8%) patients matched, leaving 334 (41.2%) unmatched.

Initial queries of ICNARC and ICU Admin identified different numbers of patients. Higher numbers for each time period were found in ICNARC, although these were significantly reduced on removal of non-trauma codes. From the remaining 433 patients, 215 (49.7%) matched to TARN. Of the 731 patients identified from ICU Admin, 477 (65.3%) matched to TARN (Table 5.4).

	Period 1	Period 2	Period 3	Period 4	Period 5	Totals
	1/2003 – 6/2006	7/2006 – 5/2008	6/2008 – 12/2008	1/2009-12/2010	1/2011-5/2011	
TARN	**Yes**	**Yes**	**Yes**	**Incomplete**	**Yes**	
Identified via query	368	245	60	121	17	811
Matched to ICNARC	N/A	128	25	52	10	215
Matched to ICU Admin	262	128	25	52	10	477
ICNARC	**No**	**Yes**	**Yes**	**Yes**	**Yes**	
Identified via query	N/A	237	67	194	37	535
Matched to TARN	N/A	128	25	52	10	215
ICU Admin	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	
Identified via free text	324	167	42	162	37	731
Matched to TARN	262	128	25	52	10	477

Table 5.4: Results of matching initial queries from TARN, ICNARC and ICU Admin

5.5.3. MATCHING ICNARC AND ICU ADMIN

The initial matching highlighted discrepancies in the patient numbers identified in ICNARC and ICU Admin. Since any admission to ICU should trigger a single entry in both databases, the numbers should theoretically match. Prior to resolving problems in non-matching with TARN, it was necessary to first resolve differences between ICNARC and ICU Admin.

The results of this process are given in Table 5.5. Discrepancies between ICNARC and ICU Admin were found to be due to duplicate entries, the presence of non-trauma patients that had not been excluded in the initial removal of non-trauma codes, or human error.

The results identified a total of 738 patients for a combined "ICU" list. 40 of these patients were subsequently found not to be trauma patients, and were excluded, leaving 698 patients which required matching to TARN. For the remainder of the matching process, the combined ICNARC/ICU Admin results will be simply termed ICU.

	Period 1	**Period 2**	**Period 3**	**Period 4**	**Period 5**	**Totals**
	1/2003 – 6/2006	**7/2006 – 5/2008**	**6/2008 – 12/2008**	**1/2009-12/2010**	**1/2011-5/2011**	
ICNARC	**No**	**Yes**	**Yes**	**Yes**	**Yes**	
Via query	N/A	237	67	194	37	535
Removal clear non-trauma	N/A	189	48	164	32	433
Removal duplicates	N/A	184	46	160	31	421
ICU Admin	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	
Via free text	324	167	42	162	37	732
Removal duplicates	311	160	41	151	35	698
Combined ICNARC/ICU Admin						
Total patient number	311	186	46	160	35	738
Not trauma	0	26	5	9	0	40
Total	311	160	41	151	35	698

Table 5.5: Results of matching trauma patients across ICNARC and ICU Admin

5.5.4. RESOLVING PATIENTS FROM TARN QUERY

Of the 811 patients identified from the initial TARN query, 334 (41.2%) could not be matched to the queries from the ICU databases. Potential reasons considered for failure to match included:

1. Admission to a critical care or high dependency area other than AICU
2. Admission to an ICU in another hospital which is part of the same NHS trust
3. Incorrect inclusion in the TARN query results
4. Omission from ICU databases

The John Radcliffe Hospital is part of the Oxford University Hospitals NHS Trust, which also includes the Churchill Hospital, the Horton General Hospital and the Nuffield Orthopaedic Centre. Prior to 2007, it also included the Radcliffe Infirmary. Other critical care areas are present in the other hospitals, but the John Radcliffe is the only receiving centre for major trauma. To establish patients that may have been admitted to other ICUs, the PAS system was used.

PAS tracks a patient's movements through wards and departments, detailing dates and locations. All of the 334 unmatched patients were located in PAS. Of these, 284 (85.0% of those unmatched) were admitted to Neuro ICU. A further 14 (4.9% unmatched patients) were admitted to other high dependency areas in the Horton hospital or the Radcliffe Infirmary. 10 patients were admitted to other critical care areas within the John Radcliffe – 9 to the Coronary Care Unit (CCU) and 1 to the Paediatric Intensive Care Unit (PICU). In total, 308 (92.2% unmatched patients) could be excluded, since they did not meet the study inclusion criteria.

Hospital notes were ordered for the 26 patients that remained unmatched. Of these, 13 were admitted to neurosciences ICU, 5 to CCU and 2 died in the ED. From the remaining 6 patients, 5 had never been admitted to AICU or any other critical care area and had been incorrectly labelled by TARN. The remaining patient was a true trauma ICU admission that had been missed in the initial query. The results of these matching procedures are shown in Figure 5.2.

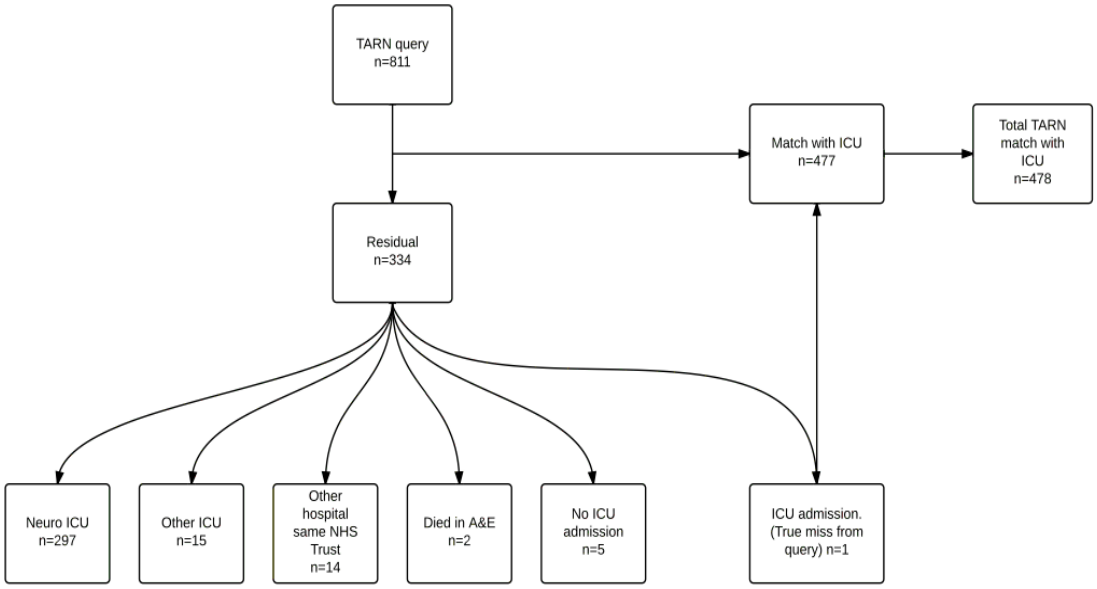


Figure 5.2: Results of matching of patients from TARN query

5.5.5. RESOLVING PATIENTS FROM ICU QUERIES

From the 698 patients identified by querying the ICU databases, 236 (33.8%) could not be matched to the list of patients from the TARN query. Potential reasons for failure to match included:

1. Not yet present in TARN due to delay in local submissions

2. Missed by TARN
3. Admitted to a different hospital in the same NHS Trust
4. Not trauma patients, or did not meet TARN eligibility criteria

103 (43.6% of unmatched) patients were from the period January 2009 – December 2010 when the national TARN database was known to be incomplete for some of the local submissions. The local files of the TARN administrator were examined for these patients. In total, 55 of the unmatched patients could be excluded. 53 did not meet the TARN criteria, which generally indicated isolated trauma with an associated medical problem i.e. fractured neck of femur with pneumonia. 1 patient was admitted to the Horton hospital, and 1 was not a trauma patient, despite being listed under a trauma consultant.

A further 70 patients could be matched to local TARN, but little information was given beyond TARN eligibility. For the remaining 111 patients, 80 were not mentioned in local TARN files and the remainder were mentioned as requiring more information to determine TARN eligibility (Figure 5.3).

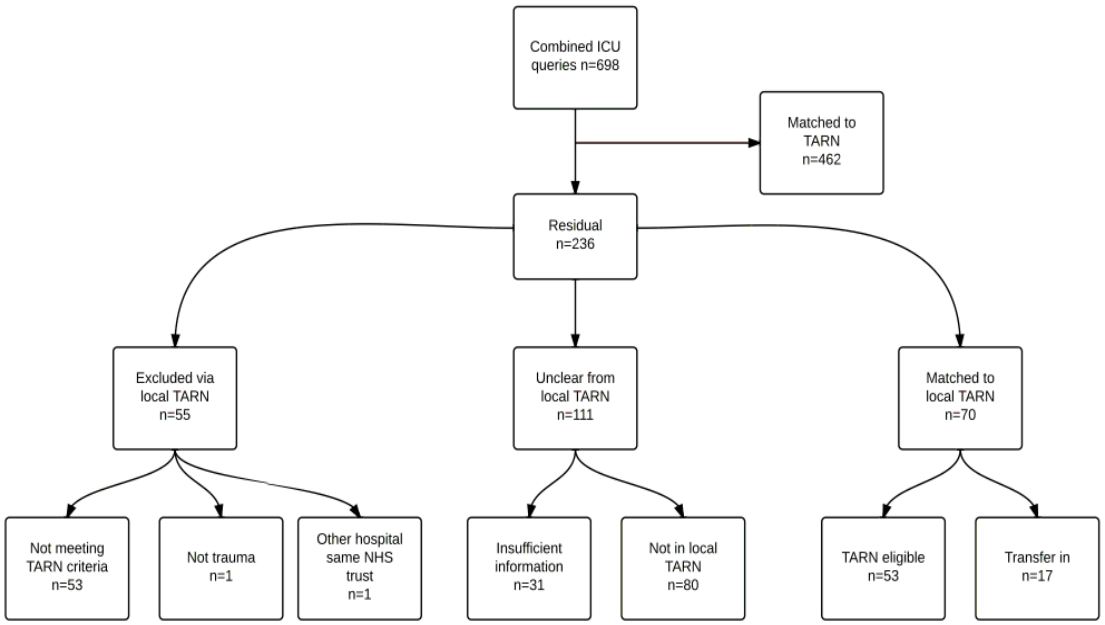


Figure 5.3: Results of matching ICU queries to local TARN files

These remaining 181 patients were then checked against the national TARN web interface. Local submissions are uploaded via the web interface and then a period of delay exists while they are checked by national TARN before inclusion in the national database. The web interface can be searched for submissions that are awaiting approval from national TARN.

81 of the 181 patients were present and approved by national TARN, and a further 51 had been rejected. This left 49 patients (20.8% of those originally unmatched) whose eligibility could not be confirmed using the TARN web interface.

Hospital notes were obtained for these remaining 49 patients. It was found that 5 of the patients should have been entered in TARN but had been missed. These patients dated from early in the TARN submission process in 2003. The remaining 44 patients were

eligible for TARN and were in the process of submission, but had stalled at various stages. Reasons for this included loss of access to hospital notes during submission (i.e. due to a readmission, a legal case involving the police or coroner, or a case undergoing a medicolegal review), a delay in information required to complete the submission (i.e. a coroner or police report) or a delay in TARN approval. Outstanding issues with these cases were resolved so that all could be authorised by national TARN and found in a repeat query of the TARN database.

The resolution of matching of the ICU queries to TARN data by the various stages employed is summarised in Figure 5.4. In total, 587 patients were matched from ICU to TARN.

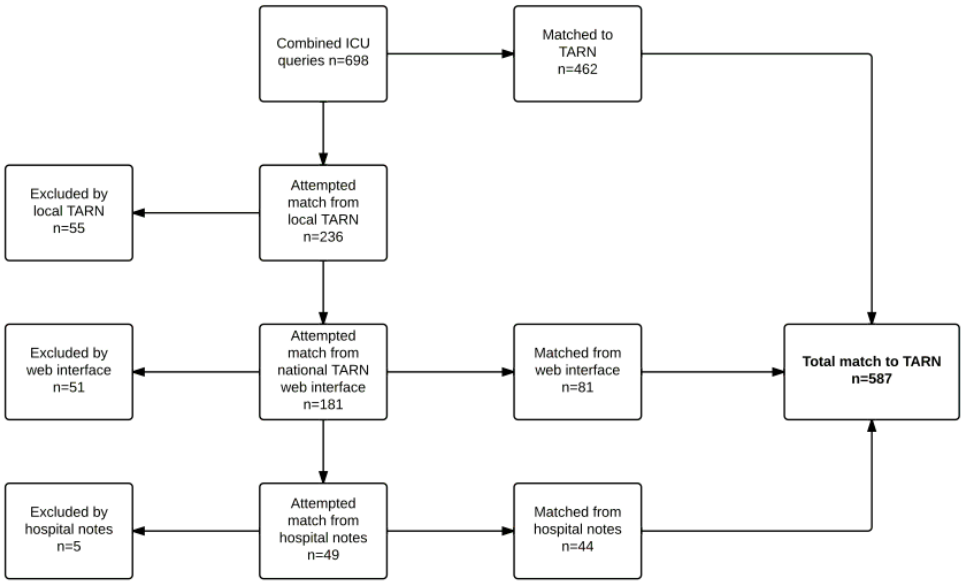


Figure 5.4: Results of matching of patients from ICU queries

5.5.6. COMPARING RESULTS FROM TARN AND ICU MATCHING

A total of 478 patients identified from TARN were matched to ICU databases. In the reverse process, 587 patients identified from the ICU databases were matched to TARN. Analysis of these two groups showed an overlap of 473 patients that had been identified by both processes.

All available data were used to resolve the residual patients. For the 5 patients identified through TARN, 2 were admitted to neurosciences ICU (known as Towler), 1 was a duplicate entry, and 2 were suitable for inclusion and subsequently found in the ICU databases.

For the 114 residual patients identified through the ICU databases, 3 were under 16 years of age despite being admitted to AICU, and 2 did not meet TARN eligibility criteria. The remaining 109 patients were suitable for inclusion, and could be identified in TARN after the outstanding issues that had been highlighted earlier in this chapter were resolved.

The overall process identified 584 trauma patients matched between TARN and ICU databases (Figure 5.5).

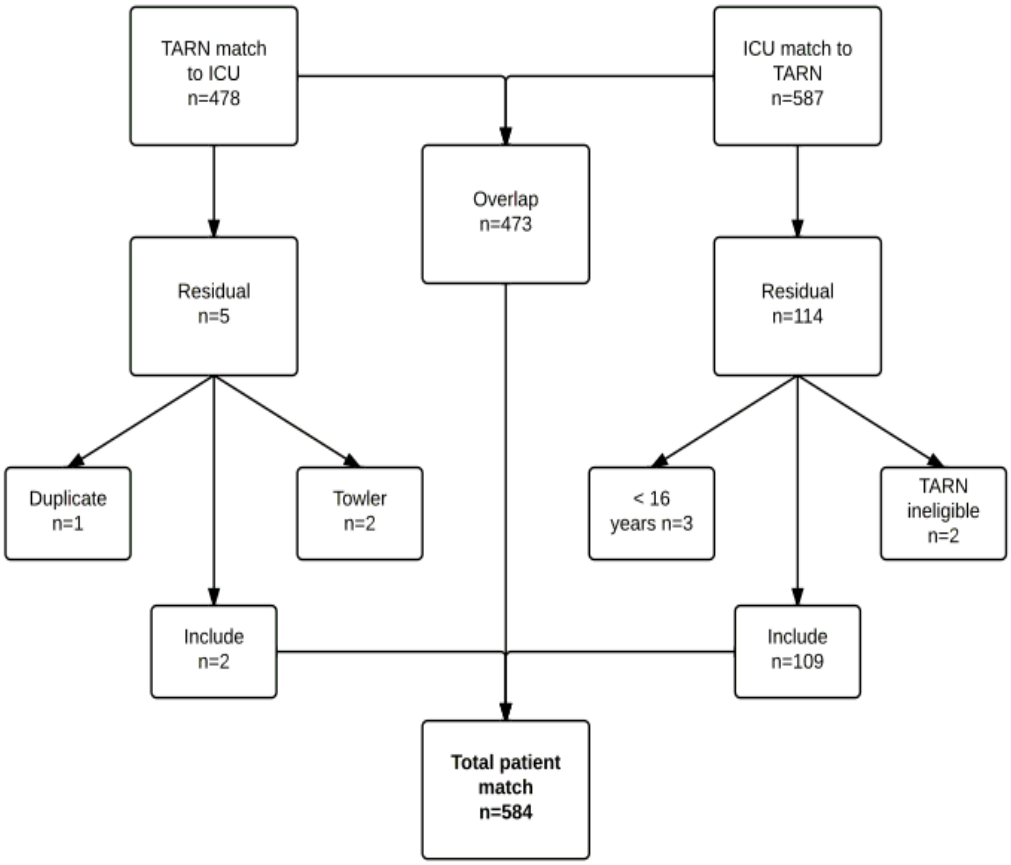


Figure 5.5: Summary of final resolution of patient matching across TARN and ICU databases

5.5.7. MATCHING TO CAREVUE

The final step in data linkage was the matching of patients to the CareVue database. CareVue is divided into two parts – the old CareVue, which is called ISM, covers the period from 1999 to June 2008, and the new CareVue, which is called CIS, covers the period from June 2008 to the present day. The structure of ISM and CIS is complex and will be discussed further in the following chapter. Due to the large size of the data contained in CIS, it is split into 7 sections, each representing 6-8 months of admissions.

The 584 patients were matched to the ISM and CIS sections of CareVue using searches of patient identifiers. Discrepancies were resolved by comparing identifiers against those from TARN, ICNARC, ICU Admin and other sources. After confirmation, patients were labelled in the CareVue database as major trauma patients by creation of a new field called TraumaDb.

Due to readmissions to ICU and the structure of CareVue data in storage, several patients had more than one entry in CareVue. These were checked and labelled to ensure all data would be able to be accessed on data extraction. 576 of the 584 patients were matched to CareVue. There appeared to be some data loss in the changeover period between the old and new CareVue systems, which accounted for 6 of the missing patients. The other 2 patients both had admission dates which fell in the crossover period between different segments of CIS.

5.5.8. FINALISING THE STUDY POPULATION

To finalise the study population, the 576 patients needed to be checked against all the inclusion and exclusion criteria, as stated in sections 2.6.2 and 2.6.3. The timing of injury and arrival at the John Radcliffe was checked for all patients. A further 59 patients were found to either have no recorded time of injury or to have arrived at the John Radcliffe more than 48 hours after injury. These patients were therefore excluded, leaving a total number of 517 patients to comprise the study population.

All patients were linked between TARN, ICNARC/ICU Admin and CareVue. Patients were labelled in CareVue as major trauma patients to aid future data extraction. Patient identifiers and details of fields required for linkage between all the data sources were

transferred into a single linkage file. Each patient was given a unique ID to allow anonymisation of data following extraction.

With the study population finalised, and the variables required for input and outcome measures confirmed, extraction of the necessary data could be started. This process will be described in the following chapter.

CHAPTER 6 DATA EXTRACTION AND ANALYSIS

6.1. OVERVIEW OF DATA EXTRACTION

6.1.1. EXTRACTING APPROPRIATE DATA

The presence and availability of potential data was assessed prior to data extraction. This involved examining each of the main data sources to determine the data contained, and the suitability of the data for use in the study. For each variable, the following process was adhered to:

1. Data sources were assessed for the presence of potentially relevant data fields
2. Potential fields were assessed for presence and utility of data
3. If data were present and relevant, extraction from the data source
4. Exact method of extraction recorded to allow future replication

Data were extracted into .csv or .xls files, depending on the data source. All were then converted to .xls files for ease of combining data from different sources. Full details on the assessment of the data sources for suitability and the field used for extraction are given in Appendix 7.

6.1.2. METHOD OF DATA EXTRACTION

The method of data extraction differed between the data sources of TARN, ICNARC and CareVue. In addition, extraction from the old version of CareVue (ISM) was different to extraction from the new version (CIS). Full details of the data extraction procedures are given for each data source later in this chapter.

Data on time of injury was not stored in CareVue. To ensure all data for the relevant time points was collected, input variable data were collected up to 48 hours after ICU admission, and outcome variable data were collected for all time periods. This was subsequently matched to timing of injury to obtain data for the first 48 hours post-injury for input variables, and from day three onwards for outcome variables.

6.1.3. DATA TRANSFORMATION AND CLEANING

To transform the raw data into data suitable for analysis the following process was used:

1. Anonymisation of raw data using the linkage file of patient identifiers
2. Creation of a data dictionary to define final variable
3. If data available from multiple sources, data combined into single file
4. Timing of data calculated from time of injury where appropriate
5. Cleaning of data to standardise units and data type, to apply collection rules as listed in data dictionary and to remove duplicates
6. Where applicable, data were divided according to the time periods required for predictive model building, and the appropriate value calculated for each time period

The full details on data collection and timing rules are given in the data dictionary. (Appendix 8)

6.2. DATA EXTRACTION FROM CAREVUE

6.2.1. OVERALL CAREVUE STRUCTURE

As previously stated, two CareVue systems have been in use in the AICU of the John Radcliffe Hospital during the time period of the study. The older system, ISM, was used

from 1999 to 1/6/2008. The newer system, CIS, has been in use from 1/6/2008 to the present day. Archiving of ISM and CIS has occurred by different processes. This has resulted in two separately stored databases, with different structures, fields, data formats and linkages.

Data is inputted into CareVue via a live system, which is markedly different to either of the archived systems. The live system could be used to identify what data might be contained in the archives, but could not be used to determine the structure of either ISM or CIS.

Archived data for both ISM and CIS were accessed via Microsoft SQL Server. Data extraction was performed using Transact-SQL (T-SQL) query language. The differences in structure between CIS and ISM did not support combining the databases. Different queries were required to extract data from CIS and ISM, reflecting the differing table structure, fields and linkages. For the purposes of data extraction, CIS and ISM were therefore considered as two different databases.

Despite the differences between CIS and ISM, efforts were made to match fields and queries were designed to extract similar data in a similar format where possible. Query design depended on detailed and accurate knowledge of the structure of the CIS and ISM databases.

6.2.2. CIS STRUCTURE

The structure of CIS was highly complex, consisting of 238 tables. The majority of these tables did not contain direct clinical data, but were required for the structure of the live interface. CIS was divided into 7 sections (CIS1-7) due to the size of the data. In essence,

this created 7 separate databases that were identical in structure, meaning the same T-SQL query could be used for extraction from each.

To break down the complex structure of CIS, groups of tables were identified that contained specific areas of relevant clinical data. These could be summarised as the following groups:

1. Patient Group
2. Observation Group
3. Intervention Group
4. Orders Group

The structure of these groups was identified separately, before assessing the relationships between groups.

PATIENT GROUP

The Patient group contained five main tables:

- PtPatientInfo
- Patient
- PtEpisode
- PtEncounter
- PtLocationStay

These tables contained details on patient demographics, background data and identifiers to allow data linkage. Hospital number was stored under the field name lifetimeNumber in the Patient table. The Patient table linked to PtEpisode which identified each

individual admission as a single episode. PtPatientInfo contained a number of patient identifiers including first and last name, and DOB. PtLocationStay contained information on date and time of ICU admission and discharge. PtEpisode, PtPatientInfo and PtLocationStay were linked via the PtEncounter table (Figure 6.1).

Data from the Patient group were required for all data extractions from CIS, so that identifiers could be obtained to match patient data. The following SQL query was used as the basis of extraction queries from CIS:

```
SELECT lifetimeNumber, firstName, lastName, startDate
FROM Patient, PtEpisode, PtEncounter, PtLocationStay, PtPatientInfo
WHERE Patient.patientId  = PtEpisode Id
AND PtEpisodeId.ptEpisodeId  = PtEncounter.ptEpisodeId
AND PtEncounter.ptEncounterId = PtLocationStay.ptEncounterId
AND PtEncounter.ptEncounterId = PtPatientInfo.ptEncounterId
```

The PtEncounter table linked to several other tables in the database, which allowed linkage of the Patient group of tables to other groups.

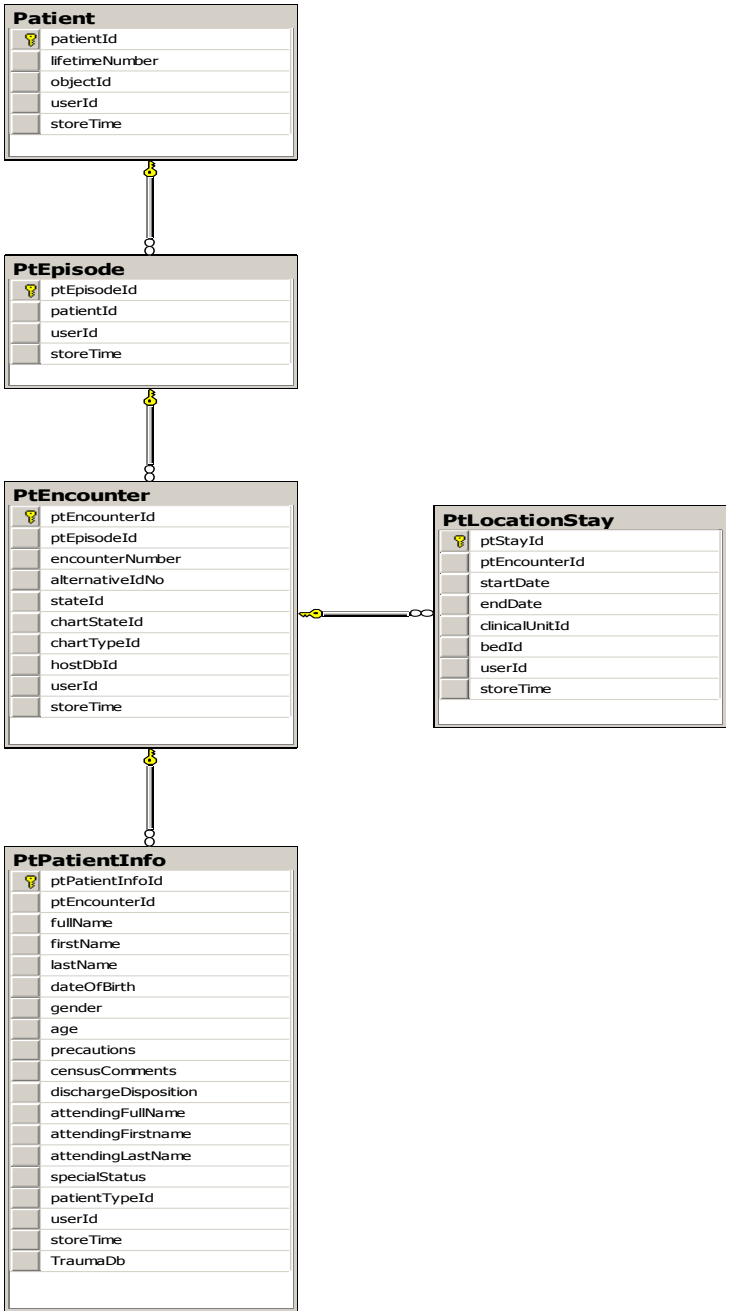


Figure 6.1: Relationships between Patient group of tables in CIS

OBSERVATION GROUP

The Observation Group contained a large amount of clinical data in two main tables – Observation and PtObservationData.

Observation contained listings of patient observations, covering a wide range of clinical areas. These could be considered to contain data that would be written on a patient chart or in patient notes. A total of 975 separate observation types were listed, each with an individual ObservationId. The Observation table linked to a number of other tables, including the PtObservationData table, which recorded individual observations for all patients. Observation and PtObservationData could be linked using the ObservationId field. PtObservationData also linked to PtEncounter via the PtEncounterId field (Figure 6.2).

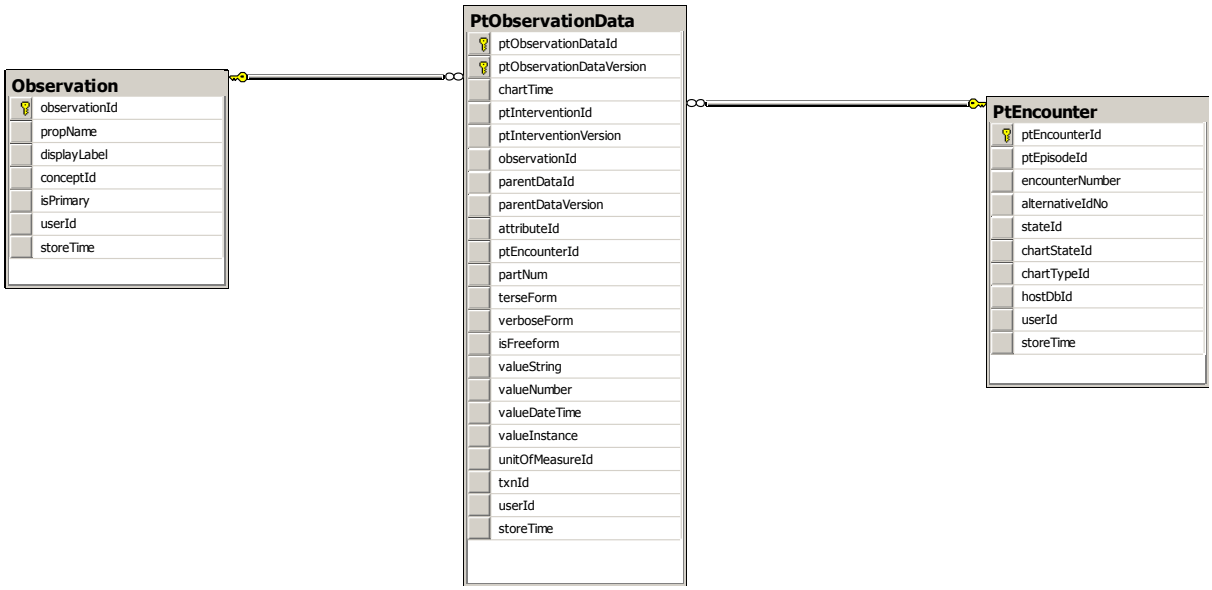


Figure 6.2: Relationships between Observation group of tables in CIS

INTERVENTION GROUP

The main tables in the Intervention Group were:

- Intervention
- PtIntervention
- PtInterventionData

The Intervention table contained 4938 distinct patient interventions, each with a unique InterventionId. The InterventionId field linked to the PtIntervention table, which contained details on the type and timing of interventions for individual patients. The PtIntervention table linked to the PtInterventionData table via the PtInternentionId field. PtInterventionData contained additional details on the intervention performed, including type, name, value and units of measurement.

Both PtIntervention and PtInterventionData linked to PtEncounter, therefore linking the Intervention group with the Patient group. The Intervention group could also be linked to the Observation group via the InterventionObservation table (Figure 6.3).



Figure 6.3: Relationships between the Intervention group of tables in CIS

ORDERS GROUP

The Orders Group contained two main tables – Orders and PtOrder. These tables contained details of prescriptions of drugs, fluids and blood products. The Orders table contained 135 distinct orders, each with a unique OrderId, but many of these had generic titles that could be populated by a range of different drugs, infusion or interventions. Further details were given in PtOrder, which contained 4576 distinct rows of Orders, and included information on drug name, dosage, frequency and prescription instructions.

Each PtOrder record linked to the PtEncounter table via the ptEncounterId field (Figure 6.4).

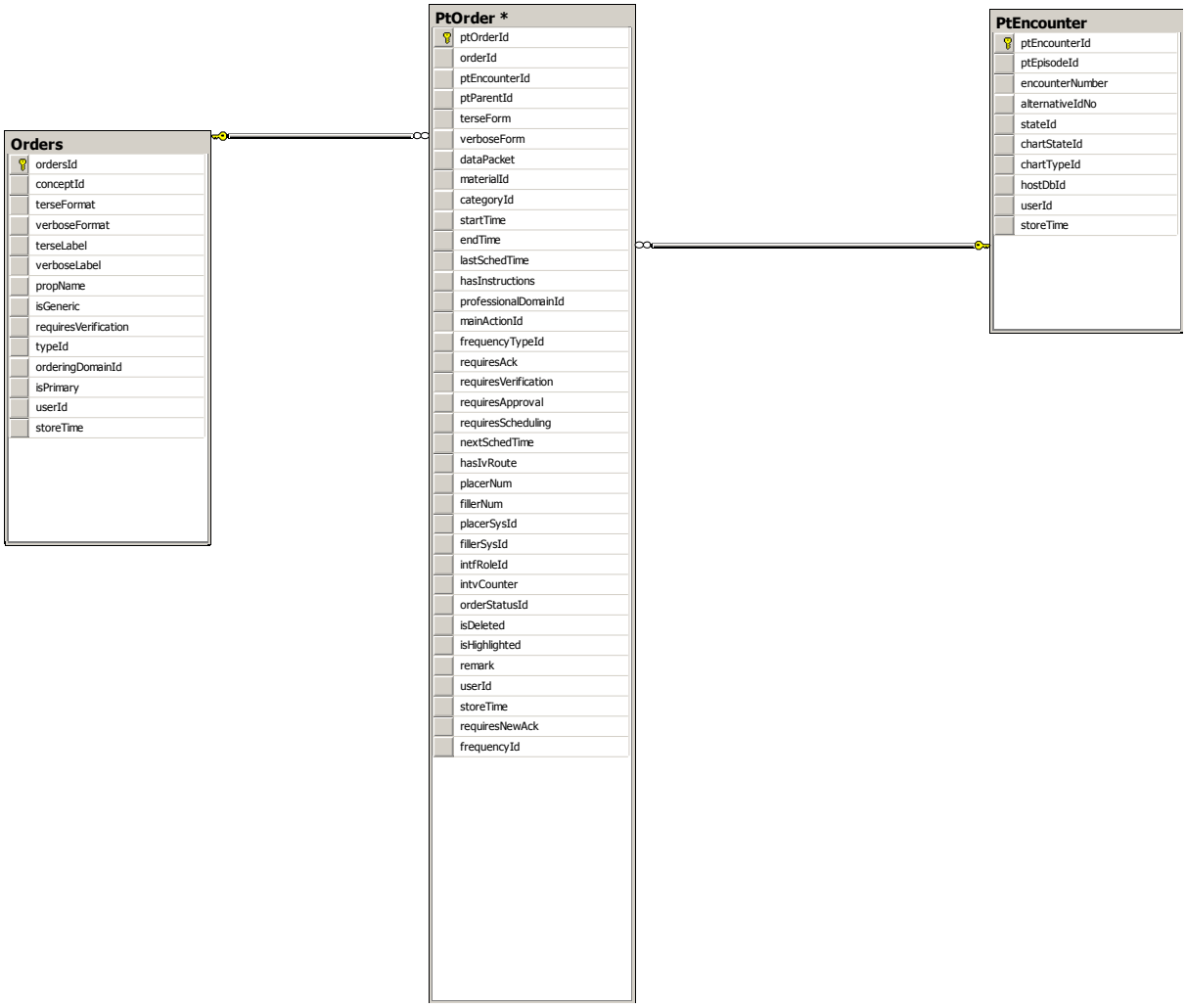


Figure 6.4: Relationships between Orders group of tables in CIS

The Orders group also linked to the Intervention group via a table called PtOrderIntervention (Figure 6.5).

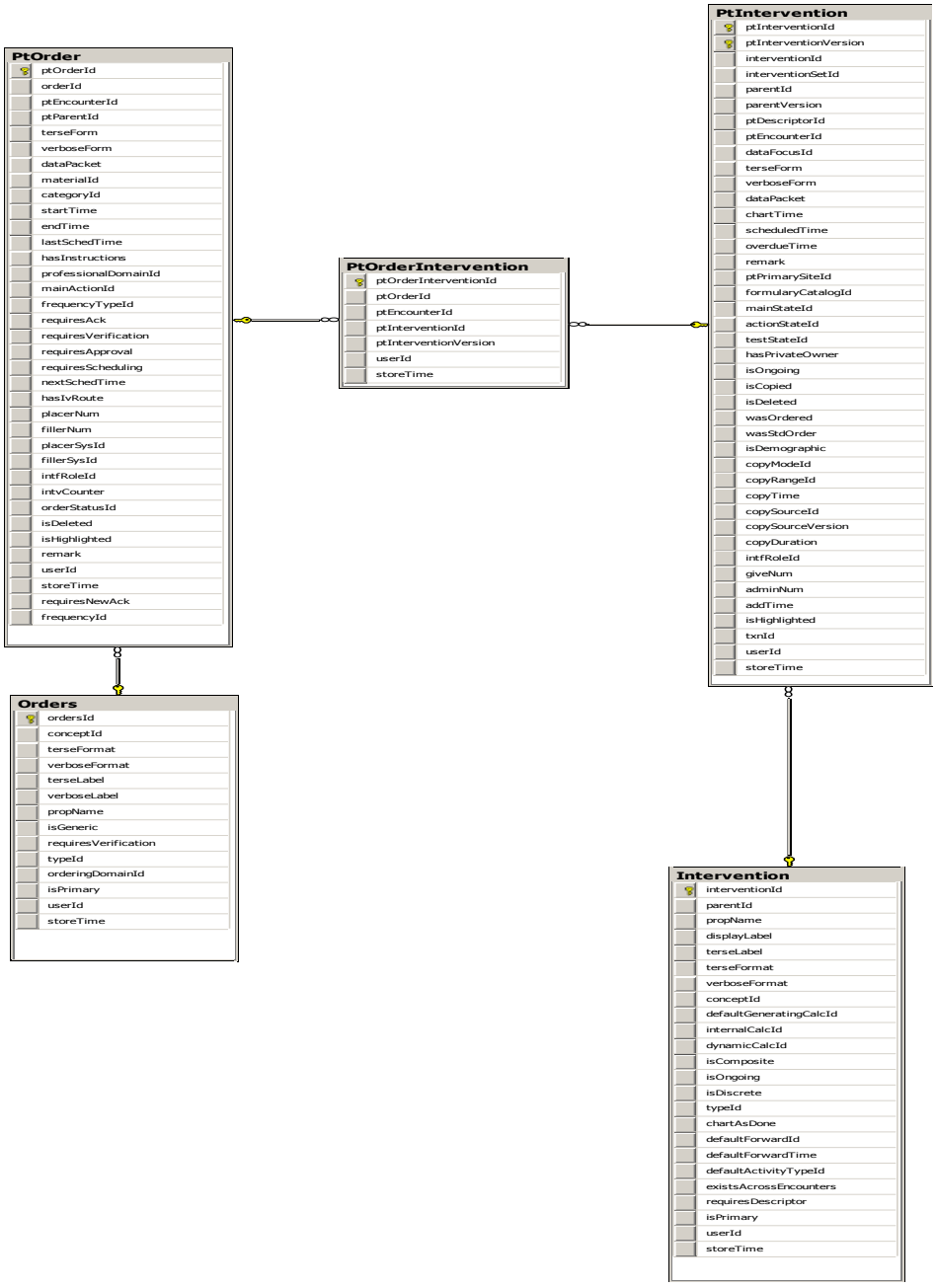


Figure 6.5: Linkage between Orders and Intervention Groups in CIS

The overall relationships between the Patient, Observation, Intervention and Orders groups are shown in schematic form in Figure 6.6.

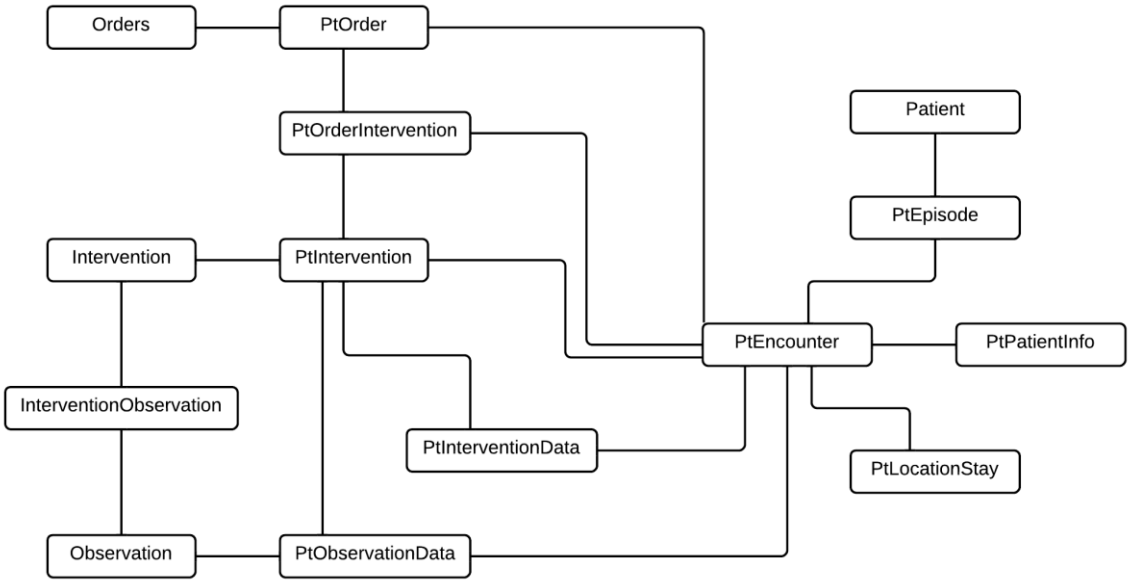


Figure 6.6: Schematic diagram of relationships between main data tables in CIS

6.2.3. CIS EXTRACTION PROCESS

To simplify the queries used for extraction, data were extracted from each group of tables separately. To ensure relevant data were obtained, the following process was used:

1. Assessment of all rows in the Observation, Intervention and Orders tables to identify those potentially relevant to the variable in question
2. Assessment of any associated fields required for data interpretation
3. Initial SQL query to establish if data stored in the tables with patient clinical data (PtObservationData, PtIntervention, PtInterventionData, PtOrders) matched to the Id numbers of the rows identified in stage 1.
4. Assessment of relevance of data
5. Extraction of relevant data

The full details of the rows identified in the Observation, Intervention and Orders tables for each variable, and their assessment for content and relevance are given in the data extraction tables in Appendix 7. For each variable extracted from CIS, a table is given which shows which rows were identified as potentially relevant, the results of assessment for content and whether data were then extracted for that row.

An extensive overlap was noted between the content of the Observation and Intervention tables, with the distinction between the two unclear. Many more Observations were listed than Interventions (4938 versus 975). However, the CareVue system has been allowed to grow organically, with users of the live system able to input new rows in certain fields. This has led to considerable redundancy with a high number of replicated rows for a single variable.

Initial SQL queries were run to assess whether potentially relevant rows identified had been used for data collection. Examples of the SQL queries for each group are given in Appendix 9. For the Orders group, different queries were run using wildcard searches, as well as specific field names to maximise the data return.

Once relevant data was established, repeat SQL queries were run to extract the data. Examples of these are similarly given in Appendix 9. Data were extracted into .csv files.

6.2.4. ISM STRUCTURE

The structure of ISM was simpler than the structure of CIS. ISM contained 35 tables, 9 of which were not used for data storage. The remaining 26 tables were assessed for linkage and structure, which had not been retained in the archive system. Unlike with CIS, it was

not possible to recreate table relationship diagrams in SQL Server, and so the linkage structure needed to be determined by hand.

Initial attempts at linkage of the 26 tables resulted in multiple errors. It was noted that several field names were repeated in different tables, but contained different information. It also became apparent that several modifications had been made to some of the tables, believed to relate to a previous attempt at database reconstruction. From the 26 tables, seven were found to contain data not relevant to this study – D_Days, M_linktable, D_Caregivers, D_CareUnits, Deliveries, D_Sources and D_Patients. No further attempts were made to link these tables, leaving 19 tables remaining.

The ISMAdmission table was a critical table for linking information as it contained patient identifiers and information on each admission. ISMAdmission contained a field called Pid (PatientId) which contained a unique identifier for each patient, and ISMAdmissionId which contained a unique identifier for each admission. The field called MRN was found to be the patient's hospital number, although a number of errors were found in this field due to the way in which the data were archived. Errors in MRN were identified by matching at least three further patient identifier fields in ISM to the details in the ICU Admin database to ensure MRN could be used for matching patients to data from other sources.

The ISMAdmission table was further modified to add a TraumaDb field to mark trauma patients included in the study, and a TraumaDuplicate field to mark trauma patients who were duplicated in ISM.

From the remaining 18 tables, three main groups were identified

1. Chart Group
2. Med Group
3. IO Group

The linkages for each of these groups were established separately.

CHART GROUP

The Chart group consisted of three tables:

- Chart Events
- D_Chartitems
- A_Chartdurations

ChartEvents contained data on all events similar to those that would be listed on a patient chart. D_Chartitems contained listings of all possible patient chart events, each with a separate ItemId. 9175 separate rows were identified covering investigations ordered, laboratory results, observations and nursing care procedures. Analysis of these 9175 rows showed around 2000 were programmed into the original database, and the other 7000 had resulted from new fields being inputted by users of the live system. There was considerable redundancy in the additional rows. A_Chartdurations contained information on the timing and duration of the chart item.

The three tables could be linked using the ItemId field. ChartEvents contained the Pid field which allowed it to be linked to ISMAdmissions. A simplified diagram of the linkage between tables is shown in Figure 6.7.

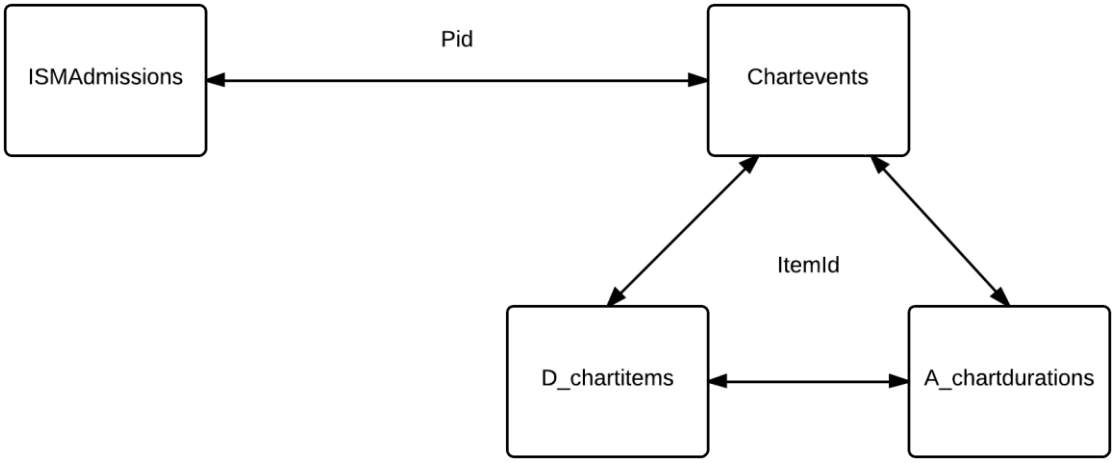


Figure 6.7: Relationships between Chart group of tables in ISM

MED GROUP

The Med group consisted of four tables:

- MedEvents
- D_meditems
- A_meddurations
- MedOrders

The Med group contained data on all medications given, and covered all routes of administration. However, many medications were also listed in the Chart group and IV medications were listed in the IO group.

MedEvents listed individual medication events for each patient. D_meditems contained 5219 rows representing different medications, each with a unique ItemId. The majority of these rows were added by users of the live interface, with replication of medications, doses and routes of administration. A_meddurations contained details on the timing and

duration of medications. MedOrders contained similar information to a prescription chart, including any additional instructions from the prescribing doctor.

The four tables of the Med group could all be linked using the ItemId field. Whilst this field had the same name as a field used to link tables in the Chart group, it contained different data. MedEvents could be linked to ISMAdmissions by the Pid and ISMAdmissionId fields (Figure 6.8).

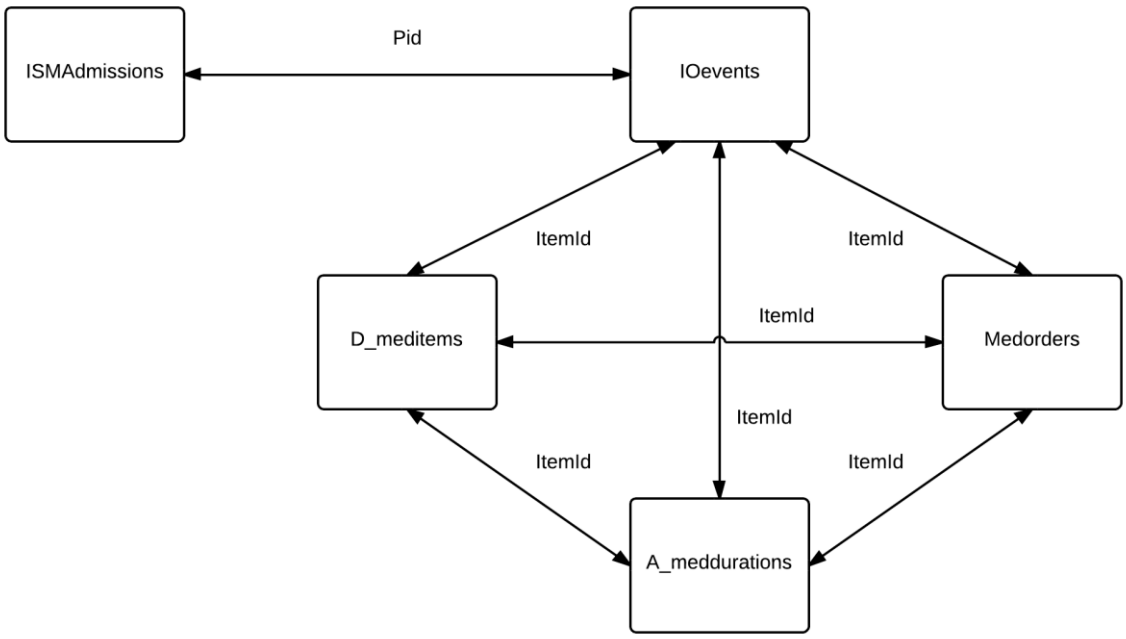


Figure 6.8: Relationships between Med group of tables in ISM

IO GROUP

The IO group consisted of the following tables, containing information on patient input and output:

- IOEvents
- D_IOItems

- A_IOdurations
- InfusionOrders
- Solutions

IOEvents listed all individual measures of fluid input and output for the ICU patients. D_IO items contained 4959 rows with a distinct ItemId, and, as with the Chart and Med groups, a great deal of duplication and redundancy was noted. The content of the tables mirrored that in the Med group, with IOEvents listing all patient events, A_IOdurations detailing timing and duration, and InfusionOrders being the database equivalent to a fluid prescription chart. Solutions was an additional table that linked to InfusionOrders and contained details on the timing, volume and route of administration of certain fluids.

The linkage structure of the tables in the IO group mirrored the Med group. IOEvents, D_IOitems, A_IOduration and InfusionOrders were all linked by the ItemId field. It should again be noted that despite the same name, this was a different field to the ItemId fields seen in the Chart and Med groups. InfusionOrders linked to Solutions by a field called IOItemId. IOEvents linked to ISMAdmissions by the Pid field (Figure 6.9).

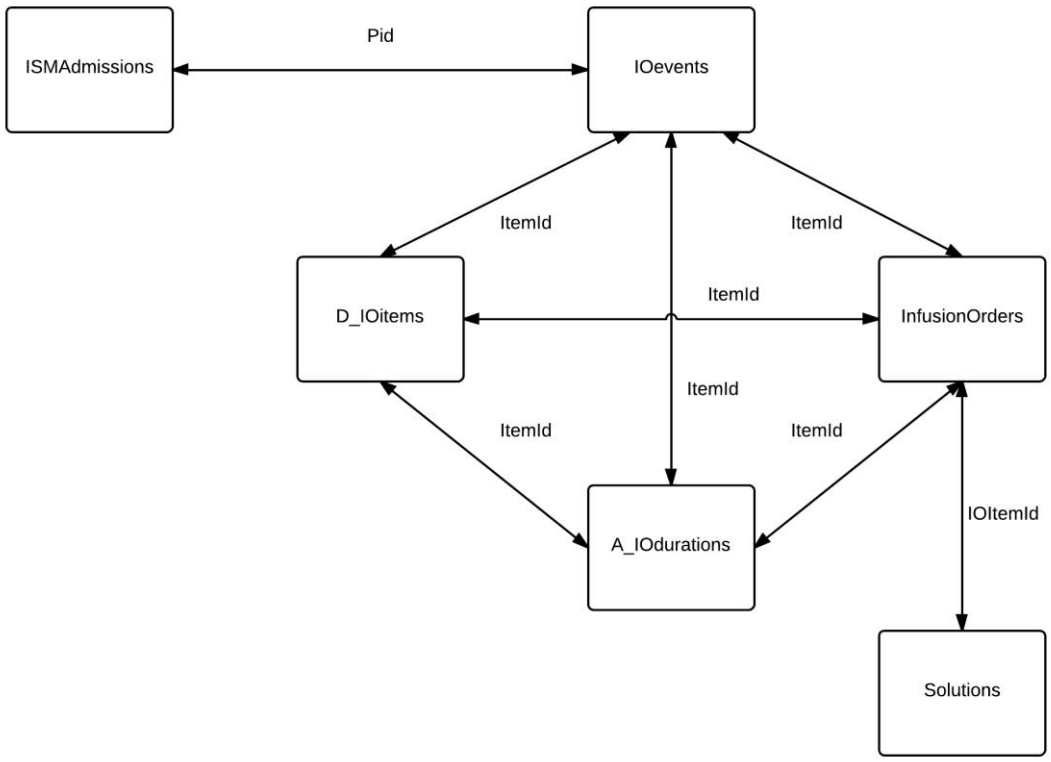


Figure 6.9: Relationships between IO group of tables in ISM

OTHER TABLES

The Chart, Med and IO groups and the ISMAdmissions table accounted for 13 of the 19 tables containing data. The remaining tables could be divided into "Event" tables and "Orders" tables.

The "Events" tables were FormEvents, Totalbalevents, Censusevents, Noteevents. All contained the Pid field and could be linked to ISMAdmissions, but could not be linked to any other tables. They appeared to contain mostly free text notes on aspects of patient care. Both FormEvents and Totalbalevents contained an ItemId field, but this was different field to the other ItemId fields already identified, and could not be matched to any other fields.

The “Order” tables were Freeformorders and Driporders. Both contained the Pid field and could be matched to ISMAdmissions. Both also contained an ItemId field, but again this did not correlate to any other fields with the same name.

Due to the problems in linkage of these tables and the nature of the data they contained, it was decided to explore them further only if data required for extraction could not be found in the other tables.

6.2.5. ISM EXTRACTION PROCESS

The process described for extraction from CIS was adapted for ISM as follows:

1. Assessment of all rows in D_chartitems, D_meditems and D_IOitems tables to identify those potentially relevant to the variable in question
2. Assessment of any associated fields required for data interpretation
3. Initial SQL query to establish if data stored in the fields related to ItemId numbers identified in Step 1
4. Relevance of data assessed
5. Extraction of relevant data

The full details of the rows identified from D_chartitems, D_meditems and D_IOitems tables for each variable, and their assessment for content and relevance are given in the data extraction tables in Appendix 7.

Examples of the initial SQL queries run to assess whether data was stored for particular rows, and the final SQL queries used for data extraction are given in Appendix 9. Queries for both specific ItemId numbers and wildcard searches were used for the Med and IO groups. Data were extracted into .csv files.

6.3. DATA EXTRACTION FROM TARN

Whilst data from local TARN files and the web interface were required for matching patients to identify the study population, all data for analysis came from the national TARN database.

A full list of the TARN Submission Ids of the study population was sent to the national TARN office, who returned all submitted data on these patients in a series of Excel spreadsheets. A complete list of field names was also supplied.

To identify the relevant data for each required variable, the field list was searched for appropriate fields. All fields were listed in tables as given in the data extraction tables in Appendix 7. Each of these fields was then assessed for data content and completeness, before extraction of data from the most relevant fields for the variable.

TARN data is divided into sections relating to the location of the patient – At Scene, Enroute, ED, Imaging, Critical Care, Operation and Ward. Many variables were entered under fields related to more than one of these sections. For the time-dependent data for physiology, laboratory and management variables, the number of relevant fields for a variable was considered for each section. Each field had an associated field with timing data, and these were also extracted.

TARN does contain fields pertaining to laboratory data, but on examining the relevant fields for content, it was found that minimal data were recorded. A decision was therefore made not to use TARN data for laboratory variables.

6.4. DATA EXTRACTION FROM ICNARC

The recreated local ICNARC database, as described in Chapter 5, was used for data extraction. The local database was recreated in Access and relevant data were extracted using SQL queries, and exported into .xls files.

ICNARC data are predominantly focussed on the first 24 hours of ICU admission, and have very specific data collection rules, detailed in the internal ICNARC data dictionary.) Data from ICNARC were used for variables that were not time-dependent, such as patient variables. ICNARC did not contain useable data on injury or management input variables. Data on physiology and laboratory variables were present in ICNARC, and this was initially extracted.

Physiology and laboratory parameters in ICNARC were defined on a daily basis, according to the ICNARC data dictionary. Most of them are defined as the most abnormal value for that 24-hour period. This meant that they could not be used for the input variables as they could not be accurately linked to time of injury. ICNARC physiology and laboratory data could be used in calculation of the Denver and SOFA scores that use the most abnormal values. For this, each daily entry was assigned the same 24-hour clock time as the patient's recorded time of admission to ICU. The data could not be used in calculation of the Marshall score, since this required representative, rather than most abnormal values.

Full details on the use of ICNARC data are given in the data dictionary. (Appendix 8)

6.5. DATA EXTRACTION FROM OTHER SOURCES

For a few pieces of data required for background information and description of the trauma population, data were required from other sources. The main example of this was the exact timing of injury, and the nature of the patient pathway, including transfers to and from other hospitals. Other sources used to obtain this information were PAS, EPR, Bluespiers, PACS and written hospital notes. Data taken from these sources were extracted and transferred to Excel spreadsheets by hand. These additional data sources were only used for data critical to the research question that could not be obtained from the main data sources.

6.6. DATA TRANSFORMATION AND CLEANING

6.6.1. ANONYMISING DATA

As soon as was practicable after data extraction, all data were anonymised and linked to the patient's unique identification number (TraumaDBID). Data from different sources was then combined using the TraumaDBID rather than any patient identifiers.

6.6.2. DATA DICTIONARY

A data dictionary was made for all input variables and all outcome variables required for calculating the Denver, SOFA or Marshall scores. The full data dictionary is given in Appendix 8.

For each variable, the fields used for data extraction were recorded for each data source. The data details were defined and included the data type, units, definition, and normal range. Rules for cleaning were defined in advance, and included hard and soft limits for

data exclusion, timing validation, and management of duplicate values. Any additional variable rules were also determined.

Reference was paid to the TARN and ICNARC data dictionaries in creating the data dictionary for this study, and to definitions employed for variables in other trauma studies.

6.6.3. DATA CLEANING

All data were cleaned according to the variable rules laid out in the data dictionary. For input variables that were time-dependent, the rules were applied to each time block to fit with model development at 2, 4, 6, 12, 24 and 48 hours following injury. To do this, data for each variable was linked with time of injury. For each patient, the data were sorted into time order, and divided into blocks to fit with the proposed model development times. The data rules as outlined in the data dictionary were then followed to obtain the appropriate value for the variable for each of the models.

Data were cleaned by hand in Excel, before importing into IBM SPSS Statistics Version 20 for analysis.

6.6.4. MISSING DATA

All missing data were categorised as missing with no attempt at imputation. The pattern of missing data was complicated, due to the nature of the study design of using retrospective electronically-held routine clinical data. A pragmatic approach was therefore taken to analyse the data that were available for each variable, whilst fully accepting the limitations of this approach in the interpretation of the results.

6.7. DATA ANALYSIS

6.7.1. OVERVIEW OF ANALYSIS

Statistical analysis of the data was performed following a pre-defined analysis plan, as given in Appendix 10 and outlined in Chapter 2 (Section 2.9.3). This was broken down into several sections:

1. Description of the whole trauma population
2. Analysis of MOF and the population at risk of MOF
3. Univariate analysis of potential MOF predictors
4. Selection of variables for multivariate analysis
5. Multivariate analysis and model building for each of the 18 models (0 – 2 hours, 0 – 4 hours, 0 – 6 hours, 0 – 12 hours, 0 – 24 hours and 0 – 48 hours for each of the Denver, SOFA and Marshall definitions of MOF)
6. Attempts to find a unifying model to predict MOF according to each of the Denver, SOFA and Marshall scores at multiple time points after injury
7. Internal validation of selected models

Further details on specific aspects of this process are given in the following sections.

6.7.2. VARIABLE SELECTION FOR MULTIVARIATE ANALYSIS

The criteria for the selection of variables for multivariate analysis were detailed in Chapter 2 (Section 2.9.3). The initial sample size calculation was based on an approximately 20 variable model, with some flexibility for an increase. However, since 84 candidate variables were described in the data dictionary, this number needed to be

reduced prior to multivariate analysis. Further details of this process are given in Chapter 7.

6.7.3. MODEL-BUILDING STRATEGY

Models were built using a backwards elimination strategy, which has been argued to be preferable to forward selection in the avoidance of over-fitting of a model.[259] The amount of missing data differed between variables, and since a pragmatic decision had been made to use only available data with no attempts at imputation, removal of a variable could potentially impact the number of patients on which a model was based. Therefore, backwards elimination was performed by analysis of the contributing variables and the extent of their missing data at each stage, rather than by a single automated approach to variable removal.

Elimination of a variable from the model at each stage was based on an assessment of the variable significance, the impact of its missing data on the population size, and its clinical relevance. Missing value analysis was used to identify patterns of variables that were linked in data collection i.e. variables that resulted from an arterial blood gas analysis were likely to be all present or all absent, depending on whether the investigation had been performed. This was taken into consideration in the elimination of variables at each stage, and variables with a large amount of missing data were eliminated earlier than those with more complete data.

For each stage of model development, a single variable was removed, taking into consideration the significance, the impact of missing data and the clinical utility of the variable. The model was then refitted, and the process repeated until minimal changes in

model parameters were seen. Variables were assessed for linearity using the Box-Tidwell procedure. Potential interaction terms were also considered in model development.

Each of the 18 models was built separately. For all variables, the coefficient (B), significance of the Wald statistic and the adjusted odds ratio (OR) were recorded. These allowed comparison of the contribution of individual variables to the model. The ORs were adjusted for the other variables included in the model, and the 95% confidence intervals (CI) for the adjusted ORs were reported. Assessment of model performance must include measures of both calibration and discrimination.[259,260]Calibration was assessed using the goodness-of-fit measures of the model χ^2, the -2 log likelihood, the Nagelkerke R^2, and the Hosmer-Lemeshow test. In broad terms, calibration measures assess the extent that observed events match with the predicted probabilities from the model. If the significance value of the model χ^2 is less than 0.05, this indicates a statistically significant model. The Hosmer-Lemeshow χ^2 is a goodness-of-fit test which compares observed to predicted probabilities by deciles of predicted probability. Low significance values (< 0.05) indicate a poorly-fitting model. The -2 log likelihood and the Nagelkerke R^2 are measures of overall model variance, and the extent to which variation in the dependent variable can be explained by the model. The lower the -2 log likelihood, the better-fitting the model is. A higher Nagelkerke R^2 value indicates more of the variance in the outcome is due to the model, indicating a better fit.

Measures of discrimination assess how accurately the model discriminates between patietns with and without the outcome. These include the category predictors of the percentage accuracy in classification (PAC), sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV), which were recorded for each stage in

model development. ROC curves were drawn for each stage, and the AUC calculated. The ROC curve is a plot of sensitivity against 1-specificity, and a higher AUC indicates improved model performance.

The selection of the most appropriate model from the multiple stages of model development was based on a combination of the measured parameters of goodness-of-fit and accuracy. Since the clinical intention was for a means of predicting patients with MOF for whom interventions were likely to have significant risks and costs, specificity and positive predictive value were considered the most important category predictors. Model selection also took into account the goodness-of-fit parameters, the calibration given by the AUC, the population size on which the model was based, and the number of variables in the model.

After all 18 models had been selected, the variables used to predict MOF at all time points were considered separately for each of the Denver, SOFA and Marshall scores to look for unifying predictors. A test model was made for each outcome measure, consisting of a limited number of variables and was applied at 2, 4, 6, 12, 24 and 48 hours post-injury to assess its potential use as a clinical predictor.

To assess the generalizability of the models, the populations included in the test models were compared with the group of patients excluded from model development due to missing data.

6.7.4. INTERNAL VALIDATION

Internal validation was performed on the final test models for each of the Denver, SOFA and Marshall scores using the Bootstrapping add-on for SPSS Version 20. Bootstrapping

was used to assess the bias and confidence intervals of coefficients for individual parameters within the models. 1000 bootstrapping samples created by sampling with replacement were analysed for each model, to obtain estimates of the bias and standard error of the coefficient (B) for each variable. Bias-corrected and accelerated confidence intervals were calculated for each coefficient.

6.8. SUMMARY OF METHODS

A list of input variables with the potential to be predictive of post-trauma MOF was drawn up following identification by literature review and questionnaire analysis of trauma professionals. Literature on MOF scoring systems was reviewed, and the Denver, SOFA and Marshall scores chosen to define MOF and allow comparative analysis. The study population was determined according to the inclusion/exclusion criteria by matching across the TARN, ICNARC/ICU Admin and CareVue databases. 517 patients were suitable for inclusion in the study.

Data on these 517 patients were extracted from TARN, ICNARC and the two parts of the archived system of CareVue – ISM and CIS. A methodical approach to data extraction was employed, and data was anonymised as soon as practicable. A formal data dictionary was created for each input variable and each variable that was a component of the outcome scores. Data was cleaned according to the rules laid out in the data dictionary. The final data were imported into SPSS Version 20 for analysis.

Data were analysed in a systematic fashion, according to a pre-defined analysis plan. First, the whole trauma population and the population at risk of MOF were defined, and the incidence of MOF examined. Next, univariate analysis was performed for potential MOF predictors. Strict criteria were applied to choose variables for multivariate analysis. Models were built at a range of times from 2 hours to 48 hours post-injury, with each of the Denver, SOFA and Marshall definitions of MOF as an outcome measure. Model

building followed a backwards elimination approach by hand, and models were assessed for calibration and discrimination.

After model selection, variables were compared across models to find a unifying set of variables to predict each of Denver, SOFA and Marshall MOF. These were tested across the full range of time points post-injury for comparison. Internal validation using bootstrapping and cross-validation techniques was performed to assess the final test models.

CHAPTER 7 RESULTS

7.1. TRAUMA POPULATION DESCRIPTION

7.1.1. PATIENT DEMOGRAPHICS

The study population consisted of 517 patients with a mean age of 42.7 years (range 16-92 years), of whom 391 (75.6%) patients were male. Age distribution for males showed a marked skew to younger age groups that was not present in females (Figure 7.1), resulting in a significant difference in mean age between sexes. (Ratio of geometric means 0.90, p = 0.027).

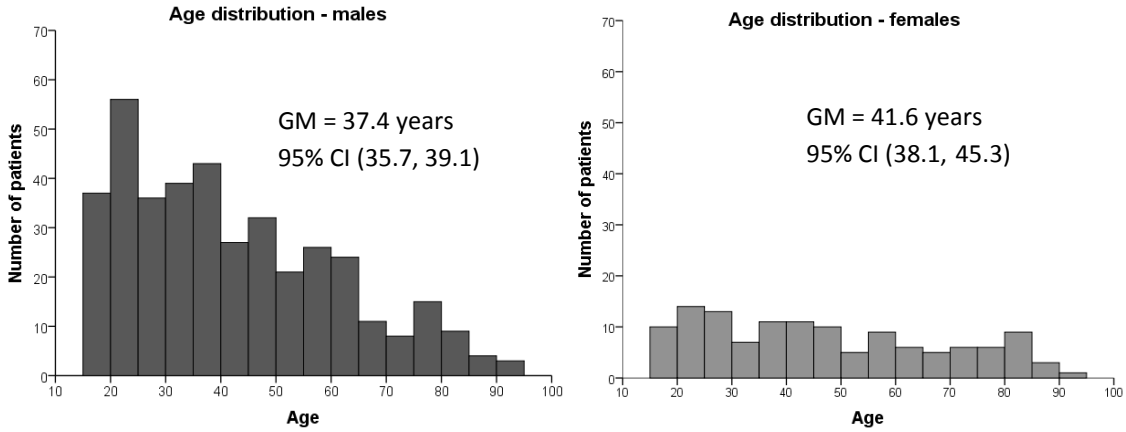


Figure 7.1: Age distribution for a) males and b) females.

(Data presented as geometric mean (GM) and 95% confidence intervals (CI).)

7.1.2. INJURY CHARACTERISTICS

Patients in the study were severely injured, with a median ISS of 25 and NISS of 34 (Table 7.1). A high proportion of patients had severe injuries to the head, chest, abdomen and pelvis or limbs, as defined by an AIS for that body region of ≥ 2. First recorded GCS was less than 15 in 338 patients (66.3%).

Injury severity n = 512 (or stated)	**Number of patients (%)**
ISS	25 (17-34)[1]
NISS	34 (25-43)[1]
AIS ≥ 2	
Head	305 (59.6%)
Face	101 (19.7%)
Chest	305 (59.6%)
Abdomen	172 (33.6%)
Limbs/Pelvis	326 (63.7%)
External	19 (3.7%)
First GCS < 15 (n = 510)	338 (66.3%)

Table 7.1: Injury severity details for trauma population

[1] Median (Interquartile range)

Blunt trauma affected 502 patients (97.1%), with RTCs the predominant mechanism (Table 7.2). Of those injured in RTCs 40% were drivers, and 24% were motorcyclists. According to the TARN criteria, 163 (31.5%) patients were trapped at scene, defined as requiring any form of specialist equipment for extrication.

Injury details n = 517 (or stated)	**Number of patients (%)**
Blunt trauma	502 (97.1%)
Mechanism	
Blast	1 (0.2%)
Blow	25 (4.8%)
Crush	7 (1.4%)
Fall < 2m	46 (8.9%)
Fall > 2m	55 (10.6%)
Shooting	2 (0.4%)
Stabbing	12 (2.3%)
RTC	354 (68.5%)
Other	15 (2.9%)
Trapped at scene	163 (31.5%)
Position in vehicle (n = 354)	
Driver	143 (40.4%)
Front seat passenger	34 (9.6%)
Rear seat passenger	33 (9.3%)
Motorbike	85 (24.1%)
Cyclist	20 (5.6%)
Pedestrian	39 (11.0%)

Table 7.2: Injury details for trauma population

7.1.3. SURVIVAL

A total of 439 (84.9%) patients were alive at 100 days post-injury. Of the 78 deaths, 26 (33.8%) occurred in the first 48 hours after injury. Only 6 (7.8%) deaths occurred later than 20 days after injury, with the latest death at day 79 post-injury (Figure 7.2).

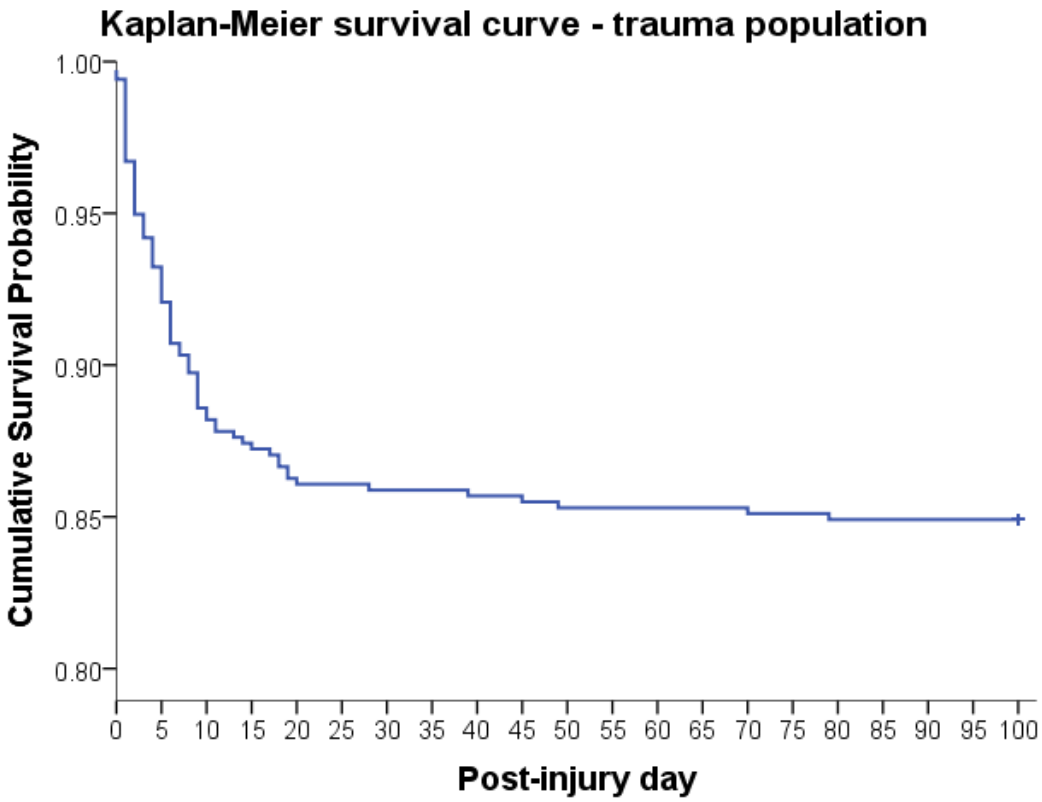


Figure 7.2: Kaplan-Meier survival curve for trauma population

In addition to the survival curve, the number of deaths each day was plotted to determine if the distribution matched the trimodal distribution (Figure 7.3). The first peak of deaths would have occurred on scene or in ED and were therefore excluded from the trauma population. A clear second peak on days 1 and 2 was seen. Whilst 52 (68.4%) deaths occurred after day 2, the timing and presence of the third peak was less clear, with small numbers of deaths occurring on most days up to day 20.

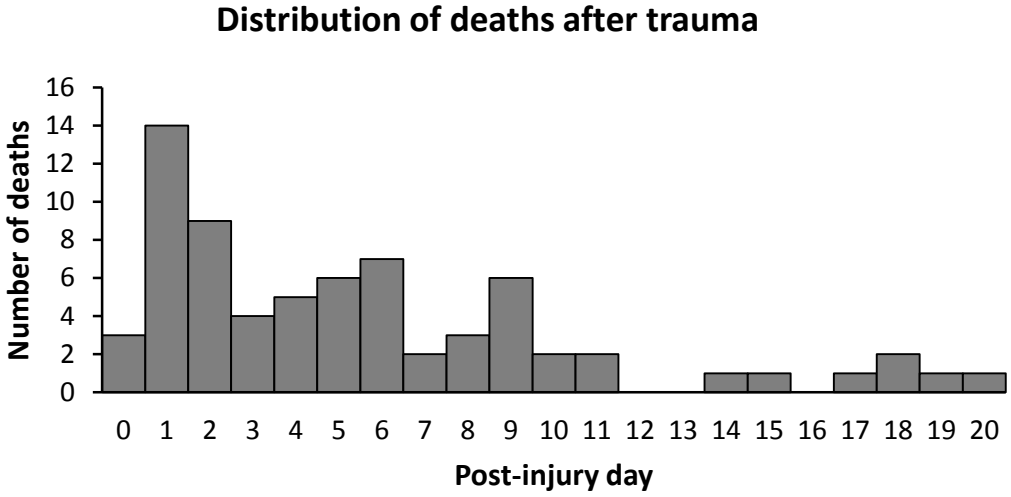


Figure 7.3: Distribution of deaths by post-injury day

Survivors had a significantly lower mean NISS of 34.4 (SD 14.7) compared to 39.4 (SD 19.2) in non-survivors ($p = 0.033$). However, whilst mean ISS was lower in survivors than non-survivors (25.2 versus 27.1) this difference was not significant ($p = 0.232$). There was no association between survival and trauma type ($\chi 2 = 0.037$, $df = 1$, $p > 0.999$). No significant relationship was found between whether a patient was admitted directly or transferred from a surrounding hospital and survival ($\chi 2 = 1.063$, $df = 1$, $p = 0.302$).

Survivors were significantly younger with a mean age of 40.5 years (SD 17.6), compared to 55.1 years (SD 24.7) in non-survivors ($p < 0.001$). Females had a higher relative incidence of death than males ($\chi 2 = 4.003$, $df = 1$, $p = 0.045$), but this association was not persistent after adjustment for age ($p = 0.204$).

7.1.4. TIMINGS

From the total population of 517 patients, 401 (77.6%) were direct admissions to the John Radcliffe Hospital. The remaining 116 (22.4%) were transferred in from surrounding

hospitals within 48 hours of injury. The median time to arrival at the first receiving hospital was 70 minutes (Inter-quartile range (IQR) 53 – 95), with 506 (97.9%) patients admitted within 12 hours of injury. The median time to arrival at the John Radcliffe for both direct admissions and transfers was 82 minutes (IQR 60 – 200), with 469 (90.7%) arriving within 12 hours of injury. 82.8% transfers were admitted within 24 hours of injury (Figure 7.4).

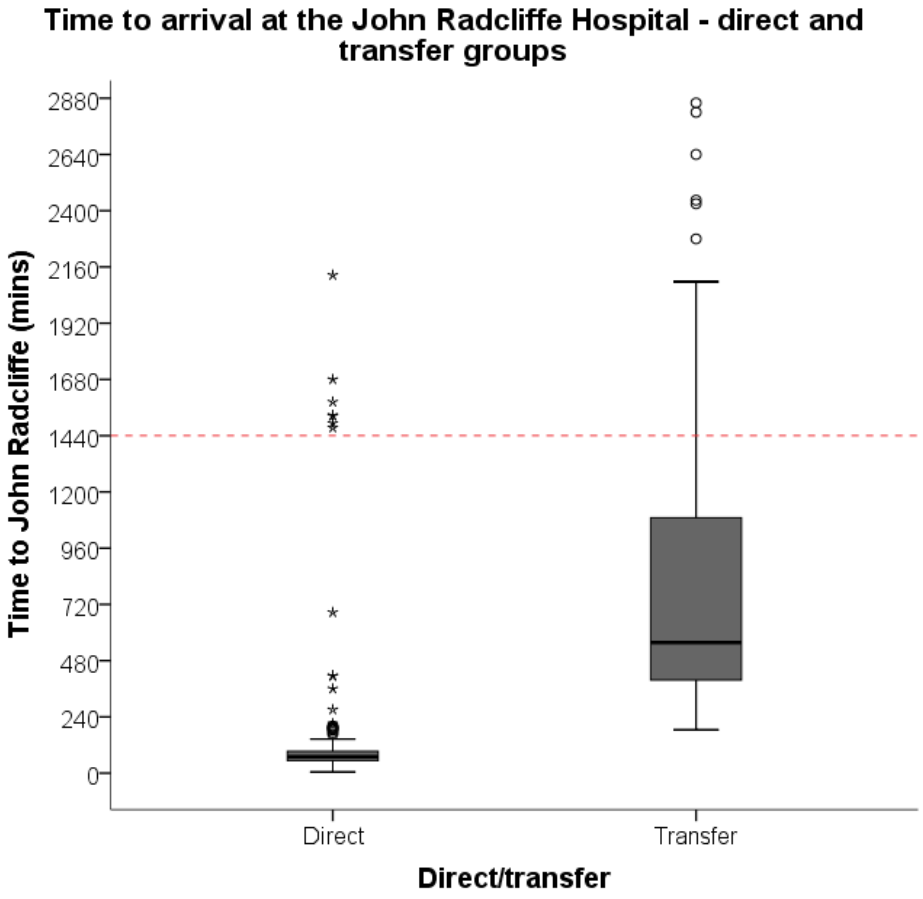


Figure 7.4: Box and whisker plot of time to arrival at the John Radcliffe for direct admissions and transfers.

(The dashed horizontal line represents 24 hours after injury.)

Admission to AICU occurred on the day of injury in 287 (55.5%) patients, with 461 (89.2%) admitted by day 2 post-injury. Nine patients (1.7%) required AICU admission after day 20 post-injury, with the latest admission on day 69. The median AICU stay was 4 days (IQR 1 – 9) which was not significantly reduced with the removal of non-survivors. Overall, 286 (55.3%) patients had an AICU stay of less than five days, and 390 (75.4%) patients had an AICU stay of less than ten days. Six patients (1.2%) required an AICU stay in excess of 40 days, with the maximum length of stay being 71 days (Figure 7.5).

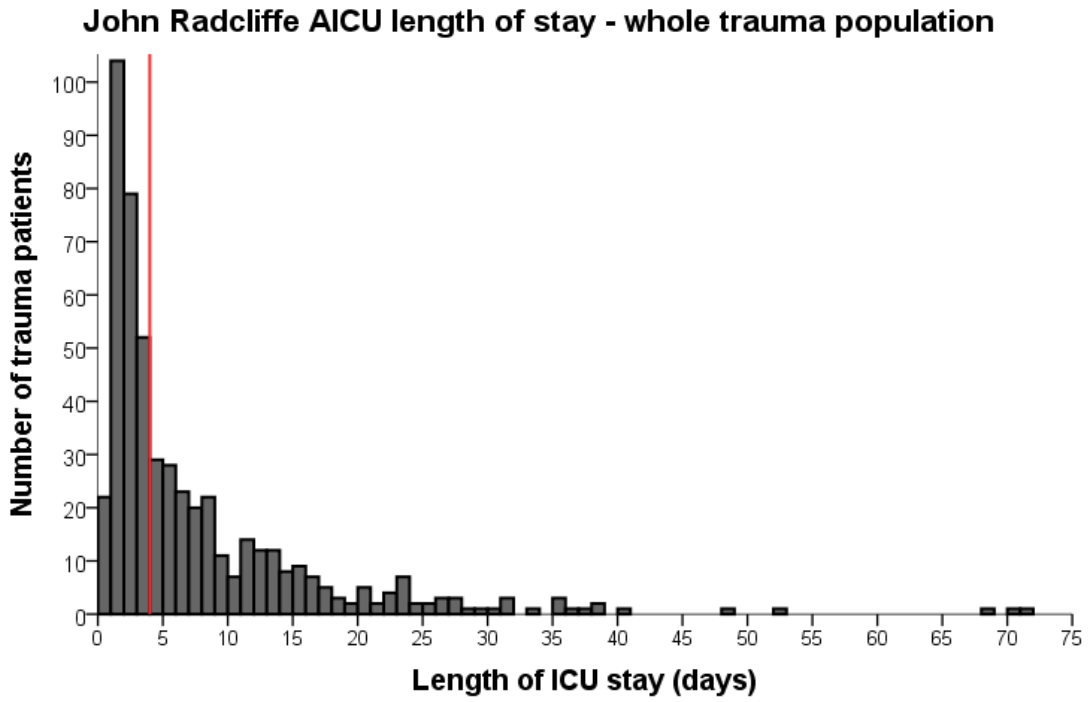


Figure 7.5: Length of AICU stay in days.

(Vertical line represents the median stay of 4 days.)

For the patients who survived, the median day of hospital discharge was day 23 post-injury (IQR 13 – 40). A total of 17 patients (3.9%) had extended hospital stays of over 100 days, with the latest discharge at Day 235 post-injury. At 28 days post-injury, 170 (38.7%) patients remained in hospital (Figure 7.6).

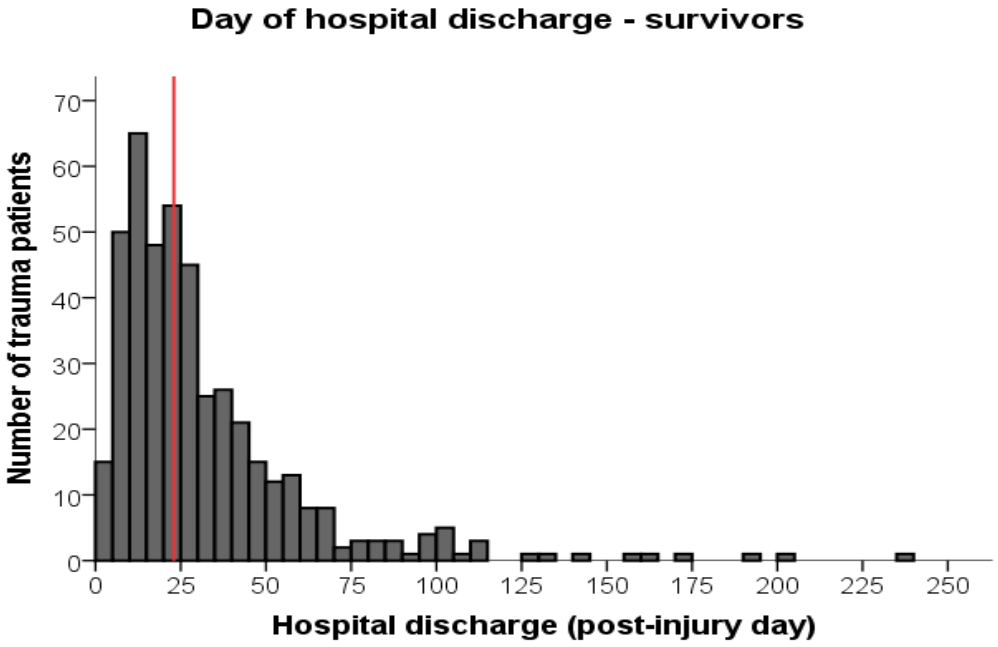


Figure 7.6: Day of hospital discharge for trauma survivors, given as post-injury day.

(Vertical line represents the median day of hospital discharge = day 23 post-injury.)

7.1.5. INTERVENTIONS

In total, 390 (75.4%) patients required an operation during their hospital day, with 233 (54.9%) undergoing a surgical procedure within the first 24 hours of admission. For the 390 patients requiring surgery, the median time to operation was 15.9 hours. From all operations performed, 29.0% took place within six hours of injury, and 70.3% within 48 hours.

Overall, 360 (69.6%) patients did not require transfusion of any blood products. From those patients who were transfused, 89 (17.2%) required 1 – 4 units, and 37 (7.2%) required a massive transfusion, defined as ≥ 10 units of blood products within the first 24 hours after injury. The maximum number of units given was 76 (Figure 7.7).

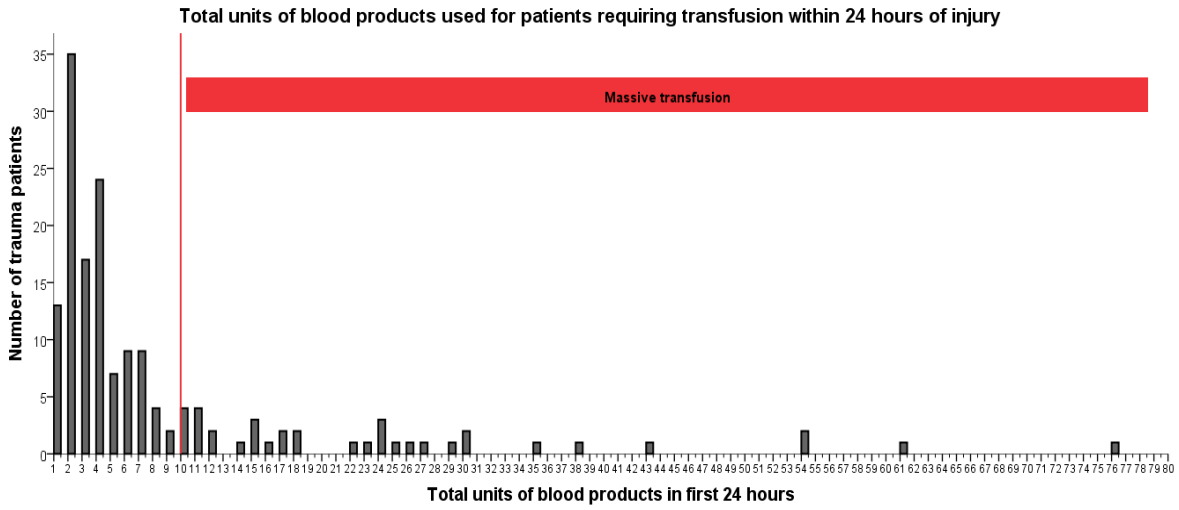


Figure 7.7: Total number of units of blood products used for patients requiring transfusion within 24 hours of injury

7.2. Population at Risk of MOF

7.2.1. Population Numbers

From the 517 trauma patients identified, 26 died within 48 hours of injury, and therefore did not survive long enough to risk developing post-trauma MOF. For all further analyses, these patients were excluded, leaving a population of 491 patients at risk of MOF (Figure 7.8).

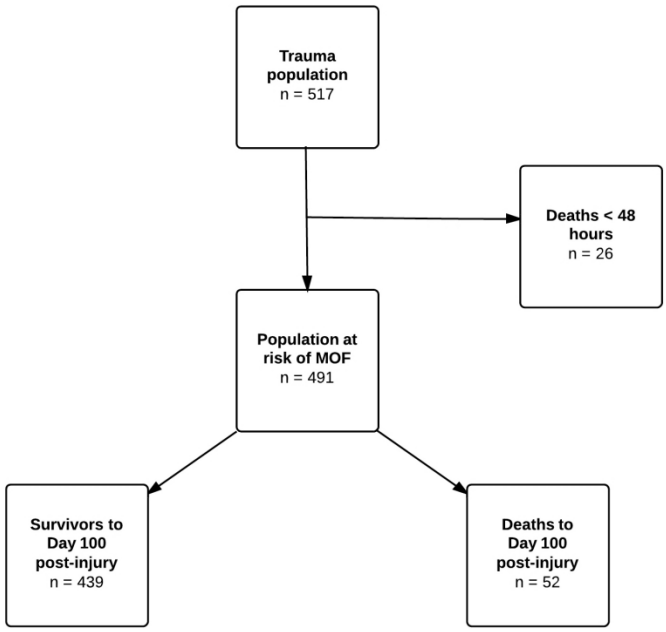


Figure 7.8: Flow diagram of patient numbers for population at risk of MOF

7.2.2. INCIDENCE OF MOF

From the 491 patients at risk of MOF, 202 (41.1%) patients did not develop MOF according to any of the Denver, SOFA or Marshall scoring systems. In total, 112 (22.8%) patients had MOF according to the Denver score, 199 (40.5%) according to the Marshall score, and 287 (58.5%) according to the SOFA score.

There were no patients classified with MOF according to the Denver score that were not also classified by the SOFA score. 2 patients were classified with MOF by the Marshall score only, and 79 with the SOFA score only. 11 patients were classified by Denver and SOFA scores, but not by the Marshall score, and 96 patients were classified by the SOFA and Marshall scores, but not by the Denver score. No patients were classified by both the Denver and Marshall scores without also being classified by the SOFA score. 101 patients were classified with MOF by all three scoring systems (Figure 7.9).

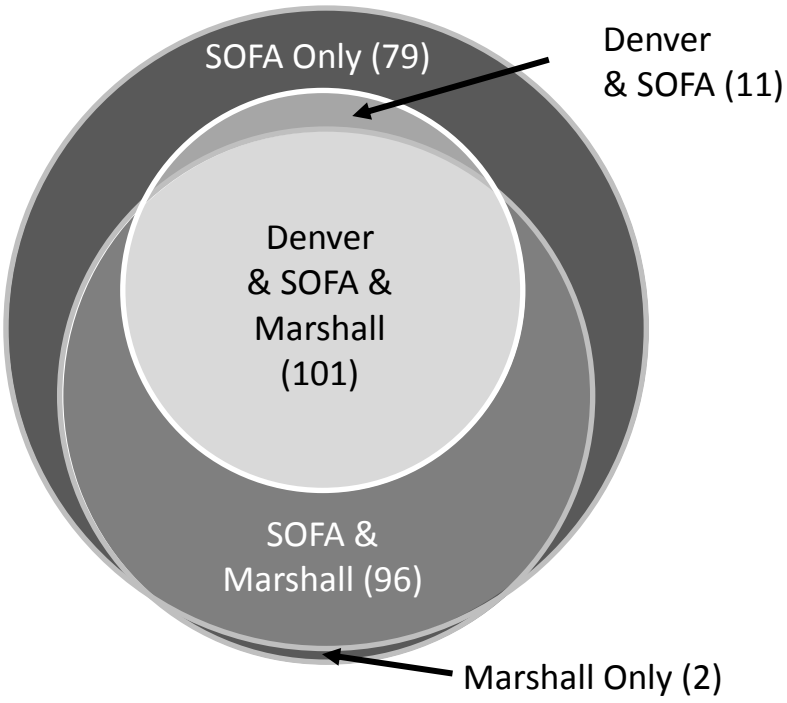


Figure 7.9: Classification of 289 patients developing MOF according to scoring system.

(Numbers in brackets indicate number of patients classified by each combination of scores.)

7.2.3. MOF AND MORTALITY

From the 491 patients who survived at least 48 hours and were at risk of MOF, a further 52 deaths were recorded in the first 100 days post-injury (10.6% patients). The relationship between mortality and MOF was examined for each MOF scoring system.

DENVER

Of the 112 patients defined as having MOF according to the Denver score, 27 (24.1%) died within the first 100 days after injury. A significant association was found between Denver MOF and mortality (χ^2 = 27.995, df = 1, $p < 0.001$), but 25 of the 379 patients without MOF (6.6%) died within 100 days of injury. The mean time to death in the Denver

MOF group was 19.1 days (95% CI 11.5, 26.7) compared to 7.4 days (95% CI 5.8, 9.0) in the no MOF group (p = 0.006). The latest death in the no MOF group occurred on day 18 post-injury, compared to day 79 in the MOF group. Figure 7.10 shows the difference in survival curves between the MOF and no MOF groups.

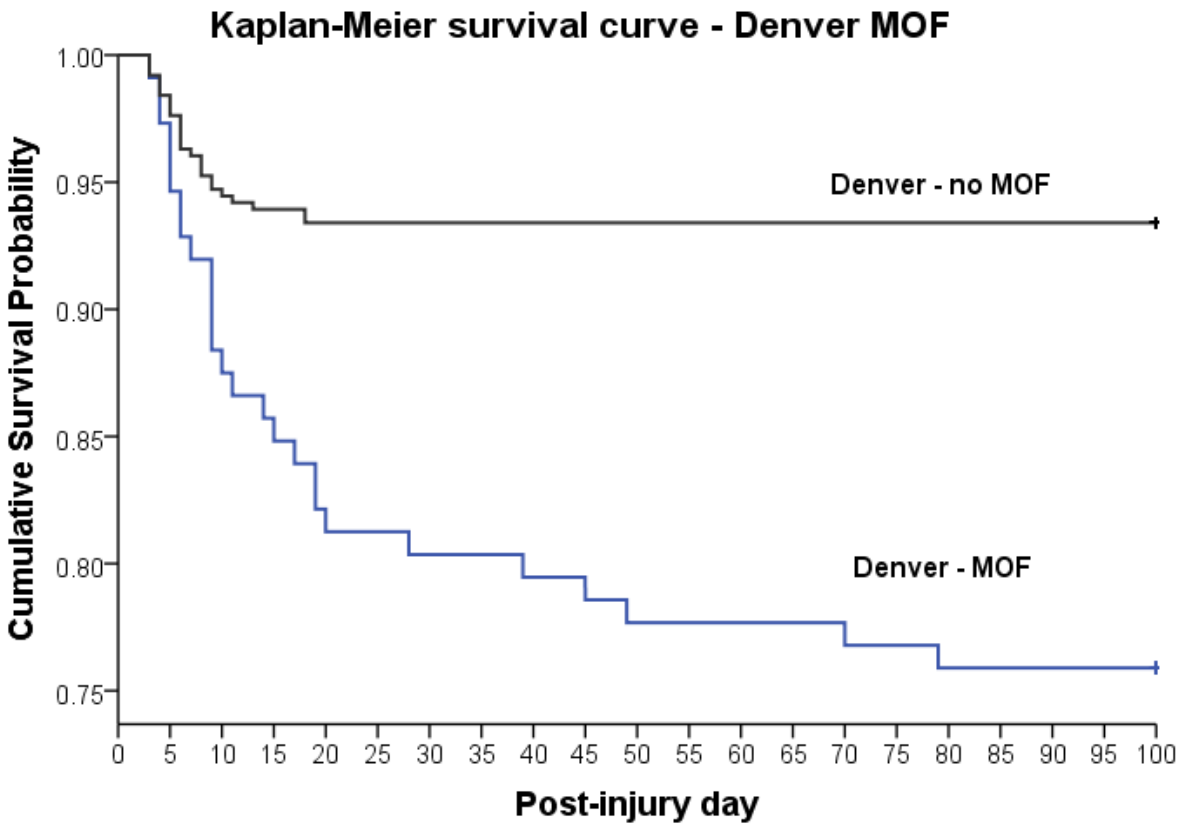


Figure 7.10: Kaplan-Meier survival curves categorised into MOF and no MOF groups according to the Denver score

SOFA

From the 287 patients defined as having MOF according to the SOFA score, 38 (13.2%) died within 100 days of injury. From the 204 patients without MOF, there were 14 deaths (6.9% patients), and a significant association was found between SOFA MOF and mortality (χ^2 = 5.122, df =1 , p = 0.024). Deaths in the SOFA MOF group occurred at a

mean of 16.9 days (95% CI 10.6, 23.1), significantly later than deaths in the no MOF group (7.6 days, 95% CI 5.3, 9.8, p = 0.013). Survival curves for the MOF and no MOF groups are given in Figure 7.11. No deaths were seen in the no MOF group after day 18 post-injury.

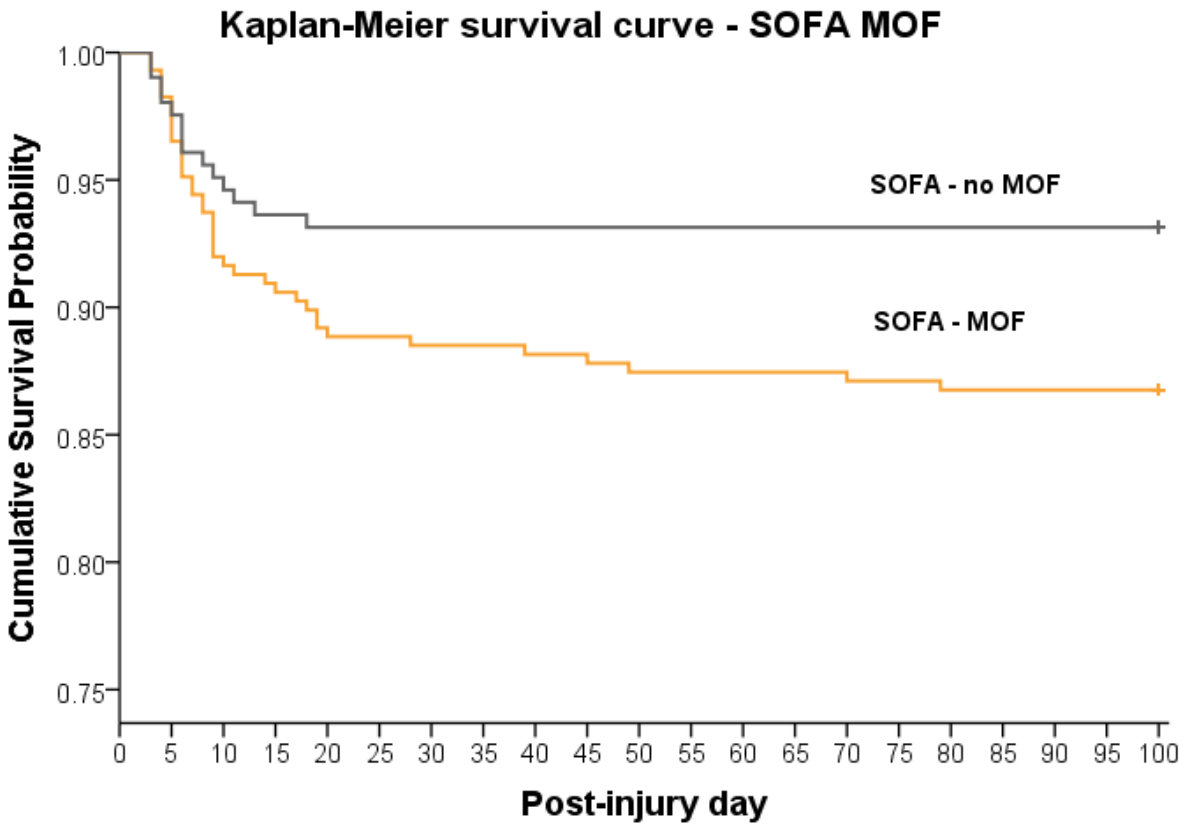


Figure 7.11: Kaplan-Meier survival curves categorised into MOF and no MOF groups according to the SOFA score

MARSHALL

From the 199 patients defined as having MOF according to the Marshall score, 34 (17.1%) died within 100 days of injury, compared to 18 (6.2%) of the 292 patients without MOF (χ^2 = 14.907, df = 1, p < 0.001). Deaths in the Marshall MOF group occurred at a mean of 16.9 days (95% CI 5.2, 8.8), compared to 7.1 days (95% CI 10.6, 23.1) in the no MOF group (p = 0.005). The latest death in the no MOF group was on day 18 post-injury, compared to

day 79 in the MOF group. The survival curves for the MOF and no MOF groups are given in Figure 7.12.

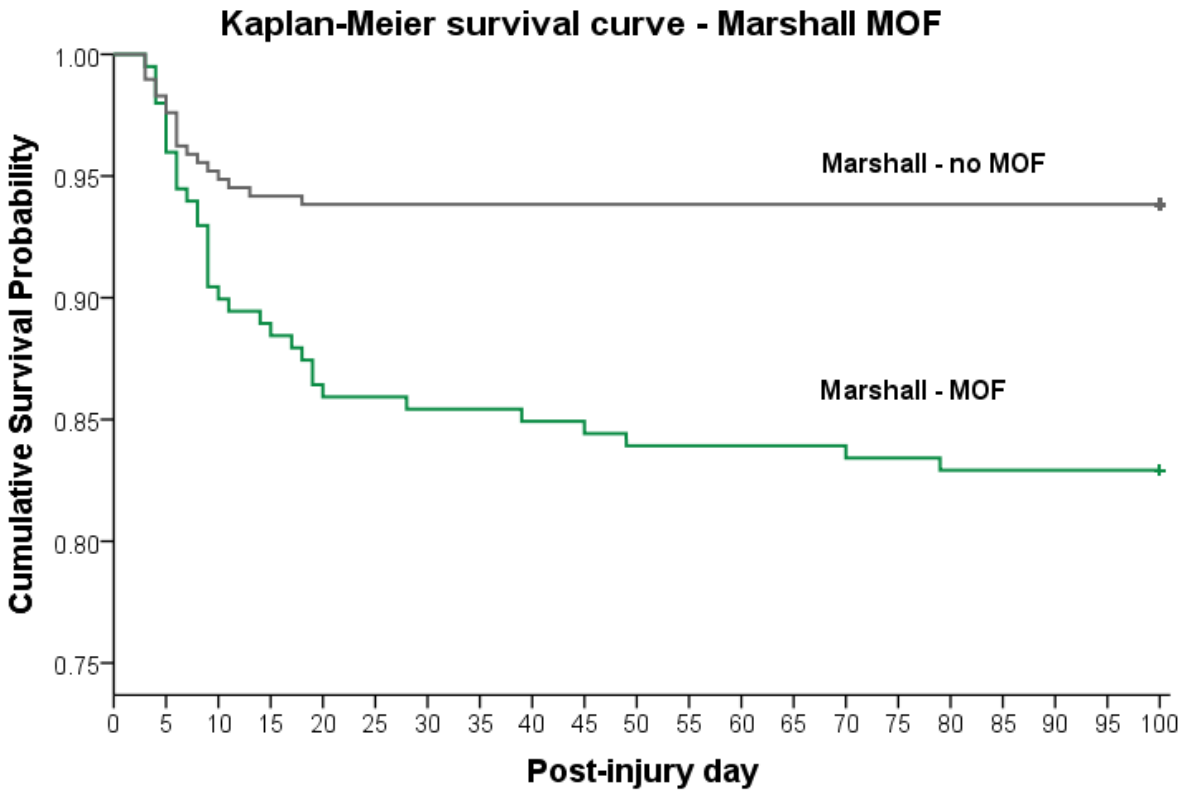


Figure 7.12: Kaplan-Meier survival curves categorised into MOF and no MOF groups according to the Marshall score

7.2.4. ONSET AND DURATION OF MOF

The onset of MOF was early for the majority of patients, with high numbers developing MOF by day 3 post-injury (Table 7.3). MOF was present in over 90% patients who developed MOF according to all three scoring systems by day 10 post-injury. A single patient developed MOF on day 69 post-injury. Excluding this patient, the latest day of onset was day 18 (Denver), day 38 (SOFA) and day 39 (Marshall).

	Number of patients with MOF	**Mean day of MOF onset** Post-injury day (95% CI)	**MOF onset by day 3 post-injury** Number of patients (%)
Denver	112	5.6 (4.3, 6.9)	61 (45.4%)
SOFA	287	4.7 (4.0, 5.4)	229 (79.8%)
Marshall	199	5.5 (4.5, 6.5)	122 (61.3%)

Table 7.3: Timing of onset of MOF

MOF defined by the SOFA score had the longest mean duration (6.5 days), 2.1 days longer than the mean duration of Denver MOF (Table 7.4). Survival analyses showed a significant proportion of patients dying without developing MOF, but the precise cause of death for these patients was difficult to ascertain even from death certification data. Therefore, since the cause of death remained unknown for some patients, no adjustment of duration of MOF was made to account for mortality.

	Number of patients with MOF	**Mean duration of MOF** Days (95%CI)	**Median** Days	**Minimum duration** Days	**Maximum duration** Days
Denver	112	4.4 (3.6, 5.3)	3	1	26
SOFA	287	6.5 (5.8, 7.3)	4	1	39
Marshall	199	5.3 (4.4, 6.1)	3	1	30

Table 7.4: Duration of MOF (unadjusted for death)

Analysis of the duration of MOF showed different patterns for each scoring system. (Figure 7.13 (a-c)). The Denver and SOFA scores both showed gradual declines in patient numbers with increasing duration of MOF. The Marshall score showed a sharp decrease in the number of patients with MOF between one and two days duration.

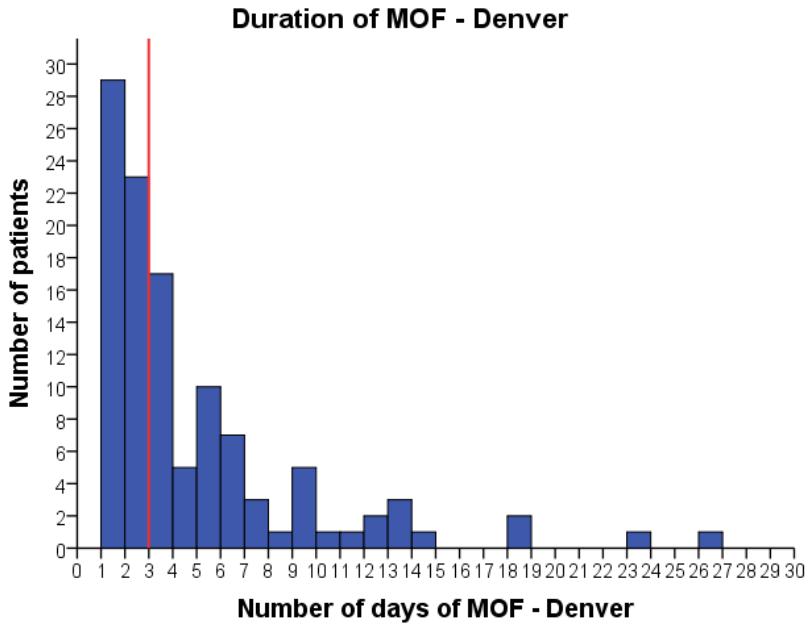


Figure 7.13 (a) Denver score – MOF duration

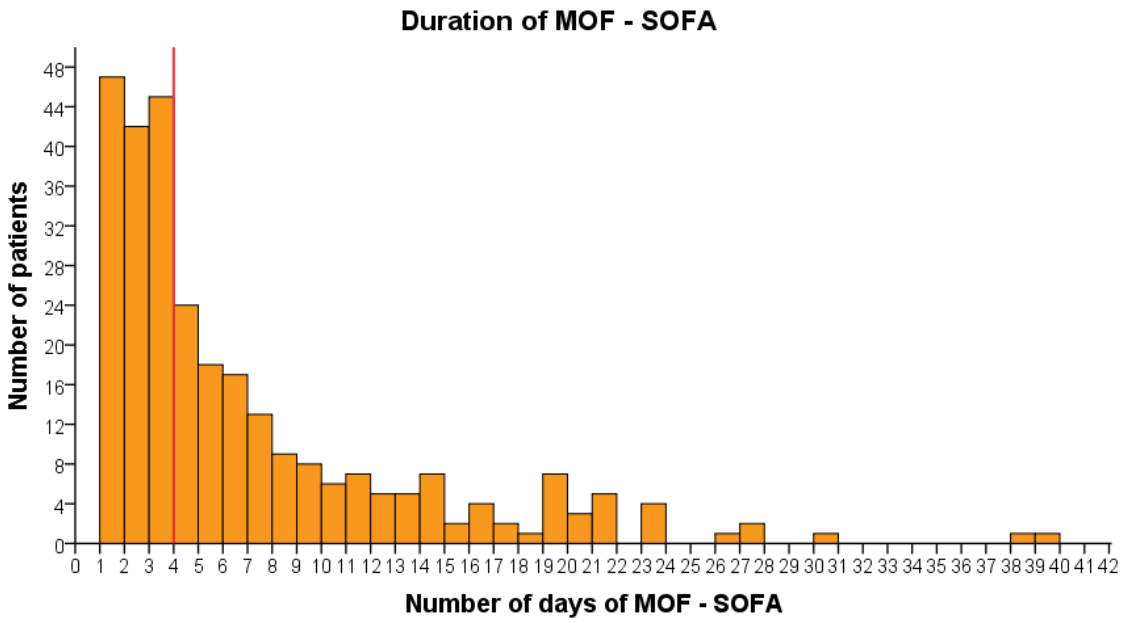


Figure 7.13 (b) SOFA score – MOF duration

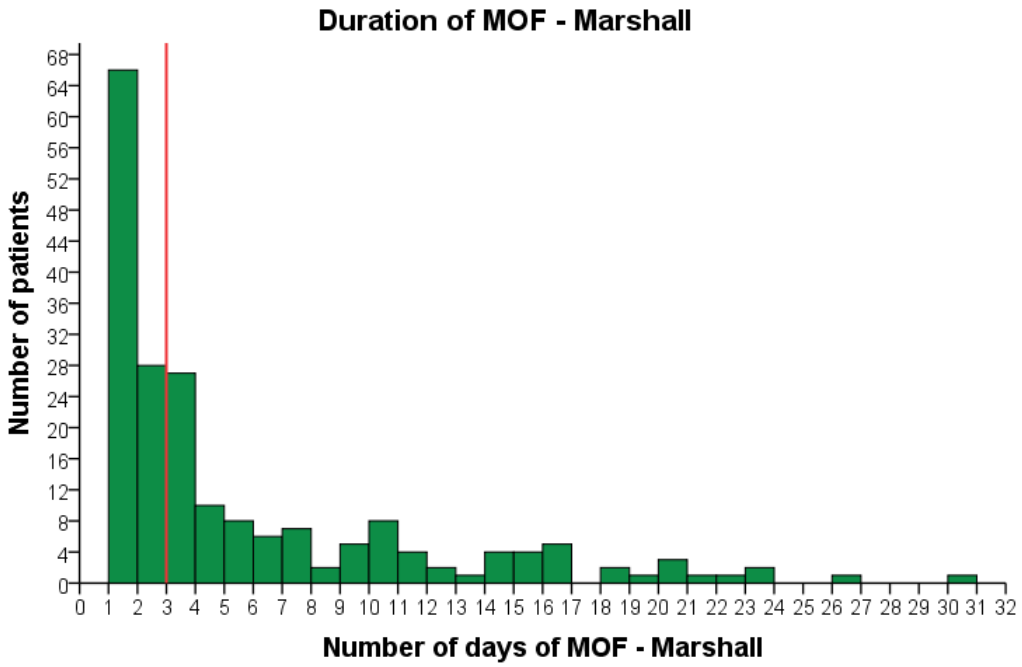


Figure 7.13 (c) Marshall score – MOF duration

Figure 7.13 (a-c): Duration of MOF for patients defined with MOF for each scoring system.

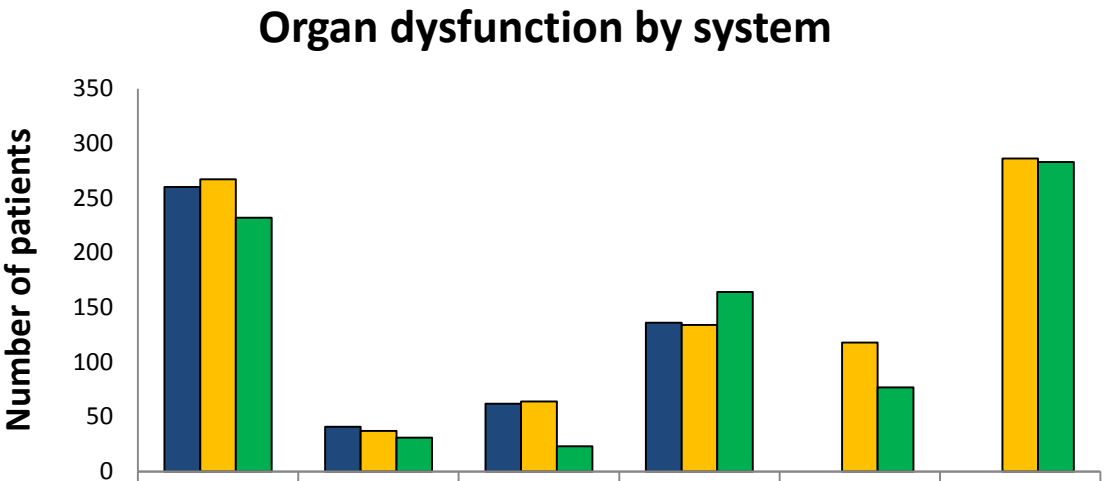
(Vertical line represents median duration. Only patients positive for MOF definition for each score included, hence minimum value = 1 day.)

7.2.5. ORGAN SYSTEM DYSFUNCTION

Individual organ dysfunction was defined as a value ≥ 1 for a single organ system on the Denver score, and ≥ 2 for the SOFA and Marshall scores. These values were designed to be indicative of organ dysfunction and not as definitions of single organ failure (SOF). The number of patients with organ dysfunction for each system was recorded (Figure 7.14).

Neurological dysfunction was most prevalent, affecting 286 (58.2%) patients by the SOFA score and 283 (57.6%) by the Marshall score. There is no neurological component to the Denver score, and respiratory dysfunction was most common, affecting 260 (53.0%) patients. Renal dysfunction was the least prevalent dysfunction by the Denver and SOFA

scores, and hepatic dysfunction the least prevalent by the Marshall score. Similar numbers were seen for organ dysfunction across the three scoring systems, with notable differences in the Marshall score for hepatic and cardiac dysfunction.



	Respiratory	Renal	Hepatic	Cardiac	Coagulation	Neurological
Denver	260	41	62	136		
SOFA	267	37	64	134	118	286
Marshall	232	31	23	164	77	283

Figure 7.14: Organ dysfunction by individual systems for Denver, SOFA and Marshall scores

Patterns in the number and type of systems with organ dysfunction were assessed (Table 7.5). The total number of patients with some form of organ dysfunction ranged from 267 (54.4%) with the Marshall score, to 296 (60.3%) with the Denver score and 308 (62.7%) with the SOFA score.

For the Denver score, a decrease in the number of patients was noted as the number of dysfunctional organ systems increased. For the SOFA score, the numbers peaked at two organ dysfunctions followed by a steady decline, and for the Marshall score, a similar peak was seen with three organ dysfunctions.

The pattern of organ dysfunctions differed depending on the scoring system used (Figure 7.14). For the Denver score, respiratory dysfunction and respiratory/cardiac dysfunction were most commonly seen. Patterns for the SOFA and Marshall scores were similar, with respiratory/neurological, respiratory/cardiac/neurological, and respiratory/cardiac/ coagulation/neurological being the most common combinations of organ dysfunction (Table 7.5)

ORGAN SYSTEM DYSFUNCTION	Denver	SOFA	Marshall
Respiratory only	128	26	21
Renal only	0	0	0
Hepatic only	3	0	0
Cardiac only	26	5	2
Coagulation only		5	2
Neurological only		20	43
TOTAL SINGLE ORGAN SYSTEM	**157**	**31**	**23**
Respiratory/Renal	4	0	0
Respiratory/Hepatic	21	0	0
Respiratory/Cardiac	62	1	6
Respiratory/Coagulation		7	5
Respiratory/Neurological		77	64
Renal/Hepatic	1	1	0
Renal/Cardiac	1	0	0
Renal/Coagulation		0	0
Renal/Neurological		1	3
Hepatic/Cardiac	3	0	0
Hepatic/Coagulation		0	0
Hepatic/Neurological		0	0
Cardiac/Coagulation		0	1
Cardiac/Neurological		15	16
Coagulation/Neurological		5	4
TOTAL TWO ORGAN SYSTEMS	**92**	**107**	**99**
Respiratory/Renal/Hepatic	3	0	0
Respiratory/Renal/Cardiac	13	1	0
Respiratory/Renal/Coagulation		0	0
Respiratory/Renal/Neurological		1	0
Respiratory/Hepatic/Cardiac	13	0	1
Respiratory/Hepatic/Coagulation		0	0
Respiratory/Hepatic/Neurological		9	2
Respiratory/Cardiac/Coagulation		0	0
Respiratory/Cardiac/Neurological		42	70
Respiratory/Coagulation/Neurological		27	13
Renal/Hepatic/Cardiac	2	0	0
Renal/Hepatic/Coagulation		0	0
Renal/Hepatic/Neurological		0	0
Renal/Cardiac/Coagulation		0	0
Renal/Cardiac/Neurological		1	1
Renal/Coagulation/Neurological		0	0
Hepatic/Cardiac/Coagulation		0	0
Hepatic/Cardiac/Neurological		0	1

Hepatic/Coagulation/Neurological		3	0
Cardiac/Coagulation/Neurological		5	12
TOTAL THREE ORGAN SYSTEMS	**31**	**89**	**100**
Respiratory/Renal/Hepatic/Cardiac	16	0	0
Respiratory/Renal/Hepatic/Coagulation		0	0
Respiratory/Renal/Hepatic/Neurological		0	0
Respiratory/Renal/Cardiac/Coagulation		0	0
Respiratory/Renal/Cardiac/Neurological		4	0
Respiratory/Renal/Coagulation/Neurological		1	0
Respiratory/Hepatic/Cardiac/Coagulation		1	0
Respiratory/Hepatic/Cardiac/Neurological		7	3
Respiratory/Hepatic/Coagulation/Neurological		13	0
Respiratory/Cardiac/Coagulation/Neurological		19	17
Renal/Hepatic/Cardiac/Coagulation		0	0
Renal/Hepatic/Cardiac/Neurological		0	1
Renal/Hepatic/Coagulation/Neurological		0	0
Renal/Cardiac/Coagulation/Neurological		0	1
Hepatic/Cardiac/Coagulation/Neurological		3	2
TOTAL FOUR ORGAN SYSTEMS	**16**	**48**	**24**
Respiratory/Renal/Hepatic/Cardiac/Coagulation		0	0
Respiratory/Renal/Hepatic/Cardiac/Neurological		4	1
Respiratory/Renal/Hepatic/Coagulation/Neurological		3	0
Respiratory/Renal/Cardiac/Coagulation/Neurological		6	8
Respiratory/Hepatic/Cardiac/Coagulation/Neurological		6	5
Renal/Hepatic/Cardiac/Coagulation/Neurological		2	0
TOTAL FIVE ORGAN SYSTEMS		**21**	**14**
Respiratory/Renal/Hepatic/Cardiac/Coagulation/Neurological		12	7
TOTAL SIX ORGAN SYSTEMS		**12**	**7**

Table 7.5: Breakdown of organ dysfunction by combinations of organ systems affected for Denver, SOFA and Marshall scores.

(Grey areas in the Denver column indicate combinations that are not possible due to the lack of neurological and coagulation components in the Denver score.)

The duration of organ dysfunction was compared for each system (Tables 7.6 (a-c)). Respiratory dysfunction had the longest duration by both Denver and SOFA scores.

Neurological dysfunction had the longest duration by the Marshall score. Cardiac dysfunction and coagulopathy had the shortest mean durations, less than the mean duration of MOF.

Denver	**Number of patients**	**Mean duration** Days (95% CI)	**Median duration** Days	**Range** Days
MOF	112	4.4 (3.6, 5.3)	3	1 – 26
Respiratory	260	7.6 (6.6, 8.6)	4	1 – 47
Renal	41	7.3 (5.1, 9.5)	5	1 – 26
Hepatic	62	5.0 (3.8, 6.3)	3	1 – 26
Cardiac	136	3.2 (2.8, 3.7)	3	1 - 16

Table 7.6 (a): Duration of organ dysfunction by Denver score

SOFA	**Number of patients**	**Mean duration** Days (95% CI)	**Median duration** Days	**Range** Days
MOF	287	6.5 (5.8, 7.3)	4	1 – 39
Respiratory	267	7.8 (6.8, 8.7)	5	1 – 47
Renal	37	7.5 (5.2, 9.8)	5	1 – 26
Hepatic	64	5.0 (3.8, 6.2)	3	1 – 26
Cardiac	134	3.2 (2.8, 3.6)	3	1 – 14
Coagulation	118	2.4 (2.0, 2.7)	2	1 – 12
Neurological	286	7.3 (6.5, 8.1)	5	1 – 35

Table 7.6 (b): Duration of organ dysfunction by SOFA score

Marshall	**Number of patients**	**Mean duration** Days (95% CI)	**Median duration** Days	**Range** Days
MOF	199	5.3 (4.4, 6.1)	3	1 – 30
Respiratory	232	6.1 (5.2, 7.1)	3	1 – 41
Renal	31	7.2 (4.7, 9.7)	4	1 – 24
Hepatic	23	5.0 (2.7, 7.3)	4	1 – 24
Cardiac	164	3.3 (2.8, 3.8)	2	1 – 18
Coagulation	77	1.9 (1.6, 2.3)	1	1 – 9
Neurological	283	9.2 (8.0, 10.4)	5	1 – 75

Table 7.6 (c): Duration of organ dysfunction by Marshall score

Table 7.6 (a-c): Duration of organ dysfunction by Denver, SOFA and Marshall scores (unadjusted for death)

The day of onset of organ dysfunction varied between systems, and between scores. (Tables 7.7 (a-c)) For the Denver score, the mean onset of respiratory and cardiac dysfunction preceded that of MOF. For the SOFA and Marshall scores, the mean onset of coagulation dysfunction preceded that of MOF. Respiratory and neurological dysfunction also occurred early in all scores, with hepatic and renal dysfunction developing late. A sharp increase in the proportion of patients with cardiac dysfunction was seen between day 3 and day 10 post-injury for the SOFA and Marshall scores.

Denver	**Number of patients**	**Mean day of onset** Post-injury day (95% CI)	**Onset by day 3** No (%)	**Onset by day 7** No (%)	**Onset by day 10** No (%)
MOF	112	5.6 (4.3, 6.9)	61 (54.5%)	90 (80.4%)	103 (92.0%)
Respiratory	260	4.8 (4.0, 5.5)	201 (77.3%)	232 (89.2%)	243 (93.5%)
Renal	41	6.0 (4.9, 7.0)	15 (36.5%)	27 (65.9%)	38 (92.7%)
Hepatic	62	6.3 (5.6, 7.0)	13 (21.0%)	42 (67.7%)	57 (91.9%)
Cardiac	136	5.2 (4.0, 6.4)	94 (69.1%)	114 (83.8%)	125 (91.9%)

Table 7.7 (a): Onset of organ dysfunction by Denver score

SOFA	**Number of patients**	**Mean day of onset** Post-injury day (95% CI)	**Onset by day 3** No (%)	**Onset by day 7** No (%)	**Onset by day 10** No (%)
MOF	287	4.7 (4.0, 5.4)	229 (79.8%)	254 (88.5%)	267 (93.0%)
Respiratory	267	4.8 (4.0, 5.5)	211 (79.0%)	237 (88.8%)	248 (92.9%)
Renal	37	6.0 (4.9, 7.1)	15 (30.7%)	23 (62.1%)	33 (89.2%)
Hepatic	64	6.2 (5.5, 6.9)	13 (20.3%)	44 (68.9%)	59 (92.2%)
Cardiac	134	5.2 (4.0, 6.4)	94 (30.1%)	112 (83.6%)	123 (91.8%)
Coagulation	118	4.3 (3.3, 5.2)	102 (86.4%)	110 (93.2%)	112 (94.9%)
Neurological	286	4.8 (4.1, 5.5)	221 (77.3%)	251 (87.8%)	263 (92.0%)

Table 7.7 (b): Onset of organ dysfunction by SOFA score

Marshall	Number of patients	Mean day of onset Post-injury day (95% CI)	Onset by day 3 No (%)	Onset by day 7 No (%)	Onset by day 10 No (%)
MOF	199	5.5 (4.5, 6.5)	122 (61.3%)	169 (84.9%)	181 (91.0%)
Respiratory	232	5.4 (4.5, 6.2)	133 (57.3%)	202 (87.1%)	212 (91.4%)
Renal	31	6.7 (5.3, 8.1)	12 (38.7%)	16 (51.6%)	26 (83.4%)
Hepatic	23	6.9 (5.6, 8.1)	4 (17.4%)	13 (56.5%)	20 (87.0%)
Cardiac	164	5.6 (4.9, 6.3)	64 (39.0%)	130 (79.3%)	152 (92.7%)
Coagulation	77	4.1 (3.5, 4.7)	59 (76.6%)	68 (88.3%)	72 (93.5%)
Neurological	283	4.9 (4.2, 5.6)	213 (75.3%)	248 (87.6%)	261 (92.2%)

Table 7.7 (c): Onset of organ dysfunction by Marshall score

Table 7.7 (a-c): Onset of organ dysfunction by Denver, SOFA and Marshall scores.

(Percentages refer to total number of patients developing organ dysfunction in the specified system. Days refer to the number of days post-injury.)

7.2.6. COMPARISON OF MOF AND NO MOF GROUPS

Before univariate analyses of all potential MOF predictors, brief comparative analyses were performed to assess differences between the MOF and no MOF populations, as defined by each score.

DENVER

Patients with Denver MOF were more severely injured, with a mean ISS of 29.2 compared to 24.2 in the no MOF group, and significantly more likely to require a transfusion in the first 24 hours after injury ($p = 0.005$). Blunt trauma was significantly associated with an increased likelihood of MOF, but numbers in the penetrating trauma group were small. Whilst more patients with MOF were transferred from surrounding hospitals, no differences were observed in the time to the first receiving hospital. Patients with MOF had extended ICU and hospital stays, and were admitted to ICU later than the no MOF group (Table 7.8).

	Denver – MOF n = 112	**Denver – No MOF** n = 379	**p value**
PATIENT CHARACTERISTICS			
Age (years)[1]	40.1 (36.7, 43.9)	37.4 (35.7, 39.2)	0.161
Sex (% male)	88 (78.6%)	283 (74.7%)	0.399
INJURY CHARACTERISTICS			
ISS	29.2 (26.5, 31.8)	24.2 (23.0, 25.4)	0.001
NISS	38.6 (35.4, 41.7)	33.6 (32.1, 35.1)	0.005
Blunt trauma (%)	112 (100%)	365 (96.3%)	0.047
ADMISSION DETAILS			
Direct (%)	76 (67.9%)	302 (79.7%)	0.009
Time to hospital (mins)	125 (69, 181)	116 (93, 138)	0.718
Day of ICU admission (post-injury day)	2.7 (1.4, 4.1)	1.3 (0.9, 1.8)	0.057
Length of ICU stay (days)	15.8 (13.1, 18.5)	5.2 (4.6, 5.7)	<0.001
Day of hospital discharge (post-injury day)	35.2 (30.0, 40.4)	28.2 (25.2, 31.3)	0.029
INTERVENTIONS			
Operation (%)	93 (83.0%)	291 (76.8%)	0.159
Operation first 24 hours (%)	47 (42.0%)	180 (47.5%)	0.302
Transfusion first 24 hours (%)	45 (40.2%)	100 (26.4%)	0.005

[1] **Geometric mean**

Table 7.8: Population comparison of Denver MOF and no MOF groups.

(Data given as mean (95% CI) or number (% of group))

SOFA

Patients with SOFA MOF were significantly older, with a mean age of 39.8 years, compared to 35.6 years in the no MOF group (p = 0.005). MOF was associated with increased severity of injury by ISS and NISS, and a significantly higher number of operations (p = <0.001). Patients with MOF had a mean ICU stay 9 days longer than patients without MOF, and a delay to hospital discharge of over 14 days. MOF was also

associated with delayed ICU admission, with a mean admission time of 2.4 days post-injury, compared to 0.7 days in the no MOF group (Table 7.9).

	SOFA – MOF n = 287	**SOFA – No MOF** n = 204	**p value**
PATIENT CHARACTERISTICS			
Age (years)[1]	39.8 (37.7, 42.1)	35.6 (33.5, 37.9)	0.005
Sex (% male)	220 (76.7%)	151 (74.0%)	0.503
INJURY CHARACTERISTICS			
ISS	28.0 (26.4, 29.5)	21.6 (20.2, 23.1)	<0.001
NISS	37.8 (36.0, 39.6)	30.4 (28.5, 32.4)	<0.001
Blunt trauma (%)	282 (98.3%)	195 (95.6%)	0.080
ADMISSION DETAILS			
Direct (%)	215 (74.9%)	163 (79.9%)	0.196
Time to hospital (mins)	118 (88,148)	117 (86, 147)	0.963
Day of ICU admission (post-injury day)	2.4 (1.6, 3.1)	0.7 (0.5, 0.9)	<0.001
Length of ICU stay (days)	11.3 (10.0, 12.6)	2.3 (2.0, 2.7)	<0.001
Day of hospital discharge (post-injury day)	35.8 (32.3, 39.4)	21.4 (17.9, 24.9)	<0.001
INTERVENTIONS			
Operation (%)	242 (84.3%)	142 (69.6%)	<0.001
Operation first 24 hours (%)	134 (46.7%)	93 (45.65)	0.809
Transfusion first 24 hours (%)	93 (32.4%)	52 (25.5)	0.098

[1] Geometric mean

Table 7.9: Population comparison of SOFA MOF and no MOF groups.

(Data given as mean (95% CI) or number (% of group))

MARSHALL

The mean age of patients with Marshall MOF was 40.1 years, compared to 36.6 years in the no MOF group (p = 0.031). MOF patients were more severely injured, reflected in higher ISS and NISS values. ICU stay was significantly longer in the MOF group, with a mean stay of 13.2 days, compared to 3.7 in the no MOF group. The mean hospital

discharge for patients with MOF was 38.4 days after injury, compared to 24.0 days in the no MOF group. MOF was associated with an overall increase in operations (p = 0.003), but there was no significant difference in the number of operations performed within 24 hours of injury (Table 7.10).

	Marshall – MOF n = 199	**Marshall – No MOF** n = 292	**p value**
PATIENT CHARACTERISTICS			
Age (years)[1]	40.1 (37.6, 42.9)	36.6 (34.7, 38.6)	0.031
Sex (% male)	157 (78.9%)	214 (73.3%)	0.156
INJURY CHARACTERISTICS			
ISS	28.7 (26.8, 30.7)	23.0 (21.7, 24.3)	<0.001
NISS	38.4 (36.1, 40.6)	32.3 (30.6, 33.9)	<0.001
Blunt trauma (%)	196 (98.5%)	281 (96.2%)	0.140
ADMISSION DETAILS			
Direct (%)	146 (73.4%)	232 (79.5%)	0.116
Time to hospital (mins)	129 (89, 169)	110 (86, 134)	0.405
Day of ICU admission (post-injury day)	2.6 (1.5, 3.6)	1.1 (0.8, 1.4)	0.006
Length of ICU stay (days)	13.2 (11.5, 14.9)	3.7 (3.3, 4.2)	<0.001
Day of hospital discharge (post-injury day)	38.4 (34.0, 42.8)	24.0 (20.9, 27.0)	<0.001
INTERVENTIONS			
Operation (%)	169 (84.9%)	215 (73.6%)	0.003
Operation first 24 hours (%)	93 (46.7%)	134 (45.9%)	0.854
Transfusion first 24 hours (%)	68 (34.2%)	77 (26.4%)	0.063

[1] **Geometric mean**

Table 7.10: Population comparison of Marshall MOF and no MOF groups.

(Data given as mean (95% CI) or number (% of group))

The significant relationship between later ICU admission and MOF was found to hold for the SOFA and Marshall scores, even after excluding an extreme outlier admitted on day

69, and adjusting for the impact of transfers on delayed admission (SOFA: p = 0.001; Marshall: p = 0.010).

7.3. Univariate Analysis of MOF Predictors

7.3.1. Candidate Input Variables

A total of 84 candidate input variables were described in the data dictionary (Appendix 8). Four of these were excluded because they either had no available data (FDPs, ethanol levels, the age of transfused blood), or were excluded from data collection for duplicating data in other fields (ventilation).

From the 80 remaining variables, two contained minimal data (cardiac index and fibrinogen levels) and three contained information duplicated in other fields (time to first operation, time to intubation and time to inotropes/vasopressors). Data were collected for minimum and maximum values of glucose and WCC, but examination of the data distributions showed the minimum values reflected values within the normal range, and these were excluded. 73 variables remained for univariate analysis. Tables 7.11 and 7.12 show the categorisation of these variables and their availability at each time point.

Analysis of the injury data led to further collapsing the AIS categories for each region into binary fields of injury (AIS ≥ 1) or no injury (AIS = 0). These are described in the univariate analysis as head injury, chest injury etc.

Category	Sub-category	Variable	Timing	0-2 hours	0-4 hours	0-6 hours	0-12 hours	0-24 hours	0-48 hours
Patient	Basic patient data	Age	Fixed	Continuous					
		BMI	Fixed						
		Sex	Fixed	Categorical - binary					
	Chronic health conditions	Cardiovascular disease	Fixed	Categorical - binary					
		Diabetes	Fixed						
		Drugs/alcohol	Fixed						
		Hepatic disease	Fixed						
		Immunocompromise	Fixed						
		Malignancy	Fixed						
		Other PMH	Fixed						
		PMH	Fixed						
		Psychiatric	Fixed						
		Renal disease	Fixed						
		Respiratory disease	Fixed						
Injury	Injury severity	Head Injury	Fixed	Categorical - 2					
		Face Injury	Fixed						
		Chest Injury	Fixed						
		Abdominal Injury	Fixed						
		Limb/pelvic Injury	Fixed						
		External Injury	Fixed						
		ISS	Fixed	Continuous					
		NISS	Fixed						
		GCS - 1st	< 2 hours	Categorical - 5					
		GCS minimum	Time blocks	Categorical - 5	Categorical - 5	Categorical - 5	Categorical - 5	Categorical - 5	Categorical - 5
	Injury details	Blunt/penetrating	Fixed	Categorical - binary					
		Mechanism of injury	Fixed	Categorical - 9					
		Protective equipment	Fixed	Categorical - binary					
		Time to hospital	Variable	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous
		Trapped	Fixed	Categorical - binary					
		Vehicle position	Fixed	Categorical - 7					
Physiology	N/A	Blood pressure - minimum SBP	Time blocks	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous
		Blood pressure - minimum MAP	Time blocks	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous
		Cardiac Index - minimum	Time blocks						
		FIO2 - maximum	Time blocks	Categorical - 4	Categorical - 4	Categorical - 4	Categorical - 4	Categorical - 4	Categorical - 4
		Heart Rate - minimum	Time blocks	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous
		PCO2 - maximum	Time blocks	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous
		PO2 - minimum	Time blocks	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous
		Respiratory Rate - minimum	Time blocks	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous
		Respiratory Rate - maximum	Time blocks	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous
		Temperature - minimum	Time blocks	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous
		Urine output- total	Time blocks	Categorical - binary	Categorical - binary	Categorical - binary	Continuous	Continuous	Continuous

Table 7.11: Categorisation of variables for univariate analysis – Patient, Injury and Physiological variables

Category	Sub-category	Variable	Timing	0-2 hours	0-4 hours	0-6 hours	0-12 hours	0-24 hours	0-48 hours
Laboratory	Haematology	APTT - maximum	Time blocks	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous
		Fibrinogen - minimum	Time blocks						
		Hct - minimum	Time blocks	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous
		Hb - minimum	Time blocks	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous
		Platelet count - minimum	Time blocks	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous
		PT - maximum	Time blocks	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous
		WCC - minimum	Time blocks						
		WCC - maximum	Time blocks	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous
	Biochemistry	Albumin - minimum	Time blocks	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous
		BD - greatest deviation from 0	Time blocks	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous
		Bilirubin - maximum	Time blocks	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous
		CK - maximum	Time blocks	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous
		Creatinine - maximum	Time blocks	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous
		Glucose - minimum	Time blocks						
		Glucose - maximum	Time blocks	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous
		Lactate - maximum	Time blocks	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous
		pH - minimum	Time blocks	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous
		Potassium - minimum	Time blocks	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous
		Potassium - maximum	Time blocks	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous
		Sodium - minimum	Time blocks	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous
		Sodium - maximum	Time blocks	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous
		Urea - maximum	Time blocks	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous
Management	Transfusion	Any transfusion	24 hours					Categorical - binary	
		Cryoprecipitate	24 hours						
		FFPs	24 hours						
		Platelets	24 hours						
		PRBCs	24 hours						
		Total blood products	24 hours					Categorical - 3	
	Emergency Surgery	Chest drain insertion	Time blocks	Categorical - binary	Categorical - binary	Categorical - binary	Categorical - binary	Categorical - binary	Categorical - binary
		ICP bolt insertion	Time blocks		Categorical - binary	Categorical - binary	Categorical - binary	Categorical - binary	Categorical - binary
		Intra-operative time	48 hours						Continuous
		Operative intervention	Time blocks	Categorical - binary	Categorical - binary	Categorical - binary	Categorical - binary	Categorical - binary	Categorical - binary
		Time to first operation	Variable						
	Emergency	CPR	Time blocks	Categorical - binary	Categorical - binary	Categorical - binary	Categorical - binary	Categorical - binary	Categorical - binary
	Intervention	Intubation	Time blocks	Categorical - binary	Categorical - binary	Categorical - binary	Categorical - binary	Categorical - binary	Categorical - binary
		IV fluids	Time blocks	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous
		Time to intubation	Variable						
		Time to vasopressors/inotropes	Variable						
		Vasopressors/inotropes	Time blocks	Categorical - binary	Categorical - binary	Categorical - binary	Categorical - binary	Categorical - binary	Categorical - binary
Total number of variables				66	67	67	67	72	73

Table 7.12:Categorisation of variables for univariate analysis – Laboratory and Management variables

7.3.2. RESULTS OF UNIVARIATE ANALYSIS

The results of univariate analysis of all variables are given in Appendices 11 and 12.

Appendix 11 contains the analyses for all variables against the outcome measures of Denver, SOFA and Marshall MOF. Unadjusted Odds Ratios (OR) with 95 % confidence intervals (CI) are given for each variable.

Appendix 12 contains the analyses for each of Denver, SOFA and Marshall MOF separately, split by time period. Since the variables with fixed timing could be used for all models, these were considered in a separate group. Time-dependent variables were then split into six groups according to the model timings. Unadjusted OR with 95% CI are given for each variable.

7.4. VARIABLES FOR MULTIVARIATE ANALYSIS

The total number of variables used for univariate analysis for each model period ranged from 66 (0 – 2 hours) to 73 (0 – 48 hours). Due to the sample size, these needed to be reduced prior to multivariate analysis, according to the criteria outlined in Chapter 2 (Section 2.9.3). Variables meeting the criteria on missing data or significance on univariate analysis are shaded in the tables in Appendix 12

The variables selected for multivariate analysis for each model are listed in Appendix 13. The number of variables per model ranged from 8 variables (Denver 0 – 2 hours) to 32 variables (SOFA 0 – 48 hours). An increase in the number of variables meeting the selection criteria for inclusion was seen with increasing time from injury. Significant efforts were made to reduce variable numbers to create viable models without excess

loss of clinical information. Variable numbers for some models remained higher than ideal for model building, and this was considered carefully in the analysis of the multivariate results and model refinement.

Multivariate analysis for the SOFA 0 – 12 model showed that the initial 29 variables were present in only 82 patients, and resulted in a poorly fitting model. The total number of variables was further reduced to 26 by examining patterns of missing data, and significance on univariate analyses.

7.5. MULTIVARIATE ANALYSIS

7.5.1. MODEL DEVELOPMENT

The full results of model development are given in Appendix 14, with six models developed for each MOF score for the time periods 0 – 2, 0 – 4, 0 – 6, 0 – 12, 0 – 24 and 0 – 48 hours. For each of the 18 models, multiple backward eliminations were performed. The results presented in this chapter represent the most accurate models. For each model, the variable lists for each stage of development, the goodness-of-fit measures, category predictors and ROC curves were recorded. For the final selected model, the coefficients, significance and unadjusted OR were detailed.

The summary measures of the number of models assessed in model development, the model chosen and the goodness-of-fit and discrimination measures for that model are given for each MOF outcome measure.

No significant interaction terms were found. The linearity assumption was shown to be valid for all variables, except for intra-operative time, which was present at 48 hours only, and was therefore excluded from further multivariate analysis.

7.5.2. DENVER MOF MULTIVARIATE MODELS

The number of models created using the backwards elimination approach to model development ranged from 4 to 21. Table 7.13 shows the selected model number (as detailed in Appendix 14), the number of patients on whom the selected model was developed and the number of variables it contained.

Denver models	Total number models during development	Selected model number	Number of patients	Number of variables
0-2 hours	4	4	257	5
0-4 hours	5	2	280	8
0-6 hours	8	3	247	9
0-12 hours	18	8	168	14
0-24 hours	18	8	231	15
0-48 hours	21	18	371	9

Table 7.13: Model selection for Denver MOF models

Tables 7.14 to 7.16 show the goodness-of-fit measures, the category predictors and the AUC values for the ROC curve for each model. Tables 7.17-7.22 give the variable coefficients, significance and unadjusted OR for each model.

Denver models	Model χ^2	Model χ^2 p value	-2 log likelihood	Nagelkerke R^2	Hosmer-Lemeshow χ^2	Hosmer-Lemeshow χ^2 p value
0-2 hours	43.237	<0.001	43.237	0.237	9.036	0.339
0-4 hours	42.534	<0.001	258.671	0.213	2.578	0.958
0-6 hours	48.045	<0.001	218.816	0.268	5.292	0.726
0-12 hours	43.816	<0.001	119.787	0.369	13.447	0.097
0-24 hours	80.922	<0.001	160.417	0.456	11.966	0.152
0-48 hours	121.451	<0.001	262.774	0.433	0.011	0.278

Table 7.14: Goodness-of-fit summary measures for selected Denver MOF models

Denver models	PAC (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
0-2 hours	80.5	28.1	95.5	64.0	82.3
0-4 hours	78.6	26.6	94.0	56.7	81.2
0-6 hours	81.0	29.8	96.4	70.8	82.1
0-12 hours	88.7	56.3	96.3	78.3	90.3
0-24 hours	85.3	54.0	93.9	71.1	88.1
0-48 hours	84.1	44.3	94.9	70.0	86.8

Table 7.15: Category predictors for Denver MOF models

Denver models	AUC	95% CI
0-2 hours	0.765	0.693, 0.837
0-4 hours	0.758	0.693, 0.822
0-6 hours	0.779	0.710, 0.848
0-12 hours	0.839	0.756, 0.922
0-24 hours	0.871	0.819, 0.923
0-48 hours	0.868	0.829, 0.909

Table 7.16: AUC and 95% CI for ROC curves for Denver MOF models

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-6.108		
CV disease	1.982	0.001	7.26 (2.20, 23.93)
Head injury	1.546	<0.001	4.69 (21.4, 10.27)
Abdominal injury	1.016	0.005	2.76 (1.36, 5.60)
HR Max	0.018	0.004	1.02 (1.01, 1.02)
RR Min	0.066	0.004	1.07 (1.02, 1.12)

Table 7.17: Denver 0 – 2 hours – model output

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-5.181		
Age	0.011	0.216	5.64 (1.78, 17.81)
CV disease	0.1729	0.003	4.47 (2.16, 9.25)
Head injury	1.498	<0.001	4.47 (2.16, 9.25)
Chest injury	0.424	0.267	1.53 (0.72, 3.23)
Abdominal injury	0.965	0.006	2.63 (1.33, 5.20)
Limb/pelvis injury	-0.393	0.307	0.68 (0.32, 1.44)
HR Max	0.016	0.014	1.02 (1.00, 1.03)
RR Min	0.004	0.851	1.00 (0.96, 1.05)

Table 7.18:Denver 0 – 4 hours – model output

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-4.268		
Age	0.004	0.696	1.00 (0.99, 1.02)
CV disease	1.492	0.022	4.45 (1.24, 15.95)
Head injury	1.361	0.001	3.90 (1.76, 8.65)
Chest injury	0.819	0.070	2.27 (0.93, 5.50)
Abdominal injury	0.711	0.064	2.04 (0.96, 4.32)
Limb/pelvis injury	0.373	0.368	1.45 (0.65, 32.7)
MAP	0.001	0.888	1.00 (0.99, 1.02)
RR min	0.028	0.279	1.03 (0.98, 1.08)
Vasopressors	1.873	0.001	6.51 (2.25, 18.80)

Table 7.19: Denver 0 – 6 hours – model output

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-1.45		
Age	-0.006	0.722	0.99 (0.96, 1.03)
CV disease	1.194	0.244	3.30 (0.44, 24.58)
Head injury	1.108	0.085	3.03 (0.86, 10.70)
Chest injury	0.958	0.166	2.61 (0.67, 10.13)
Abdominal injury	0.336	0.549	1.40 (0.48, 4.19)
Limb/pelvis injury	-0.038	0.952	0.96 (0.28, 3.32)
MAP	-0.013	0.306	0.99 (0.96, 1.01)
RR min	0.045	0.312	1.05 (0.96, 1.14)
Platelet count	-0.006	0.160	1.00 (0.99, 1.00)
PT	-0.187	0.080	0.83 (0.67, 1.02)
BD	-0.153	0.087	0.86 (0.72, 1.02)
Creatinine	0.018	0.102	1.02 (1.00, 1.04)
Lactate	-0.233	0.088	0.79 (0.61, 1.04)
Vasopressors	1.941	<0.001	6.96 (2.35, 20.61)

Table 7.20: Denver 0 – 12 hours – model output

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-4.659		
CV disease	1.817	0.047	6.15 (1.02, 37.00)
Head injury	1.653	0.001	5.23 (2.03, 13.47)
Abdominal injury	0.509	0.169	1.67 (0.81, 3.44)
Limb/pelvis injury	0.688	0.112	1.99 (0.85, 4.65)
HR	0.014	0.089	1.01 (1.00, 1.10)
RR max	0.053	0.008	1.05 (1.01, 1.10)
BD	-0.119	0.008	0.89 (0.81, 0.97)
Potassium max	-0.666	0.044	0.51 (0.27, 0.98)
IV fluids	0.001	0.090	1.00 (1.00, 1.00)
Vasopressors	1.566	<0.001	4.79 (2.28, 10.08)

Table 7.21: Denver 0 – 24 hours – model output

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-9.023		
Head injury	1.497	<0.001	4.47 (1.98, 10.09)
Limb/pelvic injury	0.525	0.158	1.69 (0.82, 3.50)
RR max	0.067	<0.001	1.07 (1.03, 1.12)
APTT	-0.012	0.053	0.99 (0.98, 1.00)
PT	0.114	0.039	1.12 (1.01, 1.25)
BD	-0.126	0.001	0.88 (0.82, 0.95)
Creatinine	0.011	0.027	1.01 (1.00, 1.02)
IV fluids	0.001	0.005	1.00 (1.00, 1.00)
Vasopressors	1.319	<0.001	3.74 (1.90, 7.36)

Table 7.22: Denver 0 – 48 hours

7.5.3. SOFA MOF MULTIVARIATE MODELS

The number of models created using the backwards elimination approach to model development ranged from 8 to 29. Tables 7.23 to 7.32 are presented in the same manner as those for the Denver MOF models.

SOFA models	Total number models during development	Selected model number	Number of patients	Number of variables
0-2 hours	8	6	251	8
0-4 hours	8	4	361	9
0-6 hours	10	6	294	10
0-12 hours	22	19	168	8
0-24 hours	25	18	218	13
0-48 hours	29	23	294	10

Table 7.23: Model selection for SOFA MOF models

SOFA models	Model χ^2	Model χ^2 p value	-2 log likelihood	Nagelkerke R^2	Hosmer-Lemeshow χ^2	Hosmer-Lemeshow χ^2 p value
0-2 hours	51.895	<0.001	292.226	0.250	7.079	0.528
0-4 hours	60.562	<0.001	435.668	0.207	15.454	0.051
0-6 hours	53.355	<0.001	348.796	0.226	12.585	0.127
0-12 hours	56.935	<0.001	166.347	0.399	7.262	0.509
0-24 hours	104.075	<0.001	189.195	0.513	8.508	0.386
0-48 hours	158.013	<0.001	233.686	0.565	8.924	0.349

Table 7.24: Goodness-of-fit summary measures for selected SOFA MOF models

SOFA models	PAC (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
0-2 hours	71.3	76.6	64.5	73.5	68.3
0-4 hours	69.0	75.5	60.9	70.6	66.7
0-6 hours	69.4	77.0	59.7	70.9	67.0
0-12 hours	76.8	84.6	64.1	79.3	71.9
0-24 hours	78.0	81.7	72.4	81.7	72.4
0-48 hours	81.0	86.2	73.5	83.9	76.9

Table 7.25: Category predictors for SOFA MOF models

SOFA models	AUC	95% CI
0-2 hours	0.760	0.700, 0.820
0-4 hours	0.730	0.678, 0.781
0-6 hours	0.741	0.684, 0.797
0-12 hours	0.823	0.759, 0.888
0-24 hours	0.869	0.822, 0.916
0-48 hours	0.888	0.850, 0.928

Table 7.26: AUC and 95% CI for ROC curves for SOFA MOF models

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-4.077		
Age	0.030	<0.001	1.03 (1.01, 1.05)
Respiratory disease	1.110	0.035	3.03 (1.08, 8.53)
Head injury	1.019	0.001	2.77 (1.52, 5.04)
External injury	-0.497	0.115	0.61 (0.33, 1.13)
Trapped	0.705	0.025	2.02 (1.09, 3.75)
SBP	-0.009	0.040	0.99 (0.98, 1.00)
HR	0.020	<0.001	1.02 (1.01, 1.03)
RR min	0.062	0.004	1.06 (1.02, 1.11)

Table 7.27: SOFA 0 – 2 hours – model output

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-3.034		
Age	0.030	<0.001	1.03 (1.02, 1.04)
Psychiatric	-0.713	0.086	0.49 (0.22, 1.11)
Respiratory disease	0.636	0.135	1.89 (0.82, 4.40)
Head injury	0.943	<0.001	2.57 (1.54, 4.27)
Abdominal injury	0.626	0.018	1.87 (1.11, 3.14)
Limb/pelvis injury	0.548	0.046	1.73 (1.01, 2.96)
Trapped	0.356	0.172	1.43 (0.86, 2.38)
SBP	-0.006	0.139	1.00 (0.99, 1.00)
HR	0.014	0.006	1.01 (1.00, 1.02)

Table 7.28: SOFA 0 – 4 hours – model output

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-3.127	<0.001	
Age	0.026	0.069	1.03 (1.01, 1.04)
Psychiatric	-0.808	0.044	0.45 (0.19, 1.06)
Respiratory disease	1.040	<0.001	2.83 (1.03, 7.77)
Head injury	1.132	0.083	3.10 (1.73, 5.57)
Abdominal injury	0.502	0.120	1.65 (0.94, 2.92)
Trapped	0.440	0.049	1.55 (0.89, 2.70)
SBP	-0.009	0.002	1.00 (0.98, 1.00)
HR	0.019	0.138	1.02 (1.01, 1.03)
RR max	0.022	0.013	1.02 (1.00, 1.05)
Intubation	-0.766	0.002	0.47 (0.25, 0.85)

Table 7.29: SOFA 0 – 6 hours – model output

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-1.160		
Respiratory disease	1.103	0.145	3.01 (0.68, 13.28)
Head injury	1.777	0.001	5.91 (2.16, 16.17)
Chest injury	1.131	0.007	3.10 (1.37, 7.00)
FiO_2		0.040	
41-60%	-0.166	0.784	0.85 (0.26, 2.78)
61-80%	1.454	0.105	4.28 (0.74, 24.87)
>80%	0.878	0.148	2.41 (0.73, 7.89)
RR max	0.053	0.019	1.06 (1.01, 1.10)
Albumin	-0.086	0.002	0.92 (0.87, 0.97)
BD	0.173	0.027	1.19 (1.02, 1.39)
Lactate	0.544	0.001	1.72 (1.26, 2.36)

Table 7.30: SOFA 0 – 12 hours model output

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-5.314		
Head injury	1.488	0.002	4.43 (1.75, 11.22)
Abdominal injury	-0.637	0.163	0.53 (0.22, 1.30)
Limb/pelvic injury	0.808	0.079	2.24 (0.91, 5.52)
External injury	-0.807	0.052	0.45 (0.20, 1.01)
FiO_2		0.031	
41 - 60%	-0.466	0.463	0.63 (0.18, 2.18)
61 - 80%	-0.178	0.842	0.84 (0.15, 4.80)
>80%	0.84	0.177	2.32 (0.68, 7.85)
HR	0.021	0.051	1.02 (1.00, 1.04)
RR max	0.031	0.128	1.03 (1.00, 1.07)
APTT	0.016	0.088	1.02 (1.00, 1.04)
Platelet count	0.015	<0.001	1.02 (1.01, 1.04)
Albumin	-0.129	<0.001	0.88 (0.83, 0.94)
Creatinine	0.018	0.031	1.02 (1.00, 1.04)
Transfusion level		0.013	
Transfusion	-0.996	0.033	0.37 (0.15, 0.92)
Massive transfusion	1.682	0.074	5.38 (0.85, 34.05)
Vasopressors	1.599	<0.001	4.95 (2.12, 11.54)

Table 7.31: **SOFA 0 – 24 hours model output**

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-24.162		
Head injury	1.035	0.009	2.81 (1.29, 6.13)
Chest injury	0.413	0.263	1.51 (0.73, 3.12)
Limb/pelvic injury	-0.502	0.243	0.61 (0.26, 1.41)
Trapped	0.355	0.368	1.43 (0.66, 3.08)
FiO_2		0.145	
41-60%	-0.694	0.228	0.50 (0.16, 1.55)
61-80%	-0.222	0.758	0.80 (0.20, 3.29)
> 80%	0.219	0.673	1.25 (0.45, 3.44)
HR	0.005	0.566	1.01 (0.99, 1.02)
RR max	0.020	0.307	1.02 (0.98, 1.06)
Albumin	-0.039	0.148	0.96 (0.91, 1.01)
IV fluids	2.616	<0.001	13.68 (5.45, 34.36)
Vasopressors	1.126	0.003	3.08 (1.46, 6.53)

Table 7.32: SOFA 0 – 48 hours model output

7.5.4. MARSHALL MOF MULTIVARIATE MODELS

The number of models created using the backwards elimination approach to model development ranged from 8 to 25. Tables 7.33 to 7.42 are presented in the same manner as previously.

Marshall models	Total number models during development	Selected model number	Number of patients	Number of variables
0-2 hours	8	4	257	10
0-4 hours	9	6	372	9
0-6 hours	8	5	380	10
0-12 hours	22	15	171	14
0-24 hours	22	15	163	16
0-48 hours	25	21	284	10

Table 7.33: Model selection for Marshall MOF models

Marshall models	Model χ^2	Model χ^2 p value	-2 log likelihood	Nagelkerke R^2	Hosmer-Lemeshow χ^2	Hosmer-Lemeshow χ^2 p value
0-2 hours	55.583	<0.001	285.092	0.265	5.548	0.698
0-4 hours	57.800	<0.001	436.885	0.196	7.980	0.435
0-6 hours	74.585	<0.001	433.484	0.242	1.472	0.993
0-12 hours	73.240	<0.001	163.550	0.465	3.864	0.869
0-24 hours	89.756	<0.001	136.204	0.565	6.304	0.613
0-48 hours	124.348	<0.001	267.654	0.474	8.475	0.389

Table 7.34: Goodness-of-fit summary measures for selected Marshall MOF models

Marshall models	PAC (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
0-2 hours	73.5	55.7	84.4	68.4	75.8
0-4 hours	71.2	45.1	87.4	68.8	72.0
0-6 hours	71.3	50.7	84.5	67.6	72.9
0-12 hours	75.4	74.4	76.4	74.4	76.4
0-24 hours	82.2	84.0	80.5	81.0	83.5
0-48 hours	81.3	78.6	83.7	80.5	82.1

Table 7.35: Category predictors for Marshall MOF models

Marshall models	AUC	95% CI
0-2 hours	0.721	0.660, 0.783
0-4 hours	0.707	0.655, 0.760
0-6 hours	0.705	0.652, 0.757
0-12 hours	0.782	0.712, 0.852
0-24 hours	0.886	0.837, 0.936
0-48 hours	0.857	0.813, 0.901

Table 7.36: AUC and 95% CI for ROC curves for Marshall MOF models

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-6.305		
Age	0.019	0.021	1.02 (1.00, 1.04)
Sex	0.333	0.335	1.40 (0.71, 2.74)
CV disease	1.074	0.072	2.93 (0.91, 9.44)
Drugs/alcohol	0.568	0.323	1.76 (0.57, 5.44)
Head injury	1.463	<0.001	4.32 (2.18, 8.55)
Abdominal injury	1.013	0.002	2.76 (1.44, 5.28)
Limb/pelvic injury	0.836	0.019	2.31 (1.15, 4.63)
External injury	-0.474	0.144	0.62 (0.33, 1.18)
HR	0.016	0.006	1.02 (1.01, 1.03)
RR min	0.062	0.003	1.06 (1.02, 1.11)

Table 7.37: Marshall 0 – 2 hours – model output

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-4.378		
Age	0.019	0.003	1.02 (1.01, 1.03)
Sex	0.431	0.112	1.54 (0.91, 2.62)
CV disease	1.091	0.028	2.98 (1.12, 7.89)
Head injury	1.281	<0.001	3.60 (2.10, 6.17)
Chest injury	0.191	0.470	1.21 (0.72, 2.03)
Abdominal injury	0.993	<0.001	2.70 (1.59, 4.57)
Limb/pelvic injury	0.815	0.004	2.26 (1.29, 3.95)
External injury	-0.369	0.139	0.69 (0.43, 1.13)
HR	0.009	0.053	1.01 (1.00, 1.02)

Table 7.38: Marshall 0 – 4 hours – model output

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-4.484		
Age	0.019	0.003	1.02 (1.01, 1.03)
Sex	0.568	0.041	1.77 (1.02, 3.04)
CV disease	1.184	0.016	3.27 (1.25, 8.54)
Head injury	1.192	<0.001	3.29 (1.92, 5.65)
Abdominal injury	0.968	<0.001	2.63 (1.58, 4.40)
Limb/pelvic injury	0.839	0.003	2.31 (1.33, 4.02)
External injury	-0.496	0.052	0.61 (0.37, 1.00)
Time to hospital	0.005	0.083	1.01 (1.00, 1.01)
HR	0.006	0.202	1.01 (1.00, 1.02)
Vasopressors	1.536	0.001	4.65 (1.92, 11.25)

Table 7.39: Marshall 0 – 6 hours – model output

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-4.764		
Age	0.019	0.104	1.02 (1.00, 1.04)
CV disease	-2.139	0.085	0.12 (0.01, 1.35)
Drugs/alcohol	0.673	0.385	1.96 (0.43, 8.95)
Head injury	2.363	<0.001	10.62 (3.44, 32.82)
Chest injury	0.842	9.095	2.32 (0.87, 6.23)
Abdominal injury	1.287	0.009	3.62 (1.38, 9.49)
Limb/pelvic injury	0.566	0.26	1.76 (0.66, 4.72)
SBP	0.007	0.423	1.01 (1.00, 1.03)
FiO_2		0.007	
41-60%	-1.072	0.089	0.34 (0.10, 1.18)
61-80%	0.855	0.318	2.35 (0.44, 12.62)
> 80%	0.393	0.54	1.48 (0.42, 5.20)
RR min	0.035	0.413	1.04 (0.95, 1.13)
APTT	0.016	0.055	1.02 (1.00, 1.03)
Haemoglobin	-0.162	0.074	0.85 (0.71, 1.02)
Lactate	0.289	0.008	1.34 (1.08, 1.66)
Vasopressors	0.635	0.168	1.89 (0.77, 4.66)

Table 7.40: Marshall 0 – 12 hours – model output

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-10.938		
Sex	1.103	0.049	3.01 (1.01, 9.04)
Respiratory disease	-0.816	0.266	0.44 (0.11, 1.86)
Head injury	3.359	<0.001	28.75 (6.89, 119.95)
Chest injury	0.857	0.088	2.36 (0.88, 6,31)
Abdominal injury	0.385	0.487	1.47 (0.50, 4.35)
Limb/pelvic injury	1.016	0.099	2.76 (0.83, 9.23)
External injury	-0.964	0.063	0.38 (0.14, 1.05)
SBP	0.026	0.045	1.03 (1.00, 1.05)
FiO_2		0.022	
41-60%	-1.485	0.060	0.23 (0.05, 1.07)
61-80%	-0.467	0.633	0.63 (0.09, 4.26)
> 80%	0.253	0.743	1.29 (0.28, 5.87)
APTT	0.037	0.120	1.04 (1.00, 1.09)
Platelet count	0.010	0.029	1.01 (1.00, 1.02)
Albumin	-0.960	0.028	0.91 (0.84, 0.990
BD	-0.247	<0.001	0.78 (0.68, 0.90)
Potassium max	0.722	0.077	2.06 (0.92, 4.58)
Transfusion level		0.138	
Transfusion	-0.977	0.128	0.38 (0.11, 1.33)
Massive transfusion	1.239	0.257	3.45 (0.41, 29.44)
IV fluids	<0.001	0.157	1.00 (1.00, 1.00)

Table 7.41:Marshall 0 – 24 hours - model output

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-3.449		
CV disease	-1.584	0.158	0.21 (0.02, 1.85)
Respiratory disease	0.698	0.135	2.01 (0.81, 5.01)
Head injury	1.365	<0.001	3.92 (1.87, 8.20)
Abdominal injury	0.652	0.056	1.92 (0.99, 3.75)
FiO_2		<0.001	
41-60%	-1.577	0.005	0.21 (0.07, 0.62)
61-80%	-0.862	0.229	0.42 (0.10, 1.72)
> 80%	0.150	0.751	1.16 (0.46, 2.95)
APTT	0.008	0.109	1.01 (1.00, 1.02)
BD	-0.064	0.084	0.94 (0.87, 1.01)
Lactate	0.162	0.083	1.18 (0.98, 1.41)
IV fluids	0.001	<0.001	1.00 (1.00, 1.00)
Vasopressors	0.846	0.010	2.33 (1.22, 4.45)

Table 7.42: Marshall 0 – 48 hours – model output

7.5.5. DENVER COMBINED MODEL

Analysis of the variables comprising each model for Denver MOF showed some common patterns. A core group of variables was developed as a test model to assess the accuracy of a small number of variables at multiple time points post-injury.

A test model was created using variables available at all time points. The variable list was refined by assessment of multiple combinations. Both maximum and minimum RR were included in the original models. Comparison showed that maximum RR had the best overall predictive power across all time points post-injury, and was included in the test model. The variable list is given below:

- Age
- Pre-existing CV disease
- Head injury
- Chest injury
- Abdominal injury
- Limb/pelvic injury
- HR
- Maximum RR

The model output for each time period is given in Figure 7.15. Whilst chest injury was not significant in any of the models, its removal resulted in a marked decrease in sensitivity and PPV. The details of the models are given in Tables 7.43 to 7.45 and Figure 7.16.

0-2 hours

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-6.197		
Age	0.012	0.213	1.01 (0.99, 1.03)
CV disease	1.649	0.009	5.20 (1.51, 17.99)
Head injury	1.608	<0.001	4.99 (2.27, 10.97)
Chest injury	0.592	0.162	1.81 (0.79, 4.14)
Abdominal injury	0.926	0.012	2.52 (1.23, 5.20)
Limb/pelvic injury	0.335	0.419	1.40 (0.62, 3.16)
HR	0.015	0.022	1.02 (1.00, 1.03)
RR max	0.025	0.197	1.03 (0.99, 1.07)

0-4 hours

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-5.561		
Age	0.011	0.21	1.01 (1.00, 1.03)
CV disease	1.733	0.003	5.66 (1.79, 17.86)
Head injury	1.498	<0.001	4.47 (2.16, 17.86)
Chest injury	0.422	0.268	1.53 (0.72, 3.22)
Abdominal injury	0.97	0.005	2.64 (1.33, 5.22)
Limb/pelvic injury	0.393	0.307	1.48 (0.70, 3.15)
HR	0.017	0.013	1.02 (1.00, 1.03)
RR max	0.002	0.926	1.00 (0.97, 1.04)

0-6 hours

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-4.926		
Age	0.008	0.331	1.01 (0.99, 1.03)
CV disease	1.633	0.005	5.12 (1.65, 15.86)
Head injury	1.422	<0.001	4.15 (2.10, 8.19)
Chest injury	0.343	0.349	1.41 (0.69, 2.89)
Abdominal injury	0.855	0.009	2.35 (1.24, 4.45)
Limb/pelvic injury	0.273	0.455	1.31 (0.64, 2.69)
HR	0.012	0.069	1.01 (1.00, 1.03)
RR max	0.013	0.456	1.01 (0.98, 1.05)

0-12 hours

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-4.583		
Age	0.006	0.394	1.01 (1.00, 1.02)
CV disease	1.532	0.003	4.63 (1.69, 12.72)
Head injury	1.278	<0.001	3.59 (1.96, 6.59)
Chest injury	0.349	0.267	1.42 (0.77, 2.63)
Abdominal injury	0.651	0.026	1.92 (1.08, 3.40)
Limb/pelvic injury	0.229	0.452	1.26 (0.69, 2.29)
HR	0.007	0.214	1.01 (1.00, 1.02)
RR max	0.027	0.065	1.03 (1.00, 1.06)

0-24 hours

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-5.416		
Age	0.010	0.138	1.01 (1.00, 1.02)
CV disease	1.231	0.014	3.43 (1.28, 9.15)
Head injury	1.207	<0.001	3.34 (1.87, 5.97)
Chest injury	0.170	0.563	1.19 (0.67, 2.11)
Abdominal injury	0.575	0.041	1.78 (1.02, 3.09)
Limb/pelvic injury	0.248	0.402	1.28 (0.72, 2.29)
HR	0.009	0.100	1.01 (1.00, 1.02)
RR max	0.036	0.009	1.04 (1.01, 1.06)

0-48 hours

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-4.850		
Age	0.010	0.136	1.01 (1.00, 1.02)
CV disease	1.229	0.01	3.42 (1.34, 8.76)
Head injury	1.201	<0.001	3.32 (1.90, 5.83)
Chest injury	0.086	0.751	1.09 (0.64, 1.85)
Abdominal injury	0.632	0.021	1.88 (1.10, 3.22)
Limb/pelvic injury	0.280	0.316	1.32 (0.77, 2.29)
HR	0.003	0.616	1.00 (0.99, 1.01)
RR max	0.050	<0.001	1.05 (1.02, 1.08)

Figure 7.15: Denver combined test model output

Denver test set	Number of patients	Model χ^2	Model χ^2 p value	-2 log likelihood	Nagelkerke R^2	Hosmer-Lemeshow χ^2	Hosmer-Lemeshow χ^2 p value
0-2 hours	257	41.763	<0.001	230.227	0.230	5.168	0.739
0-4 hours	280	42.327	<0.001	258.697	0.213	2.643	0.955
0-6 hours	301	36.699	<0.001	289.795	0.173	7.772	0.456
0-12 hours	375	35.405	<0.001	370.856	0.136	9.808	0.279
0-24 hours	410	38.305	<0.001	405.586	0.135	6.214	0.623
0-48 hours	447	44.026	<0.001	445.668	0.141	5.230	0.733

Table 7.43: Goodness-of-fit measures for Denver combined test model

Denver test set	PAC (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
0-2 hours	81.3	29.8	96.0	68.0	82.3
0-4 hours	78.6	26.6	94.0	56.7	81.2
0-6 hours	78.1	22.9	94.8	57.1	80.2
0-12 hours	77.1	10.3	97.2	52.9	78.2
0-24 hours	76.6	9.5	96.8	47.7	78.0
0-48 hours	76.3	8.5	97.4	50.0	77.4

Table 7.44: Category predictors for Denver combined test model

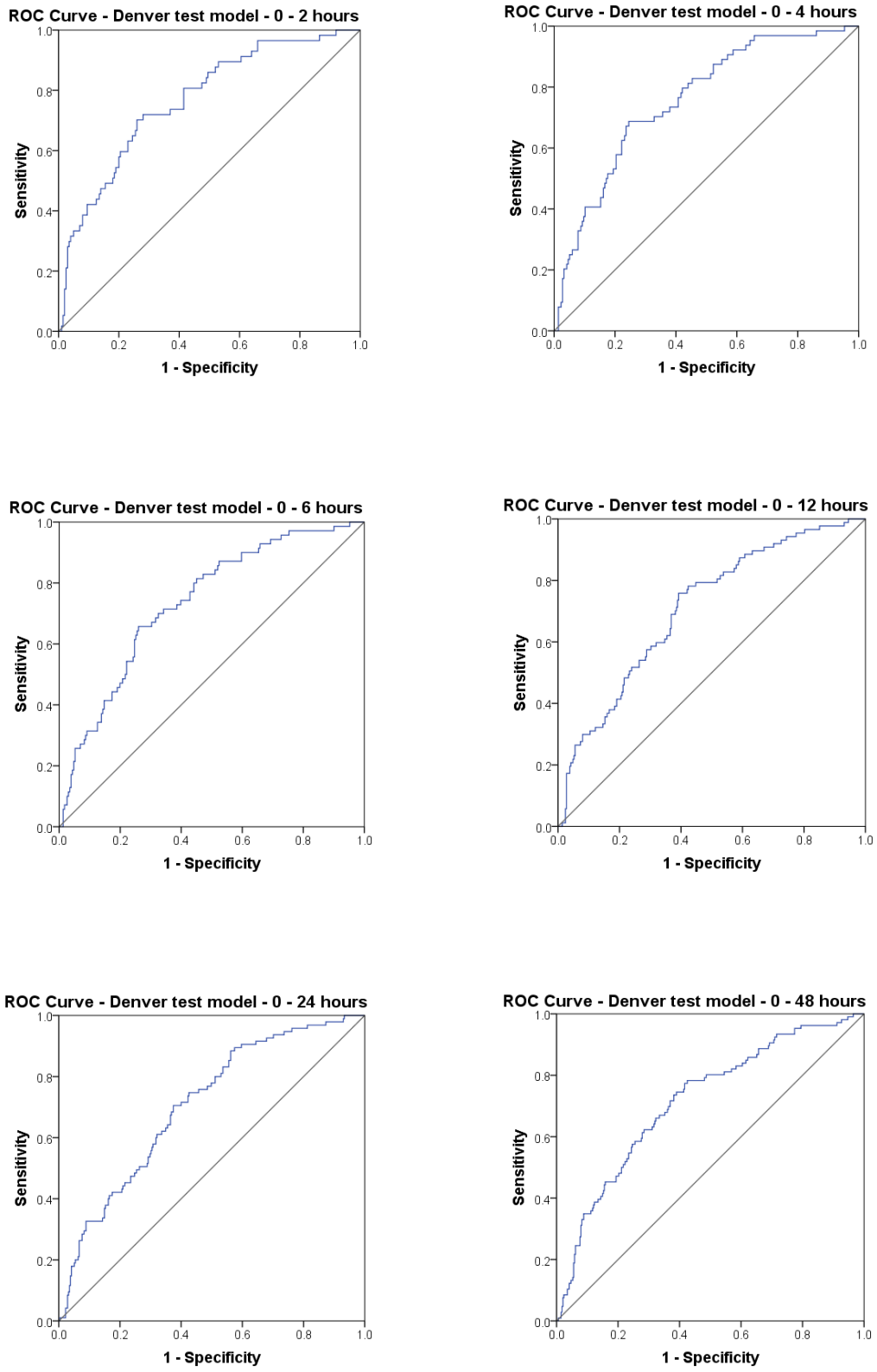


Figure 7.16: ROC curves for Denver combined test model

Denver test set	AUC	95% CI
0-2 hours	0.767	0.699, 0.836
0-4 hours	0.758	0.694, 0.823
0-6 hours	0.735	0.671, 0.799
0-12 hours	0.708	0.648, 0.768
0-24 hours	0.707	0.650, 0.763
0-48 hours	0.717	0.661, 0.772

Table 7.45: AUC for ROC curves for Denver combined test model

A marked decrease in sensitivity, PPV and AUC was observed in the 0 – 12, 0 – 24 and 0 – 48 hours models. An extended set of variables was developed, consisting of the original test set and a small number of additional variables to improve the accuracy of prediction. The following variables were added in the extended variable set:

- PT
- BD
- Creatinine
- Vasopressor use

The results for this extended variable set are shown in Figures 7.17 and 7.18 and Tables 7.46 to 7.48. Sensitivity, PPV and AUC were markedly increased for all models with the addition of these four variables to the original test set.

0-12 hours

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-6.476		
Age	0.008	0.549	1.01 (0.98, 1.04)
CV disease	0.320	0.724	1.38 (0.23, 8.15)
Head injury	0.972	0.088	2.64 (0.87, 8.08)
Chest injury	0.971	0.109	2.64 (0.73, 4.94)
Abdominal injury	0.641	0.188	1.90 (0.73, 4.94)
Limb/pelvic injury	0.324	0.553	1.38 (0.47, 4.03)
HR	0.005	0.667	1.01 (0.98, 1.03)
RR max	0.028	0.272	1.03 (0.98, 1.08)
PT	-0.077	0.343	0.93 (0.79, 1.09)
BD	-0.085	0.162	0.92 (0.82, 1.04)
Creatinine	0.017	0.062	1.02 (1.00, 1.04)
Vasopressors	1.725	<0.001	5.62 (2.26, 13.94)

0-24 hours

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-9.039		
Age	0.019	0.071	1.02 (1.00, 1.04)
CV disease	0.334	0.688	1.40 (0.27, 7.12)
Head injury	1.622	0.001	5.06 (1.96, 13.10)
Chest injury	0.273	0.534	1.31 (0.56, 3.11)
Abdominal injury	0.496	0.212	1.64 (0.75, 3.57)
Limb/pelvic injury	0.619	0.177	1.86 (0.76, 4.56)
HR	0.021	0.021	1.02 (1.00, 1.04)
RR max	0.027	0.154	1.03 (1.00, 1.09)
PT	-0.036	0.56	0.97 (0.86, 1.09)
BD	-0.085	0.064	0.92 (0.84, 1.01)
Creatinine	0.014	0.046	1.01 (1.00, 1.03)
Vasopressors	1.389	<0.001	4.01 (1.92, 8.36)

0-48 hours

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-8.661		
Age	0.006	0.509	1.01 (0.99, 1.03)
CV disease	0.243	0.791	1.28 (0.21, 7.71)
Head injury	1.792	<0.001	6.00 (2.44, 14.74)
Chest injury	0.094	0.806	1.10 (0.52, 2.32)
Abdominal injury	0.473	0.192	1.61 (0.79, 3.27)
Limb/pelvic injury	0.48	0.224	1.62 (0.75, 3.51)
HR	0.003	0.678	1.00 (0.99, 1.02)
RR max	0.052	0.004	1.05 (1.02, 1.09)
PT	0.05	0.349	1.05 (0.95, 1.17)
BD	-0.108	0.006	0.90 (0.83, 0.97)
Creatinine	0.011	0.03	1.01 (1.00, 1.02)
Vasopressors	1.368	<0.001	3.93 (1.98, 7.78)

Figure 7.17: Denver extended test model output

Denver extended set	Number of patients	Model χ^2	Model χ^2 p value	-2 log likelihood	Nagelkerke R^2	Hosmer-Lemeshow χ^2	Hosmer-Lemeshow χ^2 p value
0-12 hours	196	51.911	<0.001	146.445	0.366	6.240	0.620
0-24 hours	291	77.556	<0.001	218.601	0.366	2.178	0.975
0-48 hours	355	108.057	<0.001	258.041	0.408	8.648	0.373

Table 7.46: Goodness-of-fit measures for Denver extended test model

Denver extended set	PAC (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
0-12 hours	84.2	47.5	93.6	65.5	87.4
0-24 hours	82.8	35.0	95.2	65.6	84.9
0-48 hours	84.2	45.3	94.6	69.4	86.6

Table 7.47: Category predictors for Denver extended test model

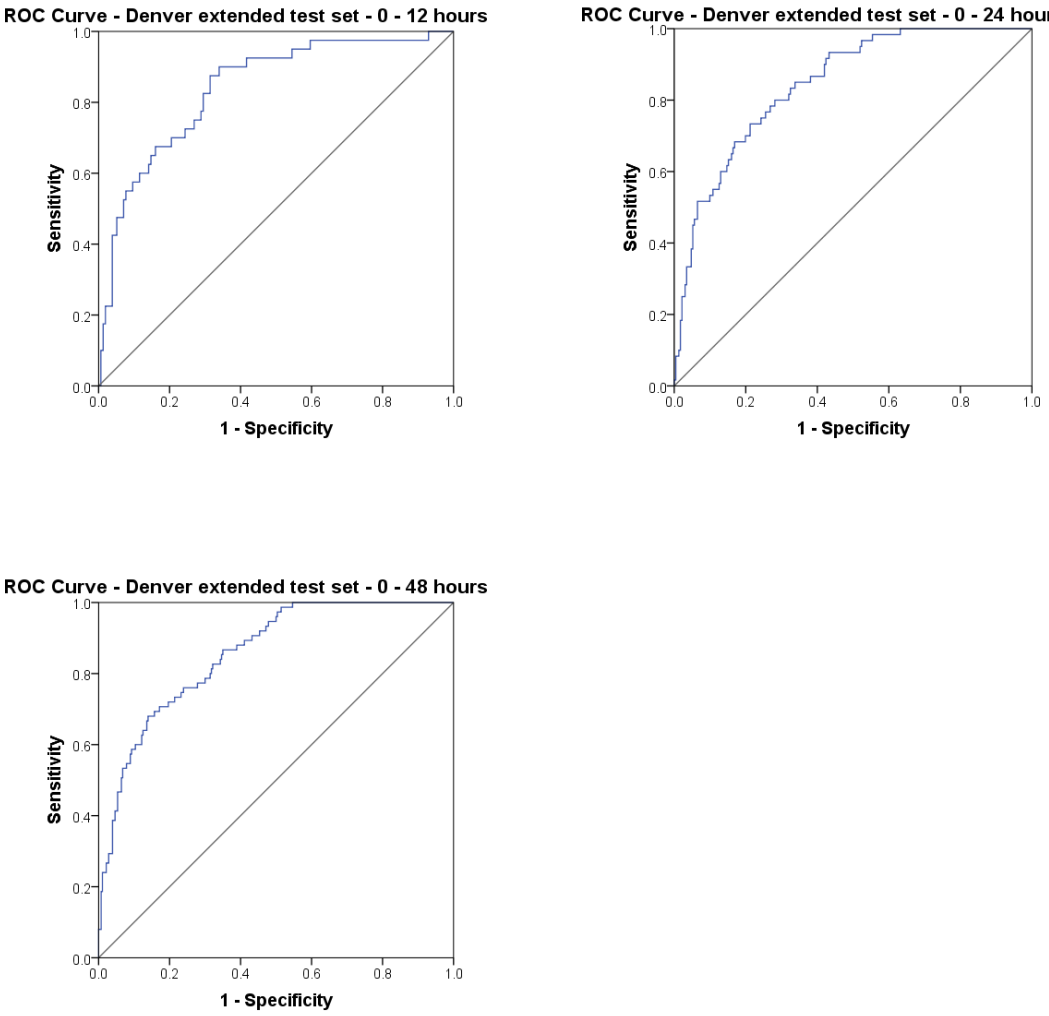


Figure 7.18:ROC curves for Denver extended test model at 0 – 12 hours, 0 – 24 hours and 0 – 48 hours post-injury

Denver extended set	AUC	95% CI
0-12 hours	0.839	0.770, 0.907
0-24 hours	0.845	0.795, 0.895
0-48 hours	0.856	0.812, 0.899

Table 7.48: AUC for ROC curves for Denver extended test model

7.5.6. SOFA COMBINED MODEL

An identical process of variable identification and refinement was performed for the results of the SOFA multivariable models. The following variables were included in a test model:

- Age
- Respiratory disease
- Head injury
- Abdominal injury
- Trapped
- SBP
- HR
- Maximum RR

The results of testing this model at multiple time periods are given in Figures 7.19 and 7.20 and Tables 7.49 to 7.51.

0-2 hours

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-3.974		
Age	0.030	<0.001	1.03 (1.01, 1.05)
Respiratory disease	1.036	0.050	2.82 (1.00, 7.95)
Head injury	1.073	0.001	2.93 (1.60, 5.36)
Abdominal injury	0.485	0.126	1.63 (0.87, 3.02)
Trapped	0.573	0.062	1.77 (0.97, 3.23)
SBP	-0.008	0.066	0.99 (0.98, 1.00)
HR	0.019	0.001	1.01 (1.00, 1.03)
RR max	0.036	0.032	1.04 (1.00, 1.07)

0-4 hours

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-3.596		
Age	0.031	<0.001	1.03 (1.02, 1.05)
Respiratory disease	1.039	0.049	2.83 (1.00, 7.95)
Head injury	1.003	0.001	2.73 (1.54, 4.81)
Abdominal injury	0.564	0.060	1.76 (0.98, 3.17)
Trapped	0.514	0.075	1.67 (0.95, 2.95)
SBP	-0.007	0.078	0.99 (0.99, 1.00)
HR	0.018	0.002	1.02 (1.01, 1.03)
RR max	0.020	0.193	1.02 (1.00, 1.05)

0-6 hours

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-3.189		
Age	0.027	<0.001	1.03 (1.01, 1.04)
Respiratory disease	0.899	0.068	2.46 (0.94, 6.45)
Head injury	0.919	0.001	2.51 (1.48, 4.27)
Abdominal injury	0.569	0.045	1.77 (1.01, 3.08)
Trapped	0.518	0.060	1.68 (0.98, 2.88)
SBP	-0.007	0.103	0.99 (0.98, 1.00)
HR	0.015	0.006	1.02 (1.00, 1.03)
RR max	0.020	0.179	1.02 (1.00, 1.05)

0-12 hours

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-3.668		
Age	0.023	<0.001	1.02 (1.01, 1.04)
Respiratory disease	0.605	0.122	1.83 (0.85, 3.94)
Head injury	0.825	0.001	2.28 (1.43, 3.65)
Abdominal injury	0.369	0.147	1.45 (0.88, 2.38)
Trapped	0.530	0.033	1.70 (1.05, 2.76)
SBP	-0.005	0.279	1.00 (1.01, 1.03)
HR	0.016	0.003	1.02 (1.01, 1.03)
RR max	0.032	0.016	1.03 (1.01, 1.06)

0-24 hours

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-3.721		
Age	0.025	<0.001	1.03 (1.01, 1.04)
Respiratory disease	0.432	0.225	1.54 (0.77, 3.10)
Head injury	0.830	<0.001	2.29 (1.46, 3.61)
Abdominal injury	0.321	0.188	1.38 (0.86, 2.23)
Trapped	0.571	0.016	1.77 (1.11, 2.81)
SBP	-0.006	0.206	1.00 (0.99, 1.00)
HR	0.017	0.002	1.02 (1.01, 1.03)
RR max	0.028	0.021	1.03 (1.00, 1.05)

0-48 hours

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-3.435		
Age	0.023	<0.001	1.02 (1.01, 1.04)
Respiratory disease	0.583	0.095	1.79 (0.90, 3.55)
Head injury	0.826	<0.001	2.29 (1.47, 3.54)
Abdominal injury	0.366	0.122	1.44 (0.91, 2.29)
Trapped	0.534	0.020	1.71 (1.09, 2.67)
SBP	-0.006	0.191	1.00 (0.99, 1.00)
HR	0.012	0.017	1.01 (1.00, 1.02)
RR max	0.033	0.006	1.03 (1.01, 1.06)

Figure 7.19: SOFA combined test model output

SOFA test set	Number of patients	Model χ^2	Model χ^2 p value	-2 log likelihood	Nagelkerke R^2	Hosmer-Lemeshow χ^2	Hosmer-Lemeshow χ^2 p value
0-2 hours	251	49.080	<0.001	295.042	0.238	9.407	0.309
0-4 hours	273	48.083	<0.001	326.848	0.216	6.144	0.631
0-6 hours	294	44.369	<0.001	358.782	0.188	7.735	0.460
0-12 hours	366	46.218	<0.001	454.853	0.159	9.667	0.289
0-24 hours	399	51.127	<0.001	494.398	0.161	10.945	0.205
0-48 hours	435	54.006	<0.001	539.876	0.157	7.860	0.447

Table 7.49: Goodness-of-fit measures for SOFA combined test model

SOFA test set	PAC (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
0-2 hours	72.1	77.3	65.5	74.1	69.2
0-4 hours	70.3	76.3	62.8	76.8	67.9
0-6 hours	66.7	73.9	57.4	68.9	62.4
0-12 hours	65.8	76.3	52.2	67.5	62.9
0-24 hours	66.2	76.7	52.3	68.0	62.9
0-48 hours	66.2	77.9	50.5	67.4	63.1

Table 7.50:Category predictors for SOFA combined test model

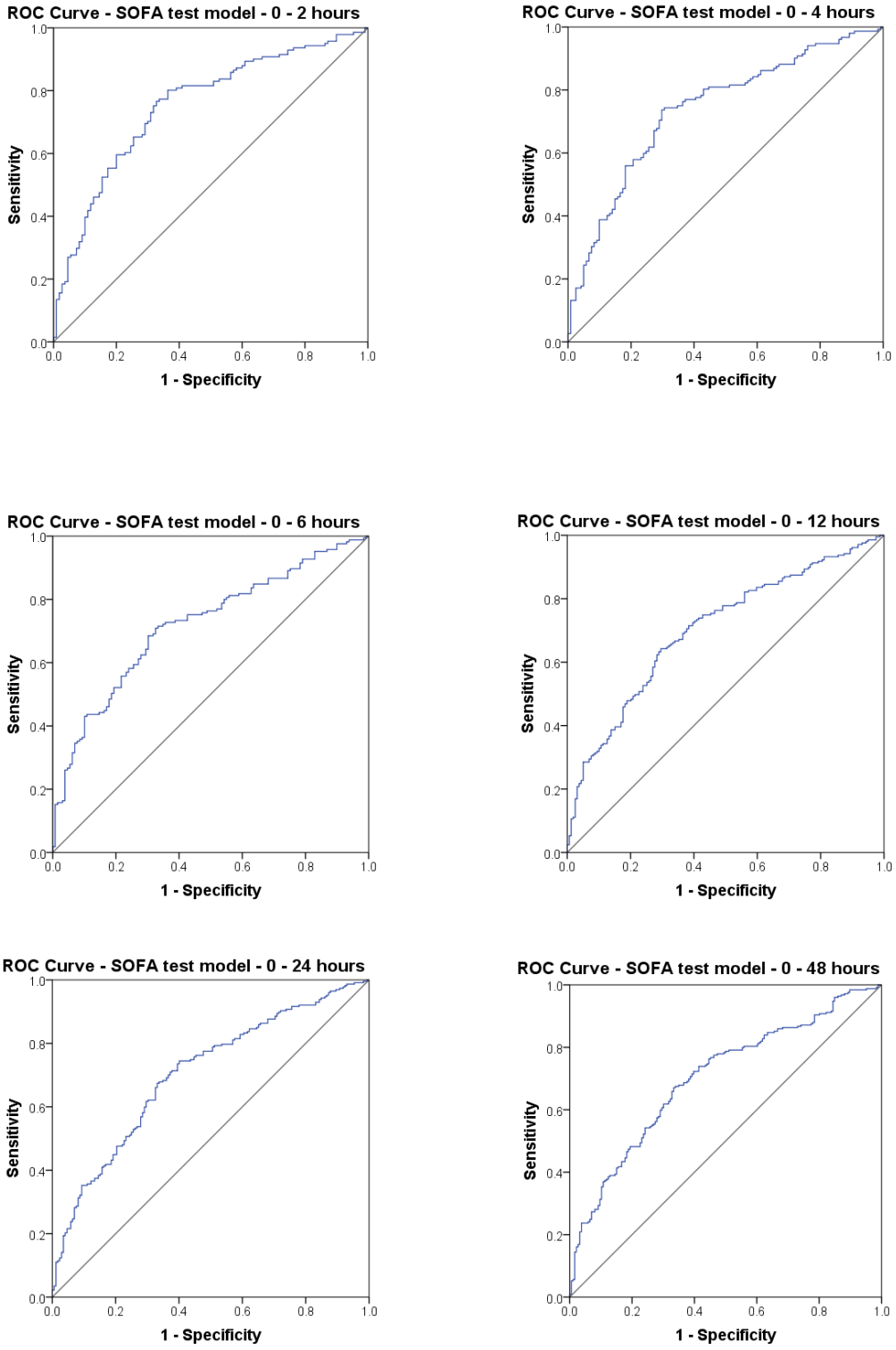


Figure 7.20: ROC curves for SOFA combined test model

SOFA test set	AUC	95% CI
0-2 hours	0.753	0.692, 0.813
0-4 hours	0.739	0.679, 0.798
0-6 hours	0.720	0.662, 0.777
0-12 hours	0.704	0.651, 0.757
0-24 hours	0.704	0.653, 0.755
0-48 hours	0.702	0.653, 0.751

Table 7.51: AUC for ROC curves for SOFA combined test model

A marked decrease in specificity and the AUC was noted for the 0 – 12, 0 – 24 and 0 – 48 hours models. An extended test set was developed for these time periods, and involved addition of the following variables:

- FiO_2
- Albumin
- Transfusion level
- Vasopressor use

The results for the extended set are given in Figures 7.21 and 7.22 and Tables 7.52 to 7.54.

0-12 hours

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-2.743		
Age	0.017	0.155	1.02 (1.00, 1.04)
Respiratory disease	1.026	0.164	2.79 (0.66, 11.83)
Head injury	1.434	0.004	4.20 (1.60, 11.00)
Abdominal injury	0.150	0.742	1.16 (0.48, 2,85)
Trapped	0.637	0.128	1.89 (0.83, 4.29)
SBP	0.009	0.304	1.01 (1.00, 1.03)
HR	0.013	0.174	1.01 (1.00, 1.03)
RR max	0.033	0.121	1.03 (1.00, 1.08)
FiO2		0.142	
41-60%	-0.472	0.418	0.62 (0.20, 1.95)
61-80%	0.566	0.522	1.76 (0.31, 9.96)
> 80%	0.500	0.396	1.65 (0.52, 5.23)
Albumin	-0.080	0.003	0.92 (0.88, 0.97)
Transfusion level		0.352	
Transfusion	-0.048	0.920	0.95 (0.37, 2.43)
Massive transfusion	1.599	0.153	4.95 (0.55, 44.33)
Vasopressors	0.951	0.042	2.59 (1.04, 6.47)

0-24 hours

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-2.390		
Age	0.015	0.153	1.02 (1.00, 1.04)
Respiratory disease	0.357	0.499	1.43 (0.51, 4.03)
Head injury	1.060	0.004	2.96 (1.41, 6.24)
Abdominal injury	0.056	0.882	1.06 (0.51, 2.21)
Trapped	0.473	0.184	1.61 (0.80, 3.23)
SBP	0.004	0.614	1.00 (0.99, 1.02)
HR	0.016	0.065	1.02 (1.00, 1.03)
RR max	0.015	0.388	1.02 (0.98, 1.05)
FiO2		0.080	
41-60%	-0.137	0.786	0.88 (0.32, 2.35)
61-80%	0.668	0.349	1.95 (0.48, 7.90)
> 80%	0.782	0.127	2.19 (0.80, 5.97)
Albumin	-0.073	0.001	0.93 (0.89, 0.97)
Transfusion level		0.480	
Transfusion	-0.228	0.555	0.80 (0.37, 1.70)
Massive transfusion	0.755	0.338	2.13 (0.45, 9.98)
Vasopressors	1.287	0.001	3.62 (1.73, 7.59)

0-48 hours

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-0.784		
Age	0.014	0.124	1.01 (1.00, 1.03)
Respiratory disease	0.611	0.246	1.84 (0.66, 5.16)
Head injury	0.996	0.006	2.71 (1.34, 5.49)
Abdominal injury	-0.285	0.421	0.75 (0.38, 1.51)
Trapped	0.406	0.237	1.50 (0.77, 2.95)
SBP	-0.001	0.860	1.00 (0.98, 1.01)
HR	0.018	0.042	1.02 (1.00, 1.04)
RR max	0.004	0.808	1.00 (0.97, 1.04)
FiO2		0.008	
41-60%	-0.524	0.331	0.60 (0.21, 1.71)
61-80%	-0.185	0.801	0.83 (0.20, 3.50)
> 80%	0.698	0.164	2.01 (0.75, 5.37)
Albumin	-0.105	<0.001	0.9 (0.86, 0.95)
Transfusion level		0.734	
Transfusion	-0.161	0.662	0.85 (0;.41, 1.75)
Massive transfusion	0.495	0.552	1.64 (0.32, 8.36)
Vasopressors	1.352	<0.001	3.87 (1.91, 7.82)

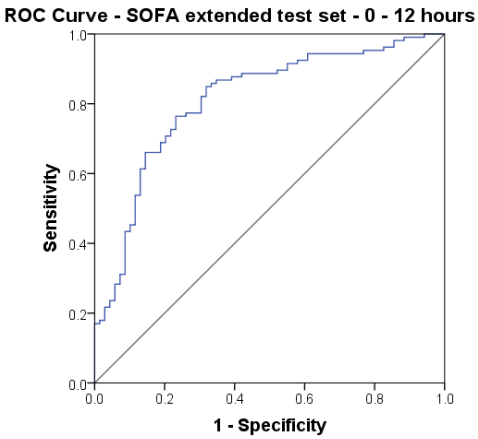
Figure 7.21: SOFA extended test model output

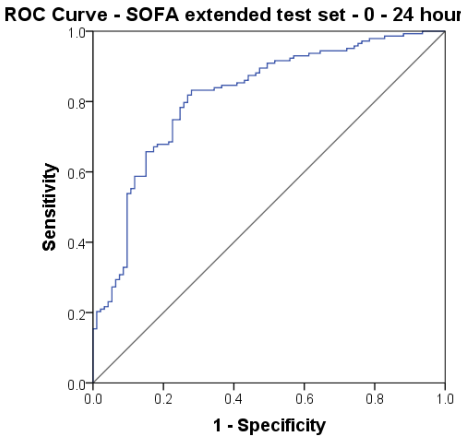
SOFA extended set	Number of patients	Model χ^2	Model χ^2 p value	-2 log likelihood	Nagelkerke R^2	Hosmer-Lemeshow χ^2	Hosmer-Lemeshow χ^2 p value
0-12 hours	175	54.285	<0.001	180.435	0.361	10.329	0.243
0-24 hours	236	73.883	<0.001	242.608	0.364	13.581	0.093
0-48 hours	279	108.679	<0.001	266.369	0.436	14.174	0.077

Table 7.52: Goodness-of-fit measures for SOFA extended test model

SOFA extended set	PAC (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
0-12 hours	78.3	85.8	66.7	79.8	75.4
0-24 hours	76.7	83.9	65.6	78.9	72.6
0-48 hours	80.6	85.1	73.9	83.1	76.6

Table 7.53: Category predictors for SOFA extended test model





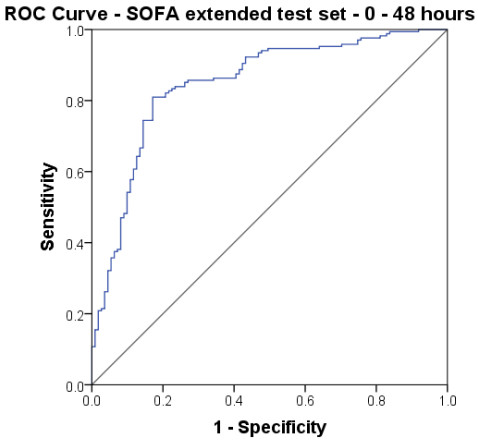


Figure 7.22: ROC curves for SOFA extended test model

SOFA extended set	AUC	95% CI
0-12 hours	0.811	0.745, 0.878
0-24 hours	0.815	0.759, 0.872
0-48 hours	0.848	0.800, 0.895

Table 7.54: AUC for ROC curves for SOFA extended test model

7.5.7. MARSHALL COMBINED MODEL

An identical process of variable identification and refinement was performed for the results of the Marshall multivariable models. The following variables were included in a test model:

- Age
- Sex
- Pre-existing CV disease
- Head injury
- Abdominal injury
- Limb/pelvic injury
- HR

The results of testing this model at multiple time periods are given in Figures 7.23 and 7.24 and Tables 7.55 to 7.57.

0-2 hours

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-4.536		
Age	0.019	0.003	1.02 (1.01, 1.03)
Sex	0.547	0.048	1.73 (1.00, 2.97)
CV disease	1.201	0.016	3.32 (1.25, 8.86)
Head injury	1.290	<0.001	3.63 (1.27, 8.86)
Abdominal injury	1.034	<0.001	2.81 (1.67, 4.73)
Limb/pelvic injury	0.764	0.006	2.15 (1.24, 3.72)
HR	0.010	0.034	1.01 (1.00, 1.02)

0-4 hours

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-4.396		
Age	0.019	0.003	1.02 (1.01, 1.03)
Sex	0.432	0.100	1.54 (0.91, 2.61)
CV disease	1.120	0.023	3.06 (1.16, 8.07)
Head injury	1.214	<0.001	2.37 (1.99, 5.70)
Abdominal injury	1.077	<0.001	2.94 (1.76, 4.89)
Limb/pelvic injury	0.787	0.004	2.20 (1.29, 3.76)
HR	0.010	0.039	1.01 (1.00, 1.02)

0-6 hours

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-4.358		
Age	0.019	0.003	1.02 (1.01, 1.03)
Sex	0.510	0.057	1.67 (0.99, 2.81)
CV disease	1.181	0.015	3.26 (1.27, 8.41)
Head injury	1.166	<0.001	3.21 (1.93, 5.35)
Abdominal injury	1.080	<0.001	2.95 (1.93, 5.35)
Limb/pelvic injury	0.773	0.004	2.17 (1.79, 4.84)
HR	0.009	0.058	1.01 (1.00, 1.02)

0-12 hours

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-3.851		
Age	0.015	0.011	1.02 (1.00, 1.03)
Sex	0.586	0.022	1.80 (1.09, 2.97)
CV disease	1.207	0.008	3.34 (1.36, 8.20)
Head injury	1.124	<0.001	3.08 (1.90, 4.98)
Abdominal injury	0.892	<0.001	2.44 (1.53, 3.89)
Limb/pelvic injury	0.695	0.005	2.00 (1.24, 3.25)
HR	0.007	0.145	1.01 (1.00, 1.02)

0-24 hours

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-4.075		
Age	0.017	0.004	1.02 (1.01, 1.03)
Sex	0.574	0.024	1.78 (1.08, 2.92)
CV disease	1.075	0.017	2.93 (1.22, 7.07)
Head injury	1.187	<0.001	3.28 (2.04, 5.27)
Abdominal injury	0.787	0.001	2.20 (1.39, 3.48)
Limb/pelvic injury	0.658	0.006	1.93 (1.20, 3.10)
HR	0.008	0.083	1.01 (1.00, 1.02)

0-48 hours

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-3.798		
Age	0.015	0.007	1.02 (1.00, 1.03)
Sex	0.620	0.012	1.86 (1.15, 3.02)
CV disease	0.972	0.025	2.64 (1.13, 6.19)
Head injury	1.183	<0.001	3.26 (2.06, 5.18)
Abdominal injury	0.752	0.001	2.12 (1.35, 3.33)
Limb/pelvic injury	0.586	0.012	1.80 (1.14, 2.84)
HR	0.007	0.139	1.01 (1.00, 1.02)

Figure 7.23: Model output for Marshall combined test model

Marshall test set	Number of patients	Model χ^2	Model χ^2 p value	-2 log likelihood	Nagelkerke R^2	Hosmer-Lemeshow χ^2	Hosmer-Lemeshow χ^2 p value
0-2 hours	358	53.610	<0.001	422.789	0.189	15.048	0.058
0-4 hours	372	55.039	<0.001	439.647	0.187	16.238	0.039
0-6 hours	382	56.650	<0.001	453.386	0.187	13.683	0.090
0-12 hours	423	50.954	<0.001	516.582	0.154	16.100	0.041
0-24 hours	442	52.042	<0.001	540.556	0.150	6.922	0.545
0-48 hours	462	48.158	<0.001	573.050	0.134	7.966	0.437

Table 7.55:Goodness-of-fit for Marshall combined test model

Marshall test set	PAC (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
0-2 hours	72.3	45.3	89.1	72.1	72.4
0-4 hours	72.3	46.5	88.3	71.0	73.6
0-6 hours	71.2	58.0	85.9	68.3	71.9
0-12 hours	68.1	43.1	84.4	64.3	69.5
0-24 hours	67.4	40.8	84.7	61.7	68.8
0-48 hours	66.2	38.0	84.9	62.5	67.4

Table 7.56: Category predictors for Marshall combined test model

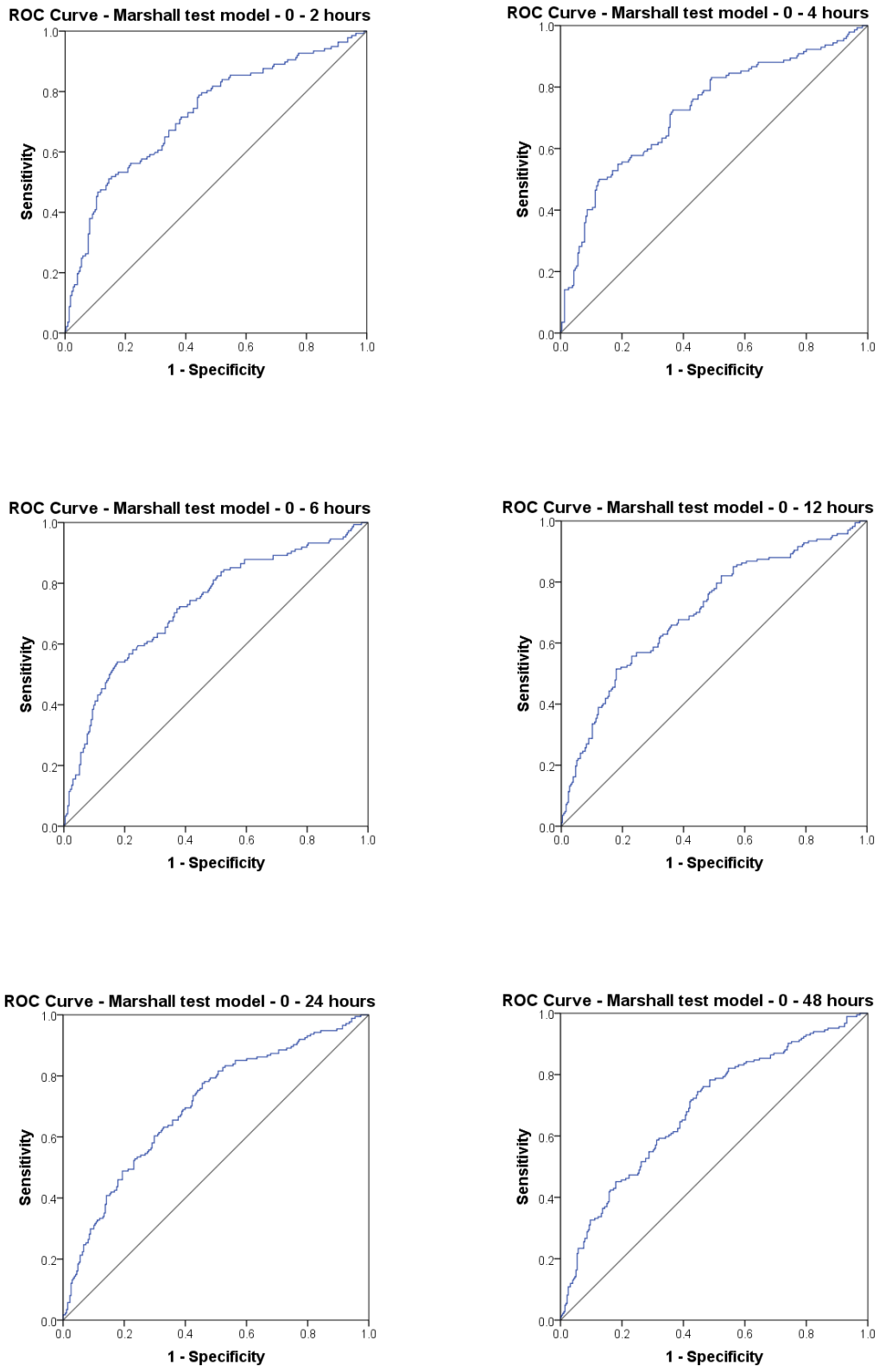


Figure 7.24: ROC curves for Marshall combined test model

Marshall test set	AUC	95% CI
0-2 hours	0.725	0.670, 0.780
0-4 hours	0.725	0.670, 0.779
0-6 hours	0.725	0.672, 0.779
0-12 hours	0.702	0.651, 0.753
0-24 hours	0.701	0.651, 0.751
0-48 hours	0.686	0.637, 0.736

Table 7.57: AUC for ROC curves for Marshall combined test model

Models using the Marshall test set were noted to be less well-fitting than for the other test sets, especially at early time points after injury.

A decrease in sensitivity, PPV and AUC were noted in the 0 – 12, 0 – 24 and 0 – 48 hours models. An extended set of variables was defined to include the following in addition to the original test set:

- FiO_2
- APTT
- Lactate
- Vasopressor use

A marked increase in sensitivity and AUC was seen with the addition of these variables (Figures 7.25 and 7.26 and Tables 7.58 to 7.60).

0-12 hours

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-5.742		
Age	0.023	0.040	1.02 (1.00, 1.05)
Sex	-0.128	0.761	0.88 (0.39, 2.01)
CV disease	-1.826	0.098	0.16 (0.02, 1.40)
Head injury	1.930	<0.001	6.89 (2.60, 18.26)
Abdominal injury	1.395	0.001	4.03 (1.78, 9.15)
Limb/pelvic injury	0.717	0.102	2.05 (0.87, 4.84)
HR	0.010	0.289	1.01 (1.00, 1.03)
FiO_2		0.017	
41-60%	-0.726	0.188	0.48 (0.16, 1.43)
61-80%	0.768	0.325	2.16 (0.47, 9.94)
> 80%	0.528	0.350	1.70 (0.56, 5.12)
APTT	0.009	0.091	1.01 (1.00, 1.02)
Lactate	0.234	0.016	1.26 (1.05, 1.53)
Vasopressors	0.708	0.091	2.03 (0.90, 4.62)

0-24 hours

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-5.593		
Age	0.020	0.041	1.02 (1.00, 1.04)
Sex	0.110	0.774	1.12 (0.53, 2.36)
CV disease	-2.174	0.053	0.11 (0.01, 1.03)
Head injury	1.935	<0.001	6.92 (2.97, 16.11)
Abdominal injury	1.166	0.002	3.21 (1.53, 6.74)
Limb/pelvic injury	0.740	0.057	2.10 (0.98, 4.49)
HR	0.008	0.343	1.01 (1.00, 1.03)
FiO_2		0.002	
41-60%	-0.962	0.061	0.38 (0.14, 1.04)
61-80%	0.444	0.487	1.56 (0.45, 5.46)
> 80%	0.448	0.368	1.57 (0.59, 4.15)
APTT	0.013	0.012	1.02 (1.00, 1.02)
Lactate	0.231	0.006	1.26 (1.07, 1.49)
Vasopressors	0.768	0.027	2.15 (1.09, 4.28)

0-48 hours

Variable	Coefficient	Wald statistic p value	Adjusted OR (95% CI)
Constant	-5.406		
Age	0.019	0.028	1.02 (1.00, 1.04)
Sex	0.450	0.201	1.57 (0.79, 3.12)
CV disease	-1.046	0.268	0.35 (0.06, 2.24)
Head injury	1.472	<0.001	4.36 (2.08, 9.13)
Abdominal injury	0.857	0.012	2.36 (1.21, 4.60)
Limb/pelvic injury	0.556	0.112	1.75 (0.88, 3.47)
HR	0.010	0.189	1.01 (1.00, 1.03)
FiO_2		0.001	
41-60%	-1.253	0.017	0.29 (0.10, 0.80)
61-80%	-0.719	0.311	0.49 (0.12, 1.96)
> 80%	0.229	0.617	1.26 (0.51, 3.10)
APTT	0.011	0.029	1.01 (1.00, 1.02)
Lactate	0.226	0.005	1.25 (1.07, 1.47)
Vasopressors	1.043	0.001	2.84 (1.53, 5.25)

Figure 7.25: Model output for Marshall extended test model

Marshall extended set	Number of patients	Model χ^2	Model χ^2 p value	-2 log likelihood	Nagelkerke R^2	Hosmer-Lemeshow χ^2	Hosmer-Lemeshow χ^2 p value
0-12 hours	187	66.669	<0.001	191.664	0.401	13.642	0.092
0-24 hours	242	88.961	<0.001	245.182	0.411	12.370	0.135
0-48 hours	285	108.853	<0.001	289.382	0.408	2.719	0.951

Table 7.58: Goodness-of-fit for Marshall extended test model

Marshall extended set	PAC (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
0-12 hours	78.1	75.9	80.0	76.7	79.2
0-24 hours	76.4	74.1	78.5	82.2	77.9
0-48 hours	74.7	71.0	77.9	73.2	75.9

Table 7.59: Category predictors for Marshall extended test model

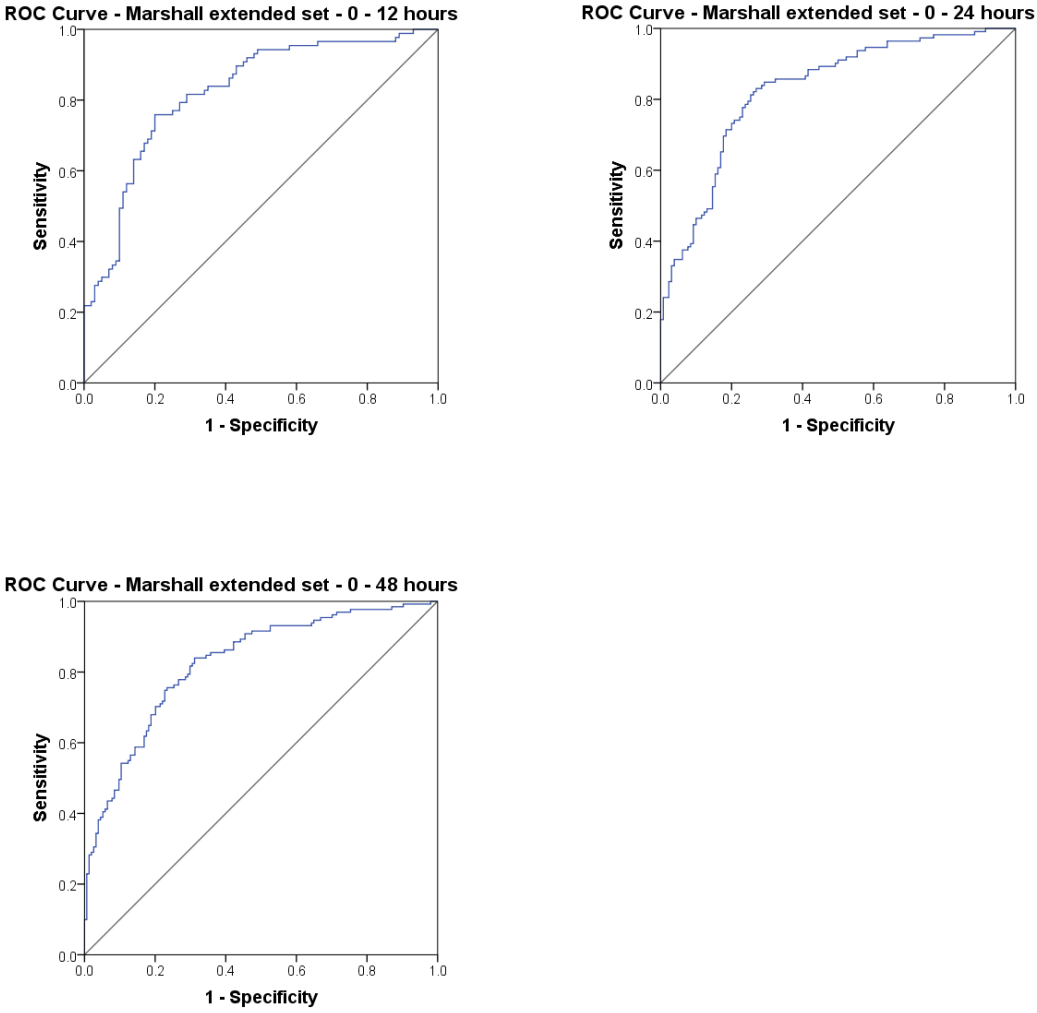


Figure 7.26: ROC curves for Marshall extended test mode

Marshall extended set	AUC	95% CI
0-12 hours	0.823	0.763, 0.884
0-24 hours	0.827	0.776, 0.879
0-48 hours	0.826	0.778, 0.874

Table 7.60: AUC for ROC curves for Marshall extended test model

7.5.8. COMPARISON OF INCLUDED AND EXCLUDED PATIENTS FROM COMBINED AND EXTENDED MODELS

Basic demographic and population characteristics were compared between the patients included in the combined and extended models, and those excluded due to missing data. The results of these comparisons are given in Appendix 15.

For the Denver combined model, significant differences were seen in age in the 0-2 and 0-4 hours combined models, with included patients being significantly older. This trend was reversed in the 0-48 hour extended model, with included patients being significantly younger. Significant variations were also seen in hospital stay (0-4 hours combined model and 0-24 hours extended model) and in deaths (0-48 hours extended model).

For the SOFA combined model, a significant difference in age was seen in the 0-2 hours model (included patients older), and in the 0-24 hours extended model (included patients younger). Hospital stay was significantly longer in the included patients for the 0-4 and 0-6 hours combined models. ISS and NISS were not significantly different for any of the combined models, but included patients showed significantly higher injury severity for all of the extended models.

For the Marshall combined model, patients in the 0-24 and 0-48 hours extended models were significantly younger than those excluded. An extended hospital stay was seen in the patients included in the 0-2, 0-4 and 0-6 hours combined model. In addition, patients included in the 0-2 and 0-4 hours combined models, and the 0-12 hours extended model had a significantly lower incidence of operative intervention. Similar to the SOFA models, ISS and NISS were not significantly different in any of the combined models, but a higher injury severity was seen in included patients for the extended models.

7.6. INTERNAL VALIDATION

7.6.1. DENVER COMBINED AND EXTENDED MODELS

The results of bootstrapping for individual variables are given in Figures 7.27 and 7.28. Variation was observed in the coefficients obtained on bootstrapping, with confidence intervals crossing zero for a number of variables. Head injury and abdominal injury were consistently shown to be highly significant predictors, with maximum HR showing good predictive value in the early phase following injury. Bias estimates were low, with the exception of cardiovascular disease which was greatly impacted by sampling, since it was present in only a small number of patients.

0-2 hours

Variable	Coefficient (95% BCa CI)	Bias	Standard Error	p value
Constant	-6.197 (-8.236, -5.141)			
Age	0.012 (-0.009, 0.034)	0.001	0.010	0.211
CV disease	1.649 (-0.647, 4.197)	0.126	1.072	0.020
Head injury	1.608 (0.756, 2.967)	0.143	0.446	0.002
Chest injury	0.592 (-0.38, 1.741)	0.069	0.471	0.169
Abdominal injury	0.926 (0.205, 1.776)	0.032	0.393	0.012
Limb/pelvic injury	0.335 (-0.604, 1.434)	0.043	0.482	0.450
HR	0.015 (0.001, 0.035)	0.001	0.008	0.025
RR max	0.025 (-0.014, 0.067)	0.001	0.019	0.162

0-4 hours

Variable	Coefficient (95% BCa CI)	Bias	Standard Error	p value
Constant	-5.561 (-7.566, -4.289)			
Age	0.011 (-0.007, 0.031)	0.001	0.009	0.200
CV disease	1.733 (0.152, 3.309)	0.059	0.707	0.004
Head injury	1.498 (0.699, 2.692)	0.091	0.415	0.001
Chest injury	0.422 (-0.523, 1.473)	0.053	0.432	0.289
Abdominal injury	0.970 (0.260, 1.756)	0.032	0.351	0.003
Limb/pelvic injury	0.393 (-0.424, 1.432)	0.035	0.450	0.329
HR	0.017 (0.004, 0.032)	0.001	0.007	0.012
RR max	0.002 (-0.039, 0.044)	0.001	0.020	0.913

0-6 hours

Variable	Coefficient (95% BCa CI)	Bias	Standard Error	p value
Constant	-4.926 (-7.019, -3.153)			
Age	0.008 (-0.10, 0.025)	0.000	0.009	0.330
CV disease	1.633(-0.143, 3.294)	0.076	0.711	0.006
Head injury	1.422 (0.694, 2.328)	0.061	0.358	0.001
Chest injury	0.343 (-0.466, 1.245)	0.042	0.395	0.333
Abdominal injury	0.855 (0.157, 1.536)	0.025	0.329	0.009
Limb/pelvic injury	0.273 (-0.509, 1.190)	0.038	0.391	0.443
HR	0.012 (-0.002, 0.026)	0.000	0.007	0.072
RR max	0.013 (-0.024, 0.051)	0.000	0.019	0.488

0-12 hours

Variable	Coefficient (95% BCa CI)	Bias	Standard Error	p value
Constant	-4.583 (-6.753, -2.941)			
Age	0.006 (-0.011, 0.023)	0.000	0.008	0.397
CV disease	1.532 (0.211, 3.008)	0.062	0.601	0.010
Head injury	1.278 (0.739, 2.041)	0.042	0.306	0.001
Chest injury	0.349 (-0.280, 1.035)	0.008	0.335	0.279
Abdominal injury	0.651 (0.081, 1.264)	0.018	0.287	0.013
Limb/pelvic injury	0.229 (-0.389, 0.911)	0.009	0.326	0.463
HR	0.007 (-0.007, 0.023)	0.000	0.007	0.266
RR max	0.027 (-0.002, 0.058)	0.001	0.015	0.066

0-24 hours

Variable	Coefficient (95% BCa CI)	Bias	Standard Error	p value
Constant	-5.146 (-6.770, -4.158)			
Age	0.010 (-0.004, 0.028)	0.001	0.007	0.124
CV disease	1.231 (-0.024, 2.394)	0.019	0.596	0.023
Head injury	1.207 (0.633, 1.961)	0.044	0.286	0.001
Chest injury	0.170 (-0.524, 0.915)	0.020	0.315	0.547
Abdominal injury	0.575 (-0.006, 1.119)	0.001	0.283	0.036
Limb/pelvic injury	0.248 (-0.464, 0.997)	0.019	0.330	0.426
HR	0.009 (-0.004, 0.024)	0.001	0.006	0.116
RR max	0.036 (0.005, 0.076)	0.001	0.015	0.012

0-48 hours

Variable	Coefficient (95% BCa CI)	Bias	Standard Error	p value
Constant	-4.850 (-6.679, -3.394)			
Age	0.010 (-0.003, 0.023)	0.000	0.007	0.127
CV disease	1.229 (-0.101, 2.521)	0.046	0.599	0.028
Head injury	1.201 (0.638, 1.942)	0.049	0.284	0.001
Chest injury	0.086 (-0.456, 0.657)	0.003	0.287	0.759
Abdominal injury	0.632 (0.055, 1.316)	0.038	0.292	0.019
Limb/pelvic injury	0.280 (-0.359, 0.911)	0.013	0.301	0.330
HR	0.003 (-0.008, 0.014)	0.000	0.006	0.618
RR max	0.050 (0.021, 0.090)	0.002	0.016	0.002

Figure 7.27: Bootstrapping results for Denver combined test model

0-12 hours

Variable	Coefficient (95% BCa CI)	Bias	Standard Error	p value
Constant	-6.197 (-8.236, -5.141)			
Age	0.008 (-0.025, 0.043)	0.000	0.018	0.563
CV disease	0.320 (-20.378, 2.903)	-0.721	4.501	0.725
Head injury	0.972 (-0.241, 3.275)	0.195	0.678	0.076
Chest injury	0.971 (-0.559, 19.201)	0.244	1.349	0.094
Abdominal injury	0.641 (-0.384, 2.093)	0.075	0.566	0.198
Limb/pelvic injury	0.324 (-0.864, 1.701)	0.063	0.591	0.518
HR	0.005 (-0.028, 0.043)	0.001	0.015	0.722
RR max	0.028 (-0.051, 0.117)	0.004	0.034	0.342
PT	-0.077 (-0.261, 0.040)	-0.026	0.105	0.355
BD	-0.085 (-0.257, 0.044)	-0.010	0.081	0.214
Creatinine	0.017 (-0.005, 0.059)	0.004	0.012	0.091
Vasopressors	1.725 (0.505, 4.169)	0.211	0.595	0.001

0-24 hours

Variable	Coefficient (95% BCa CI)	Bias	Standard Error	p value
Constant	-9.039 (-12.702, -7.381)			
Age	0.019 (-0.004, 0.047)	0.001	0.012	0.080
CV disease	0.334 (-19.429, 2.093)	-0.753	4.296	0.664
Head injury	1.622 (0.680, 3.310)	0.166	0.505	0.001
Chest injury	0.273 (-0.688, 1.534)	0.070	0.480	0.537
Abdominal injury	0.496 (-0.476, 1.496)	0.016	0.448	0.212
Limb/pelvic injury	0.619 (-0.346, 1.857)	0.054	0.499	0.160
HR	0.021 (0.000, 0.048)	0.002	0.011	0.037
RR max	0.027 (-0.022, 0.085)	0.003	0.024	0.198
PT	-0.036 (-0.201, 0.090)	-0.007	0.076	0.571
BD	-0.085 (-0.210, 0.023)	-0.002	0.056	0.091
Creatinine	0.014 (-0.003, 0.034)	0.001	0.008	0.047
Vasopressors	1.389 (0.461, 2.603)	0.113	0.428	0.001

0-48 hours

Variable	Coefficient (95% BCa CI)	Bias	Standard Error	p value
Constant	-8.661 (-11.165, -7.480)			
Age	0.006 (-0.014, 0.026)	0.000	0.010	0.510
CV disease	0.243 (-17.955, 1.414)	-1.108	4.428	0.708
Head injury	1.792 (0.884, 3.248)	0.138	0.452	0.001
Chest injury	0.094 (-0.781, 1.038)	0.022	0.409	0.792
Abdominal injury	0.473 (-0.379, 1.418)	0.053	0.389	0.197
Limb/pelvic injury	0.480 (-0.316, 1.415)	0.023	0.436	0.226
HR	0.003 (-0.012, 0.021)	0.001	0.008	0.652
RR max	0.052 (0.004, 0.112)	0.005	0.021	0.006
PT	0.050 (-0.114, 0.178)	-0.008	0.077	0.483
BD	-0.108 (-0.201, -0.033)	-0.005	0.049	0.016
Creatinine	0.011 (0.000, 0.026)	0.001	0.005	0.028
Vasopressors	1.368 (0.591, 2.393)	0.076	0.396	0.001

Figure 7.28: Bootstrapping results for Denver extended model

7.6.2. SOFA COMBINED AND EXTENDED MODELS

The results of bootstrapping for variables in the SOFA combined test and extended models are given in Figures 7.29 and 7.30. In the test model variables, bias estimates were low, with the exception of respiratory disease that was present only in a small number of patients and was therefore greatly affected by sampling. The presence of head injury, maximum heart rate and maximum respiratory rate were consistently shown to be good predictors of SOFA MOF at all time periods, with confidence intervals for coefficients not crossing zero.

The additional variables in the extended model, with the exception of vasopressor use, had confidence intervals crossing zero. Vasopressor use was consistently associated with SOFA MOF although bias estimates were higher than for other variables.

0-2 hours

Variable	Coefficient (95% BCa CI)	Bias	Standard Error	p value
Constant	-3.974 (-5.688, -2.541)			
Age	0.030 (0.012, 0.055)	0.001	0.009	0.002
Respiratory disease	1.036 (-0.147, 2.967)	0.175	1.439	0.067
Head injury	1.073 (0.455, 1.926)	0.037	0.334	0.002
Abdominal injury	0.485 (-0.137, 1.052)	-0.008	0.324	0.111
Trapped	0.573 (-0.053, 1.365)	0.038	0.314	0.051
SBP	-0.008 (-0.016, 0.000)	0.000	0.004	0.059
HR	0.019 (0.006, 0.034)	0.001	0.006	0.001
RR max	0.036 (0.003, 0.081)	0.001	0.019	0.046

0-4 hours

Variable	Coefficient (95% BCa CI)	Bias	Standard Error	p value
Constant	-3.596 (-5.625, -1.979)			
Age	0.031 (0.016, 0.050)	0.001	0.008	0.001
Respiratory disease	1.039 (-0.267, 2.972)	0.113	1.260	0.042
Head injury	1.003 (0.409, 1.745)	0.021	0.321	0.002
Abdominal injury	0.564 (-0.052, 1.199)	0.019	0.304	0.051
Trapped	0.514 (-0.093, 1.190)	0.029	0.299	0.070
SBP	-0.007 (-0.016, 0.000)	0.000	0.004	0.061
HR	0.018 (0.006, 0.030)	0.000	0.006	0.003
RR max	0.020 (-0.014, 0.065)	0.002	0.018	0.231

0-6 hours

Variable	Coefficient (95% BCa CI)	Bias	Standard Error	p value
Constant	-3.189 (-5.228, -1.660)			
Age	0.027 (0.011, 0.045)	0.001	0.008	0.001
Respiratory disease	0.899 (-0.198, 2.350)	0.087	0.575	0.076
Head injury	0.919 (0.404, 1.596)	0.028	0.292	0.001
Abdominal injury	0.569 (-0.027, 1.237)	0.009	0.302	0.051
Trapped	0.518 (-0.155, 1.201)	0.041	0.286	0.059
SBP	-0.007 (-0.017, 0.002)	0.000	0.005	0.104
HR	0.015 (0.004, 0.029)	0.001	0.005	0.003
RR max	0.020 (-0.008, 0.055)	0.001	0.016	0.187

0-12 hours

Variable	Coefficient (95% BCa CI)	Bias	Standard Error	p value
Constant	-3.668 (-5.660, -2.018)			
Age	0.023 (0.011, 0.041)	0.001	0.007	0.001
Respiratory disease	0.605 (-0.224, 1.532)	0.022	0.426	0.131
Head injury	0.825 (0.327, 1.415)	0.022	0.261	0.002
Abdominal injury	0.369 (-0.159, 0.920)	0.013	0.258	0.135
Trapped	0.530 (0.021, 1.088)	0.016	0.253	0.032
SBP	-0.005 (-0.015, 0.004)	0.000	0.004	0.273
HR	0.016 (0.004, 0.029)	0.001	0.006	0.005
RR max	0.032 (0.008, 0.066)	0.001	0.014	0.018

0-24 hours

Variable	Coefficient (95% BCa CI)	Bias	Standard Error	p value
Constant	-3.721 (-5.961, -1.930)			
Age	0.025 (0.013, 0.040)	0.001	0.006	0.001
Respiratory disease	0.432 (-0.237, 1.298)	0.021	0.379	0.242
Head injury	0.830 (0.309, 1.442)	0.029	0.247	0.001
Abdominal injury	0.321 (-0.111, 0.818)	0.004	0.246	0.178
Trapped	0.571 (0.080, 1.080)	0.008	0.249	0.018
SBP	-0.006 (-0.017, 0.003)	0.000	0.005	0.226
HR	0.017 (0.006, 0.027)	0.000	0.005	0.002
RR max	0.028 (0.002, 0.058)	0.001	0.013	0.026

0-48 hours

Variable	Coefficient (95% BCa CI)	Bias	Standard Error	p value
Constant	-3.435 (-5.481, -1.796)			
Age	0.023 (0.012, 0.039)	0.001	0.006	0.001
Respiratory disease	0.583 (-0.077, 1.352)	0.006	0.375	0.100
Head injury	0.826 (0.332, 1.393)	0.022	0.240	0.001
Abdominal injury	0.366 (-0.106, 0.872)	0.014	0.234	0.114
Trapped	0.534 (0.069, 1.046)	0.013	0.234	0.024
SBP	-0.006 (-0.015, 0.003)	0.000	0.005	0.223
HR	0.012 (0.002, 0.023)	0.000	0.005	0.012
RR max	0.033 (0.007, 0.063)	0.001	0.013	0.007

Figure 7.29: Bootstrapping results for SOFA combined test model

0-12 hours

Variable	Coefficient (95% BCa CI)	Bias	Standard Error	p value
Constant	-2.743 (-7.085, 0.179)			
Age	0.017 (-0.013, 0.060)	0.003	0.015	0.183
Respiratory disease	1.026 (-1.218, 21.504)	1.037	4.279	0.153
Head injury	1.434 (0.068, 3.920)	0.209	0.656	0.007
Abdominal injury	0.150 (-0.914, 1.292)	-0.004	0.556	0.751
Trapped	0.637 (-0.487, 1.964)	0.075	0.525	0.151
SBP	0.009 (-0.015, 0.030)	0.000	0.010	0.282
HR	0.013 (-0.011, 0.052)	0.004	0.012	0.179
RR max	0.033 (-0.019, 0.117)	0.005	0.028	0.175
FiO2				
41-60%	-0.472 (-1.820, 0.735)	-0.049	0.695	0.414
61-80%	0.566 (-2.662, 21.014)	0.344	2.644	0.499
> 80%	0.500 (-0.904, 2.074)	0.062	0.678	0.404
Albumin	-0.080 (-0.150, -0.039)	-0.010	0.037	0.007
Transfusion level				
Transfusion	-0.048 (-1.239, 1.123)	-0.013	0.660	0.939
Massive transfusion	1.599 (-0.681, 2.806)	5.972	9.160	0.058
Vasopressors	0.951 (-0.088, 2.806)	0.148	0.528	0.030

0-24 hours

Variable	Coefficient (95% BCa CI)	Bias	Standard Error	p value
Constant	-2.390 (-6.840, 1.237)			
Age	0.015 (-0.009, 0.043)	0.001	0.012	0.188
Respiratory disease	0.357 (-1.087, 2.059)	0.050	0.719	0.551
Head injury	1.086 (0.175, 2.498)	0.116	0.436	0.010
Abdominal injury	0.056 (-0.731, 0.750)	-0.040	0.413	0.871
Trapped	0.473 (-0.365, 1.460)	0.058	0.420	0.219
SBP	0.004 (-0.018, 0.022)	0.000	0.010	0.596
HR	0.016 (-0.004, 0.043)	0.002	0.010	0.074
RR max	0.015 (-0.027, 0.062)	0.002	0.020	0.410
FiO2				
41-60%	-0.137 (-1.270, 0.911)	-0.027	0.594	0.788
61-80%	0.668 (-1.798, 3.997)	0.120	1.261	0.414
> 80%	0.782 (-0.397, 2.143)	0.066	0.593	0.157
Albumin	-0.073 (-0.119, -0.043)	-0.007	0.026	0.001
Transfusion level				
Transfusion	-0.228 (-1.066, 0.566)	-0.022	0.440	0.577
Massive transfusion	0.755 (-1.473, 20.956)	0.885	4.072	0.313
Vasopressors	1.287 (0.349, 2.894)	0.127	0.424	0.001

0-48 hours

Variable	Coefficient (95% BCa CI)	Bias	Standard Error	p value
Constant	-0.784 (-4.627, 2.610)			
Age	0.014 (-0.007, 0.037)	0.001	0.010	0.123
Respiratory disease	0.611 (-0.589, 2.252)	0.073	0.694	0.292
Head injury	0.996 (0.080, 2.220)	0.112	0.406	0.005
Abdominal injury	-0.285 (-0.983, 0.380)	-0.041	0.380	0.426
Trapped	0.406 (-0.495, 1.373)	0.049	0.402	0.262
SBP	-0.001 (-0.018, 0.016)	0.000	0.008	0.866
HR	0.018 (-0.001, 0.040)	0.002	0.009	0.032
RR max	0.004 (-0.034, 0.047)	0.001	0.020	0.811
FiO2				
41-60%	-0.524 (-1.654, 0.456)	-0.038	0.583	0.310
61-80%	-0.185 (-2.100, 1.887)	0.106	0.823	0.795
> 80%	0.698 (-0.348, 1.972)	0.038	0.502	0.128
Albumin	-0.105 (-0.155, -0.079)	-0.010	0.028	0.001
Transfusion level				
Transfusion	-0.161 (-0.959, 0.669)	0.002	0.417	0.680
Massive transfusion	0.495 (-2.343, 20.678)	1.209	4.657	0.543
Vasopressors	1.352 (0.515, 2.571)	0.084	0.414	0.001

Figure 7.30: Bootstrapping results for SOFA extended model

7.6.3. MARSHALL COMBINED AND EXTENDED MODELS

The results of bootstrapping for individual variables contained in the Marshall combined test and extended models are given in Figures 7.31 and 7.32. Age, head injury, abdominal injury and limb/pelvic injury were all associated with Marshall MOF at all time-points in the combined test model, with robust confidence intervals for the coefficient and low levels of bias. However, more variability in parameter estimates was seen in the extended model. CV disease had wide confidence intervals and high levels of bias, suggesting a lack of reliability in the estimates for this variable in the original model.

0-2 hours

Variable	Coefficient (95% BCa CI)	Bias	Standard Error	p value
Constant	-4.536 (-6.215, -3.262)			
Age	0.019 (0.006, 0.036)	0.001	0.007	0.006
Sex	0.547 (-0.010, 1.190)	0.023	0.290	0.053
CV disease	1.201 (-0.004, 2.571)	0.038	0.593	0.024
Head injury	1.290 (0.728, 2.058)	0.047	0.291	0.001
Abdominal injury	1.034 (0.511, 1.684)	0.029	0.275	0.001
Limb/pelvic injury	0.764 (0.171, 1.421)	0.014	0.295	0.007
HR	0.010 (0.001, 0.021)	0.000	0.005	0.029

0-4 hours

Variable	Coefficient (95% BCa CI)	Bias	Standard Error	p value
Constant	-4.396 (-5.994, -3.255)			
Age	0.019 (0.005, 0.034)	0.001	0.007	0.003
Sex	0.432 (-0.194, 1.048)	0.014	0.280	0.107
CV disease	1.120 (0.019, 2.749)	0.077	0.557	0.023
Head injury	1.214 (0.656, 1.984)	0.041	0.282	0.001
Abdominal injury	1.077 (0.538, 1.675)	0.022	0.265	0.001
Limb/pelvic injury	0.787 (0.212, 1.434)	0.015	0.288	0.004
HR	0.010 (0.000, 0.021)	0.001	0.005	0.034

0-6 hours

Variable	Coefficient (95% BCa CI)	Bias	Standard Error	p value
Constant	-4.358 (-5.988, -3.071)			
Age	0.019 (0.004, 0.035)	0.001	0.007	0.002
Sex	0.510 (-0.034, 1.133)	0.024	0.287	0.019
CV disease	1.181 (0.060, 2.538)	0.041	0.559	0.001
Head injury	1.166 (0.620, 1.838)	0.036	0.271	0.001
Abdominal injury	1.080 (0.546, 1.645)	0.023	0.251	0.004
Limb/pelvic injury	0.773 (0.255, 1.323)	-0.002	0.271	0.059
HR	0.009 (0.000, 0.020)	0.000	0.005	0.001

0-12 hours

Variable	Coefficient (95% BCa CI)	Bias	Standard Error	p value
Constant	-3.851 (-5.509, -2.587)			
Age	0.015 (0.003, 0.028)	0.000	0.007	0.018
Sex	0.586 (0.051, 1.214)	0.022	0.272	0.032
CV disease	1.207 (0.191, 2.331)	0.013	0.518	0.010
Head injury	1.124 (0.619, 1.746)	0.032	0.248	0.001
Abdominal injury	0.892 (0.415, 1.380)	0.010	0.236	0.001
Limb/pelvic injury	0.695 (0.209, 1.290)	0.018	0.262	0.011
HR	0.007 (-0.004, 0.018)	0.000	0.005	0.161

0-24 hours

Variable	Coefficient (95% BCa CI)	Bias	Standard Error	p value
Constant	-4.075 (-5.762, -2.668)			
Age	0.017 (0.005, 0.030)	0.000	0.006	0.006
Sex	0.574 (0.022, 1.131)	0.010	0.271	0.028
CV disease	1.075 (0.150, 2.184)	0.013	0.506	0.021
Head injury	1.187 (0.697, 1.722)	0.016	0.245	0.001
Abdominal injury	0.787 (0.310, 1.299)	0.019	0.232	0.002
Limb/pelvic injury	0.658 (0.160, 1.200)	0.018	0.254	0.007
HR	0.008 (-0.001, 0.020)	0.000	0.005	0.106

0-48 hours

Variable	Coefficient (95% BCa CI)	Bias	Standard Error	p value
Constant	-3.798 (-5.285, -2.671)			
Age	0.015 (0.003, 0.030)	0.001	0.006	0.009
Sex	0.620 (0.158, 1.144)	0.008	0.258	0.014
CV disease	0.972 (-0.031, 2.028)	0.016	0.486	0.028
Head injury	1.183 (0.716, 1.753)	0.028	0.237	0.001
Abdominal injury	0.752 (0.313, 1.256)	0.015	0.245	0.002
Limb/pelvic injury	0.586 (0.134, 1.092)	0.003	0.240	0.014
HR	0.007 (-0.004, 0.019)	0.001	0.005	0.152

Figure 7.31: Bootstrapping results for Marshall combined test model

0-12 hours

Variable	Coefficient (95% BCa CI)	Bias	Standard Error	p value
Constant	-5.742 (-8.631, -4.532)			
Age	0.023 (-0.002, 0.058)	0.003	0.013	0.042
Sex	-0.128 (-1.190, 1.011)	0.040	0.505	0.753
CV disease	-1.826 (-21.109, -0.584)	-2.426	7.008	0.016
Head injury	1.930 (0.811, 4.483)	0.208	0.624	0.001
Abdominal injury	1.395 (0.445, 2.961)	0.112	0.497	0.001
Limb/pelvic injury	0.717 (-0.293, 1.932)	0.066	0.482	0.088
HR	0.010 (-0.12, 0.036)	0.001	0.011	0.300
FiO2				
41-60%	-0.726 (-1.911, 0.552)	-0.010	0.647	0.214
61-80%	0.768 (-1.087, 2.084)	0.176	1.223	0.285
> 80%	0.528 (-0.685, 2.179)	0.087	0.625	0.328
APTT	0.009 (-0.008, 0.075)	0.003	0.012	0.078
Lactate	0.234 (0.025, 0.620)	0.034	0.125	0.016
Vasopressors	0.708 (-0.220, 1.732)	0.033	0.481	0.095

0-24 hours

Variable	Coefficient (95% BCa CI)	Bias	Standard Error	p value
Constant	-5.593 (-8.058, -4.335)			
Age	0.020 (-0.003, 0.050)	0.002	0.012	0.055
Sex	0.110 (-0.765, 1.163)	0.047	0.425	0.785
CV disease	-2.174 (-21.366, -0.875)	-2.347	6.349	0.006
Head injury	1.935 (0.848, 3.492)	0.161	0.485	0.001
Abdominal injury	1.166 (0.267, 2.279)	0.077	0.426	0.002
Limb/pelvic injury	0.740 (0.019, 1.572)	0.009	0.416	0.047
HR	0.008 (-0.010, 0.030)	0.001	0.009	0.360
FiO2				
41-60%	-0.962 (-2.239, 0.203)	-0.010	0.615	0.080
61-80%	0.444 (-0.932, 1.991)	0.065	0.681	0.459
> 80%	0.448 (-0.826, 1.855)	0.052	0.592	0.394
APTT	0.013 (0.002, 0.076)	0.004	0.012	0.019
Lactate	0.231 (0.038, 0.602)	0.022	0.112	0.015
Vasopressors	0.768 (-0.013, 1.702)	0.048	0.385	0.032

0-48 hours

Variable	Coefficient (95% BCa CI)	Bias	Standard Error	p value
Constant	-5.406 (-7.600, -4.018)			
Age	0.019 (-0.001, 0.042)	0.001	0.010	0.043
Sex	0.450 (-0.366, 1.468)	0.072	0.392	0.207
CV disease	-1.046 (-8.083, 1.058)	-0.988	4.731	0.304
Head injury	1.472 (0.609, 2.635)	0.093	0.427	0.001
Abdominal injury	0.857 (0.102, 1.778)	0.049	0.378	0.012
Limb/pelvic injury	0.556 (-0.141, 1.400)	0.021	0.369	0.106
HR	0.010 (-0.005, 0.027)	0.000	0.008	0.189
FiO2				
41-60%	-1.253 (-2.540, -0.179)	-0.031	0.602	0.023
61-80%	-0.719 (-2.041, 0.520)	-0.030	0.680	0.241
> 80%	0.229 (-0.913, 1.512)	0.061	0.527	0.637
APTT	0.011 (0.001, 0.034)	0.002	0.007	0.019
Lactate	0.226 (0.072, 0.512)	0.021	0.102	0.010
Vasopressors	1.043 (0.319, 1.919)	0.054	0.346	0.002

Figure 7.32: Bootstrapping results for Marshall extended model

7.7. SUMMARY OF RESULTS

The study population of 517 patients was predominantly composed of young, severely injured males. Overall survival was 84.9% at 100 days post-injury, with 26 patients dying before 48 hours and the possible onset of MOF. Of the 491 patients at risk of MOF, incidence of MOF ranged from 22.8% (Denver score) to 58.5% (SOFA score). The mean onset of MOF was between 4.7 (SOFA) and 5.6 (Denver) days post-injury, with a mean duration of 4.4 (Denver) to 6.5 (SOFA) days. The respiratory and neurological systems were the most commonly affected. MOF was associated with increased injury severity, regardless of the score used to define MOF. Patients with MOF had a later mean time to ICU admission, and increased ICU and hospital stays.

A large number of variables were associated with MOF on univariate analysis. Models were built to find combinations of variables that could most accurately predict MOF at 2, 4, 6, 12, 24 and 48 hours post-injury. The variables used to predict MOF differed depending on whether Denver, SOFA or Marshall definitions of MOF were used as the outcome measure. Models for the three MOF definitions also differed in predictive value, calibration and discrimination.

For all three of the Denver, SOFA and Marshall MOF definitions, common variable sets were identified that could be used to predict MOF at a range of times post-injury. Addition of further variables to the models was required to improve accuracy of prediction from 12 to 48 hours after injury. Comparison between patients included in the models, and those excluded due to missing data showed relatively few consistent

differences, with the exception of an increase in injury severity for the patients included in the SOFA and Marshall extended models. Internal validation demonstrated that levels of bias varied, with some variables showing consistent association with MOF, and others showing wide confidence intervals on bootstrapping. The potential uses and implications of these variable combinations will be discussed in Chapter 8.

CHAPTER 8 DISCUSSION

8.1. PREDICTION OF MOF FOLLOWING MAJOR TRAUMA

Major trauma remains a significant global health problem, causing substantial mortality and morbidity. MOF is associated with a high proportion of late trauma deaths, and with poor long-term health outcomes in survivors. The incidence, aetiology and pathophysiology of MOF have been widely debated, and whilst areas of consensus have formed, no single model of MOF development has been accepted.

Literature review showed a number of variables to be associated with post-trauma MOF in previous studies. Attempts to identify combinations of variables predictive of MOF development have been made for cohorts from the USA and Australia, but no equivalent UK studies have been performed. Despite these attempts, in clinical practice, assessment of a patient's risk of MOF is based predominantly on the knowledge and judgement of individual doctors.

As outlined in Section 1.5, this study aimed to assess if a predictive model could be developed to predict post-trauma MOF, using routine clinical data collected within the first 48 hours following injury. Secondary objectives included describing the current composition of a major trauma cohort in the UK, an assessment of the use of routinely collected data in trauma research, and the identification of potentially modifiable risk factors for MOF.

8.2. DESCRIBING THE CURRENT TRAUMA POPULATION

The study period (2003-2011) preceded the official designation of the John Radcliffe Hospital as a MTC in April 2012. However, during the study period, tertiary referrals were taken from across the Thames Valley region, reflecting a similar pattern to that formalised with the official MTC designation. Whilst trauma patterns in the Oxford region may vary from highly urbanised centres, the study population gives insight into the current composition of severe trauma in the UK.

In 2005 TARN analysed data for blunt trauma patients of all severities, providing a useful comparator for the study population since penetrating trauma was uncommon.[261] Similar to the TARN study, major trauma patients were predominantly male, with a high incidence of head, chest and lower limb or pelvic injuries. With a median ISS of 25, and median NISS of 34, the John Radcliffe cohort represented a severely injured group of patients. TARN divided patients into groups for analysis based on ISS, with the groups with an ISS of 17-25 and 26-75 most closely reflecting the John Radcliffe study cohort. For both of these TARN groups, RTCs were the predominant cause of injury, with an incidence rising to 63% in the most severely injured group, similar to the 69% observed in this study.

For the more severely injured groups, TARN observed an in-hospital mortality of 14.8% for ISS 17-25 patients and 32.3% for ISS 26-70. Mortality for the John Radcliffe group, which spanned these ISS ranges was comparatively low, with 15.1% patients dying within

100 days of injury. As previously discussed (Section 1.2.1.), the trimodal distribution of post-trauma deaths has been increasingly debated.[39,41] Whilst a clustering of deaths was observed in the first 48 hours after injury, no clear third peak of deaths was observed in the study population. Patients who died were more likely to be female, older and more severely injured.

Almost a quarter of the trauma patients admitted to the John Radcliffe were transferred in from surrounding centres. The majority of transfers were received early after injury, with hospital arrival times reflecting those seen in the TARN study.[261] Similarly, AICU admission occurred early in the bulk of the cohort, with nearly 90% admitted within the first two days after injury. However, in the TARN study, only 61.2% of patients with ISS 17-25 and 79.1% with ISS 26-75 were admitted to an ICU, which was felt to reflect a lack of critical care capacity. Since AICU admission was an inclusion criteria for this study, no comparison with those not admitted was possible.

Length of AICU stay reflected the findings of the TARN study with a median stay of 4 days. Hospital stays were prolonged, with a small but significant number of patients having extended stays of over 100 days. Operative rates were higher than seen in the TARN data, with over 75% patients requiring surgical intervention. Using NHS reference costs from 2004, TARN estimated the mean cost of treatment for a patient with ISS > 15 to be £16,776.[261] Costs were higher in survivors, and in the more severely injured. It is clear that this patient group exert a significant financial burden on the NHS.

The incidence of MOF in the study population ranged from 23 to 59%, depending on the MOF scoring system used. MOF was associated with mortality, but up to 48% of deaths that occurred later than 48 hours after injury were seen in patients who did not develop

MOF. Late deaths from patients who did not develop MOF occurred at a mean of 7 days from injury, in comparison to a mean of 16.9 days for patients with SOFA or Marshall MOF and 19.1 days for patients with Denver MOF. It is therefore arguable that some of the group of patients who died without developing MOF may not have survived sufficiently long for MOF to occur.

In patients who did develop MOF, mortality ranged from 13 to 24%, depending on the MOF definition, which was consistent with previous studies.[93,208] MOF occurred early, with over 90% cases seen within 10 days of injury. The mean duration of MOF was between 4 and 6 days, depending on MOF definition, but a small number of patients had a protracted course, with a duration measured in weeks.

Patterns of organ dysfunction showed that respiratory, neurological and cardiac failure were most prevalent. Coagulopathy and neurological dysfunction occurred early, commonly preceding MOF development; respiratory and cardiac dysfunction occurred around the time of MOF onset, and renal and hepatic dysfunction occurred late. Coagulopathy and cardiac failure were commonly short-lived in duration, with neurological, renal and respiratory dysfunction taking the longest time to resolve. Ciesla et al examined the role of respiratory dysfunction in post-trauma MOF and found it occurred slightly earlier than cardiac failure, with renal and hepatic failure occurring late.[102] Respiratory failure was found to be the greatest contributor to the MOF Denver score, and was felt to be a major driver of subsequent sequential organ failure. Similarly, Durham et al found respiratory failure occurred first in the majority of patients and was responsible for all single organ failure.[103] Neither of these studies examined the role of

neurological dysfunction, and haematological failure was defined in a different manner to that in the SOFA or Marshall scores.

Patients with MOF were more likely to be more severely injured, require operative intervention, and to have longer ICU and hospital stays, regardless of the definition of MOF. Associations were also seen with age (SOFA and Marshall MOF), and the need for transfusion (Denver MOF). Of note, the timing of ICU admission in relation to the time of injury was found to be significantly later in the SOFA and Marshall MOF groups, with a trend to significance in the Denver MOF groups ($p = 0.057$). This may reflect the current inaccuracy of clinical prediction of patient risk of MOF in the early stages following injury.

8.3. Defining MOF After Major Trauma

The incidence of post-trauma MOF varied widely depending on the scoring system used for MOF definition. A core group of patients (around 20%) were defined as having MOF by all three of the Denver, SOFA and Marshall scores. Conversely, around 40% patients did not have MOF by any of the three scores. For the remaining 40% patients, diagnosis of MOF was dependent on the score used. Sauaia et al reported incidences of 22.2% for Denver MOF and 49.7% for Marshall MOF,[208] similar to those seen in this study of 22.8% and 40.5% respectively. Ciesla et al reported a Denver MOF incidence of 25%.[93] However, a number of more recent studies have reported lower incidences, with Minei et al showing a Marshall MOF incidence of 29%,[172] and Dewar et al showing a Denver MOF incidence of 15%.[173] The exact population inclusion criteria are likely to greatly affect the precise MOF incidence.

Whilst MOF defined by any of the three scores was significantly associated with mortality, differences were seen in the relationships between MOF definition and deaths. The cumulative survival probability was worst in those patients defined with MOF by the Denver score. The SOFA and Marshall scores appeared much less sensitive for mortality in a trauma setting. Whilst all three scores were originally validated against mortality, this may reflect the nature of the validation populations, with the Denver score developed on trauma patients, and the SOFA and Marshall scores developed on general ICU populations. Sauaia et al reported a higher case-fatality rate with the Denver score, but that both the Denver and Marshall scores were highly predictive for mortality,[208] which tallies with the results seen in this study.

Analysis of the patient cohorts given by each of the Denver, SOFA and Marshall MOF scores in this study shows variations in the populations described. Differences were observed in timing of MOF onset, duration and in the organ systems affected. These population differences were additionally reflected in the variables found to be significant in univariate and multivariate analyses, and in the variables included in the final models. Areas of consensus were seen, but it was clear that the definition of MOF affected the clinical variables found to be significantly associated with MOF.

What are the implications of these differences? It is clear that a gold-standard definition of MOF would greatly improve consistency in both research and clinical settings, and allow meaningful comparison of studies and populations from different institutions. The low mortality seen in this study supports the position that more discriminatory measures of clinical outcome are required in the early stages after trauma. MOF has the potential

to be a useful outcome measure in such circumstances, but a unifying definition must be adopted.

Discussion of the most appropriate scoring system for post-trauma MOF must consider the results from model development, and will be discussed further in the following sections.

8.4. PREDICTING MOF

Model selection was based on a combination of goodness-of-fit, calibration, population size and the number of included variables. This strategy resulted in different variable combinations being selected at different time-points after injury. These variable lists were then examined to look for consensus variables that could be used to form a model that could be used at multiple time-points. This process creates an inevitable trade-off between model accuracy and clinical utility.

Overall, predictive models for Denver MOF had the highest specificity and lowest PPVs. Predictive models for SOFA MOF conversely had the lowest specificity and highest PPVs, with the lowest PAC values of any models. Models for Marshall MOF fell between the two in terms of category predictor values, although AUC values were comparable across models for all three scoring systems. Similarities were seen in the variables identified in the combined models, with age, head injury, abdominal injury and maximum heart rate being included in all. In the variables added from 12-48 hours to improve model accuracy, vasopressor use was a strong predictor of all definitions of MOF. However, the remaining

variables differed depending on the MOF definition, emphasising the likelihood that different MOF scores describe different patient populations.

Comparison with models developed by other researchers is limited by the available reported data given in publications of previous studies. Sauaia et al developed Denver MOF prediction models based on additive models at a series of time-points – admission to hospital, 12 hours, 24 hours and 48 hours after injury. The admission model of ISS, red blood cell count (RBC), platelet count and age had a sensitivity of 54% and a specificity of 78%. Addition of inotrope use at 12 hours improved sensitivity to 63% and specificity to 84%. Further addition of lactate levels at 24 hours and APACHE III score at 48 hours resulted in an increase in sensitivity to 71% at both times, but with no increase in specificity. AUC values were not quoted by the authors, but the ROC curves showed a significant increase in predictive value from admission to 12 hours, with minimal further changes at later times after injury.[97]

Initial models of Denver MOF prediction were less sensitive than the values quoted by Sauaia et al,[97] with a maximum sensitivity of 56.3% at 12 hours post-injury. However, the overall PAC at this time-point was high at 88.7%, with a specificity of 96.3% which far exceeded the values seen in the Sauaia study. However, the model at 12 hours involved the highest number of variables and the smallest patient population of any of the Denver models. Reduction to the eight variables of the Denver test set resulted in extremely low sensitivity values after 12 hours. Modelling using the extended test set increased the sensitivity, PPV and NPV, with a small associated loss of specificity.

Dewar et al reported a predictive model for Denver MOF at hospital admission using only patient age and platelet count with an AUC of 0.71. The addition of maximum creatinine

and minimum bilirubin levels at 24 hours increased the AUC to 0.84.[173] Goodness-of-fit measures and other category predictors were not reported. However, Dewar et al utilised a dichotomous approach to variables, with scoring points awarded if the threshold for a variable was exceeded. This differs from the approach used in this study and makes positive identification of patients with MOF highly likely, since the sickest patients will be the ones who reach the defined thresholds.

Additionally, it was noted by Dewar et al [96] that the significant variables in their models differed from those identified in the previous study by Sauaia et al.[97] Identified variables in this study supported elements of both, and highlighted some additional variables not used in either previous study. From the variables used by Dewar et al, age and creatinine levels were included in the final models. From the variables used by Sauaia et al, age, measures of injury severity and vasopressor use were included. However, a number of variables that significantly contributed to the models developed were not identified in previous studies, including pre-existing cardiovascular disease, maximum heart rate, maximum respiratory rate, PT and BD. Notably, platelet count was included in both Sauaia et al and Dewar et al's models, but was not found to significantly contribute to a unified model in this study.

The Inflammation and the Host Response to Injury study by Minei et al attempted to derive a predictive model for Marshall MOF after trauma.[172] The exact timing of the model was not clear, although it included variables collected up to 24 hours after injury. Goodness-of-fit and category predictors for the model were not published, and therefore comparison is only possible with the variables identified. Similar to this study, Minei et al identified sex, pre-existing cardiovascular disease, limb injury, vasopressor use and

lactate levels as independent predictors of MOF defined by the Marshall score. However, instead of age, head injury, abdominal injury, heart rate, FiO_2 and APTT levels identified in the extended test set in this study, Minei et al found pre-existing hepatic disease, chest injury, crystalloid infusion volume, FFP and BD levels to predict Marshall MOF.[172]

Variables shown to be associated with MOF in this study included both those fixed at the time of the traumatic event (patient demographics and injury patterns) and those that can be potentially altered by clinical interventions (laboratory and physiological parameters). A strong association between vasopressor use and MOF was seen universally across all combined models from 12 hours onwards. However, the relationship between MOF and parameters that might have been used by clinicians to determine a need for vasopressor treatment, such as hypotension and evidence of hypoperfusion, were less clear-cut. Whilst patients receiving vasopressors are clearly suffering from cardiovascular dysfunction, this suggests that there may be an additional iatrogenic element to MOF development when vasopressors are used.

Additionally of interest from a surgical perspective, was the lack of association between operative intervention and MOF. Surgery has been considered an important "second hit" in the two-hit hypothesis of MOF development.[37] However, no clear evidence was found from this study that the timing of surgical intervention contributed to MOF development.

8.5. STUDY LIMITATIONS

The limitations of using retrospective data are well-documented.[262,263] Despite these problems, in certain circumstances, observational routinely-collected data sources can

have significant value. In February 2012, the Equator (Enhancing the Quality and Transparency of Health Research) Network, produced the RECORD (Reporting of studies conducted using observational routinely-collected data) statement. This summary document declared an intention to develop reporting guidelines for observational studies using health data collected for non-research purposes.[264] These guidelines are intended to be a development of the STROBE (Strengthening the reporting of observational studies in epidemiology) statement,[265] to assess the specific problems with using data collected without a specific research question in mind.

Without specific guidelines currently in place, similar studies in other medical areas and the STROBE guidelines were consulted in the design and reporting of this project. The STROBE guidelines emphasise the need for clear definition of all variables, including the data sources, measurement, definitions and categorisation. In this study, an extensive data dictionary (Appendix 8) was created to define terms and collection rules for all variables used in analysis.

There is an inherent risk of data bias in using pre-existing data collected for a primary purpose other than to answer the chosen research question. Part of this bias exists in data availability, since only those variables found in the data source can be used for analysis. The use of multiple data sources increased the range of available variables. A wide-ranging approach was taken to input variable identification, but the inclusion of an excess number of variables can lead to statistical inaccuracy, and a sensible approach to variable reduction had to be employed.

A similarly pragmatic approach had to be utilised for missing data, since the use of retrospective data increases the likelihood of data gaps. Attempts were made to reduce

the extent of missing data by using multiple data sources, and by specific attempts to obtain additional data from hand searches. However, a certain amount of missing data is an inevitable consequence of the study design.

Additionally, this missing data is unlikely to occur at random, since certain clinical variables are inter-related, and specific patient groups may have increased likelihood to have certain variables measured in clinical practice. These factors were considered in the formation of the data dictionary and in the data definitions for each variable, but a high proportion of missing data remained inevitable. Arguments for imputation of missing variables were complicated by this lack of a random pattern, and by the fact that the study population contained predominantly patients with abnormal clinical parameters, making an estimation of "normal" in this group difficult. Analysis of missing data in the Canadian National Trauma Data Bank led to the development of a complex multiple imputation model to complete data gaps.[266,267] However, this model was found to apply only to physiological data, and to rely on underlying assumptions of data normality and the randomness of missing data. The development of such a model relies on access to a large amount of pre-existing data, which was not available in these circumstances.

Given the difficulties outlined, a pragmatic approach was taken to missing data. Models were developed only on those patients with data available for all variables. Whilst aware that this greatly reduced that number of patients in the model-building population, and therefore the number of variables that could be considered, this was felt to be the most practical response to prevent errors involved in missing value imputation. The number of patients on which each model was based was one of the considerations in the selection of the best-fitting models.

Comparison was made between the patients included in the final models, and those excluded due to missing data. This showed that only a few differences in demographics and clinical parameters were seen, and that the models could be generally considered to have been developed on a sample representative of the overall trauma population. A notable exception to this was the differences in injury severity in the SOFA and Marshall extended models, with the patients included being more severely injured than those excluded due to missing data. This has the potential to increase overfitting of the models, by including a group at higher risk of MOF than the overall population.

The selection of the most parsimonious models was based on a range of considerations including measures of model performance, the number of variables and its relation to the number of patients in the development population, and the potential clinical utility of the variables contained in the model. A critical consideration when examining category predictors of the models was the potential clinical use. Early identification of patients at risk of MOF would, in clinical practice, lead to early interventions to support such patients and reduce the risk of MOF development. Since such interventions are likely to have high costs and risks associated, models with PPV and high specificity would have greatest clinical utility to target patients most effectively. However, maximising these parameters involved an inevitable trade-off with other measures of model performance.

To assess the performance of any prognostic model, validation of the model is required.[268,269] Due to the reduction in population size when patients with missing data were excluded from model development, the division of the data into development and validation sets was not feasible. Internal validation using bootstrapping is an accepted method of validation in the absence of a separate test set, and has been used in other

similar studies.[97,268] Bootstrapping was performed on the combined variable sets only, to assess the performance of a single model across different time points. Since by definition, these were not the best-fitting models at each time-point, variability in the bias estimates for each variable would be expected. In general, bias estimates were low for individual variables, but the bootstrapping emphasised that some variables were consistently associated with MOF for models at all time-points, and some showed wide confidence intervals, indicating a less clear-cut association. To fully assess the performance of a prognostic model, external validation is required on a dataset entirely separate to the original. Models may perform less well on external datasets due to differences in the population composition or deficiencies in the original model.[268] However, even a model with reduced calibration and discrimination on external validation may have clinical utility, based on context.[270] External validation on an additional population was outside the scope of this thesis.

In the absence of external validation, how useful are the results of this study? Hayden et al developed the Quality in Prognosis Studies (QUIPS) tool to provide a framework for the assessment of study biases.[271]This identified six main areas of potential bias – study participation, study attrition, prognostic factor measurement, outcome measurement, confounding, and statistical analysis.

The cohort used in this study was drawn from a single centre only. This impacts not only on the data available for analysis, and the range of variables measured, but also on the potential for data to reflect local practice, and therefore reduce applicability to external populations. TARN data has consistently audited the John Radcliffe as having high survival levels after major trauma, and the population in this study will have received specialist

care. However, with the development of the MTCs, trauma care is now increasingly specialised, and the findings of the study will be most applicable to the group of severely injured patients who require MTC care. Whilst variations will exist in clinical practice, it is likely that the treatment received by the study population would be broadly comparable to that at other MTCs.

Extensive attempts were made to maximise the available data, and to include all possible patients. Unfortunately, a group of patients had to be excluded due to poor record-keeping as to the time of injury. Patients were also lost to model development if they had missing data, as previously discussed. However, the included and excluded populations were broadly similar in terms of demographics and basic population characteristics.

Whilst with this study design, there is an inevitable risk of bias, attempts were made throughout to minimise these risks, with the intention of creating models that were generalizable to other major trauma populations. The implications of this will be further discussed with reference to future uses of the models.

8.6. Recommendations for Future Work

The models created in this study provide a statistical overview of variables found to be associated with different definitions of MOF. At this stage, these results do not have a great deal of utility for the clinician at the bedside, since they provide a statistical model for a population, rather than a clinical score for use in individual patients. To transfer the results of this study into clinical practice, a number of secondary stages are required. To fully ascertain the generalizability of the models, these should be tested and validated on

a separate population. Ideally, external validation should occur on a prospective cohort to reduce the problems with missing data seen in this study. Given the numbers needed for meaningful validation, multi-centre data collection would be required to obtain a sufficient population size.

Prior to further validation and model development, the question of MOF definition needs to be addressed. Whilst the use of multiple MOF scores in this project allowed comparison between the scores, it highlighted the differences seen depending on the MOF definition used. To standardise future trauma research, a single definition of MOF is required to allow meaningful study comparisons. The Denver score was developed on trauma patients, has been used in a range of previous trauma studies, and showed a clear association with trauma mortality. Additionally, it imposes the smallest data collection burden, and is not impacted by GCS, which in trauma patients may be affected for a range of reasons other than MOF. Models based on the Denver score had high specificity, although PPV was lower than with other outcome measures. Whilst there are arguments for use of each of the MOF scoring systems, for the above reasons, the Denver score would seem to be most suitable as the standard post-trauma MOF outcome measure.

For longer-term medical use, development of a clinical scoring system of MOF prediction would be of great value. A more complex statistical model-based system could be used for clinical trial stratification or performance-based audit, but computerised models have limited value in day-to-day clinical practice. The development of a simple bedside scoring system to be used by clinicians would involve a loss of precision when compared to complex models, but would provide a valuable tool for patient assessment. In its current form, the models identified are not yet suitable for use as a clinical tool for bedside use.

Development and validation of such a score involves a range of factors, including the exact nature of the variables included, variable categorisation, ease of data collection, and the human and psychological factors involved in score use and application.[272]

Dewar et al attempted to initiate the process of creating a clinical score from their prediction model of post-trauma MOF by dichotomising variables to ease use in clinical practice. They used classification and regression trees to determine the optimal thresholds for each variable, which were then rounded to fit with clinically acceptable levels.[173] Vogel et al attempted to create a score for MOF prediction in the emergency department, and converted the results of regression analysis to a numerical score by simply rounding the regression coefficients of the component variables to the nearest integer.[170]

A variety of methods have been proposed for the creation of predictive scores, including the use of classification and decision trees with differing methods to classify the nodes or end-points that signify optimal cut-off levels for each variable.[273,274] Extrapolation of population risk to individual patient risk has been shown to create difficulties across a range of medical fields, with problems noted with a lack of precision at individual level, and discrepancies in predicted and observed outcomes.[275,276] Therefore, to ensure a useful model at an individual level, the variable combinations identified in this study would need to be tested and validated on an external population to maximise precision and generalizability, before a model for individual patient use could be created.

In their current form, the models identified could be used for comparison of trauma populations, and to allow standardisation of patient groups for future studies. However, at an individual level, further work would be required to create a bedside score. The ideal

use for such a score would be to be able to stratify an individual patient into risk groups, and then to appropriately target interventions dependent on the level of risk, aiming to reduce the risk of MOF development. One potential intervention identified would be the early transfer of high-risk patients to ICU, given the finding that delay to ICU admission was associated with MOF development.

Despite the lack of suitability for current use at individual level, this does not mean that the developed models do not have immediate clinical use. Knowledge of the variables identified as associated with MOF could greatly assist in clinical assessment, without being used for formal quantification of risk.

8.7. CONCLUSIONS

Major trauma continues to require extensive healthcare provision, resulting in extended ICU and hospital stays, and high surgical intervention rates. Mortality in the severely injured was low, and MOF development is a more discriminatory measure of patient outcome. However, the score used to define MOF affects the population described, and standardisation of the definition of post-trauma MOF is critical for future research and population comparisons. From the three MOF scores compared, the Denver score appears the most suitable for use in measuring post-trauma MOF.

Models created from routine clinical data could predict future MOF development from an early stage after injury. Prediction was possible from 2 hours post-injury, with accuracy improved from 12 hours with the addition of extra variables to the model. Age, head

injury, abdominal injury, maximum heart rate and the use of vasopressors were consistent predictors of all definitions of MOF.

Knowledge of the variables associated with post-trauma MOF assists in clinical awareness of patient risk factors. External validation of the identified models, and analysis of the factors involved in clinical score development could lead to a bedside score to assist clinicians in early trauma management.

.

BIBLIOGRAPHY

1. WHO The global burden of disease: 2004 update . *WHO*.

2. WHO Global Burden of Disease 2008 *WHO*.

3. CDC. CDC - Injury - WISQARS (Web-based Injury Statistics Query and Reporting System). *Inj Prev Control Data Stat*. 2004. Available at: http://www.cdc.gov/injury/wisqars/index.html.

4. Toroyan T. Global status report on road safety. *Inj Prev*. 2009;15(4):286.

5. Martin M, Oh J, Currier H, et al. An analysis of in-hospital deaths at a modern combat support hospital. *J Trauma*. 2009;66(4 Suppl):S51-S60; discussion S60-S61.

6. Robinson PM, O'Meara MJ. The Thomas splint: its origins and use in trauma. *J Bone Joint Surg Br*. 2009;91(4):540-544.

7. Pruitt B a. The symbiosis of combat casualty care and civilian trauma care: 1914-2007. *J Trauma*. 2008;64(2 Suppl):S4-8.

8. Poulos R, Hayen A, Finch C, Zwi A. Area socioeconomic status and childhood injury morbidity in New South Wales, Australia. *Inj Prev*. 2007;13(5):322-327.

9. Naumann RB, Dellinger AM, Zaloshnja E, Lawrence BA, Miller TR. Incidence and total lifetime costs of motor vehicle-related fatal and nonfatal injury by road user type, United States, 2005. *Traffic Inj Prev*. 2010;11(4):353-360.

10. Office for National Statistics. Mortality statistics. Available at: http://www.statistics.gov.uk/hub/health-social-care/health-of-the-population/causes-of-death.

11. Loo BPY, Tsui KL. Factors affecting the likelihood of reporting road crashes resulting in medical treatment to the police. *Inj Prev*. 2007;13(3):186-189.

12. Anil Bhagat, Francis L, Kilbey P, et al. *Reported Road Casualties Great Britain: 2008 Annual Report*.; 2008:201.

13. Department of Health. Saving Lives: Our Healthier Nation. 1999.

14. Department for Transport. *Valuation of Road Accidents. Highway Economics Note.*; 1996.

15. Anderson ID, Woodford M, de Dombal FT, Irving M. Retrospective study of 1000 deaths from injury in England and Wales. *Br Med J (Clin Res Ed)*. 1988;296(6632):1305-1308.

16. West JG, Trunkey DD, Lim RC. Systems of trauma care. A study of two counties. *Arch Surg*. 1979;114(4):455-460.

17. Cales RH. Trauma mortality in Orange County: the effect of implementation of a regional trauma system. *Ann Emerg Med*. 1984;13(1):1-10.

18. West JG, Cales RH, Gazzaniga AB. Impact of regionalization. The Orange County experience. *Arch Surg*. 1983;118(6):740-744.

19. Shackford SR, Hollingworth-Fridlund P, Cooper GF, Eastman AB. The effect of regionalization upon the quality of trauma care as assessed by concurrent audit before and after institution of a trauma system: a preliminary report. *J Trauma*. 1986;26(9):812-820.

20. Nicholl J, Turner J. *Effectiveness of a regional trauma system in reducing mortality from major trauma: before and after study.*; 1997:1349-1354.

21. Oakley PA, MacKenzie G, Templeton J, Cook AL, Kirby RM. Longitudinal trends in trauma mortality and survival in Stoke-on-Trent 1992-1998. *Injury*. 2004;35(4):379-385.

22. Parr MJ, Nolan J. Effectiveness of regional trauma systems: Wrong comparisons were made. *BMJ*. 1998;316:1383-1384.

23. Royal College of Surgeons of England. Better Care for the Severely Injured. 2003.

24. Lecky FE, Woodford M, Bouamra O, Yates DW. Lack of change in trauma care in England and Wales since 1994. *Emerg Med J*. 2002;19(6):520-3.

25. Nathens AB, Maier R V. The relationship between trauma center volume and outcome. *Adv Surg*. 2001;35:61-75.

26. Celso B, Tepas J, Langland-Orban B, et al. A systematic review and meta-analysis comparing outcome of severely injured patients treated in trauma centers following the establishment of trauma systems. *J Trauma*. 2006;60(2):371-378; discussion 378.

27. National Confidential Enquiry into Patient Outcomes and Death. *Trauma: Who cares? A report of the National Confi dential Enquiry into Patient Outcome and Death.* 2007.

28. National Audit Office. *Major trauma care in England*. 2010.

29. NHS Choices. Major trauma services. Available at: http://www.nhs.uk/NHSEngland/AboutNHSservices/Emergencyandurgentcareservices/Pages/Majortraumaservices.aspx.

30. Baker C, Oppenheimer L, Stephens B, Lewis F, Trunkey D. Epidemiology of trauma deaths. *Am J Surg*. 1980;140(1):144-150.

31. Bickell WH, Wall MJ, Pepe PE, et al. Immediate versus delayed fluid resuscitation for hypotensive patients with penetrating torso injuries*. New Engl J Med*. 1994:1105-1109.

32. Dutton RP, Mackenzie CF, Scalea TM. Hypotensive resuscitation during active hemorrhage: impact on in-hospital mortality. *J Trauma.* 2002:1141-1146.

33. Borgman MA, Spinella PC, Perkins JG, et al. The ratio of blood products transfused affects mortality in patients receiving massive transfusions at a combat support hospital. *J Trauma*. 2007;63(4):805-813.

34. Maegele M, Lefering R, Paffrath T, Tjardes T, Simanski C, Bouillon B. Red blood cell to plasma ratios transfused during massive transfusion are associated with mortality in severe multiply injury: a retrospective analysis from the Trauma Registry of the Deutsche Gesellschaft für Unfallchirurgie. *Vox Sang*. 2008;95(2):112-119.

35. Fry D, Pearlstein L, Fulton R, Polk H. Multiple system organ failure. The role of uncontrolled infection. *Arch Surg*. 1980;115(2).

36. Faist E, Baue AE, Dittmer H, Heberer G. Multiple organ failure in polytrauma patients. *J Trauma*. 1983;23(9):775-87..

37. Moore FA, Sauaia A, Moore EE, Haenel JB, Burch JM, Lezotte DC. Postinjury multiple organ failure: a bimodal phenomenon. *J Trauma*. 1996;40(4):501-510; discussion 510-512.

38. Nast-Kolb D, Aufmkolk M, Rucholtz S, Obertacke U, Waydhas C. Multiple organ failure still a major cause of morbidity but not mortality in blunt multiple trauma. *J Trauma*. 2001;51(5):835-41; discussion 841-2.

39. Demetriades D, Kimbrell B, Salim A, et al. Trauma deaths in a mature urban trauma system: is "trimodal" distribution a valid concept? *J Am Coll Surg*. 2005;201(3):343-348.

40. Sauaia A, Moore FA, Moore EE, et al. Epidemiology of trauma deaths: a reassessment. *J Trauma*. 1995;38(2):185-93.

41. De Knegt C, Meylaerts S, Leenen LPH. Applicability of the trimodal distribution of trauma deaths in a Level I trauma centre in the Netherlands with a population of mainly blunt trauma. *Injury*. 2008;39(9):993-1000.

42. Pang J-M, Civil I, Ng A, Adams D, Koelmeyer T. Is the trimodal pattern of death after trauma a dated concept in the 21st century? Trauma deaths in Auckland 2004. *Injury*. 2008;39(1):102-106.

43. Pfeifer R, Tarkin IS, Rocos B, Pape H-C. Patterns of mortality and causes of death in polytrauma patients—Has anything changed? *Injury*. 2009;40(9):907-911.

44. Regel G, Lobenhoffer P, Grotz M, Pape HC, Lehmann U, Tscherne H. Treatment results of patients with multiple trauma: an analysis of 3406 cases treated between 1972 and 1991 at a German Level I Trauma Center. *J Trauma*. 1995;38(1):70-8.

45. Pape HC, Remmers D, Rice J, Ebisch M, Krettek C, Tscherne H. Appraisal of early evaluation of blunt chest trauma: development of a standardized scoring system for initial clinical decision making. *J Trauma*. 2000;49(3):496-504.

46. Holbrook TL, Anderson JP, Sieber WJ, Browner D, Hoyt DB. Outcome after major trauma: 12-month and 18-month follow-up results from the Trauma Recovery Project. *J Trauma*. 1999;46(5):765-771; discussion 771-773.

47. Cameron P, Gabbe BJ, McNeil JJ. The importance of quality of survival as an outcome measure for an integrated trauma system. *Injury*. 2006;37(12):1178-84.

48. Butcher N, Balogh ZJ. The definition of polytrauma: the need for international consensus. *Injury*. 2009;40 Suppl 4:S12-22.

49. Butcher N, Balogh ZJ. AIS>2 in at least two body regions: A potential new anatomical definition of polytrauma. *Injury*. 2012;43(2):196-199.

50. American Association for Automotive Medicine. AIS (C) 2005 Update 2008. Available at: www.aaam1.org/ais/.

51. Baker SP, O'Neill B, Haddon W, Long WB. The injury severity score: a method for describing patients with multiple injuries and evaluating emergency care. *J Trauma*. 1974;14(3):187-196.

52. Semmlow JL, Cone R. Utility of the injury severity score: a confirmation. *Health Serv Res*. 1976;11(1):45-52.

53. Bull JP. Measures of severity of injury. *Injury*. 1978;9(3):184-187.

54. Copes WS, Champion HR, Sacco WJ, Lawnick MM, Keast S, Bain LW. The Injury Severity Score Revisited. *J Trauma*. 1988;28(1):69-77.

55. Linn S. The injury severity score--importance and uses. *Ann Epidemiol*. 1995;5(6):440-446.

56. Schwab W, Frankel HL, Rotondo MF, et al. The impact of true partnership between a university Level I trauma center and a community Level II trauma center on patient transfer practices. *J Trauma*. 1998;44(5):815-819; discussion 819-820.

57. Osler T, Baker SP, Long W. A modification of the injury severity score that both improves accuracy and simplifies scoring. *J Trauma*. 1997;43(6):922-925; discussion 925-926.

58. Lavoie A, Moore L, LeSage N, Liberman M, Sampalis JS. The New Injury Severity Score:a more accurate predictor of in-hospital mortality than the Injury Severity Score. *J Trauma*. 2004;56(6):1312-1320.

59. Balogh ZJ, Varga E, Tomka J, Süveges G, Tóth L, Simonka JA. The new injury severity score is a better predictor of extended hospitalization and intensive care unit admission than the injury severity score in patients with multiple orthopaedic injuries. *J Orthop Trauma*. 2003;17(7):508-512.

60. Copes WS, Champion HR, Sacco WJ, et al. Progress in characterizing anatomic injury. *J Trauma*. 1990;30(10):1200-1207.

61. Osler T, Rutledge R, Deis J, Bedrick E. ICISS: an international classification of disease-9 based injury severity score. *J Trauma*. 1996;41(3):380-386; discussion 386-388.

62. Rutledge R, Osler T. The ICD-9-based illness severity score: a new model that outperforms both DRG and APR-DRG as predictors of survival and resource utilization. *J Trauma*. 1998;45(4):791-799.

63. Willis CD, Gabbe BJ, Jolley D, Harrison JE, Cameron PA. Predicting trauma patient mortality: ICD [or ICD-10-AM] versus AIS based approaches. *ANZ J Surg*. 2010;80(11):802-806.

64. Champion HR, Sacco WJ, Hannan DS, et al. Assessment of injury severity: the triage index. *Crit Care Med*. 1980;8(4):201-208.

65. Champion HR, Sacco WJ, Carnazzo AJ, Copes W, Fouty WJ. Trauma score. *Crit Care Med*. 1981;9(9):672-676.

66. Champion HR, Sacco WJ, Copes WS, Gann DS, Gennarelli TA, Flanagan ME. A revision of the Trauma Score. *J Trauma*. 1989;29(5):623-629.

67. Boyd C, Tolson M. Accuracy and relationship of mechanisms of injury, trauma score and injury severity score in identifying major trauma. *Am J Surg*. 1987;153(448):514.

68. Koehler JJ, Baer LJ, Malafa SA, Meindertsma MS, Navitskas NR, Huizenga JE. Prehospital Index: a scoring system for field triage of trauma victims. *Ann Emerg Med*. 1986;15(2):178-82.

69. Vassar MJ, Wilkerson CL, Duran PJ, Perry CA, Holcroft JW. Comparison of APACHE II, TRISS, and a proposed 24-hour ICU point system for prediction of outcome in ICU trauma patients. *J Trauma*. 1992;32(4):490-499; discussion 499-500.

70. Kuhls DA, Malone DL, McCarter RJ, Napolitano LM. Predictors of mortality in adult trauma patients: the physiologic trauma score is equivalent to the Trauma and Injury Severity Score. *J Am Coll Surg*. 2002;194(6):695-704.

71. Guzzo JL, Bochicchio G V, Napolitano LM, Malone DL, Meyer W, Scalea TM. Prediction of outcomes in trauma: anatomic or physiologic parameters? *J Am Coll Surg*. 2005;201(6):891-897.

72. Boyd CR, Tolson MA, Copes WS. Evaluating trauma care: the TRISS method. Trauma Score and the Injury Severity Score. *J Trauma*. 1987;27(4):370-378.

73. Champion HR, Sacco WJ, Lepper RL, Atzinger EM, Copes WS, Prall RH. An anatomic index of injury severity. *J Trauma*. 1980;20(3):197-202.

74. Demetriades D, Chan LS, Velmahos G, et al. TRISS methodology in trauma: the need for alternatives. *Br J Surg*. 1998;85(3):379-384.

75. Davis EG, MacKenzie EJ, Sacco WJ, et al. A new “TRISS-like” probability of survival model for intubated trauma patients. *J Trauma*. 2003;55(1):53-61.

76. Kaufmann CR, Maier R V, Kaufmann EJ, Rivara FP, Carrico CJ. Validity of applying adult TRISS analysis to injured children. *J Trauma*. 1991;31(5):691-697; discussion 697-698.

77. Lavoie A, Moore L, LeSage N, Liberman M, Sampalis JS. The Injury Severity Score or the New Injury Severity Score for predicting intensive care unit admission and hospital length of stay? *Injury*. 2005;36(4):477-483.

78. Hannan EL, Waller CH, Farrell LS, Cayten CG. A comparison among the abilities of various injury severity measures to predict mortality with and without accompanying physiologic information. *J Trauma*. 2005;58(2):244-251.

79. Wong DT, Barrow PM, Gomez M, McGuire GP. A comparison of the Acute Physiology and Chronic Health Evaluation (APACHE) II score and the Trauma-Injury Severity Score (TRISS) for outcome assessment in intensive care unit trauma patients. *Crit Care Med*. 1996;24(10):1642-1648.

80. Champion H, Copes W, Sacco W, et al. The Major Trauma Outcome Study: establishing national norms for trauma care. *J Trauma*. 1990;30(11):1356.

81. Champion HR, Copes WS, Sacco WJ, et al. Improved predictions from a severity characterization of trauma (ASCOT) over Trauma and Injury Severity Score (TRISS): results of an independent evaluation. *J Trauma*. 1996;40(1):42-8; discussion 48-9.

82. Markle J, Cayten CG, Byrne DW, Moy F, Murphy JG. Comparison between TRISS and ASCOT methods in controlling for injury severity. *J Trauma*. 1992;33(2):326-332.

83. Gormican SP. CRAMS scale: field triage of trauma victims. *Ann Emerg Med*. 1982;11(3):132-135.

84. Ornato J, Mlinek EJ, Craren EJ, Nelson N. Ineffectiveness of the trauma score and the CRAMS scale for accurately triaging patients to trauma centers. *Ann Emerg Med*. 1985;14(11):1061-1064.

85. Bouamra O, Wrotchford A, Hollis S, Vail A, Woodford M, Lecky F. A new approach to outcome prediction in trauma: A comparison with the TRISS model. *J Trauma*. 2006;61(3):701-710.

86. Lefering R. Development and validation of the revised injury severity classification score for severely injured patients. *Eur J Trauma Emerg Surg*. 2009;35(5):437-447.

87. Sartorius D, Le Manach Y, David J-S, et al. Mechanism, glasgow coma scale, age, and arterial pressure (MGAP): a new simple prehospital triage score to predict mortality in trauma patients. *Crit Care Med*. 2010;38(3):831-837.

88. Kondo Y, Abe T, Kohshi K, Tokuda Y, Cook EF, Kukita I. Revised trauma scoring system to predict in-hospital mortality in the emergency department: Glasgow Coma Scale, Age, and Systolic Blood Pressure score. *Crit Care*. 2011;15(4):R191.

89. Eiseman B, Beart R, Norton L. Multiple organ failure. *Surg Gynecol Obstet*. 1977;144(3):323-326.

90. Baue AE, Durham R, Faist E. Systemic inflammatory response syndrome (SIRS), multiple organ dysfunction syndrome (MODS), multiple organ failure (MOF): are we winning the battle? *Shock*. 1998;10(2):79-89.

91. Baue A. MOF, MODS, and SIRS: what is in a name or an acronym? *Shock*. 2006;26(5):438.

92. Goris R, Boekhorst T, Nuytinck J, Gimbrère J. Multiple-organ failure. Generalized autodestructive inflammation? *Arch Surg*. 1985;120(10).

93. Ciesla D, Moore EE, Johnson J, Burch J, Cothren C, Sauaia A. A 12-year prospective study of postinjury multiple organ failure: has anything changed? *Arch Surg*. 2005;140(5):432.

94. Regel G, Grotz M, Weltner T, Sturm JA, Tscherne H. Pattern of organ failure following severe trauma. *World J Surg*. 1996;20(4):422-429.

95. Cryer H, Leong K, McArthur D, et al. Multiple organ failure: by the time you predict it, it's already there. *J Trauma*. 1999;46(4):597.

96. Dewar DC, Mackay P, Balogh Z. Epidemiology of post-injury multiple organ failure in an Australian trauma system. *ANZ J Surg*. 2009;79(6):431-436.

97. Sauaia A, Moore F, Moore E, Norris J, Lezotte D, Hamman R. Multiple organ failure can be predicted as early as 12 hours after injury. *J Trauma*. 1998;45(2):291-303.

98. Moore FA, Sauaia A, Moore EE, Haenel JB, Burch JM, Lezotte DC. Postinjury multiple organ failure: a bimodal phenomenon. *J Trauma*. 1996;40(4):501-510; discussion 510-512.

99. Nast-Kolb D, Waydhas C, Gippner-Steppert C, et al. Indicators of the posttraumatic inflammatory response correlate with organ failure in patients with multiple injuries. *J Trauma*. 1997;42(3):4464-54; discussion 454-455.

100. Dewar D, Moore FA, Moore EE, Balogh Z. Postinjury multiple organ failure. *Injury*. 2009;40(9):912-918.

101. Ciesla DJ, Moore EE, Johnson JL, et al. Decreased progression of postinjury lung dysfunction to the acute respiratory distress syndrome and multiple organ failure. *Surgery*. 2006;140(4):640-647; discussion 647-648.

102. Ciesla DJ, Moore EE, Johnson JL, Burch JM, Cothren CC, Sauaia A. The role of the lung in postinjury multiple organ failure. *Surgery*. 2005;138(4):749-757; discussion 757-758.

103. Durham RM, Moran JJ, Mazuski JE, Shapiro MJ, Baue AE, Flint LM. Multiple organ failure in trauma patients. *J Trauma*. 2003;55(4):608-616.

104. Proulx F, Joyal JS, Mariscalco MM, Leteurtre S, Leclerc F, Lacroix J. The pediatric multiple organ dysfunction syndrome. *Pediatr Crit Care Med*. 2009;10(1):12-22.

105. Calkins CM, Bensard DD, Moore EE, et al. The injured child is resistant to multiple organ failure: a different inflammatory response? *J Trauma*. 2002;53(6):1058-1063.

106. Proulx F, Gauthier M, Nadeau D, Lacroix J, Farrell CA. Timing and predictors of death in pediatric patients with multiple organ system failure. *Crit Care Med*. 1994;22(6):1025-1031.

107. Wood JH, Partrick DA, Johnston RB. The inflammatory response to injury in children. *Curr Opin Pediatr*. 2010;22(3):315-320.

108. Probst C, Zelle BA, Sittaro NA, Lohse R, Krettek C, Pape HC. Late death after multiple severe trauma: when does it occur and what are the causes? *J Trauma*. 2009;66(4):1212-1217.

109. Ulvik A, Kvåle R, Wentzel-Larsen T, Flaatten H. Multiple organ failure after trauma affects even long-term survival and functional status. *Crit Care*. 2007;11(5):R95.

110. Livingston D, Tripp T, Biggs C, Lavery R. A fate worse than death? Long-term outcome of trauma patients admitted to the surgical intensive care unit. *J Trauma*. 2009;67(2):341.

111. Eddleston JM, White P, Guthrie E. Survival, morbidity, and quality of life after discharge from intensive care. *Crit Care Med*. 2000;28(7):2293-2299.

112. Broomhead LR, Brett SJ. Clinical review: Intensive care follow-up--what has it told us? *Crit Care*. 2002;6(5):411-417.

113. Meakins J. Etiology of multiple organ failure. *J Trauma*. 1990;30:165.

114. Waydhas C, Nast-Kolb D, Trupka A, et al. Posttraumatic inflammatory response, secondary operations, and late multiple organ failure. *J Trauma*. 1996;40(4):624-628.

115. Moore EE, Moore FA, Harken AH, Johnson JL, Ciesla D, Banerjee A. The two-event construct of postinjury multiple organ failure. *Shock*. 2005;24 Suppl 1:71-74.

116. Carrico CJ, Meakins JL, Marshall JC, Fry D, Maier R V. Multiple-organ-failure syndrome. *Arch Surg*. 1986;121(2):196-208.

117. Deitch EA, Ma WJ, Ma L, Berg R, Specian RD. Endotoxin-induced bacterial translocation: a study of mechanisms. *Surgery*. 1989;106(2):292-299; discussion 299-300.

118. Cerra FB, Maddaus MA, Dunn DL, et al. Selective gut decontamination reduces nosocomial infections and length of stay but not mortality or organ failure in surgical intensive care unit patients. *Arch Surg*. 1992;127(2):163-167; discussion 167-169.

119. Reidy JJ, Ramsay G. Clinical trials of selective decontamination of the digestive tract: review. *Crit Care Med*. 1990;18(12):1449-1456.

120. Dunham CM, Damiano AM, Wiles CE, Cushing BM. Post-traumatic multiple organ dysfunction syndrome--infection is an uncommon antecedent risk factor. *Injury*. 1995;26(6):373-378.

121. Moore FA, Moore EE. Evolving concepts in the pathogenesis of postinjury multiple organ failure. *Surg Clin North Am*. 1995;75(2):257-277.

122. Smail N, Messiah A, Edouard A, et al. Role of systemic inflammatory response syndrome and infection in the occurrence of early multiple organ dysfunction syndrome following severe trauma. *Intensive Care Med*. 1995;21(10):813-816.

123. Bone R, Balk R, Slotman G, et al. Adult respiratory distress syndrome. Sequence and importance of development of multiple organ failure. *Chest*. 1992;101(2):320-326.

124. Napolitano LM, Ferrer T, McCarter RJ, Scalea TM. Systemic inflammatory response syndrome score at admission independently predicts mortality and length of stay in trauma patients. *J Trauma*. 2000;49(4):647-652; discussion 652-653..

125. Malone DL, Kuhls D, Napolitano LM, McCarter R, Scalea T. Back to basics: validation of the admission systemic inflammatory response syndrome score in predicting outcome in trauma. *J Trauma*. 2001;51(3):458-463.

126. Lausevic, Trbokevic-Stankovic, Mirjana T-S, Krstic, Lausevic S, Slobodan S. Predicting multiple organ failure in patients with severe trauma. *Can J Surg*. 2008;51(2):97-102.

127. Tschoeke SK, Ertel W. Immunoparalysis after multiple trauma. *Injury*. 2007;38(12):1346-1357.

128. Stahel PF, Smith WR, Moore EE. Role of biological modifiers regulating the immune response after trauma. *Injury*. 2007;38(12):1409-1422.

129. Giannoudis P V. Current concepts of the inflammatory response after major trauma: an update. *Injury*. 2003;34(6):397-404.

130. Kapur MM, Jain P, Gidh M. The effect of trauma on serum C3 activation and its correlation with injury severity score in man. *J Trauma*. 1986;26(5):464-466.

131. Gutierrez G, Reines HD, Wulf-Gutierrez ME. Clinical review: hemorrhagic shock. *Crit Care*. 2004;8(5):373-381.

132. Visser T, Pillay J, Koenderman L, Leenen LPH. Postinjury immune monitoring: can multiple organ failure be predicted? *Curr Opin Crit Care*. 2008;14(6):666-672.

133. Flohé SB, Flohé S, Schade FU. Invited review: deterioration of the immune system after trauma: signals and cellular mechanisms. *Innate Immun*. 2008;14(6):333-344.

134. Partrick D, Moore EE, Moore F, Biffl W, Barnett C. Release of anti-inflammatory mediators after major torso trauma correlates with the development of postinjury multiple organ failure. *Am J Surg*. 1999;178(6):564-568.

135. Jastrow KM, Gonzalez EA, McGuire MF, et al. Early cytokine production risk stratifies trauma patients for multiple organ failure. *J Am Coll Surg*. 2009;209(3):320-331.

136. Cohen MJ, Call M, Nelson M, et al. Critical role of activated protein C in early coagulopathy and later organ failure, infection and death in trauma patients. *Ann Surg*. 2012;255(2):379-385.

137. Law MM, Cryer HG, Abraham E. Elevated levels of soluble ICAM-1 correlate with the development of multiple organ failure in severely injured trauma patients. *J Trauma*. 1994;37(1):100-109; discussion 109-110.

138. Gando S, Kameue T, Matsuda N, et al. Combined activation of coagulation and inflammation has an important role in multiple organ dysfunction and poor outcome after severe trauma. *Thromb Haemost*. 2002;88(6):943-949.

139. Partrick D, Moore EE, Silliman C, Barnett C, Kuypers F. Secretory phospholipase A2 activity correlates with postinjury multiple organ failure. *Crit Care Med*. 2001;29(5):989.

140. Svoboda P, Kantorová I, Ochmann J. Dynamics of interleukin 1, 2, and 6 and tumor necrosis factor alpha in multiple trauma patients. *J Trauma*. 1994;36(3):336-340.

141. Rabinovici R, John R, Esser KM, Vernick J, Feuerstein G. Serum tumor necrosis factor-alpha profile in trauma patients. *J Trauma*. 1993;35(5):698-702.

142. Biffl WL, Moore EE, Zallen G, et al. Neutrophils are primed for cytotoxicity and resist apoptosis in injured patients at risk for multiple organ failure. *Surgery*. 1999;126(2):198-202.

143. Ogura H, Tanaka H, Koh T, et al. Priming, second-hit priming, and apoptosis in leukocytes from trauma patients. *J Trauma*. 1999;46(5):774-781; discussion 781-783.

144. Frink M, Pape H-C, van Griensven M, Krettek C, Chaudry IH, Hildebrand F. Influence of sex and age on mods and cytokines after multiple injuries. *Shock*. 2007;27(2):151-156.

145. Spruijt NE, Visser T, Leenen LP. A systematic review of randomized controlled trials exploring the effect of immunomodulative interventions on infection, organ failure, and mortality in trauma patients. *Crit Care*. 2010;14(4):R150.

146. Tsukamoto T, Chanthaphavong RS, Pape H-C. Current theories on the pathophysiology of multiple organ failure after trauma. *Injury*. 2010;41(1):21-26.

147. Sauaia A, Moore F, Moore EE, Haenel J, Read R, Lezotte D. Early predictors of postinjury multiple organ failure. *Arch Surg*. 1994;129(1):39-43.

148. Sauaia A, Moore FA, Moore EE, Lezotte DC. Early risk factors for postinjury multiple organ failure. *World J Surg*. 1996;20(4):392-400.

149. Ciesla D, Moore EE, Johnson J, et al. Multiple organ dysfunction during resuscitation is not postinjury multiple organ failure. *Arch Surg*. 2004;139(6):590.

150. Callaway DW, Shapiro NI, Donnino MW, Baker C, Rosen CL. Serum lactate and base deficit as predictors of mortality in normotensive elderly blunt trauma patients. *J Trauma*. 2009;66(4):1040-1044.

151. Schulman AM, Claridge JA, Young JS. Young versus old: factors affecting mortality after blunt traumatic injury. *Am Surg*. 2002;68(11):942-947; discussion 947-948.

152. Milzman DP, Rothenhaus TC. Resuscitation of the geriatric patient. *Emerg Med Clin North Am*. 1996;14(1):233-244.

153. Angele MK, Schwacha MG, Ayala A, Chaudry IH. Effect of gender and sex hormones on immune responses following shock. *Shock*. 2000;14(2):81-90.

154. Offner PJ, Moore EE, Biffl WL. Male gender is a risk factor for major infections after surgery. *Arch Surg*. 1999;134(9):935-938; discussion 938-940.

155. Sperry JL, Frankel HL, Vanek SL, et al. Early hyperglycemia predicts multiple organ failure and mortality but not infection. *J Trauma*. 2007;63(3):487-493.

156. Keel M, Trentz O. Pathophysiology of polytrauma. *Injury*. 2005;36(6):691-709.

157. Krishna G, Sleigh JW, Rahman H. Physiological predictors of death in exsanguinating trauma patients undergoing conventional trauma surgery. *Aust N Z J Surg*. 1998;68(12):826-829.

158. Davis JW, Parks SN, Kaups KL, Gladen HE, O'Donnell-Nicol S. Admission base deficit predicts transfusion requirements and risk of complications. *J Trauma*. 1996;41(5):769-774.

159. Kincaid EH, Miller PR, Meredith JW, Rahman N, Chang MC. Elevated arterial base deficit in trauma patients: a marker of impaired oxygen utilization. *J Am Coll Surg*. 1998;187(4):384-392.

160. Blow O, Magliore L, Claridge J, Butler K, Young J. The golden hour and the silver day: detection and correction of occult hypoperfusion within 24 hours improves outcome from major trauma. *J Trauma*. 1999;47(5):964-968.

161. Crowl AC, Young JS, Kahler DM, Claridge JA, Chrzanowski DS, Pomphrey M. Occult hypoperfusion is associated with increased morbidity in patients undergoing early femur fracture fixation. *J Trauma*. 2000;48(2):260-267.

162. Brohi K, Singh J, Heron M, Coats T. Acute traumatic coagulopathy. *J Trauma*. 2003;54(6):1127-1130.

163. Moore F, Moore EE, Sauaia A. Blood transfusion: an independent risk factor for postinjury multiple organ failure. *Arch Surg*. 1997;132(6):620-622.

164. Watson GA, Sperry JL, Rosengart MR, et al. Fresh frozen plasma is independently associated with a higher risk of multiple organ failure and acute respiratory distress syndrome. *J Trauma*. 2009;67(2):221-227; discussion 228-230.

165. Tscherne H, Regel G. Care of the polytraumatised patient. *J Bone Joint Surg Br*. 1996;78(5):840-852.

166. Shapiro MB, Jenkins DH, Schwab CW, Rotondo MF. Damage control: collective review. *J Trauma*. 2000;49(5):969-978.

167. Pape HC, van Griensven M, Rice J, et al. Major secondary surgery in blunt trauma patients and perioperative cytokine liberation: determination of the clinical relevance of biochemical markers. *J Trauma*. 2001;50(6):989-1000.

168. Pape H-C, Griensven M V, Hildebrand FF, et al. Systemic inflammatory response after extremity or truncal fracture operations. *J Trauma*. 2008;65(6):1379-1384.

169. Lasanianos NG, Kanakaris NK, Dimitriou R, Pape HC, Giannoudis P V. Second hit phenomenon: existing evidence of clinical implications. *Injury*. 2011;42(7):617-629.

170. Vogel J, Liao M, hopkins E, et al. Prediction of post-injury multiple organ failure: Derivation and internal validation of the Denver Trauma Organ Failure Score. *Acad Emerg Med*. 2011;18(5, Suppl1):S144-S145.

171. SAEM Abstracts, Plenary Session. *Acad Emerg Med*. 2012;19:S4-S393.

172. Minei JP, Cuschieri J, Sperry J, et al. The changing pattern and implications of multiple organ failure after blunt injury with hemorrhagic shock. *Crit Care Med*. 2012;40(4):1129-1135.

173. Dewar DC, Tarrant SM, King KL, Balogh ZJ. Changes in the epidemiology and prediction of multiple-organ failure after injury. *J Trauma Acute Care Surg*. 2013;74(3):774-9.

174. Vincent J, Moreno R, Takala J, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. *Intensive Care Med*. 1996;22(7):707-710.

175. Marshall J, Cook D, Christou N, Bernard G, Sprung C, Sibbald W. Multiple organ dysfunction score: a reliable descriptor of a complex clinical outcome. *Crit Care Med*. 1995;23(10):1638-1642.

176. Moons KGM, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? *BMJ*. 2009;338:b375.

177. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstem AR. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol*. 1996;49:1373-1379.

178. Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996;15:361-387.

179. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and cox regression. *Am J Epidemiol*. 2007;165:710-718.

180. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ*. 2009;339:b2535.

181. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ*. 2009;339:b2700.

182. Jüni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA J Am Med Assoc*. 1999;282(11):1054-1060.

183. Aldrian S, Nau T, Koenig F, Vécsei. Geriatric polytrauma. *Wien Klin Wochenschr*. 2005;117(4):145-149.

184. Balogh Z, Offner P, Moore EE, Biffl W. NISS predicts postinjury multiple organ failure better than the ISS. *J Trauma*. 2000;48(4):624-627.

185. Beilman G, Blondet J, Nelson T, et al. Early hypothermia in severely injured trauma patients is a significant risk factor for multiple organ dysfunction syndrome but not mortality. *Ann Surg*. 2009;249(5):845-848.

186. Borgman MA, Spinella PC, Holcomb JB, et al. The effect of FFP: RBC ratio on morbidity and mortality in trauma patients based on transfusion prediction score. *Vox Sang*. 2011.

187. Brattstrom O, Granath F, Rossi P, Oldner A. Early predictors of morbidity and mortality in trauma patients treated in the intensive care unit. *Acta Anaesthesiol Scand*. 2010;54(8):1007-1017.

188. Ciesla DJ, Moore EE, Johnson JL, Burch JM, Cothren CC, Sauaia A. Obesity increases risk of organ failure after severe trauma. *J Am Coll Surg*. 2006;203(4):539-45.

189. Cotton B, Au B, Nunez T, Gunter O, Robertson A, Young P. Predefined massive transfusion protocols are associated with a reduction in organ failure and postinjury complications. *J Trauma*. 2009;66(1):41-43.

190. Edmonds R, Cuschieri J, Minei J, et al. Body Adipose Content is Independently Associated With a Higher Risk of Organ Failure and Nosocomial Infection in the Nonobese Patient Postinjury. *J Trauma Acute Care Surg*. 2011;70(2):292-298.

191. Gando S, Nanzaki S, Kemmotsu O. Disseminated intravascular coagulation and sustained systemic inflammatory response syndrome predict organ dysfunctions after trauma: application of clinical decision analysis. *Ann Surg*. 1999;229(1):121-127.

192. Johnson J. Effect of blood products transfusion on the development of postinjury multiple organ failure. *Arch Surg*. 2010;145(10):973-977.

193. Maegele M, Lefering R, Yucel N, et al. Early coagulopathy in multiple injury: An analysis from the German Trauma Registry on 8724 patients. *Injury*. 2007;38(3):298-304.

194. Moore FA, Moore EE, Sauaia A. Blood transfusion. An independent risk factor for postinjury multiple organ failure. *Arch Surg*. 1997;132(6):620-624; discussion 624-625.

195. Oberholzer A, Keel M, Zellweger R, Steckholzer U, Trentz O, Ertel W. Incidence of septic complications and multiple organ failure in severely injured patients is sex specific. *J Trauma*. 2000;48(5):932-935.

196. Sperry JL, Nathens AB, Frankel HL, et al. Characterization of the gender dimorphism after injury and hemorrhagic shock: Are hormonal differences responsible?. *Crit Care Med*. 2008;36(6):1838-1845.

197. Trupka A, Waydhas C, Nast-Kolb D, Schweiberer L. Early intubation in severely injured patients. *Eur J Emerg Me*. 1994;1(1):1-4.

198. Tuttle MS, Smith WR, Williams AE, et al. Safety and efficacy of damage control external fixation versus early definitive stabilization for femoral shaft fractures in the multiple-injured patient. *J Trauma Inj Infect Crit Care*. 2009;67(3):602-605.

199. Winfield RD, Delano MJ, Lottenberg L, et al. Traditional resuscitative practices fail to resolve metabolic acidosis in morbidly obese patients after severe blunt trauma. *J Trauma Inj Infect Crit Care*. 2010;68(2):317-330.

200. Zallen G, Offner P, Moore EE, et al. Age of transfused blood is an independent risk factor for postinjury multiple organ failure. *Am J Surg*. 1999;178(6):570-572.

201. Zeckey C, Dannecker S, Hildebr F, et al. Alcohol and multiple trauma—is there an influence on the outcome? *Alcohol*. 2011;45(3):245-251.

202. Sperry JL, Frankel HL, Vanek SL, et al. Early hyperglycemia predicts multiple organ failure and mortality but not infection. *J Trauma Inj Infect Crit Care*. 2007;63(3):487-494.

203. Ciesla DJ, Moore EE, Johnson JL, Burch JM, Cothren CC, Sauaia A. Obesity increases risk of organ failure after severe trauma. *J Am Coll Surg*. 2006;203(4):539-545.

204. Moore F, Sauaia A, Moore EE, Haenel J, Burch J, Lezotte D. Postinjury multiple organ failure: a bimodal phenomenon. *J Trauma*. 1996;40(4):501-504.

205. Durham RM, Moran JJ, Mazuski JE, Shapiro MJ, Baue AE, Flint LM. Multiple organ failure in trauma patients. *J Trauma Inj Infect Crit Care*. 2003;55(4):608-616.

206. Frink M, Pape H-C, van Griensven M, Krettek C, Chaudry IH, Hildebrand F. Influence of sex and age on MODS and cytokines after multiple injuries. *Shock*. 2007;27(2):151-156.

207. Watson GA, Sperry JL, Rosengart MR, et al. Fresh Frozen Plasma is independently associated with a higher risk of multiple organ failure and Acute Respiratory Distress Syndrome. *J Trauma Inj Infect Crit Care*. 2009;67(2):221-230.

208. Sauaia A, Moore EE, Johnson JL, Ciesla DJ, Biffl WL, Banerjee A. Validation of postinjury multiple organ failure scores. *Shock*. 2009;31(5):438-447.

209. Warren H, Elson C, Hayden D, et al. A genomic score prognostic of outcome in trauma patients. *Mol Med*. 2009;15(7-8).

210. Skillman JJ, Bushnell LS, Goldman H, Silen W. Respiratory failure, hypotension, sepsis, and jaundice. A clinical syndrome associated with lethal hemorrhage from acute stress ulceration of the stomach. *Am J Surg*. 1969;117(4):523-530.

211. Tilney N, Bailey G, Morgan A. Sequential system failure after rupture of abdominal aortic aneurysms: an unsolved problem in postoperative care. *Ann Surg*. 1973;178(2):117.

212. Pine R, Wertz M, Lennard E, Dellinger E, Carrico C, Minshew B. Determinants of organ malfunction or death in patients with intra-abdominal sepsis. A discriminant analysis. *Arch Surg*. 1983;118(2).

213. Bell R, Coalson J, Smith J, Johanson W. Multiple organ system failure and infection in adult respiratory distress syndrome. *Ann Intern Med*. 1983;99(3):293-298.

214. Marshall W, Dimick A. The natural history of major burns with multiple subsystem failure. *J Trauma*. 1983;23(2):102.

215. Stevens L. Gauging the severity of surgical sepsis. *Arch Surg*. 1983;118(10):1190-1192.

216. Manship L, McMillin RD, Brown JJ. The influence of sepsis and multisystem and organ failure on mortality in the surgical intensive care unit. *Am Surg*. 1984;50(2):94-101.

217. Knaus W, Draper E, Wagner D, Zimmerman J. Prognosis in acute organ-system failure. *Ann Surg*. 1985;202(6):685-687.

218. Norton L. Does drainage of intraabdominal pus reverse multiple organ failure? *Am J Surg*. 1985;149(3).

219. Maetani S, Nishikawa T, Tobe T, Hirakawa A. Role of blood transfusion in organ system failure following major abdominal surgery. *Ann Surg*. 1986;203(3):275-277.

220. Shen P, Zhang S. Acute renal failure and multiple organ system failure. *Arch Surg*. 1987;122(10).

221. Bihari D, Smithies M, Gimson A, Tinker J. The effects of vasodilation with prostacyclin on oxygen delivery and uptake in critically ill patients. *N Engl J Med*. 1987;317(7):397-403.

222. Marshall J, Christou N, Horn R, Meakins J. The microbiology of multiple organ failure. The proximal gastrointestinal tract as an occult reservoir of pathogens. *Arch Surg*. 1988;123(3).

223. Darling GE, Duff JH, Mustard RA, Finley RJ. Multiorgan failure in critically ill patients. *Can J Surg*. 1988;31(3):172-176.

224. Bumaschny E, Doglio G, Pusajó J, et al. Postoperative acute gastrointestinal tract hemorrhage and multiple-organ failure. *Arch Surg*. 1988;123(6).

225. Rubin D, Wiener-Kronish J, Murray JF, et al. Elevated von Willebrand factor antigen is an early plasma predictor of acute lung injury in nonpulmonary sepsis syndrome. *J Clin Invest*. 1990;86(2):474-476.

226. Dorinsky PM, Gadek JE. Multiple organ failure. *Clin Chest Med*. 1990;11(4):581-591.

227. Tran D, Groeneveld AB, Meulen J, Nauta JJ, Strack S, Thijs LG. Age, chronic disease, sepsis, organ system failure, and mortality in a medical intensive care unit. *Crit Care Med*. 1990;18(5):474-478.

228. Bone RC. Let's agree on terminology: definitions of sepsis. *Crit Care Med*. 1991;19(7):973-976.

229. Villar J, Manzano JJ, Blazquez MA, Quintana J, Lubillo S. Multiple system organ failure in acute respiratory failure. *J Crit Care*. 1991;6(2):75-80.

230. Ruokonen E, Takala J, Kari A, Alhava E. Septic shock and multiple organ failure. *Crit Care Med*. 1991;19(9):1146-1149.

231. Moore F, Moore EE, Poggetti R, et al. Gut bacterial translocation via the portal vein: a clinical perspective with major torso trauma. *J Trauma*. 1991;31(5):629-633.

232. Henao F, Daes J, Dennis R. Risk factors for multiorgan failure: a case-control study. *J Trauma*. 1991;31(1):74-77.

233. Shoemaker W, Appel P, Kram H. Role of oxygen debt in the development of organ failure sepsis, and death in high-risk surgical patients. *Chest*. 1992;102(1):208-215.

234. Martin L, Booth F V, Reines H, et al. Stress ulcers and organ failure in intubated patients in surgical intensive care units. *Ann Surg*. 1992;215(4):332-334.

235. Dobb G. Multiple organ failure: Words mean what I say they mean. *Intensive Care World*. 1992;14:61-63.

236. Hebert P, Drummond AJ, Singer J, Bernard GR, Russell JA. A simple multiple system organ failure scoring system predicts mortality of patients who have sepsis syndrome. *Chest*. 1993;104(1):230-235.

237. Kollef M. Ventilator-associated pneumonia. A multivariate analysis. *JAMA*. 1993;270(16):1965-1970.

238. Fagon JY, Chastre J, Novara A, Medioni P, Gibert C. Characterization of intensive care unit patients using a model based on the presence or absence of organ dysfunctions and/or infection: the ODIN model. *Intensive Care Med*. 1993;19(3):137-144.

239. Saffle J, Sullivan J, Tuohig G, Larson C. Multiple organ failure in patients with thermal injury. *Crit Care Med*. 1993;21(11):1673-1676.

240. Gall J, Klar J, Lemeshow S. How to assess organ dysfunction in the intensive care unit? The logistic organ dysfunction (LOD) system. *Sepsis*. 1997;1(1):45-47.

241. Bernard G. The Brussels score. *Sepsis*. 1997;1(1):43-44.

242. Lemeshow S, Le Gall JR. Modeling the severity of illness of ICU patients. A systems update. *JAMA*. 1994;272(13):1049-1055.

243. Strand K, Flaatten H. Severity scoring in the ICU: a review. *Acta Anaesthesiol Scand*. 2008;52(4):467-478.

244. Rowan KM, Angus DC. Don’t let perfection be the enemy of the good: it's time for optimism over the role of severity scoring systems in intensive care unit performance measurement. *Curr Opin Crit Care*. 2000;6(3):153-154.

245. Le Gall J-R. The use of severity scores in the intensive care unit. *Intensive Care Med*. 2005;31(12):1618-1623.

246. Gunning K, Rowan K. ABC of intensive care: outcome data and scoring systems. *BMJ*. 1999;319(7204):241-244.

247. Marshall J. The multiple organ dysfunction (MOD) score. *Sepsis*. 1997;1(1):49-52.

248. Marshall J. Multiple Organ Dysfunction Syndrome (MODS). In: Sibbald WJ, Vincent JL, eds. *Clinical Trials for the Treatment of Sepsis*.; 1995:122-138.

249. Peres Bota D, Melot C, Lopes Ferreira F, Nguyen Ba V, Vincent J-L. The Multiple Organ Dysfunction Score (MODS) versus the Sequential Organ Failure Assessment (SOFA) score in outcome prediction. *Intensive Care Med*. 2002;28(11):1619-1624.

250. Vincent JL, de Mendonça A, Cantraine F, et al. Use of the SOFA score to assess the incidence of organ dysfunction/failure in intensive care units: results of a multicenter, prospective study. *Crit Care Med*. 1998;26(11):1793-800.

251. Moreno R, Vincent J, Matos R, et al. The use of maximum SOFA score to quantify organ dysfunction/failure in intensive care. Results of a prospective, multicentre study. *Intensive Care Med*. 1999;25(7):686-696.

252. Ferreira F, Bota D, Bross A, Mélot C, Vincent J. Serial evaluation of the SOFA score to predict outcome in critically ill patients. *JAMA J Am Med Assoc*. 2001;286(14):1754-1758.

253. Minne L, Abu-Hanna A, de Jonge E. Evaluation of SOFA-based models for predicting mortality in the ICU: A systematic review. *Crit Care*. 2008;12(6):R161.

254. Cook R, Cook D, Tilley J, Lee K, Marshall J. Multiple organ dysfunction: baseline and serial component scores. *Crit Care Med*. 2001;29(11):2046-2050.

255. Winfield RD, Delano MJ, Dixon DJ, et al. Differences in outcome between obese and nonobese patients following severe blunt trauma are not consistent with an early inflammatory genomic response. *Crit Care Med*. 2010;38(1):51-58.

256. Antonelli M, Moreno R, Vincent JL, et al. Application of SOFA score to trauma patients. Sequential Organ Failure Assessment. *Intensive Care Med*. 1999;25(4):389-394.

257. Ulvik A, Kvåle R, Wentzel-Larsen T, Flaatten H. Multiple organ failure after trauma affects even long-term survival and functional status. *Crit Care*. 2007;11(5):R95.

258. Young JD, Goldfrad C, Rowan K. Development and testing of a hierarchical method to code the reason for admission to intensive care units: the ICNARC Coding Method. Intensive Care National Audit & Research Centre. *Br J Anaesth*. 2001;87(4):543-8.

259. Royston P, Moons KGM, Altman DG, Vergouwe Y. Prognosis and prognostic research: Developing a prognostic model. *BMJ*. 2009;338:b604.

260. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*. 2010;21:128-138.

261. Christensen MC, Ridley S, Lecky FE, Munro V, Morris S. Outcomes and costs of blunt trauma in England and Wales. *Crit Care*. 2008;12(1):R23.

262. Hess DR. Retrospective studies and chart reviews. *Respir Care*. 2004;49(10):1171-1174.

263. vonKoss Krowchuk H, Moore ML, Richardson L. Using health care records as sources of data for research. *J Nurs Meas*. 1995;3(1):3-12.

264. Equator Network. Available at: http://www.equator-network.org/library/reporting-guidelines-under-development/#10.

265. Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ*. 2007;335(7624):806-808.

266. Moore L, Hanley JA, Lavoie A, Turgeon A. Evaluating the validity of multiple imputation for missing physiological data in the national trauma data bank. *J Emerg Trauma Shock*. 2009;2(2):73-79.

267. Moore L, Hanley JA, Turgeon AF, Lavoie A, Emond M. A multiple imputation model for imputing missing physiologic data in the national trauma data bank. *J Am Coll Surg*. 2009;209(5):572-579.

268. Altman DG, Vergouwe Y, Royston P, Moons KGM. Prognosis and prognostic research: validating a prognostic model. *BMJ*. 2009;338:b605.

269. Mourouga P, Goldfrad C, Rowan KM. Does it fit? Is it good? Assessment of scoring systems. *Curr Opin Crit Care*. 2000;6(3):176-180.

270. Altman DG, Royston P. What do we mean by validating a prognostic model? *Stat Med*. 2000;19(4):453-473.

271. Hayden J, van der Windt D, Cartwright JL, Côté P, Bombardier C. Assessing bias in studies of prognostic factors. *Ann Intern Med*. 2013;158:280-6.

272. Moore LJ, Jones SL, Kreiner LA, et al. Validation of a screening tool for the early identification of sepsis. *J Trauma*. 2009;66:1539-1546; discussion 1546-1547.

273. Kingsford C, Salzberg SL. What are decision trees? *Nat Biotechnol*. 2008;26:1011-1013.

274. Podgorelec V, Kokol P, Stiglic B, Rozman I. Decision trees: An overview and their use in medicine. *J Med Syst*. 2002;26:445-463.

275. Apfel CC, Kranke P, Eberhart LHJ, Roos A, Roewer N. Comparison of predictive models for postoperative nausea and vomiting. *Br J Anaesth*. 2002;88:234-240.

276. Spiegelman D, Colditz GA, Hunter D, Hertzmark E. Validation of the gail et-al model for predicting individual breast-cancer risk. *J Natl Cancer Inst*. 1994;86:600–607.

APPENDICES

Please see separate volume for Appendices 1 – 15

Appendix 1 TARN Entry Criteria

Appendix 2 Data management plan

Appendix 3 Systematic review protocol

Appendix 4 Literature review results tables

Appendix 5 Questionnaire

Appendix 6 MOF scoring systems

Appendix 7 Data extraction tables

Appendix 8 Data dictionary

Appendix 9 SQL example queries

Appendix 10 Statistical analysis plan

Appendix 11 Univariate results by variable

Appendix 12 Univariate results by time period

Appendix 13 Variables selected for multivariate analysis

Appendix 14 Model development

Appendix 15 Comparison of model populations