Kirsty Hélène Elizabeth Challen
80110408

Thesis in support of application for the degree of Doctor of Philosophy

Clinical scoring system to identify high-acuity patients from information available in the Emergency Department

June 2014

Section of Health Services Research, School of Health and Related Research, University of Sheffield

Supervised by Prof S Goodacre, Prof J Nicholl, Mr M Bradburn**Contents**

Figures

Chapter 8 Operationalisation and implications

Tables

List of appendices

Abstract

Background

Acute illness results in millions of hospital admissions per year. Assessment of illness severity can guide the intensity and location of care provided but although multiple clinical decision aids exist for large numbers of conditions no well-validated clinical decision aid exists for the assessment of patients with unselected medical emergencies. It cannot be assumed that tools which predict death will accurately identify those patients with the most potential to benefit from urgent care.

Methods

A prospective cohort was analysed using logistic regression to develop bedside scores to identify patients at high risk of death, and those where emergency care has the potential to affect survival. Consensus methodology was used to develop threshold responses for the Emergency Department.

Results

7 variables and one interaction (age, respiratory rate, diastolic blood pressure, oxygen saturation, temperature, GCS, pre-existing respiratory disease, respiratory disease by temperature interaction) predicted death in 7 days with AUROC 0.753 (derivation set: n=2437) and 0.719 (validation set: n=2322). Other scores showed AUROC 0.658 – 0.762.

3 variables (pulse, systolic blood pressure, GCS) predicted potentially preventable or potentially prevented death with AUROC 0.737 (derivation set: n=398) and 0.686 (validation set: n=227). Other scores showed AUROC 0.559 – 0.684.

Consensus was reached on four thresholds of clinical response on the 0-27 point score for potentially preventable or potentially prevented death.

Conclusions

No published tool exists to identify the patient most likely to benefit from emergency department care. Variables predicting death do not necessarily predict potential to benefit from care, and existing scores have only moderate discrimination for this. The tool developed here shows potential but ongoing research should address which patients will benefit from time-critical interventions and the complexities of ED prioritisation.

Chapter 1
Background

1.1 Introduction

"Patients die not of their disease, they die of the physiological abnormalities of their disease."

Sir William Osler (1849-1919)

As Osler described, the deleterious effects of a disease process on a patient are usually mediated by a disruption in homeostasis. Recognition of this underpins the traditional "TPR" (temperature, pulse and respiration) nursing observations. This thesis will explore the extent to which easily-identifiable markers of homeostatic disruption, such as those obtained in a standard set of nursing observations, can be used to identify patients at risk of further physiological decompensation. It will attempt to construct a tool which can be used at the bedside to help medical and nursing staff recognise patients who are likely to benefit from high-acuity care as provided in intensive care and high dependency areas.

1.2 Acute illness

1.2.1 Prevalence of acute illness

There were 18.3 million attendances at Emergency Departments in England in 2012-3, 3.8 million of which resulted in a hospital admission (1).

1.2.2 Prevalence of severe acute illness

Of 5 million emergency hospital admissions in 2008-9, 203,790 resulted in inpatient death (2). Similarly, in 2010-1, only 71,801 (45%) of 160,460 critical care admissions were planned (3). Thus it is clear that acute severe illness has significant human and resource implications for the NHS.

1.2.3 Acuity assessment: the context of triage

Triage (from the French *trier*, to sort) has been a part of battlefield medicine since the early 19th century, when Napoleon's surgeon Baron Dominique-Jean Larrey

recognised that “those who are dangerously wounded should receive the first attention, without regard to rank or distinction” (4).

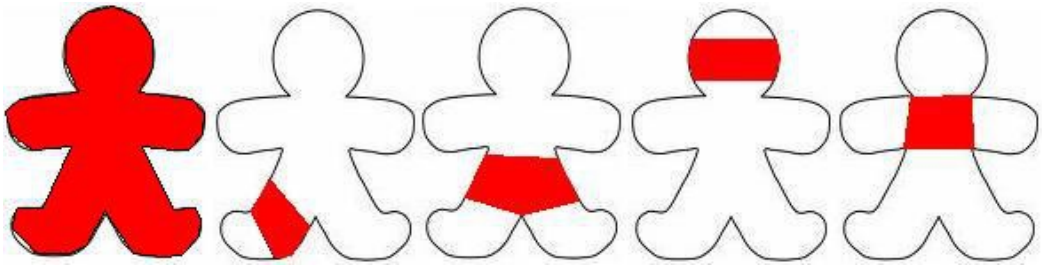
Triage in civilian emergency medicine in the developed world is usually a matter of ranking the severity of patient illness or injury to determine the order in which patients are seen, rather than limiting care, as for the most part resources and demand are roughly matched *in toto* (although there may be a temporal mismatch). In a typical Emergency Department, therefore, triage aims to identify levels of patient acuity (severity of illness/injury) and match the patient to an appropriately resourced and staffed area of the department (or for some low-acuity patients, to an alternative healthcare provider). A number of triage scales exist for triage within the ED; these are more fully explored in section 1.2.5b below.

The importance of casualty sorting in civilian disasters was suggested following the Harrow and Wealdstone railway disaster in which 85 died and 170 were hospitalised (5). More recently the Millian utilitarian philosophy of the “greatest happiness for the greatest number”(6) has been integrated into civilian practice (usually misquoted as “the greatest *good* for the greatest number”); system requirements for “maximal casualty survival” (ie the greatest possible number of survivors in the best possible functional condition) were identified two decades ago (7). To illustrate, the 5 patients in figure 1 might present in quick succession.

If resources (in blue) are allocated purely on a first-come, first-served basis, the first patient, who has suffered an unsurvivable burn, receives intubation, intravenous fluids and rapid transport to an intensive care bed despite his poor prognosis; he dies anyway. The patient with an uncomplicated lower limb fracture who would survive without intervention is cannulated, given intravenous fluids and transported to theatre for non-life-saving surgery. No resources are left for the last three patients, who despite being potentially salvageable all die (black lines), resulting in 80% mortality (fig 2).

Figures 1-4: the value of triage

Figure 1: initial presentations



90% burns | Lower limb fracture | Intra-abdominal bleeding | Allergic reaction with airway compromise | Acute ST-elevation myocardial infarction

Figure 2: first-come, first-served resource allocation (80% mortality)

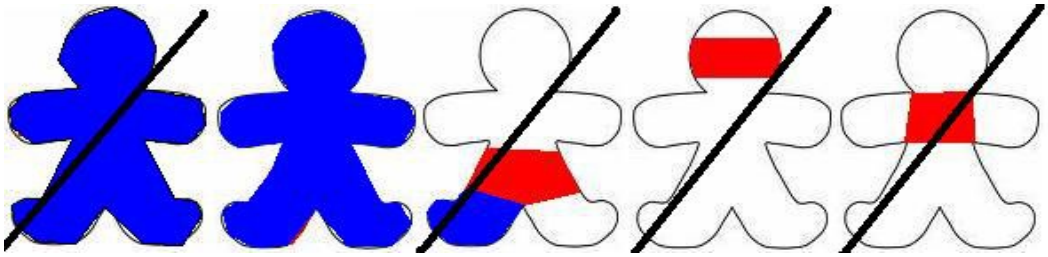


Figure 3: triaged resource allocation

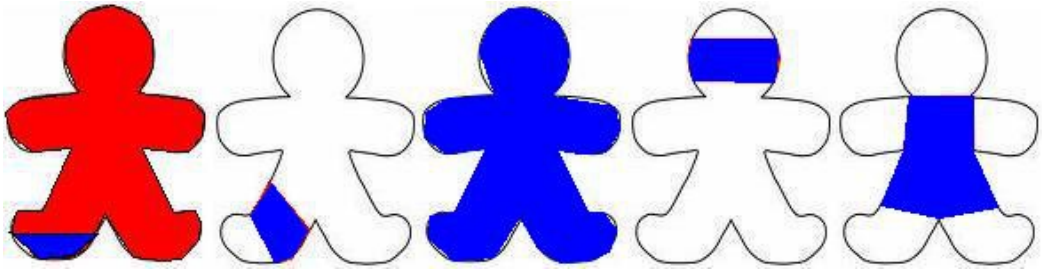
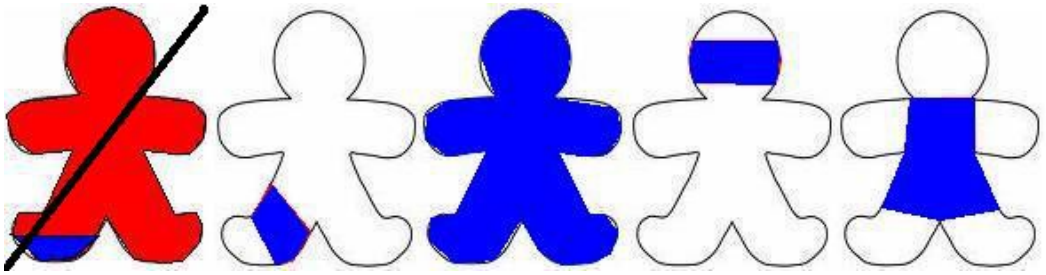


Figure 4: effects of triage (20% mortality)



If patients are triaged resources are triaged, the burn patient, who is irretrievably injured, receives less resource-intensive palliative care. The patient with a limb

fracture receives immobilisation, the patient with abdominal bleeding is treated for shock and transferred to an operating theatre early, the patient with the allergic reaction has her airway secured, and the patient with an acute ST elevation myocardial infarction has thrombolysis or primary percutaneous coronary intervention. This consumes the same resources as previously (fig 3). However, the targeting of treatment using triage results in a reduction in mortality to 20% (fig 4) (8).

Accurate assessment of illness severity has implications both for the individual patient and the health care system. Patients admitted to critical care areas via non-critical care areas (such as standard wards) have higher rates of mortality than those admitted directly from the ED (9-11), with significant numbers of patients (23/122 (9) and 144/343 (11)) admitted to critical care more than 24 hours after their ED attendance. The observational studies supporting this are of course subject to many confounders; the clinician who thinks of critical care for a patient may also be the clinician who understands the implications of delays in antibiotic therapy; thus early admission to critical care may be a proxy for higher quality care globally. Conversely, unnecessary admission, whether simply to hospital, or to a critical care area, is likely to expose a patient and family to extra stress and the health care system to excessive expenditure. There is also the potential for morbidity or even mortality caused by unnecessary interventions and healthcare associated infection.

If one assumes that a gold standard for defining patient acuity exists (this is discussed further in sections 1.2.4 and 1.4 below), undertriage occurs when a critically ill patient is not triaged to immediate or urgent care. Conversely, overtriage involves the allocation of a non-critical patient to high intensity care (table 1). The overtriage rate is calculated as true positive/(true positive plus false positive), analogous to one minus the positive predictive value of a test. The undertriage rate is calculated as false negative/true positive plus false negative, analogous to one minus the specificity. Although undertriage may intuitively seem much worse (a sick patient is sent to an area with inadequate resources), it has

been demonstrated in the major incident literature that the critical mortality of any incident (the proportion of patients surviving to hospital who then die as a result of the incident) is directly related to the rate of overtriage to that facility during that incident (12). The same is likely to be true outside the major incident arena; even if overtriaged patients (false positives) may come to no individual harm from unnecessary interventions, there is a system-wide opportunity and financial cost in using resources where they are not required. This has been recently highlighted in the context of the US debate over affordable care, with arguments advanced that "waste avoidance" in healthcare provision is a professional obligation (13). It is also possible that there is a "grey" group not included in this conceptualisation; those who will survive in a lower acuity area but whose morbidity would be reduced by higher acuity care.

Table 1: Under- and over-triage

	Critically ill patient	Non-critical patient
High acuity area	Correct high priority triage (True positive)	Incorrect high priority triage = OVERTRIAGE (False positive)
Low acuity area	Incorrect low priority triage = UNDERTRIAGE (False negative)	Correct low priority triage (True negative)

This represents a huge challenge in terms of patient assessment, with many EDs still staffed by relatively junior doctors and struggling to provide 24 hour senior shop-floor cover. ED staff are required to make decisions in a limited time and often with incomplete information in order to maintain patient flow through the ED. The National Confidential Enquiry into Patient Outcomes (NCEPOD) has documented widespread failings in the identification of sick patients and their escalation to appropriately senior staff (14), and the National Early Warning Score has been launched by the Royal College of Physicians in response to the problem (15).

1.2.4 Acuity assessment: what is patient acuity?

Severity of illness could be defined in a number of ways – risk of death in either the short or long term, magnitude of symptoms such as pain or nausea, effect on functional status and deviation from either clinical or laboratory "norms". I will

argue that patient acuity is not necessarily concurrent with illness severity. I suggest that the point of an emergency care system is to provide prompt care to those patients likely to benefit in a time-sensitive manner from interventions. These are not inherently the same patients who are at highest risk of death – some of these patients will progress to death irrespective of interventions. Nor are they the patients with the highest overall benefit from healthcare interventions – the young man with a testicular teratoma has massive potential for benefiting from treatment, but this will not be affected by whether he is seen within one hour or six hours in the Emergency Department. A measure of patient acuity should therefore reflect patients whose outcome will be improved with prompt care and/or those whose outcome will worsen without this care.

Unfortunately within emergency care little evidence exists as to the time-sensitive nature of many interventions (16). Although, as has been argued in the context of critical care outreach, it is intuitively appealing to define a deteriorating patient and respond rapidly (17), meta-analysis of tools to do this in an in-patient setting has failed to identify a benefit in terms of patient outcome (18-19). The introduction of a medical emergency team in one hospital was associated with a decrease in mortality amongst surgical patients but a sustained increase in mortality in medical patients, highlighting the non-congruence of risk of death with the potential to benefit from early medical intervention (20). Recent commentary on clinical decision rules has drawn attention to the disconnect between the identification of patients at risk of a particular outcome and the potential of those patients to benefit from available interventions (21). The evidence available for specific time-sensitive interventions will be discussed in section 1.4 below.

1.2.5 Acuity assessment: application in the clinical setting

One potential role of formalised severity assessment tools is to identify patients who need higher levels of care, in an attempt to prevent the pattern where a patient's physiological decompensation is carefully documented but no action taken in the hours before the patient's cardiac arrest (22). Of 383 hospitals surveyed by NCEPOD in 2009, 376 reported some kind of "early warning system" to identify the

deteriorating patient; 365 of these formally linked the early warning score to an escalation policy with a mandated medical and nursing response (14). Some hospitals have even adopted the paradigm to the extent that an automatic bleep to the critical care team is generated by the electronic patient record system when deranged observations are recorded.

However, it is possible that a similar philosophy could be adopted at the opposite end of the severity spectrum. The Pneumonia Severity Index was originally developed to identify patients with low risk of complications to support clinicians in discharging these patients home (23). There is particular potential value to this approach in situations where a significant proportion of patients presenting are the "worried well", who have symptoms that cause them concern but usually represent benign disease, such as headache or chest pain in the young. A tool to identify the low-risk patient would help to standardise care provided purely due to clinical risk aversion and potentially liberate resources to treat other patients who might benefit more.

In some cases severity of illness can provide guidance on the extent to which aggressive care is appropriate. If an inevitable death can reliably be identified, ceilings can be placed on care and a dignified death with adequate palliation achieved. Equally, some treatments are only justifiable in terms of risk-benefit ratio within particular severity parameters; notably stroke thrombolysis may have a window of benefit in moderate severity stroke in which the benefits justify the risks; this benefit is not seen in severe stroke (24).

1.2.5 Current severity assessment in acute illness

1.2.5a Ambulance dispatch assessment

Emergency ambulances in the UK are dispatched at three levels of urgency: categories A (patients who are or may be immediately life threatened), B (patients who require urgent face to face clinical attention) and C (patients who do not require an immediate or urgent response by blue light). Expectations for speed of response vary depending on the category to which the call is assigned, and in fact

ambulance dispatch is not required for all category C calls, which can be reassigned to an alternative care provider.

The allocation of dispatch category depends on a severity assessment obtained by asking the caller a standardised set of questions about the patient's condition. For example a patient with chest pain would be allocated to category A if clammy or vomiting, but category C if breathing normally and aged under 35; a patient who has apparently fainted is allocated category C if fully recovered and aged under 35, category B if female of reproductive age with abdominal pain and category A if still unconscious or having had multiple episodes.

These dispatch categories require only minimal assessment by the caller, as no medical training can be assumed on the part of the caller, and are therefore biased towards overtriage, the assumption that a patient is more ill and requires higher acuity care than is actually the case. The evidence for accuracy of dispatch priority systems is, however, mixed (25-26) and there is little data to assess whether patient outcomes are influenced by the use of such systems (27).

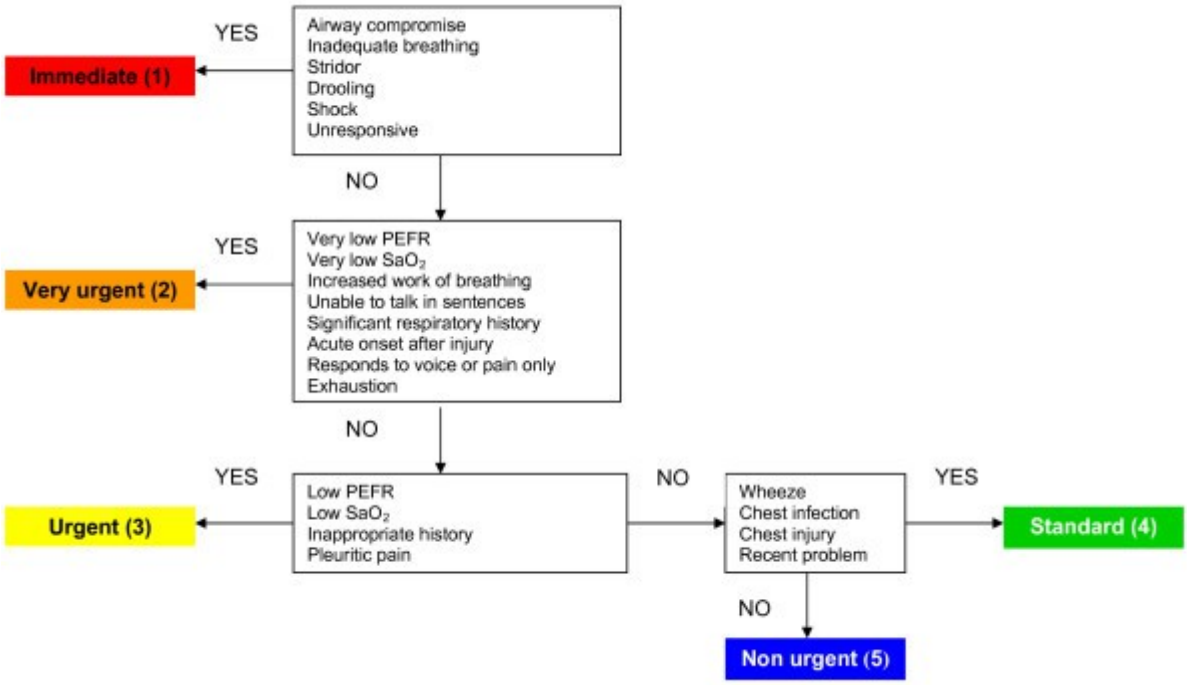
1.2.5b Emergency Department triage

Almost all UK EDs use some form of nurse-led triage system. This involves patients being seen within a short time-frame, as close to arrival as possible, to undergo a brief assessment. This triage assessment is then used to allocate patients to a specific area of the ED (resuscitation, majors, minors, minor injuries, primary care etc) and, if the patient is found to be in need of emergency treatment, to alert appropriate senior nursing and medical staff. A number of standardised triage systems exist.

The Manchester Triage System (MTS) (28) was developed and revised by a consensus group in the UK. It uses a set of presentational flowcharts (eg "abdominal pain – adult" or "unwell adult") with various discriminators (eg airway compromise, severe pain, abnormal pulse, recent onset) to allocate patients to five categories (red, orange, yellow, green and blue) which correlate to the urgency with

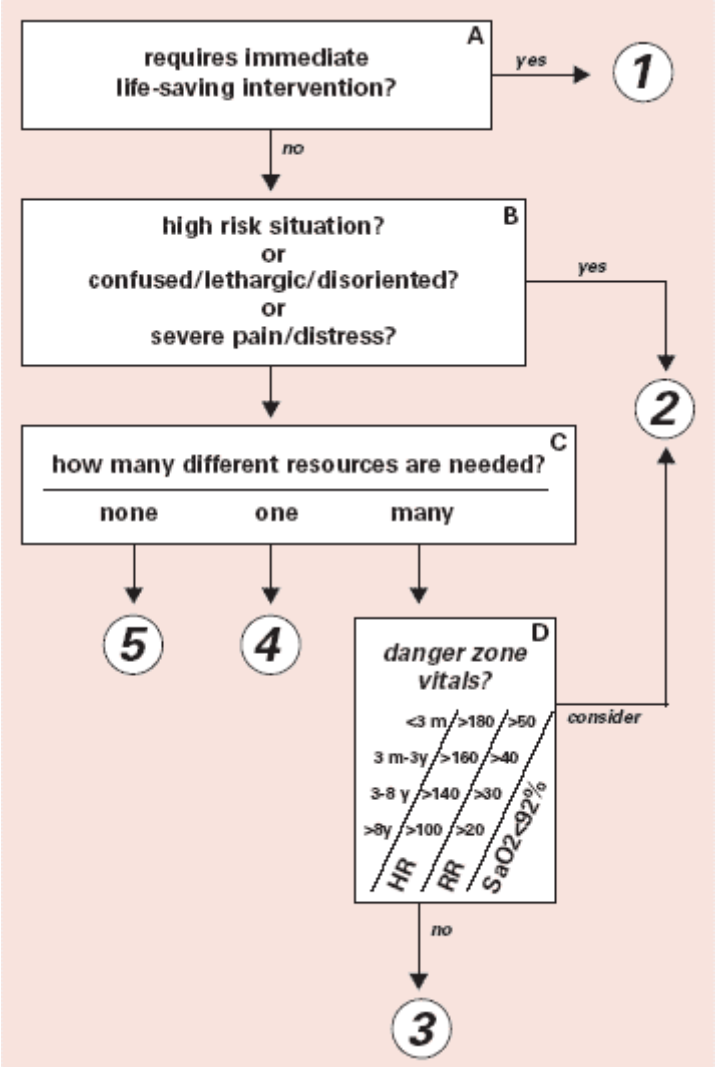
which they should be seen by a clinical decision-maker (immediately, under 10 minutes, under 2 hours, under 4 hours, non-urgent). Figure 5 shows an example MTS flowchart.

Figure 5: Manchester Triage flowchart for Shortness of Breath in Children



The Emergency Severity Index (ESI) (29), widely used in the US, combines the triage nurse's assessment of severity of illness (for categories 1 and 2), defined physiological variables (for categories 2 and 3) and the triage nurse's estimate of resource requirement (categories 3 to 5). The ESI flowchart is shown in figure 6.

Figure 6: Emergency Severity Index flowchart (from Agency for Healthcare Research and Quality)



The Canadian Triage and Acuity Scale (CTAS) (30) initially stratifies patients according to their presenting complaint, which allocates a minimum triage category (from 1 to 5) defined by the presenting complaint. This is then modified by vital signs (haemodynamic compromise, temperature and consciousness) and pain level, followed by modifiers specific to the condition (such as mechanism of injury for trauma or blood glucose in diabetics).

These various triage systems have been assessed in terms of inter-rater reliability (31) and their ability to identify high-risk patients (32-34), but analysis of the effect of their implementation on patient-oriented outcomes is minimal (35-36).

Some institutions have adopted RATS (Rapid Assessment and Treatment System) whereby a senior clinician is allocated to "eyeball" patients on or soon after arrival to initiate investigations and immediate treatment. This has been demonstrated to reduce the number of patients waiting (37) and time spent waiting at various stages of the process (38-39) but has not been shown to influence non-time-related outcomes.

1.2.5c Admission and in-patient severity of illness scoring

As stated above, almost all UK hospitals now have some form of physiological Early Warning Score in use in in-patient areas (14). This was originally based on the work of McQuillan et al, in a confidential inquiry into care preceding intensive care admission (10), and Goldhill et al, in analysing physiological data in the 24 hours preceding intensive care admission (22). Both these groups identified a pattern of clearly documented worsening physiological derangement which was either not recognised, not communicated appropriately, or not acted upon. It is now a national recommendation that all acutely ill patients should have an early warning score recorded at the point of the decision to admit to hospital, and that such a score should form part of routine observations, the frequency of which should be stipulated in the care plan (40). Despite this the NCEPOD investigation into cardiac arrest management found widespread failure to recognise acutely ill patients and to involve appropriately senior personnel in their care (14); it appears that simply the existence of a scoring system is insufficient to have an impact on patient outcomes.

1.3 Clinical decision aids

Clinical decision aids exist in a number of contexts; they can be used to guide diagnosis and inform on prognosis, they can take the form of cumulative scores or absolute decision rules and they can be applicable to broad populations or restricted to specific disease sets.

1.3.1 The role of clinical decision aids in the clinical environment

It might be argued that clinical decision aids would be unnecessary were clinicians appropriately trained and competent. Indeed, it has been demonstrated that in

identifying patients at low risk of pulmonary embolism, the gestalt of a senior clinician is as good as any validated decision aid (41). However, care is not provided universally by senior clinicians, and the value of safety checklists in high-stakes situations has been demonstrated following the adoption of the WHO surgical safety checklist (42). A recent CEPOD analysis identified recurrent failures of clinical staff, particularly those with less experience, to recognise and act on severe illness (14), and similar patterns have been identified in the Emergency Department triage process (43). There is also an argument that in terms of fairness and medicolegally it is beneficial to objectify the decision-making process as this may have potential to standardise both the process and its recording.

1.3.2 Diagnostic and prognostic decision aids

The difference between diagnostic and prognostic decision aids is probably best illustrated in the context of pulmonary embolism. The Wells and Geneva diagnostic scores use a variety of features of the patient's history and examination to calculate a probability that the patient is suffering from a pulmonary embolism, and therefore to guide the investigative strategy. They provide no information, however, about the severity of the embolism if that is the final diagnosis. Conversely, the Pulmonary Embolism Severity Index (44) calculates a risk of poor outcome in patients known to have pulmonary embolism, but is of no value in making the initial diagnosis.

1.3.3 Clinical scores and decision rules

Clinical decision aids can supplement risk assessment in two different ways: a decision rule will provide an (almost) absolute rule-in or rule-out based on a number of criteria; once ruled-in or ruled-out no further conclusions can be drawn about the relative risks of two different patients fulfilling the criteria. A clinical score will provide ranking information on the relative risk between two patients, and can be repeated over time to establish trends in improvement or deterioration, but tends to involve more variables (often weighted) and be more cumbersome. This difference is well illustrated by the two stages of calculating the Pneumonia Severity Index (23). Stage 1 is a clinical decision rule: if the patient is negative for 3

criteria (1. age over 50; 2. history of neoplasm, congestive heart failure, cerebrovascular disease, renal disease and liver disease; 3. abnormalities on examination: mental status, pulse, respiratory rate, blood pressure and temperature) he is considered to be class I (very low risk). A patient not fitting into class I is then assessed using a score, with weighted values for various demographic and historical factors and examination and laboratory findings.

1.3.4 General and disease-specific decision aids

Clinical decision aids can be applicable across a wide unselected population, or specific to patients with a particular diagnosis. Specific decision aids may be more accurate within their patient population, but all require an accurate diagnosis to be made; their role may therefore be limited in the ED where specific diagnosis is often not possible given time constraints and incomplete information, and are vulnerable to misdiagnosis. A patient who is low-risk according to the CURB-65 pneumonia severity score is likely to be at high risk if his pneumonia is in fact a misdiagnosed dissecting thoracic aneurysm, while a generic score such as APACHE II might have identified his severity-of-illness without providing specific information as to the source of that illness. Generic scores also allow (to some extent) the prioritisation of patients with a variety of conditions; this is likely to be of particular relevance in resource-limited situations such as pandemics, and the application of scores to that purpose has been discussed in the pandemic planning literature (45-46).

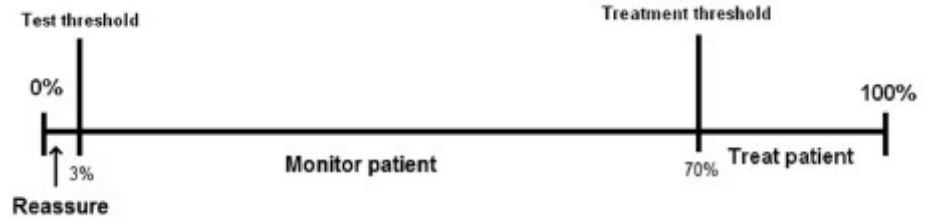
1.3.5 The ideal decision aid

I suggest that an ideal clinical decision aid has three major characteristics – ease of use, accuracy and reproducibility. These are reflected in Stiell and Wells' processes for the development of a decision aid discussed below. In order to be adopted and useful in the clinical environment, it needs to be easily and readily recalled and applied, using variables that are readily accessible in that particular environment. Thus an aid requiring knowledge of the pulmonary artery wedge pressure is unlikely to be useful in an Emergency Department but may be of great value in a cardiothoracic critical care area. The aid should be as accurate as possible; failures to predict poor outcomes (false negatives) will not only risk compromising patient

care but may also reduce staff confidence in the tool, while large numbers of false alarms (false positives) are likely to result in the tool being ignored. Finally, an ideal aid will produce the same prediction whether applied at 10am by the senior sister or 3am by the FY2 doctor; this is particularly important in emergency care with its high level of both patient and staff turnover.

It has been suggested that in order to be useful, medical information should be relevant, valid and easy to access (47). Decision rules that relate to diagnosis or prognosis are related to the concept of test and treatment thresholds; ideally the rule is defined in the clinical context and provides cut points that coincide with risk levels where the risk-benefit balance is clearly in favour of treatment or not (see figure 11 from Ebell (48)). Thus a rule with a cut-point based on maximising the sum of sensitivity and specificity may not be clinically helpful as it may not coincide with the clinical treatment threshold (see further discussion in 1.9.2 below).

Figure 7: test and treatment thresholds



Data from implementation studies of computerised clinical decision support systems suggests that they can positively influence practitioner behaviour, but evidence for benefit in terms of patient outcome is more limited (49), and calls have been made for further investigation of their integration into real-world practice situations (50). It appears that decision support systems that are integrated into existing data management practices (whether paper or computerised) are more likely to influence patient outcomes, as are those that require a clinician to provide a justification for deviation from system recommendations (51).

1.3.6 Development of a decision aid

The processes for the development of clinical decision rule have been elucidated by Stiell and Wells, both authors of rules which are widely used in emergency medicine (52). They describe six steps (table 2).

Table 2: Stiell and Wells processes for development of a clinical decision rule

1. Identifying a need	a. Prevalence of relevant condition – common enough for a rule to have a reasonable impact. b. Current use of the related diagnostic test – the currently used test has low enough specificity for a rule to improve on this. c. Variation in practice – similar institutions and practitioners vary in their use of the diagnostic test. d. Attitudes of physicians – adoption of a rule is more likely if physicians suspect that their current practice involves unnecessary testing. e. Clinical accuracy – potential for clinical examination to make or exclude a diagnosis.
2. Deriving the rule with appropriate standards	a. Defining outcome – clinically important and clearly defined; assessed by a blinded observer. b. Defining predictor variables – clearly defined and ideally collected in a prospective standardised manner. c. Reliability of predictor variables – good intra- and inter-rater reliability. d. Selection of subjects – subjects well-defined; subjects and setting clearly described so that generalisability of the rule can be assessed. e. Sample size f. Mathematical techniques g. Sensibility of the decision rule – rules are more likely to be adopted if clinicians consider them to have content validity, ie if they include plausible variables and have not excluded any variables widely felt to be important. h. Accuracy – presented in terms of sensitivity and specificity.
3. Prospective validation	a. Prospective validation – new patient population, ideally with different clinicians in a different setting. b. Selection of subjects – ideally an unselected set of patients with the condition in question. c. Application of the rule – including adequate training of clinicians in its use. d. Outcomes – ideally all patients would undergo the gold standard test as used in the derivation study, but proxies can be substituted. e. Accuracy of the rule f. Reliability of the rule – its intra- and inter-rater reliability in

	the validation population. g. Physicians' interpretation – accuracy of clinicians' interpretation, and ease of use. h. Refinement – this phase may result in a modified or streamlined version of the rule. i. Potential effect
4. Successful implementation into clinical practice	a. Clinical trial – ideally a randomised controlled trial, but before-and-after studies may also be used. b. Effect on use c. Rule accuracy d. Acceptability
5. Cost-effectiveness of using the rule	
6. Dissemination and implementation of the rule	

McGinn further quantified the process of validation and the clinical confidence that could be placed in a decision aid (53) (table 3).

Table 3: McGinn levels of clinical confidence in a decision aid

Level 1: Rule can be used in a wide variety of settings with confidence it can change clinician behaviour and improve patient outcomes.	At least 1 prospective validation in a different population and 1 impact analysis demonstrating change in clinician behaviour with beneficial consequences.
Level 2: Rule can be used in various settings with confidence in its accuracy.	Demonstrated accuracy in either 1 large prospective study including a broad spectrum of patients and clinicians or validated in several smaller settings that differ from one another.
Level 3: Rule that clinician may consider using with caution and only if patients in the study are similar to those in the clinician's clinical setting.	Validated in only one narrow prospective sample.
Level 4: Rule that needs further evaluation before it can be applied clinically.	Derived but not validated or validated only in split samples, large retrospective databases or by statistical techniques.

The issues of identifying a need (Stiell stage 1) are addressed in section 1.2 above. Topics underlying rule derivation and validation (stages 2 and 3) are in sections 1.4 to 1.9 and 1.10 respectively. Stages 4 and 6, of implementation and dissemination,

are outwith the scope of this thesis, but issues around operationalisation are explored in chapter 10.

1.4 Outcome measures

Prognostic clinical decision aids can predict a variety of outcomes; the most frequently used is patient death, but non-death events and composite outcomes are also used. This section will explore the different properties of outcome measures and their relevance in emergency care.

1.4.1 Death

Death is arguably the outcome that matters most to patients and their families. However, there are a number of issues in using only death as an outcome measure. In some patients, death may reflect the end stage of a chronic disease process, rather than an acute event or derangement. There are some occasions where clinician, patient and family have agreed that interventions aiming to extend life are futile and/or not in the best interests of the patient; there may be many benefits to identifying these occasions early in order to facilitate good palliative care. Apart from these cases, discrimination between "inevitable" and "preventable" death is notoriously observer-dependent (54) and clinicians have not yet demonstrated the ability to prognosticate reliably even in restricted specialist fields such as intensive care (55).

In health care systems with well developed critical care, the use of death as an outcome measure for early warning systems may become a self-denying prophesy; a patient initially at high risk of death is identified by the early warning system, treated adequately and promptly and therefore becomes a patient with a low risk of death. As addressed in 1.2.4 above, patient acuity does not necessarily equate to longer-term probability of death. This has been further explored in the trauma literature; patients with low probabilities of death overall may still require emergent care. Trauma scoring can include the use of an Abbreviated Injury Score (AIS), which rates the injuries in each anatomical area on a scale from 1 (minor) to 5 (critical) and 6 (unsurviveable) (56). A chest wall injury causing tension

pneumothorax may well kill a patient rapidly but would only be coded with an Abbreviated Injury Score of 2 (57).

1.4.2 Non-death adverse events

Non-death adverse events may reflect either process or outcome.

It has previously been assumed by some authors that immediate acuity can be retrospectively identified using long-term mortality (this is the premise of injury severity scoring). However, a relatively low mortality condition, (eg facial angioedema) may cause a patient to be of high acuity, requiring an emergent intervention (airway management) to prevent death. It has therefore been more recently suggested that acuity could be better quantified using process outcomes of various urgency, such as emergency airway interventions or surgery (58-59).

Process events might include admission to critical care areas or procedures such as emergency surgery or cardioversion that eventually result in a good outcome for the patient. The use of these events as outcome measures has been criticised on the grounds that they are harder to define than death, and may be subject to individual clinician variation (for example if using intensive care admission as an outcome, one clinician may be unduly pessimistic about the benefits of ventilation for patients with COPD, so will admit very few, whereas a different clinician has a particular interest in long-term ventilatory weaning so will admit significantly more). They may also be subject to variation in external resources; a process measure of emergency surgery within six hours may vary between a hospital with on-site vascular surgery where the emergency theatre is occupied by a patient undergoing an aneurysm repair and another hospital with no vascular surgery where the theatre is immediately available for a patient with the same severity of illness. A major flaw in the use of process measures is their potential circularity; if, for example, patients with a low blood pressure receive aggressive fluid resuscitation, low blood pressure will become a predictor of aggressive intervention without there necessarily being any evidence of benefit from that intervention. This is difficult to illustrate from the literature as many emergency interventions have been inferred

from basic principles to be effective without rigorous supporting evidence (this is discussed further in section 3.7.2 below). Intuitively it seems that requiring evidence for the efficacy of identifying patients with airway obstruction is as perverse as requiring evidence for the efficacy of parachutes when jumping out of planes (60). However amongst mid-range acuity conditions this does not hold; it is not clear that identification of small bowel obstruction in four hours rather than two changes outcome, and it is unlikely that primary research will ever address this point. Some examples exist, however, including a Scottish series which argued that although depressed consciousness is taught to be an indication for endotracheal intubation, some patients with severely reduced consciousness can be managed safely without (61).

Non-death outcome measures include measures of global outcome (the Glasgow Outcome Score), functional status (the Barthel index), psychological health (Hospital Anxiety and Depression Scale) and measures of health-related quality of life (EQ5D). It is arguable that these may reflect quality of health care with more subtlety than crude death rates, but they are more complex to administer, require patient (or relative) co-operation and are more vulnerable to criticisms of response and expectation bias than hard outcomes such as death. The usual process of collecting quality of life data using patient self-reporting further potentiates bias as certain patient sets (for example the unconscious or acutely critically ill) may be less likely or able to respond.

1.4.3 Patients with the potential to benefit

Overlapping with both death and non-death groups as above are patients with the potential to benefit from health interventions. Within this thesis the potential to benefit will be framed in terms of avoidable premature mortality. Although it must be recognised that many health interventions also aim to improve quality as well as quantity of life, assessment of quality of life outcomes is outwith the scope of this study.

Patients with the potential to benefit are those in whom mortality could have been prevented had an intervention been provided, or those in whom mortality has been prevented by the provision of said intervention. Thus not all patients who die had potential to benefit, as for some death may have represented an inevitable end point of an unalterable disease process. This "preventability" concept is explored by Hillman, whose structure for identifying preventable adverse events amongst in-patients excludes those patients in whom death is felt to be inevitable (those deemed "not for resuscitation") (62) (table 4).

Table 4: Hillman structure for preventable adverse events

			Potential preventability
Deaths	Minus all 'not for resuscitation' patients	Potentially unexpected deaths	All events preceded by pre-defined criteria which have not been appropriately acted on
Cardiorespiratory arrests	Minus all 'not for resuscitation' patients	Potentially unexpected cardiorespiratory arrests	
Intensive care admissions	Minus all elective or expected admissions from operating rooms, recovery area, ED	Unanticipated admissions to ICU	

It is possible, even likely, that the overlap between risk of death and potential to benefit from life-saving treatment varies between subgroups of patients. Some patients with a high risk of death (many previously young fit patients with major trauma) have a high likelihood of benefiting from timely care (left in figure 8); a different patient with high risk of death (the elderly patient with metastatic cancer and little cardiopulmonary reserve) has minimal chance of his or her life being saved, however prompt the treatment (right in figure 8).

Figure 8: Overlap between risk of death and potential to benefit

1.4.4 Composite outcome measures

Probably the most widely known composite outcome measure is MACE – major adverse cardiac events – comprised of death, acute myocardial infarction or urgent coronary revascularisation, and sometimes including life-threatening arrhythmia and new onset heart failure, within a given timeframe. Composite outcome measures may seem attractive, in that they include all outcomes that patients are likely to wish to avoid, and may aggregate individually rare outcomes to a point where they are more statistically manageable. However, they risk all the flaws of their individual components, and assume that all component measures are equally unattractive (ie that it would be as unwelcome for a patient to undergo percutaneous angioplasty as to die). Use of a composite outcome measure to develop a prognostic rule also implies an assumption that the risk factors for each of the components are the same.

1.4.5 Timescale for outcome measures

The most widely used timeframe for analysis of outcome is 30 days, or the length of the inpatient stay; over half of the literature analysed in chapter 2 used these timeframes. It is unclear whether this is the most appropriate timeframe to use when analysing events and care in the emergency department as it is so open to confounding by the quality of ongoing care, the patient's underlying chronic state and unrelated chance occurrences. It may be that longer timescales are required for some patient subsets where issues of quality take longer to manifest.

1.5 Potential predictor variables

1.5.1 Demographic variables

Age is included as a variable in a number of well-known prognostic scoring systems; ABCD2, which predicts risk of stroke after transient ischaemic attack, gives a point for age ≥60 (63), APACHE II, predicting intensive care mortality, uses a cut-off of 45 (64), and CURB-65, adopted by the British Thoracic Society for mortality prediction

in pneumonia (65) and TIMI, for predicting major adverse cardiac events (66) both include age 65 or over. It is not clear whether increasing age is associated with increasing disease acuity or simply with a globally increased risk of all-cause mortality, and this will be explored in the thesis.

Sex is not widely included in prognostic scores, although it has been found that females have a higher risk of mortality after acute myocardial infarction (67) and of cardiogenic shock after cardiac thrombolysis (68), and that males have a higher risk of ICU admission with pneumonia (69). The prognostic effect of gender may be confounded by a higher proportion of females presenting with "atypical" symptomatology or by a preponderance of women amongst the oldest old, and this will be further explored.

1.5.2 Chronic health and functional state

Chronic illness is included as a variable in many prognostic aids including APACHE II (hepatic, cardiovascular, renal, respiratory and immunosuppressive (64)), the Blatchford score for upper gastrointestinal bleeding (hepatic and cardiac (70)), the Pulmonary Embolism Severity Index (cardiopulmonary (71)) and the Pneumonia Severity Index (neoplasia (23)). However, an apparently standardised diagnosis may not represent the same additional risk in all populations – a patient newly diagnosed with diabetes in a deprived resource-poor area with low health literacy is likely to have suffered more pathophysiological damage before diagnosis than a patient who is newly diagnosed in an affluent area – the "constant risk fallacy" (72).

The use of a global functional state to predict prognosis is well-recognised in other medical specialities, particularly oncology, where the Karnofsky (73) and other scores are used to assess fitness for chemotherapy and other interventions. This is less prevalent in emergency care, although the MEDS (Mortality in ED sepsis) allocates points for nursing home residence (74), and the PMEWS score includes points for limitation to any activity other than strenuous work (75).

1.5.3 Acute physiology

If one assumes that physiological derangement in response to illness represents the same risk across all populations, a measure of acute physiological derangement would provide a more robust tool for prognostication. This assumption is likely to be defensible as biologically-based physiological measurements can be expected to have a relatively constant association with risk, unlike diagnostic labels which are more likely to be setting-specific. This does not guarantee that physiological measures are constant predictors of risk across all populations; it is clear from paediatrics that the association varies with patient age and it may well be that this variability continues across adulthood. However, physiological predictors will hopefully be more robust across healthcare settings.

There are, however, potential problems with the use of physiological data to prognosticate. The DaVROS study showed that even physiology is subject to a constant risk fallacy – ie that the same physiological derangement does not confer the same risk across different settings; this may reflect variations in underlying population "normality", different practices in the measurement and documentation of physiological variables, or the effect of varying caregiver responses to physiology (76).

It has generally been assumed in the medical literature that most physiological variables are distributed in a Gaussian (normal) distribution, with a mean in the healthy population from which deviation is correlated with severity of illness. There are potential flaws in this assumption; firstly that the ranges quoted in the medical literature as "healthy" accurately represent the population. This has been questioned in the paediatric field, where meta-analysis of observational studies produced centile ranges that differed substantially from reference ranges quoted in life support courses (77). Tarassenko et al, in analysing a large adult vital signs database, found mean values amongst acute inpatients of 84.2 for pulse rate, 18.6 for respiratory rate, 96% for SaO2 and 128.5 for systolic blood pressure (78).

Secondly, it is unclear whether the "normal" range in the well population is the same as that associated with the best prognosis in the acutely ill population; it may

be, for example, that some degree of tachycardia during acute illness represents an appropriate physiological response and the lack of this is in fact a marker of poor prognosis. The published literature includes cut-offs for pulse rates that carry a higher risk of adverse outcome at 80, 90, 100, 110, 113, 115, 120, 125 and 130 beats per minute, neatly dissecting both the "normal" and "abnormal" ranges. It has often been assumed in the existing literature that a linear relationship exists between physiological variables and adverse outcome (ie, the faster the pulse, the greater the risk of death). This is not supported where the assumption has been explored and relationships exist in the form of U-shaped curves (for pulse in acute coronary syndrome and blood pressure in stroke), J-shaped curves (blood pressure in stroke), inverse exponentials (blood pressure in acute coronary syndrome) and interactions between two variables (pulse and blood pressure in sepsis).

1.6 Selection of subjects

Prognosis researchers must define *a priori* subjects to be included in the data set. This involves a decision about whether the inclusion and exclusion criteria should be narrow or pragmatic. For example, almost all prognostic studies of community-acquired pneumonia exclude patients who are immunosuppressed, and require identification of infiltrates on chest Xray by a radiologist. This limits generalisability to routine practice, where many patients are unknowingly HIV positive or taking steroids, and where radiologists rarely review all ED chest Xray films. It would appear therefore that for the development of tools to use in the ED setting, broad pragmatic inclusion criteria are to be preferred. This risks, however, confusion where a particular variable has different effects in different subgroups; a study of all patients seen in the resuscitation area might conclude that a systolic blood pressure of 180 has prognostic benefit, based on patients with acute coronary syndrome or pulmonary oedema, when for the subgroup with stroke it is a marker of poor prognosis.

1.6.1 Prospective versus retrospective data collection

Prospective data collection has the benefits of specific standardised measurement of both variables and outcomes, and overcomes the requirement for blinding during

the extraction of retrospective data. However, data collection may be a confounder in itself, where the data collection tool may trigger measurement of variables which would otherwise not have been carried out and which then alter treatment decisions.

In some rare conditions it may not be practically feasible to collect purely prospective data, and retrospective disease registries may be of value. This is unlikely to be applicable to this study.

1.6.2 Source of data set

It is generally accepted that the optimal data set from which to derive a clinical decision aid is a cohort study or the placebo arm of a randomised controlled trial (both arms can be combined if the treatment had no significant effect). Use of data from a randomised controlled trial may offer longer follow-up times (if these are relevant) for fewer resources, but there are likely to be more problems with generalisability given the usually tighter recruitment restrictions for randomised controlled trials. It is also conceivable that willingness (or not) to participate in a randomised controlled trial could be a prognostic feature in itself. Although a case-control model can be used this is limited by lack of knowledge of the size of the source population (79).

1.7 Sample size

Statistical simulation studies have suggested that for use of logistic regression analysis, a cohort including ten outcomes events for interest per variable assessed provides an acceptable precision of regression coefficients (80). This rule has been challenged more recently (81), but this thesis will adopt the more conservative estimate. This will be discussed in more detail in section 3.10 below.

1.8 Mathematical techniques

Clinical decision aids can be developed using a variety of statistical methods, the two most common being logistic regression analysis and recursive partitioning. Logistic regression generates β-coefficients that simultaneously weight the value of variables in predicting a dichotomous outcome. These coefficients can then be rounded to create a clinically-useable score. Logistic regression models can provide good overall accuracy but may lack the ability to generate very high sensitivity or specificity if this is what is required. Recursive partitioning sequentially identifies dichotomous predictor variables that divide a population into subgroups with only a particular binary outcome of interest. This is particularly effective for generating decision rules with very high sensitivity, and generates a rule that does not require the user to make calculations, but requires arbitrary dichotomisation of continuous and categorical data. This thesis will therefore use regression techniques.

1.9 Assessment of accuracy

1.9.1 Discrimination and calibration

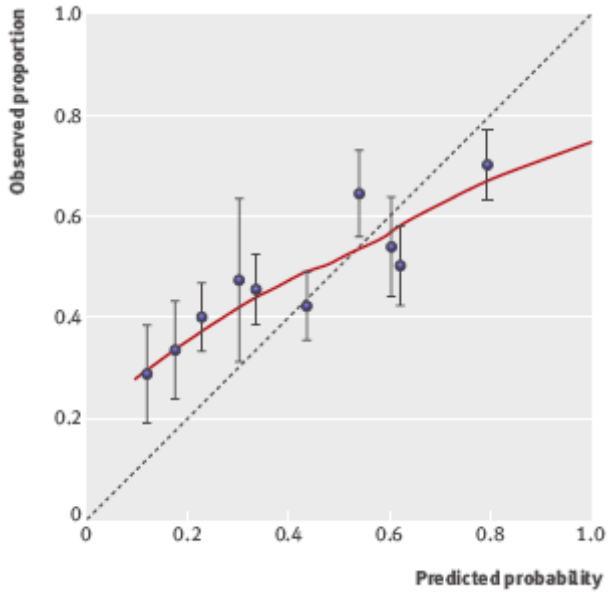
Decision aids providing a probability of an adverse outcome can be assessed in two main ways: discrimination and calibration.

Discrimination is the ability of a decision aid accurately to identify that a patient who suffers the outcome is at higher risk of the designated outcome than a patient who does not suffer the outcome (ie the probability that a randomly selected patient with the designated outcome will have a higher score than a randomly selected patient without the designated outcome). This is usually presented in terms of the area under ROC (receiver operator characteristic) curve, also known as

the c-statistic. The ROC curve plots sensitivity at various threshold values against 1-specificity in order to quantify the trade-off between increasing sensitivity and falling specificity (or vice versa). An ideal decision aid would have a c-statistic of 1.0. A c-statistic of 0.5 is no better than chance, while realistically a good model achieves a c-statistic of 0.8 or above.

Calibration is a measurement of the agreement between the predicted risk of death in subgroups of the population (usually demarcated by deciles of the risk score or deciles of risk of death) and observed rates of death in the same subgroups. The most frequently presented measure of this is the Hosmer-Lemeshow test, a modification of the chi-squared test. Figure 9 below is a graphical presentation of a calibration test from Altman (82), where the observed and predicted proportions of an outcome are plotted – in this example calibration (shown by the red line) is poor.

Figure 9: graphical representation of calibration



1.9.2 Trade-off of sensitivity and specificity

In order to be clinically relevant and useful, a decision aid must make an acceptable trade-off between sensitivity (ensuring that all patients with the condition have a positive test result) and specificity (ensuring that patients without the condition have a negative test result). The ideal test would be 100% sensitive and 100%

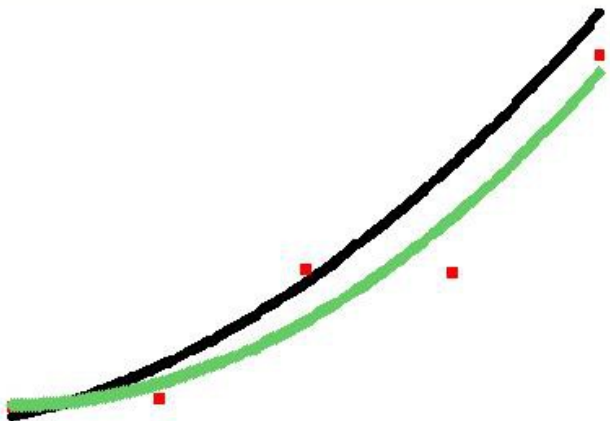
specific (the definition of a gold standard or reference test) but such tests remain elusive in the near-patient environment.

An acceptable trade-off between sensitivity and specificity will depend on the potential benefit of identifying true positive outcomes (a function of the severity of the outcome and its amenability to treatment), the potential harm (in health and economic terms) of treating false positive outcomes, and the individual patient's preferences. For example a decision rule to dictate CT brain scanning in patients with head injury would need to have high sensitivity if the outcome measure was immediate neurosurgical intervention (as a false negative rule result could be disastrous), whilst lower sensitivity would be tolerated in an attempt to avoid radiation exposure if the outcome measure was radiologic findings not requiring any intervention. Conversely a decision rule for the diagnosis of appendicitis would need to have high specificity (to minimise unnecessary surgical procedures) but might tolerate lower sensitivity as it could be reapplied to patients with an evolving presentation.

1.10 Validation

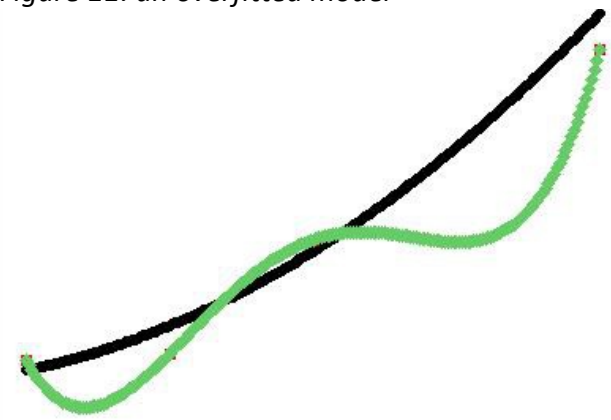
The validation of clinical decision aids addresses two potential problems. Firstly, clinical decision aids tend to perform better in the data set from which they are derived – known as "over-fitting". In figure 10, the black line represents the true relationship between two variables. The red dots represent a sample from which the relationship is estimated, generating the green line.

Figure 10: a well fitted model



In an overfitted model, shown in figure 11, the green line fits the sampled data more closely (obscuring the red dots) but is a less accurate estimate of the true relationship.

Figure 11: an overfitted model



Internal validation addresses this by providing an alternate set of sampled data.

Secondly, decision aids perform better in populations similar to those from which they were derived (the similarity may be explicit in terms of the variables included in the decision aid or occult in terms of unmeasured variables). External validation addresses the problem of generalisability to other institutions and health care settings.

1.10.1 Selection of data set

Internal validation involves the splitting of a single data set into two subsets before derivation of the decision aid from one subset, with subsequent validation on the second subset. This can be combined with "bootstrapping", where resampling is randomly undertaken from the full data set. This has a number of practical benefits in that the data collection infrastructure only has to be established once and piloting issues are minimised. However, it risks overestimating the performance of the decision aid in other settings, as the validation population should be very similar to the derivation population. A partial response to this is temporal validation, where a temporally separate dataset is collected at the same institution(s) as the

derivation dataset. This does not, however, address the issue of similarities in the patient population likely to present to the same setting.

External validation involves testing the rule against a dataset from a different institution. This dataset can reflect closely the inclusion and exclusion criteria from the derivation study, or can be much broader. Broader datasets, however, tend to demonstrate poorer performance of the decision aid; a model derived to predict complications in young ambulatory patients with pneumonia is likely to suffer if applied to a population where a significant number of patients are bedbound in nursing homes.

1.11 Summary

- Acute illness and acute severe illness present a significant logistical challenge to the NHS.
- Severity assessment of acute illness has a role in guiding immediate care and in case-mix adjustment for research and audit.
- Clinical decision aids exist in a number of forms and have been used to predict a number of outcomes.
- Multiple issues of quality exist in the development and testing of clinical decision aids.

Chapter 2
Literature review

2.1 Introduction

The literature review was undertaken to inform the empirical work in terms of identifying structures of decision aids suitable for emergency care, potential variables which are both feasibly collected in the emergency situation and may predict relevant outcomes, and gaps in the current evidence base around decision aids in emergency care. Specifically, it aimed to identify existing tools to identify critical illness in the acutely ill, ideally in an undifferentiated population.

During the conduct of this research, the Royal College of Physicians launched a National Early Warning Score developed by a multidisciplinary consensus panel. This was in response to "the multiplicity of early warning systems used in different hospitals in the UK …causing a lack of consistency in detecting deterioration of patients' conditions and calling for urgent medical help". The NEWS is advocated, for the purposes of standardisation "during the initial prehospital and/or hospital assessment of a patient and throughout the patient's hospital stay", although the development group did not include emergency physicians or nurses as stakeholders and the group was unable to identify any relevant literature relating to Emergency Department patients (15).

2.1.1 Background

The literature review aimed to address three broad points:

1. Format and scope of existing clinical decision aids, with particular attention to an undifferentiated patient group.
2. The role of physiological variables in existing clinical decision aids;
3. Evidence to support the predictive value of individual physiological variables.

2.1.2 Format of clinical decision aids

As discussed previously, clinical decision aids can take a number of forms. There is no widely accepted standard format and it may be that a variety of formats are required for different situations; it may be reasonable to trade off some reduction

in performance for simplicity of application where a decision aid is to be applied to a large number of patients, while in a patient set with low numbers but high-stakes decision-making a more complex model is appropriate. The most appropriate format of aid may also be influenced by the availability of technology; the increasing availability of near-patient and smartphone apps will enable the real-time use of more complex modelling.

The first aim was therefore to summarise the various formats of clinical decision aid reported.

2.1.3 Scope of existing clinical decision aids

Attempts to implement risk-prediction methods for both clinical decision-making and audit and research are hampered by the substantial range and number of risk scores available. There are so many potential scores for non-trauma patients that deciding which score should be used and which outcome is most relevant presents a challenge in itself.

The aim was to summarise the scope of clinical decision aids which related to short-term outcomes and were applicable at the point of patient presentation to unscheduled healthcare services (excluding trauma, paediatrics and purely obstetric or psychiatric presentations). This would support the underlying aim of the thesis, to identify prognostic factors in emergency patients and construct a bedside tool to support their care.

2.1.4 Existing clinical decision aids for general medical emergencies

As discussed in chapter 1, the applicability of disease-specific decision aids is limited by the impracticality of always reaching a definitive diagnosis in the Emergency Department. They can also be misleading if misdiagnosis has occurred, and do not allow for comparison of severity of illness between patients with different diagnoses. The aim was therefore to identify the extent to which any decision aids relating to general unselected medical emergencies had been developed and validated.

2.1.5 Physiological variables in existing clinical decision aids

It appears that inclusion of physiological variables in existing scoring systems is inconsistent, both in the variables included and in their weighting and cut-off values for "abnormality". For example, if a patient with sepsis secondary to pneumonia is assessed using a sepsis score (MEDS), a respiratory rate over 20 is a marker of severity, while if the same patient is assessed using a pneumonia score (CURB-65), his respiratory rate needs to be over 30 before it is considered a marker of severity.

The aim was therefore to summarise the inclusion of physiological variables in existing clinical decision aids, and the various cut-points at which physiological derangement was considered significant.

2.1.6 Evidence for the use of individual physiological variables

Although it appears self-evident that deviation from physiological norms is likely to be associated with adverse outcome, there is in fact little evidence to support widely-used definitions of "normal", particularly in the acutely unwell patient. For example, although various life support courses may teach that the "normal range" of pulse rate in adults is 60 – 100 beats per minute, a pulse of over 90 is one of the criteria for systemic inflammatory response syndrome (SIRS), while one pneumonia score (PSI) does not include tachycardia as a marker of severity until the rate reaches 125 beats per minute.

Non-physiological markers such as pre-existing diagnoses and treatment have been identified as predictors of outcome in a number of situations. These may be problematic in the emergency situation, where the patient may be too unwell to communicate, may have forgotten information previously conveyed to them or may not have had all relevant investigations (not all patients with clinically diagnosed heart failure have undergone echocardiographic quantification of ejection fraction, for example). Pre-existing diagnoses are also subject to the "constant risk fallacy" (72) whereby the same diagnosis does not confer the same risk in different populations; a diagnosis of diabetes in a population which generally does not

engage with health care is likely to be worse prognostically than the "same" diagnosis in a population with well established early screening programmes and widespread community engagement. There is little evidence as to whether physiological variables are equally applicable across all patient populations, as has previously been assumed, although the DAVROS study indicated that they too exhibited non-constant risk (76).

The aim was therefore to summarise the evidence supporting the value of individual physiological variables as predictors of various adverse outcomes.

2.2 Methods

2.2.1 Search strategy

The literature review used a deliberately inclusive strategy, which aimed to identify literature relating specifically to severity scores and also that relating to prognostic indicators, whether or not incorporated into a score.

The first strategy was to search generically for scores or aids applicable to undifferentiated patients in emergency or acute care. This involved searching Medline 1946 – 2013 using the strategy as below:

1 clinical prediction aid.mp. (0)
2 clinical prediction rule.mp. or Decision Support Techniques/ (10582)
3 clinical decision aid.mp. (17)
4 clinical decision rule.mp. (211)
5 score.mp. (238805)
6 risk.mp. or Risk/ (1317067)
7 prognosis.mp. or Prognosis/ (437345)
8 outcome.mp. (1029477)
9 acute.mp. (821130)
10 emergency.mp. or Emergencies/ (185645)
11 1 or 2 or 3 or 4 or 5 (248848)
12 6 or 7 or 8 (2397572)

13 9 or 10 (973362)

14 11 and 12 and 13 (18135)

15 limit 14 to (english language and humans and "all adult (19 plus years)") (13242)

The second strategy was to search for previously identified severity scores by name and commonly-used abbreviation. The scores were identified from personal knowledge, informal discussion with colleagues, and snowballing as the literature review developed. The final search covered Medline 1950 to October 2009 (in title, keywords, abstract or text): Altona (peritonitis), Alvarado (appendicitis), APACHE, Balthazar (pancreatitis), Blatchford (GI bleed), Canadian Triage and Acuity (CTAS), Emergency Severity Index (ESI), Essen, Early Warning Score (EWS), Geneva (pulmonary embolism) Glasgow Coma Scale/Score (GCS), Glasgow (pancreatitis), Goldman (acute coronary syndromes/ACS), GRACE (ACS), Hardman (abdominal aortic aneurysm/AAA), Manchester Triage (MTS), Mannheim (peritonitis), MEDS (sepsis), Mainz (unselected), MELD (cirrhosis), Mortality Probability Model (MPM), Norris (ACS), peritonitis severity score, POSSUM (emergency surgery), PURSUIT (ACS), Ranson (peritonitis), Rapid Acute Physiology Score (RAPS), Rapid Emergency Medicine Score (REMS), RISC (cerebrovascular event), Rockall (GI bleed), ROSE (syncope), San Francisco (syncope), Simplified Acute Physiology Score (SAPS), Scorten (toxic epidermal necrolysis), Sequential Organ Failure Assessment (SOFA), TIMI (ACS), Therapeutic Intervention Severity Score (TISS) and Wells (pulmonary embolism).

The third strategy was to search for prognostic indicators by condition. This involved searching Medline 1950 to October 2009 for: Prognosis/ OR “Severity of illness index”/ OR severity.mp OR risk/. This was then cross-referenced to the MeSH terms of: acute coronary syndrome, aneurysm (including dissecting, false, iliac, infected, ruptured), aortic aneurysm, arachnoiditis, arsenic poisoning, arterial occlusive disease, asthma (exp), bacteremia, brain abscess, brain infarction, chronic bronchitis, bronchopneumonia, cardiomyopathy (alcoholic, chagas, dilated, hypertrophic, takotsubo), chronic obstructive pulmonary disease, cirrhosis, CNS infections (bacterial, fungal, parasitic, viral), confusion, coronary aneurysm,

delirium, dermatitis (including herpetiformis), dermatomyositis, diabetic coma, diabetic ketoacidosis, encephalitis, encephalomyelitis, endocarditis (including bacterial and subacute), epidural abscess, fungaemia, gastrointestinal haemorrhage (exp), heart aneurysm, heart failure, heat exhaustion (exp), heat stroke (exp), hematemesis (exp), hepatic encephalopathy, hepatic insufficiency, hepatitis, hyperglycaemic hyperosmolar nonketotic coma, hypothermia (exp), intracranial aneurysm, intracranial embolism, intracranial thrombosis, liver failure (including acute), melaena (exp), meningitis (including aseptic, bacterial, fungal, viral), meningoencephalitis, mesenteric vascular occlusion, myocardial infarction, myocarditis, necrotising fasciitis, pancreatitis (including acute necrotising, alcoholic), peptic ulcer haemorrhage (exp), peritonitis (including tuberculous), pleuropneumonia, pneumonia (including aspiration, bacterial, pneumocystis, viral), poisoning (arsenic, cadmium, carbon monoxide, carbon tetrachloride, ciguatera, fluoride, food, gas, heavy metal, lead, manganese, mercury, MPTP, mushroom, plant, salmonella, staphylococcal), pulmonary embolism, pulmonary infarction, renal artery obstruction, sepsis, septic shock, skin diseases (including eczematous, infectious, metabolic), soft tissue infection, status asthmaticus, stroke, subarachnoid haemorrhage, subdural empyema, subphrenic abscess, suppuration, syncope (including vasovagal), toxaemia, transient ischaemic attack, urinary tract infaction and ventricular dysfunction (including left and right). The MeSH terms were selected by manual scanning of the MeSH subject headings to include all conditions that could reasonably present in the acute setting. All searches were limited to English language, humans and adults.

2.2.2 Article selection

The initial searches identified 13242 (strategy 1), 14659 (strategy 2) and 46605 (strategy 3) titles. A significant number of titles were identified by more than one search. Titles and abstracts were screened initially and where they appeared relevant or where relevance could not be established from the title and/or abstract the full paper was obtained. 47 (method 1), 682 (method 2) and 1661 (method 3) abstracts were screened.

To assess formalised scores or decision aids, search output was limited by title, abstract or full paper review to those papers including a wholly or predominantly clinical assessment (ie not biomarkers or specialist tests not available in the majority of EDs such as myocardial scintigraphy), to those which assessed an adult population and to those who assessed an outcome measure within 30 days of presentation. Also excluded were assessment tools requiring a specialist algorithm not freely available, or those which were applied only to patients in a critical care setting. 225 papers were deemed to fit the inclusion criteria.

To assess specific potential predictor variables, search output was limited to those papers assessing physiological variables at presentation and describing outcome measures up to 30 days after presentation. Data sets consisting purely or predominantly (over 50%) of trauma, paediatric, obstetric and psychiatric patients were excluded. 542 papers were deemed to fit the inclusion criteria.

Figure 12: Flow diagram of literature review

2.2.3 Data extraction

In the assessment of formalised scores or decision aids, the following data were extracted from each article selected for inclusion: the name and/or acronym of the score, the target condition or conditions, the patient groups included in the target condition(s), the main outcomes measured and the discriminant value of the score, expressed as the area under the receiver-operator characteristic curve (AUROC) or sensitivity and specificity.

In the assessment of potential predictor variables, scores identified as above were examined for the cut-off points of the variables deemed significant. From the search specifically for these variables, data were extracted from each article selected for inclusion was: the physiological variable(s) under analysis, the target condition or conditions, the patient groups included in the target condition(s), the

main outcomes measured and the relationship between the variable and outcome (difference in means, odds or risk ratios, non-linear relationship).

2.3 Results

2.3.1 Format of existing clinical decision aids

Clinical decision aids published to date appear to adopt one of five formats: the "all or nothing" rule, the unweighted simple summative score, the weighted summative score, the decision tree and the nomogram. Table 5 shows the number of decision aids identified in each format and the broad patient categories to which they have been applied.

Table 5: Format of existing clinical decision aids

Format	Number of tools	Patient categories
All or nothing	7	Acute coronary syndrome GI bleed Influenza Pneumonia Poisoning Syncope
Unweighted summative score	33	Aortic aneurysm Acute coronary syndrome Asthma/COPD GI bleed Heart failure Influenza Pancreatitis Pneumonia Pulmonary embolism Sepsis Surgical Syncope TIA Unselected
Weighted summative score	78	Aortic aneurysm Acute coronary syndrome Asthma/COPD GI bleed Heart failure Hypothermia Influenza Pancreatitis Pneumonia Pulmonary embolism Sepsis Stroke Surgical Syncope Unselected
Decision tree	5	Acute coronary syndrome Asthma/COPD Heart failure Sepsis Unselected
Nomogram	0	

2.3.1a The "all or nothing" rule

These rules are generally to be found as "rule-out" systems in common conditions that occasionally indicate sinister underlying disease processes, and require a consistent answer (usually "no") to a set of questions. For example, the San Francisco Syncope rule requires the patient to have none of: a. congestive heart failure history, b. haematocrit under 30%, c. abnormal ECG, d. history of shortness of breath and e. systolic blood pressure under 90mmHg in order to be deemed low risk (83).

2.3.1b Unweighted summative scores

These scores include a number of variables, scoring one point for each variable. The number of points is then summed and risk of adverse outcome related to the total score. For example, the CURB-65 pneumonia score allocates a point each for: a. confusion, b. urea over 7.0, c. respiratory rate over 30 breaths per minute, d. blood pressure of under 90mmHg (systolic) or 60mmHg (diastolic) and e. age over 65. A score of 0 or 1 is deemed low risk, 2 intermediate risk and 3 high risk (65).

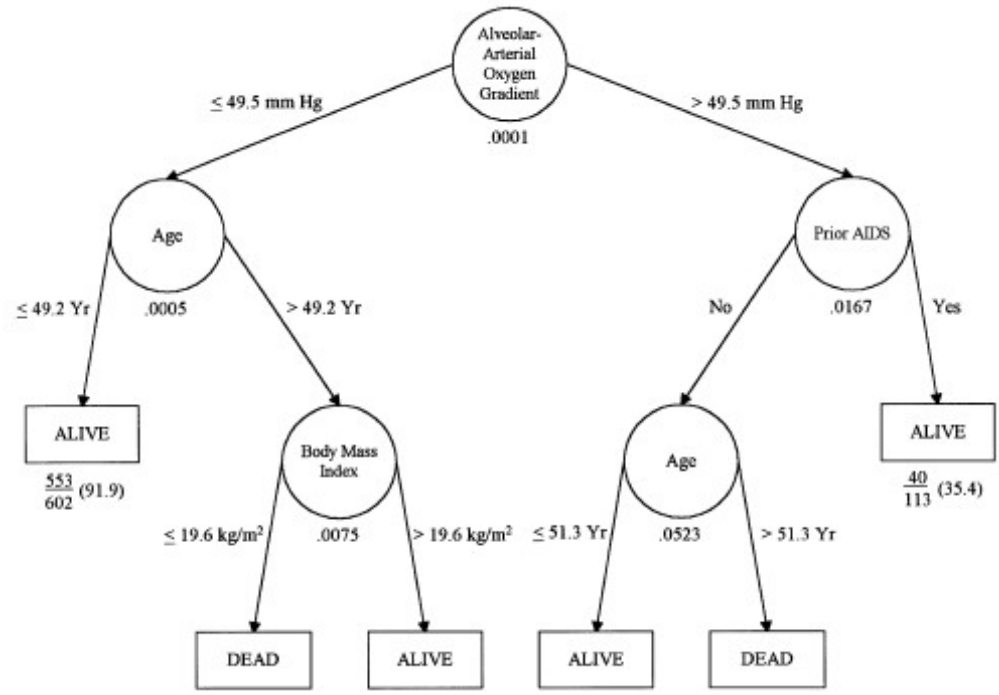
2.3.1c Weighted summative scores

These scores also sum points from individual variables to create a score which relates to risk. The weighting can be inter-variable (such as the PSI, where a respiratory rate of over 29 breaths per minute scores 20 points, but a pulse of over 124 beats per minute scores only 10 points), or intra-variable (in APACHE II a pulse of 120 beats per minute scores 2 points but one of 165 beats per minute scores 3 points).

2.3.1d Decision trees

Decision trees subdivide the population at risk according to various nodal points; the next nodal point is dependent on the response to the previous one. For example, in the decision tree shown below, which predicts probability of death or survival in patients with AIDS-related pneumocystis pneumonia, the two possible values at the initial node (alveolar-arterial oxygen gradient) lead to different secondary decision nodes (age or prior diagnosis of AIDS).

Figure 13: Example of decision tree

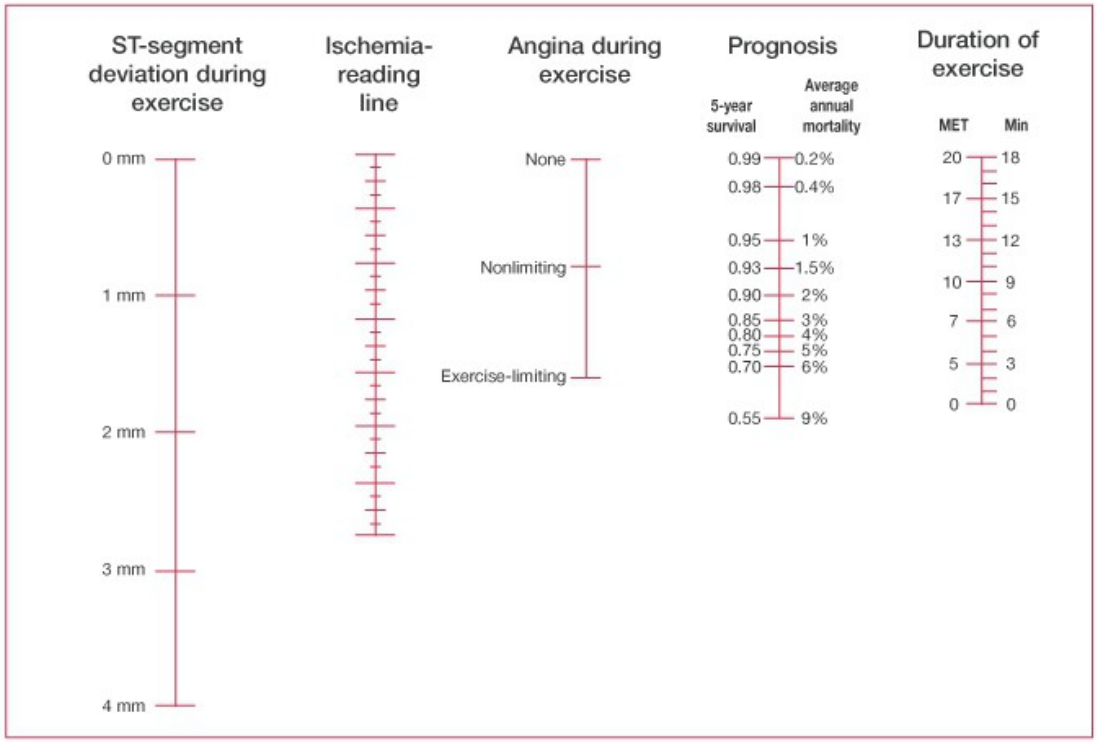


Reproduced from Yarnold et al (84).

2.3.1e Nomograms

Nomograms diagrammatically represent the effect of a particular variable conditional on the values existing for other variables. For example, in the nomogram reproduced in figure 14, which predicts 1-year mortality based on the results of a Bruce protocol exercise ECG, the ischaemia reading is conditional on the results of ST-segment deviation and angina, while the prognosis is conditional on the results of the ischaemia reading and the duration of exercise. No nomograms were identified applicable to the patient group of interest to this thesis.

Figure 14: Example of nomogram



Reproduced from Mark et al (85).

2.3.2 Decision aids for undifferentiated patient populations

Full details are given in appendix 1. Briefly, 23 manuscripts related to decision aids or risk stratification in undifferentiated patient populations. 2 manuscripts could not be obtained despite interlibrary loan requests and attempts to contact the authors. A further study described a decision aid for use in the prehospital environment (86). The structure of the identified decision aids is summarised in table 6.

Table 6: Structure and variables of decision aids for undifferentiated patient populations

	Structure	Variables included
Bispebjerg (87)	Weighted summative score	Consciousness (AVPU) Heart rate Respiratory rate Systolic blood pressure Temperature
Hillerød (88)	Multiple level all-or-nothing	Airway problem Chest pain Comorbidity Consciousness (GCS) Dyspnoea ECG changes Heart rate Oxygen saturation Respiratory rate Systolic blood pressure Temperature
HOTEL (89)	Unweighted summative score	ECG changes Mobility Oxygen saturation Systolic blood pressure Temperature
PREEMPT-2 (90)	Weighted summative score	Age Consciousness (AVPU) H^+ Oxygen saturation $PaCO_2$ Respiratory rate Systolic blood pressure
Rapid Emergency Medicine Score (91)	Weighted summative score	Consciousness (GCS) Heart rate Mean arterial pressure Oxygen saturation Respiratory rate Temperature
Simple Clinical Score (92)	Weighted summative score	Age Breathlessness Consciousness Diabetes ECG changes Functional status Heart rate Mobility Oxygen saturation Stroke

		Systolic blood pressure Temperature
South African Triage Score (93) and Cape Triage Score (94)	Weighted summative score plus all-or-nothing discriminators	Consciousness (AVPU) Heart rate Mobility Pain Respiratory rate Systolic blood pressure Target time to treatment Temperature Trauma as presenting complaint Multiple features of clinical history
Sun score (95)	Weighted summative score	Age Arrival by ambulance Comorbidity Racial group Recent hospitalisation or ED visit Unstructured triage nurse assessment
Vital Signs Score (96)	Unweighted summative score	Airway problems Consciousness (GCS) Heart rate Oxygen saturation Respiratory rate Seizures Systolic blood pressure
VitalPAC™ EWS (ViEWS) (97)	Weighted summative score	Consciousness (AVPU) Heart rate Inspired oxygen Oxygen saturation Respiratory rate Systolic blood pressure Temperature

The various scores have been evaluated as in table 7 below. Notable is the general lack of external validation.

Table 7: Evaluation of decision aids for undifferentiated patient populations

Title and year	Design of study	Patients and setting	Outcome measures	Main results
Acute illness severity score				
Mikulich 2011 (98)	Database analysis	25883 patients; 49	30-day in-hospital mortality	OR (vs group 1) for group 2: 1.72, group 3: 4.66, group 4: 6.51, group 5: 10.00, group 6: 19.92.

		337admissions. Admission unit, HDU or ICU		
Canadian Triage and Acuity Scale (30)				

Lee 2011 (99)	Retrospective chart review	1903 patients aged ≥65 years presenting	Life-saving intervention in ED	49% CTAS 1 and 2 2% CTAS 3 0 CTAS 4 and 5
			ICU admission	75% CTAS 1 40% CTAS 2 3% CTAS 3 0 CTAS 4 and 5
			Death	4% CTAS 1 1% CTAS 2 0 CTAS 3, 4 and 5

		to ED.		
Prah Ruger 2007 (100)	Retrospective chart review	77709 adult patients undergoing E	Hospital admission or ED death	70% CTAS 1 43% CTAS 2 10% CTAS 3
			Admitted to ICU or theatre	24% CTAS 1 3% CTAS 2 0.4% CTAS 3

		D tria ge.		
Cape Triage Score				
Bruijns 2008 (94)	Prospective database with post hoc simulation.	798 patients (including trauma & ob	Admission or ED death	Undertriage (green) 24%, overtriage (orange/red) 25%.

		st e tr ic) p r e s e n ti n g t o E D in o ff ic e h o u rs .		
Early Warning Score				
Armagan	Prospective cohort study	3	Hospital admission	57% admitted score >4 vs 37% score <5.

2008 (101)		09 patients presenting to ED	ICU admission	OR for score >4 1.95
			Death	OR for score >4 14
Groarke 2008 (102)	Retrospective chart review	225 patients	ICU admission	OR for each rise in score category 3.35.
			Death	OR for each rise in score category 2.19.

		admitted to medical ward via MAU		
Burch 2008 (103)	Retrospective chart review	790 medic	Hospital admission	Score 0-2 45%, score 3-4 59%, score >4 79%.
			Inhospital death	Score 0-2 5%, score 3-4 16%, score >4 26%.

		al patients presenting to ED		
Heitz 2010 (104)	Retrospective chart review	280 adult pati	Higher level care in 24h	AUROC of maximum EWS in ED 0.73. AUROC of first EWS at presentation 0.668.

		entsadmittedviaED,excludingtraumaandd		

		cardiology		
Perera 2011 (105)	Prospective cohort study	242 adult patients admitted to	HDU or ICU admission, cardiorespiratory emergency or resuscitation, inhospital death.	AUROC 0.68 +/- 0.06

		MAU		
HAPT				
Barfod 2012 (88)	Retrospective database review	6279 admissions to ED observation	ICU admission	Vital signs OR Red: 38.6, Orange: 10.9, Yellow 4.3. Final triage Red 40.3, Orange 8.5, Yellow 3.5
			Hospital mortality	Vital signs OR Red: 20.1, Orange 3.9, Yellow 2.2 Final triage OR Red: 24.0, Orange 8.0, Yellow 2.8

		ward general ward		
HOTEL				
Kellett 2008 (89)	Prospective dataset	10290 acute medi	Death 15min – 24h	AUROC 0.865 derivation, 0.854 validation.

		cal admissions (6947 derivation, 3343 val		

		id a ti o n)		
PREEMPT-2				
Carmichael 2011 (90)	Prospective dataset	4 3 1 4 a c u t e m e di c al a d m is si o n s	ICU admission	AUROC 0.89.
REMS				

Olsson 2004 (91)	Prospective dataset	1175 1 acute non surgical non trauma pre	Inhospital death	AUROC 0.852.

		s e n t a t i o n s t o E D		
Simple Clinical Score				
Li 2012 (106)	Retrospective chart review	4 1 7 a d m i s s i o n s t o A c u	Length of stay	<0.001
			MET call	p =0.1
			ICU admission	p = 0.7
			Hospital mortality	p <0.001

		te Admission Unit Excluded direct admission		

		s from ED to ICU, CCU or stroke unit		
Kellett 2011 (107)	Prospective cohort	1165	Inhospital mortality	OR 10.10 where SCS was increased 24h after admission.

		unselected medical admissions (aged 14 or		

		over)		
Subbe 2010 (108)	Retrospective review	1098 patients admitted to acute me	Death in 48 hours	Very low risk 0, low risk 1%, average risk 0.5%, high risk 2.5%, very high risk 8%.
			Death in 7 days	Very low risk 0, low risk 2%, average risk 1.5%, high risk 5%, very high risk 18.5%.

		dicine		
Subbe 2010 (109)	Prospective data	281 patients admitted to medical emergency	Death in 7 days	AUROC 0.85
			ICU admission	Very low risk 1%, low risk 2%, average risk 1.5%, high risk 16.5%, very high risk 19%.

		r g e n ci e s a c r o s s 2 1 c e n tr e s U K, E u r o p e a		

		nd New Zealand		
Kellett 2006 (92)	Prospective cohort	6736 (derivation) and 32	Death in 24h	AUROC for derivation 0.90; for validation 0.91.
			Death in 30 days	AUROC for derivation 0.86; for validation 0.86.

		28 (validation) unselected adult medical adm		

		is si o n s		
South African Triage Score				
Rosedale 2011 (93)	Prospective dataset	5 8 9 p a ti e n ts p r e s e n ti n g t o E D in o	Admission or ED death	Undertriage (green) 4.4%, overtriage (orange/red) 4.3%.

		ffice hours		
Sun score				
Sun 2011 (95)	Retrospective database analysis	3175 81 ED visits (60% deriva tio	Hospital admission	AUROC 0.85 in derivation set AUROC 0.849 in validation set

		n, 40% validation)		
ViEWS				
Prytherch 2010 (97)	Prospective dataset	35585 patients admit	Hospital mortality within 24h	AUROC 0.88.

		edtoMedicalAssessmentUnit.Multiplescoo		

		r e s f o r e a c h p a ti e n ts a s vi t al si g n s r e c o r		

		dingswererepeated		
Vital Signs Score				
Merz 2011 (96)	Prospective dataset	43888patientsadm	Inhospital death	AUROC 0.72.

		it t e d vi a t h e E D		

Notably, the absence of reporting of AUROC or at least sensitivity and specificity limits interpretation of the utility in clinical practice of many of these tools.

2.3.3 Scope of existing clinical decision aids

Clinical decision aids existed for 13 broad diagnostic groups: abdominal aortic aneurysm (7 aids), diagnosed or potential acute coronary syndrome (30), asthma and chronic obstructive pulmonary disease (6), gastrointestinal bleeding (7), heart failure and pulmonary oedema (9), influenza (9), pancreatitis (11), pneumonia (26), pulmonary embolism (4), sepsis (10), surgical emergencies (13), syncope (3), transient ischaemic attack (2). 13 tools were available for wider diagnostic groups (such as patients in ED resuscitation areas or those arriving by ambulance). Single manuscripts addressed prognostication in poisoning, hypothermia, meningitis, myxoedema coma, symptomatic atrial fibrillation and Fournier's gangrene.

Multiple sets of inclusion criteria were used, even within diagnostic groups. This was particularly notable within actual or potential acute coronary syndrome patients, and less so amongst patients with abdominal aortic aneurysm and pancreatitis.

Data sets were variously purely derivation sets, mixed derivation and validation sets (including split sample methodology and temporal validation), external validation sets and secondary analysis of other data sets (including disease registries).

Detailed information on patient sets, levels of validation and outcomes considered are shown in appendix 2. Briefly, of 136 tools identified, 24 were not disease-specific. Several tools were assessed in multiple disease categories. Of these tools, 35 had been widely (more than 2 studies) externally validated (McGinn level 2), 72 had been externally validated in one or two studies (McGinn level 3), 16 had only been internally validated (including split sample or temporal validation) and 13 presented data only on derivation (McGinn level 4) (53).

2.3.4 Physiological variables in existing clinical decision aids

This analysis scrutinised 103 scores. They apply to the same 17 broad diagnostic categories as in 2.3.2 above. Only 31 of the scores were purely clinical; 61 of the others required some element of laboratory testing and a further 12 required ECG analysis. The use of physiological variables in these scores is summarised in table 8 (full details in appendix 3).

Table 8: Physiological variables in existing scores

Variable	Number of scores with single cutpoint	Number of scores with multiple cutpoints	Number of scores using continuous variable	Number of scores with non-linear association
Age	29	19	10	1 ($(Age/10)^2$)
AVPU consciousness	3	4	-	-
Diastolic BP	10	1	-	1 (Drop of 20mmHg from baseline)
Glasgow Coma Scale	3	11	3	-
Mean arterial pressure	2	5	-	2 (1/MAP, log(MAP))
Oxygen saturation	5	7	-	-
Pulse rate	18	21	6	3 ("shock", related to SBP)
Respiratory rate	16	13	2	-
Systolic BP	28	26	4	5 ("shock", squared, 1/SBP)
Temperature	4	13	-	-

The use of peripheral oxygen saturation in scoring is more complex, as it can be recorded with or without supplemental inspired oxygen. One score stipulated that the SaO2 had to be recorded whilst breathing room air, another rated the use of any supplemental oxygen as a risk factor in itself and a third allowed the substitution of a PaO2/FiO2 ratio for saturations.

2.3.5 Evidence for the use of individual physiological variables

Within the literature there was minimal standardisation of the timing of physiological measurement. It might be assumed that the variables were measured at the earliest possible opportunity, but no studies using hospital data also analysed prehospital information. A significant number also ignored Emergency Department observations and treatment, collecting data for assessment at the point of admission to a ward or assessment unit. This is potentially very problematic as little or no allowance is made for the potential effect of prehospital and Emergency Department interventions on the patient's level of illness.

The reported predictive ability of individual physiological variables is shown in table 9 and in detail in appendices 4 and 5. Where raw data was available, L'Abbé plots are shown in appendix 4 in accordance with recommendations from the Cochrane group (110). A L'Abbe plot shows the event rate (in this case proportion of adverse outcomes) in one group against the event rate in another group (111). These are often intervention and control groups but in this case the groups are those deemed low (x axis) and high (y axis) risk by dichotomisation of a physiological variable. Point size represents the size of the trial. From the L'Abbé plots it appears that there is consistently an increase in risk of poor outcomes with increasing age, falling consciousness, falling oxygen saturations, increasing pulse rate, increasing respiratory rate and falling systolic blood pressure although its magnitude and cut points are inconsistent. Diastolic blood pressure appears to be associated with poor outcomes at both extremes; of note there is no exploration of the potential association between high systolic blood pressures and poor outcome. There appears to be no consistent association between temperature and poor outcome.

Where odds ratios were reported without raw data these are summarised in appendix 5. Findings are generally in line with those where raw data was available.

It is impossible to report with any accuracy the proportions of studies where a particular variable was found to be predictive, as the majority of studies only report positive associations found without providing information on all potential predictor variables that were assessed.

Table 9: Predictive ability of individual physiological variables

Variable	Raw data available (studies; no of patients) Appendix 4	Odds ratio (studies) Appendix 5	Non-linear relationship reported (studies)	Means, medians or p values reported (studies)
Age	120; 730269	115	35	164
AVPU consciousness	3; 2396	3	0	0
Diastolic BP	5; 11300	11	5	43
GCS	11; 8916	6	5	10
Mean arterial pressure	2; 524	3	3	17
Oxygen saturation	12; 9362	12	5	11
Pulse	34; 32420	50	19	63
Respiratory rate	26; 22587	20	8	26
Systolic BP	50; 98473	57	20	61
Temperature	22; 11527	9	16	33

2.4 Conclusions

2.4.1. Format and scope of existing clinical decision aids

There is clearly a wide variation in the patient groups to which scoring systems are applied, and an equally wide variation in patient outcomes considered relevant. The sheer number of available tools makes it impossible for the working clinician to use more than a few in daily practice. The discriminant value of the scores (expressed as an AUROC or sensitivity and specificity) often varies between studies and is poor in many cases. Many scores have undergone no external validation, ie they have only been tested in the population in which they were developed. The combination of poor discriminant value and lack of external validation is likely to reduce their value in day-to-day clinical practice.

Review of risk scores is limited by the structure and the lack of information in many included papers. Few were precise about the timing of the assessment, leaving potential for lead-time bias. The majority focus on hospital-specific outcomes and it is often unclear to what extent patient-relevant out-of-hospital outcomes have been investigated. The often restricted nature of patient sets (for example, requiring consultant radiologist confirmation for the diagnosis of pneumonia) limits the generalisability of many of the results to the day-to-day ED population where formal diagnosis is often not known initially.

2.4.2. The role of physiological variables in existing clinical decision aids

The origin of the cut-points and variables included in existing decision aids is often difficult to extract from the published literature. In many cases cut-points appear to have been selected *a priori*. Even where the cut-points were generated from the data the methods for this are scantily described and frequently appear to have assumed linearity in the relationship between variable and outcome.

Although there is obviously a huge amount of primary data relating to risk scores there have been few attempts to systematically evaluate these data and draw broader conclusions for clinical practice. Indeed, one of the characteristics of the literature relating to risk scores is that each risk score seems to be developed *de novo* with very little reference to previous studies or other scores. This may reflect the tendency for studies developing risk scores to be secondary analyses of existing datasets rather than studies undertaken for the primary purpose of developing a risk score.

2.4.3 Existing clinical decision aids for unselected medical emergencies

It is clear that the literature relating to severity assessment tools for unselected medical emergency patients is lacking. This is in sharp contrast to the cornucopia of scores developed in patient groups of interest to specialist practitioners.

2.4.4. Evidence to support the predictive value of individual physiological variables

As above, much of the literature is limited in its description of the methodology underlying selection of significant predictor variables. Explicit exploration of the (non-)linearity of the predictor variable-risk relationship is rare, but when it occurs does not support the assumption of linearity seen in the remaining literature. The assumptions discussed in 1.5.3 of a linear or Gaussian distribution of risk have not been fully explored or challenged.

It is notable that the literature on death as an outcome greatly exceeds all other outcomes. Although it is easy to see the attraction of this as an incontrovertible outcome measure, it cannot reasonably be argued that (if developed in an

established healthcare system), measures predicting death equate to measures of potential to benefit from treatment. Acutely ill patients may reliably be aided to avoid death if identified and high-acuity treatment offered in a timely fashion.

2.5 Objectives arising from the literature review

1. To generate a list of interventions which could be used to define potential to benefit from emergency (time-critical) care: section 3.7.3b.
2. To identify variables predicting potential to benefit from emergency care in a broader population than disease-specific scores: chapter 5.
3. To identify variables predicting death in emergency patients and explore whether these are the same as those predicting potential to benefit: chapter 4.
4. To develop near-patient scoring systems reflecting the variables identified in 2 and 3: sections 4.10 and 5.12.
5. To explore whether the scoring systems generated in 4 as well as those already identified perform adequately in the prediction of death and potential to benefit in a broader patient population: sections 6.8 and 7.10.

2.6 Summary

- Clinical decision aids exist in a number of formats.
- Clinical decision aids exist for large numbers of conditions and have been derived and validated in multiple data sets from varying sources with varying inclusion criteria.
- No well-validated clinical decision aid exists for the assessment of patients with medical emergencies.
- Individual physiological variables are used inconsistently in existing clinical decision aids.
- Individual physiological variables have performed inconsistently in predicting adverse outcomes.

Chapter 3
Methodology and Methods

3.1 Introduction

A vast array of literature exists around clinical aids for risk prediction of various adverse outcomes in patients presenting for emergency care. As discussed in the previous chapter, this is diverse in its applicability, generalisability, and methodology and in the performance characteristics reported.

The literature surrounding prediction of adverse outcome from individual patient variables is equally vast and limited in its applicability to the problem posed in this thesis. Minimal evidence exists to underpin risk prediction in unselected patient populations, with diverse presenting complaints and partial and unclear diagnoses. A prospective study in an unselected Emergency Department is therefore required.

Ideally this would involve a multicentre cohort recruited specifically to derive and validate a decision aid, or at least analysis of prospective registry data in the manner of many cardiology studies. Pragmatically, the infrastructure is not currently established to collect such data within the ED; large datasets of physiological variables do exist (for example that used to develop the ViEWS score (97)) but data collection only begins once the patient is admitted to a ward area. Thus the data accumulated for the DAVROS project provides the best proxy dataset for this study although not designed specifically for the purpose.

3.2 Aims and objectives

The study will aim to address whether patient variables available at or immediately after arrival at the Emergency Department can reliably be used, either singly or in combination, to predict short-term adverse outcomes in unselected emergency patients. The central premise underlying this is that existing scoring systems based primarily on risk of death may not function well in emergency as what is in fact required is a score that identifies an immediate need for intervention (see 1.4.3).

3.2.1 Epistemology and assumptions underlying the study

This study assumes a positivist standpoint. I would argue that conceptions of disease and illness and in particular their causation and symptomatology should be approached in a post-positivist manner, as their construction and effects are likely to be significantly related to culture and personal experience. I would however consider that societal perceptions of physiological variables and of critical illness are similar enough across populations that they can be considered as analogous to a "truth" and analysed within a positivist paradigm. The study will therefore assume the existence of some physiological variables which are constantly and generalisably associated with critical illness.

The main premise underlying the study is that patient-level variables (such as age, diagnosis or blood pressure) have a consistent mathematical relationship with, and can therefore predict, the likelihood of a particular outcome (such as death or successful response to a treatment). This premise underlies case-mix adjustment, which assumes that each patient's risk of death is manifested in consistent measurable parameters, and clinical decision rules, which assume that the chance of the outcome of interest is also predictable from measurable parameters. It is unlikely that the full complexity of Emergency Medicine can be reflected in a mathematical model; however it is possible that identifiable characteristics may contain enough to predictive value to be clinically useful. To quote the statistician George Box "all models are wrong; the practical question is how wrong do they have to be to not be useful" (112).

In this study the outcome of interest is the potential to benefit from a time-dependent intervention. The rationales of case-mix adjustment and clinical decision rule development are similar enough that it is reasonable to use an overlapping data set for both; however the differences in outcome (death versus benefit from urgent treatment) and thus the potential differences in parameters of interest require analysis specific to each.

3.2.2 Specific aims

The following will therefore be performed:

a. Identification of patient variables which accurately predict short-term (7-day) mortality;

b. Combination of these variables into a near-patient scoring system;

c. Assessment of the performance in terms of discrimination and calibration of this scoring system;

d. Identification of patient variables which accurately predict critical illness in terms of potential to benefit from time-dependent treatment (this is discussed further in 3.8 below);

e. Combination of these variables into a near-patient scoring system;

f. Assessment of the performance in terms of discrimination and calibration of this scoring system.

3.2.2a Originality of this research compared with DAVROS study

This study differs from the main DAVROS work in the following ways:

a. The DAVROS score for death at 7 days was developed for casemix adjustment and therefore includes variables (particularly ICD-10) not available to the treating clinician in the ED, whereas this study will develop a score to predict death at 7 days useable by the bedside clinician.
b. This study will also consider potential to benefit (see 3.8 below), an outcome not considered in the DAVROS study and therefore requiring specific collection of outcome data within the study cohort; a score will then be developed for this outcome.

3.3 Setting

3.3.1 Northern General Hospital, Sheffield

The Northern General Hospital, Sheffield, houses the only adult Emergency Department serving the half million population of Sheffield. It is an acute teaching hospital with 1100 beds which, in conjunction with other hospitals in the Sheffield Teaching Hospitals Trust and Sheffield Children's Hospital NHS Trust, offers care from all acute specialities. The Emergency Department at the time of the study had 98,000 attendances per year. Coronary Care (including cardiac catheterisation

facilities), High Dependency and Intensive Care units and acute theatres are all available on the Northern General site. The Northern General patient set was selected as the largest cohort within the derivation phase of the DAVROS study (which also included Barnsley and Rotherham) which was complete at the time of commencement of this study.

3.3.2 Yorkshire Ambulance Service

Yorkshire Ambulance Service (YAS) responds to around 700,000 emergency calls per year from a population of 5 million throughout Yorkshire, using 500 vehicles operating from 62 ambulance stations.

3.3.3 DAVROS

The DAVROS project, of which this study is an ancillary, was funded by the Medical Research Council. It aimed to identify variables that predict 7-day mortality in patients presenting to hospital by emergency ambulance and to develop and validate a risk-adjustment tool applicable across a range of settings. Phases 2 and 3 (derivation) took place from February to May 2008 and involved collection of data in the prehospital and ED phases of care for patients transported by emergency ambulance to Northern General Hospital (Sheffield), Barnsley, and Rotherham Hospitals. Phase 4 (validation) collected a further cohort from the same hospitals from October to December 2008. Phase 5 and 6 extended validation of the model and assessed its role in risk adjustment across new UK and international sites. The specific DAVROS methodology and its strengths and weaknesses for this study are addressed in detail below.

3.4 Subjects

The project relates to unselected adult non-trauma patients with potentially critical illness.

3.4.1 Inclusion criteria

Patients were included if they were transported by an emergency ambulance and were alive and not in cardiac arrest at the point of ambulance arrival, then either died in the ambulance or emergency department or were admitted to hospital.

This ensures a broad range of patients, the use of which has been lacking in the literature as discussed above. It will hopefully facilitate some form of comparison of illness severity across different diagnostic groups.

3.4.2 Exclusion of ambulatory patients

Ambulatory patients (ie those who self-present to hospital rather than call for an ambulance) have been excluded within the DAVROS dataset. This was justified in terms of ensuring robust data collection at a standardised point in time, and minimising the potential confounding effects of emergency treatment.

Ideally, self-presenting patients would be studied in a similar manner as they arrive at the ED, and their absence will be a limitation of this study. It is likely that self-presenting patients represent a different population in terms of illness severity; however, presenting independently of an ambulance does not preclude the possibility of serious illness. It seems logical that self-presenting patients would have the same physiological processes and therefore the same predictors of adverse outcome as in the ambulance-transported population. However, it may be that having the physiological reserve to allow self-presentation is in itself a prognostic factor and therefore this population merits study in its own right.

3.4.3 Exclusion of patients discharged from the Emergency Department

Patients who were not admitted to hospital were excluded from the initial DAVROS dataset as the aim was to develop a risk-adjustment tool for emergency admissions to hospital. This limits this substudy in terms of developing a clinical score as there is no analysis of patients who were discharged from the ED, which may restrict generalisability.

Ideally a full cohort of presenting patients would be studied and those discharged from the ED followed up as outpatients to ensure the absence of post-discharge adverse events. However within the scope of this study that was logistically unfeasible. Rates of short-term death after discharge from the ED have been reported as 30/100000 (113) to 50/10000 (114), so the required cohort size to investigate this fully would have been impractical.

3.4.4 Exclusion of patient sub-groups

Patients who were considered dead (ie had no vital signs) at the time of ambulance arrival to the patient (even if resuscitation was attempted) were excluded.

Patients aged under 65 transported following trauma were excluded as risk prediction in the trauma population has been extensively studied; it is apparent from the existing literature that prognostication in trauma is strongly guided by anatomical site, type of injury (blunt or penetrating), and severity of individual injuries, variables which are not replicated in non-trauma patients. Increasingly systems of care for trauma patients are also differently structured than those for non-trauma patients, with prehospital triage to major trauma centres and ambulance bypass of smaller units. Patients aged over 65 transported following apparent trauma were included in the initial DAVROS data set as many have an underlying medical cause for their trauma, but were retrospectively excluded for this thesis if found to be admitted for trauma (including fractured neck of femur)..

Children and patients with purely obstetric presentations were excluded as their physiology is different and therefore predictors of adverse outcome are likely to be different, as are mortality and critical illness rates. The systems providing emergency care in paediatrics and obstetrics are also significantly different from those used for most adult patients.

Patients with purely psychiatric problems were excluded as systems of care are almost entirely different for them and outcomes of interest are also different.

3.5 Ethical issues

3.5.1 Risks to patients

This research carries minimal risk to patients, as it is purely observational and involves no extra or novel interventions. The main risk would be of breach of confidentiality in the handling of their identifiable patient information. This risk has been minimised by the pseudonymisation of the data within the working project data set. Linkage data which might be used to identify patients is held only on password protected computers at the Northern General Hospital. All data processing has been carried out in compliance with the Data Protection Act under the auspices of the University of Sheffield which is a registered data handling body.

3.5.2 Risks to researchers

This study involves minimal researcher risk, as there will be no additional patient contact and no use of biological substances or ionising radiation. It was possible that the case note review process outlined in sections 3.7 and 3.8 could have identified instances where the patient care provided raised concerns about ongoing patient safety. The project plan was to feed these back to the Clinical Director of the relevant specialty in accordance with standard NHS clinical governance procedures. In fact no such cases were identified.

3.5.3 Societal risks and benefits

This research primarily raises the issues of using identifiable patient data without explicit consent. This is regulated under section 251 of the NHS Act 2006 (administered at the time of data collection by the National Information Governance Board, which subsequently became the Confidentiality Advisory Group of the National Research Ethics Service), and data must be held in accordance with the System Level Security Policy approved by the NIGB. This provides for password-controlled access to the DAVROS database held at the Clinical Trials Research Unit at SCHARR and for password-controlled access to a separate database at the Northern General Hospital linking DAVROS reference numbers to NGH patient record numbers.

3.5.4 Ethical approval

Ethical approval both for the original DAVROS study and for this study has been obtained from the Leeds (East) REC. Approval to collect use patient identifiable data without specific consent was gained from the Patient Information Advisory group for the original DAVROS study. Approval for further data collection and analysis in this study has been gained from the National Information Governance Board.

3.6 Variable data obtained from the DAVROS database

Variable data was extracted by DAVROS research staff from the Emergency Department records and entered directly into the online DAVROS database. These data were linked to data from scanned YAS patient report forms (PRFs) by incident number and date. Data entry staff were non-medically-trained clerical staff who were provided with specific training on medical abbreviations. Random sampling and rechecking of entered data was undertaken by DAVROS co-ordinating staff (Richard Wilson and Martina Santarelli) to ensure data quality.

3.6.1 Presenting complaint

Data were recorded regarding the complaint as described by the YAS crew at handover to the Emergency Department staff.

3.6.2 Physiological variables

Data were available on age, diastolic blood pressure, fingerstick blood glucose (BM), Glasgow coma scale, oxygen saturation (breathing air and/or breathing supplemental oxygen), pulse rate, respiratory rate and temperature as first recorded in the Emergency Department. These were chosen as they are already widely recorded in a relatively standardised manner, so implementation of their collection is unlikely to be problematic. Pulse pressure was calculated from systolic and diastolic blood pressures immediately prior to data analysis. As discussed in 2.1.6 above it is likely that patient physiology will be less sensitive to socio-cultural confounders, such as access to health care, than other potentially predictive variables.

3.6.3 Historical variables

Also recorded as extracted from the Emergency Department notes was any history of heart or chronic respiratory disease (including asthma), malignancy, diabetes and epilepsy and use of warfarin or steroid therapy, all as recorded by the attending clinician, whatever the source of that information. If numerical values were missing no entry was made on the database and this was later coded as "missing". For categorical values such as those relating to patient history these were assumed to be negative unless specifically marked as positive; in a small number of cases the presence of comorbidity was inferred from medication lists (eg the presence of diabetes if insulin was a prescribed medication).

The working diagnosis in the Emergency Department was not collected as it was felt to be inconsistently recorded and subject to substantial change during the course of a hospital admission. Although diagnostic ICD-10 codes were available in the initial DAVROS data set, they are coded retrospectively after all diagnostic data from the patient admission is available and are therefore not pertinent to this study.

3.7 Outcome measures collected for this study

3.7.1 Outcome measures collected for DAVROS study

The original DAVROS study examined mortality in the seven days after hospital admission. Mortality was selected as a clearly-measurable non-biased outcome, and avoidance of mortality was felt in most circumstances to be a positive outcome. 7 days was chosen as a timescale suitable for emergency care as it was hoped to reflect the management of an acute crisis rather than the progress of a chronic disease process.

3.7.2 Additional outcome measures for this study

The patient group of interest in this study is those where urgent intervention had the potential to affect survival, ie where a patient death was or could have been prevented. These patients occupy a "Goldilocks zone" (neither too well nor too ill) as per figure 15.

Figure 15: the 'Goldilocks zone' of potential patient benefit

Some deaths (in the black zone) will be the inevitable outcome of chronic deterioration and emergent high-intensity therapy would not be appropriate for these patients as it would be futile. Identification of these patients is an imperfect science but early institution of terminal care (see 3.7.3a below) will be used as a proxy. In other cases high-intensity intervention might ensure survival to 7 days but this would not be considered in the patient's interests due to poor quality of life. Thus not all patients who died will be considered to have fulfilled an outcome of interest.

Equally, we would wish to identify patients where emergent therapy has potentially prevented death. The definition of such patients is difficult and still under debate; many routinely used urgent therapies have an incomplete evidence base and a number of interventions apparently well founded in physiological principles have recently been questioned in terms of actual patient benefit (for example, high-flow oxygen in myocardial ischaemia (115) and blood transfusion in anaemia induced by gastrointestinal bleeding (116)). However, there does appear to be some measure of consensus; when Haukoos et al attempted to identify patients benefiting from emergency care, the Spearman-Brown coefficient between different raters was between 0.83 and 0.87 in different patient sets (117).

The use of preventable or prevented deaths, with inevitable deaths excluded, as the primary outcome measure is a significant departure from the bulk of the published literature, where all deaths have been included in the outcome measures whether or not affected by the intensity of care.

3.7.3 Collection of outcome data

3.7.3a Death

7-day mortality was recorded by DAVROS research staff using hospital record analysis. Patients who died were assumed to have the potential to benefit from emergency care unless a "Do Not Resuscitate" order or terminal care pathway (initiated at or before the time of first contact with an inpatient consultant) is present in the hospital notes during scrutiny as described below.

3.7.3b Potential to benefit amongst survivors

It was assumed that all patients admitted to intensive care received potentially life-saving intervention(s). There is evidence to support the mortality benefit of admission to intensive care (118) and also that this may be time-dependent (119).

Notes of other patients were scrutinised for interventions defined *a priori* as potentially life-saving on the basis of their inclusion in evidence-based guidelines for acute care listed below by patient group. These guidelines were identified by manual searching of the websites of relevant bodies, supplemented by informal discussion with colleagues.

- Acute illness: NICE clinical guideline 50
- Acute onset atrial fibrillation: NICE clinical guideline 36, SIGN 94 and 129
- Alcohol-related disease: NICE clinical guideline 100
- Anaphylaxis: College of Emergency Medicine
- Arrhythmia: SIGN 94
- Asthma: SIGN 101
- Cardiac arrest: SIGN 94
- COPD: NICE clinical guideline 101
- DVT: NICE clinical guideline 144 and SIGN 122
- Headache: College of Emergency Medicine and American College of Emergency Physicians
- Hypertension: American College of Emergency Physicians
- LVF: American College of Emergency Physicians and European Society of Cardiology
- NSTEMI/Unstable angina: NICE clinical guideline 94, SIGN 93, American College of Emergency Physicians and European Society of Cardiology
- PE: NICE clinical guideline 144, SIGN 122, American College of Emergency Physicians and European Society of Cardiology.
- Pneumonia: American College of Emergency Physicians and British Thoracic Society.
- Pneumothorax (spontaneous): British Thoracic Society

- Poisoning: NICE clinical guideline 16, American College of Emergency Physicians and College of Emergency Medicine.
- Seizure: American College of Emergency Physicians and College of Emergency Medicine.
- Sepsis: NICE clinical guideline 151.
- STEMI: SIGN 93 and European Society of Cardiology.
- Stroke/TIA: NICE clinical guideline 68, SIGN 108.
- Transient loss of consciousness: NICE clinical guideline 109.
- Type 1 diabetes: NICE clinical guideline 15.
- Upper GI bleeding: NICE clinical guideline 141, SIGN 105.
- Urinary tract infection: SIGN 88.

Having scrutinised the above guidelines, the interventions identified as having the potential to save patient life were as listed below, and if a patient received any of these interventions he or she was considered as having a potentially prevented death.

Airway interventions

- Use of airway adjunct or procedure to maintain patent airway.
- Use of intravenous/intramuscular adrenaline to treat or prevent airway compromise.

Breathing interventions

- Bag-valve-mask ventilation (unless during procedural sedation), intermittent positive pressure ventilation, or non-invasive ventilation.
- Decompression of tension pneumothorax.
- Drainage of significant pleural effusion (>1 litre).
- Insertion of chest drain for pneumothorax in patients with pre-existing lung disease.
- Intravenous therapy except steroids for asthma.

Circulation interventions

- Cardioversion (chemical or DC) of ventricular tachycardia or supraventricular tachycardia or atrial fibrillation with accessory pathway.
- CPR.
- Emergency endoscopy or surgery for upper GI bleed or use of Sengstaken tube or use of vasopressin/terlipressin.
- Infusion of >2 litres of fluid or transfusion for haemodynamic instability.
- Laparotomy for GI bleed/gynaecological bleed (including ectopic)/AAA.
- Sepsis care bundle.
- Thrombolysis for AMI or PE, or percutaneous revascularisation.
- Therapeutic (not diagnostic) pericardiocentesis.
- Transcutaneous or external pacing or administration of atropine (except in theatre).
- Vasopressor use (except bolus dosing in theatre).

Disability interventions

- Administration of naloxone or flumazenil (unless related to procedural sedation).
- Administration of 10%/50% dextrose.
- Administration of >1 dose benzodiazepines/other anticonvulsants for fitting.
- Neurosurgical intervention.

Other interventions

- Active rewarming (not including Bair hugger warming blanket).
- Laparotomy for sepsis/infarction/obstruction.
- New initiation of renal replacement therapy.
- Specific poisons antidotes including N-acetylcysteine.

It must be noted that many of these interventions are not supported by high quality evidence. They are, however, similar to the list of potentially life-saving interventions developed concurrently by another research group (120).

3.8 Pilot phase of data collection

The pilot phase of data collection was carried out to establish the practicality of the data collection plan. Initially it was planned to examine ED records and only access hospital casenotes if evidence of interventions as in 3.7.1 above was not found. It became clear during the pilot that the ED notes had been microfiched and were only accessible on a single machine in the ED. This machine was generally only accessible out-of-hours due to the competing demands of clinical care and medico-legal requirements. Thus for the main study this process was abandoned and full hospital casenotes (including photocopied ED records) were examined.

During the pilot phase 21 cases were identified where a patient received an acute intervention not identified *a priori* as above. These cases were summarised into brief vignettes which were examined by two independent experts in Emergency Medicine (SG and FM) to adjudicate as to potential to benefit. However as the kappa value for agreement between SG and FM was under 0.1 this process was abandoned and the interventions in 3.7.1 adhered to. The 21 cases were not classified as receiving a potentially life-saving intervention.

3.9 Data collection process

The final data collection process involved 4 steps:

1. Translation of the DAVROS database number to the NGH reference number using the secure DAVROS database at NGH.
2. Identification of the new STH casenote number from the NGH reference number and identifying subsequent patient death from STH patient administration system. This was required as the Trust had undergone a casenote reorganisation and renumbering process in 2010. Date of death was required as notes for subsequently deceased patients were stored in a different archive.
3. Examination of casenotes at the Medical Records Department at NGH; some had to be retrieved from offsite storage, and where sets of notes were unavailable in Medical Records due to clinic or inpatient attendance they were sought on multiple occasions.
4. Extraction of outcome data as above and entry into secure database.

It was estimated from the pilot phase described in 3.8 above that this process took around 30 minutes per patient.

3.10 Sample size

Statistical simulation studies have suggested that for use of logistic regression analysis, a cohort including ten outcomes events for interest per variable assessed provides an acceptable precision of regression coefficients (80). This rule has been challenged more recently (81), but this thesis will adopt the more conservative estimate.

To develop a score predicting potential to benefit it was planned to analyse 8 potential predictor variables, with an assumed loss of at least one in multivariate analysis, thus requiring a minimum of 70 outcomes of interest (prevented or preventable death). We therefore planned to collect information sequentially until 100 outcomes of interest were included. From piloting it appeared that around 1 in 5 patients had an outcome of interest, requiring a sample size of around 500, which was logistically achievable given the time requirement identified in piloting.

To develop a score predicting death in 7 days it was expected that around 18 potential predictor variables would be analysed. Assuming some would be eliminated at the multivariate stage, the Northern General Hospital DAVROS phase 2 dataset, with 128 outcomes of interest (deaths) was deemed to be adequately sized. Using the dataset from the same setting as the subset above enhanced comparability.

An alternative power calculation would be to find the number of outcomes of interest required to estimate the sensitivity of the predictive tool to within 5%. Conservatively assuming the model to be 75% sensitive and using the approximate standard error of a proportion as:

$$SE = \sqrt{[p*1-p/n]}$$

where n denotes the number of outcomes of interest. This then gives

$$0.05 = \sqrt{[0.75*0.25/n]}$$

which can be solved to give a minimum number of outcomes of interest of 75, which would also be acceptable for both the above data sets. The event rate can safely be assumed to lie below 50%, meaning that the specificity will be estimated to within a standard error less than 5%.

3.11 Variable analysis

As discussed in 1.8 above, we wished to avoid arbitrary dichotomisation of data, particularly as there was little evidence to support the linear and monotonic relationship between predictor and outcome variables. Analysis therefore was structured around logistic regression rather than recursive partitioning.

3.11.1 Descriptive data

Basic descriptive data is reported, including frequencies for categorical data. For continuous data, a graphical analysis is made of normality of distribution and data presented in terms of mean and standard deviation or median and inter-quartile range, whichever is appropriate. The proportion of missing data for each variable is also presented.

3.11.2 Relationship with adverse outcome and categorisation

Where outcomes were unknown the case was excluded from the data set.

The relationship between each potential predictor variable and the primary outcome is examined graphically to identify non-linear and non-monotonic relationships. Where a relationship is found to be non-linear, the variable is categorised to describe the relationship. The strength of relationship will be described in terms of the odds ratio generated by various categories of the variable.

Specifically, the predictive value of missing variables is explored. The fact that data is missing can be predictive in itself. Missing data could be “unmeasureable”, for example a systolic blood pressure that was so low it could not be detected by health care providers. Alternatively, missing data could be “unmeasured”, either because the patient was felt to be so ill that all resources were being expended on care

rather than documentation, or because the patient was so "obviously" well that the practitioner deemed it irrelevant or unnecessary to record specific variables. It cannot be assumed that the characteristics of "missingness" will be the same for all variables within one data set; when analysis was undertaken of a trauma registry, imputed values for GCS were significantly lower (median 12) than those actually observed (median 15), implying that GCS was less likely to be recorded in sicker patients. However, imputed and recorded systolic blood pressures did not differ, implying failure to record, rather than inability to measure.

If data is missing because it was assumed to be normal (ie the "missing" category predicts good outcome), then it can be replaced with a normal value. If data is missing systematically (ie because the patient was unwell, so that the "missing" category predicts poor outcome), it is possible to impute the missing values. This involves calculating the most likely value of the missing data based on the data that has been recorded. This risks overestimation of the predictive power of the decision aid. This is explored in chapters 4 and 5.

The aim of this study is to develop a decision aid that can be used in real time, rather than retrospectively. There is therefore less of an issue in terms of data that is missing because it is "unmeasured" once the decision aid exists, as required variables will be measured. Some data may still be missing as "unmeasurable" and the predictive value of this is explored below.

3.11.3 Analysis of univariate significance

The univariate association between each potential predictor variable and the primary outcome of interest is determined. The form of the relationship between the variable and outcome, as explored in 3.9.2 above is used to generate clinically-relevant categories for continuous variables. The first order interaction of variables which are significantly related to outcome at $p<0.15$ is examined.

3.11.4 Multivariate analysis

Variables found to be significantly predictive of outcome at $p<0.1$ are entered into the multivariate analysis. A subset of variables independently predictive of

outcome are selected by removing terms at p>0.05. Linear coefficients are then recalculated for those independently predictive variables and an equation to predict poor outcome generated from those coefficients using the general formula p(outcome)= $e^a/1+e^a$, where a is the sum of coefficients.

3.12 Development of a point-of-care score

As demonstrated in 2.3.2 above, almost no evidence exists to guide the use of illness severity scoring in the unselected ED population. It is hoped that a point-of-care score might aid resource allocation in and after the ED by enabling the comparison of illness severity across multiple diagnostic groups.

3.12.1 Variable selection and weighting

Variables found to be independently predictive of outcome in multivariate analysis and included in the equation as at 3.11.4 are included in the score. The β-coefficient generated in the multivariate analysis is approximated to generate an integer value to weight each variable in terms of its explanatory value for poor outcome. The integer values for each variable are summed to obtain the bedside score.

This is an attractive principle for clinicians who are unlikely to wish to perform unnecessarily complex mental arithmetic at the bedside. It is however statistically flawed, as the predictions generated by logistic regression are not linear (hence the need for exponentiation as in 3.11.4 above). As a hypothetical example, if the risk of death from ST elevation myocardial infarction could be explained by two factors: anterior MI (yes/no) and hypotension (yes/no), where:

P(death)=-3 + (anterior MI)*1 + (hypotension)*4

then the additional risk conferred by hypotension is:

expit(-3+1+4) – expit(-3+1+0) = 0.76 (if the MI is anterior)

or:

expit (-3+4) – expit(-3+0) = 0.64 (if the MI is not anterior)

where expit(a)=e^a/1+e^a

Thus the effect of each predictor variable on the probability of the outcome is dependent on the other predictor variables and simple summation represents a linear approximation to a non-linear equation. However, this approximation is in common use in the development of bedside scoring and has been shown in elsewhere not markedly to reduce predictive power (121).

3.13 Score validity

The validity of the score in the development dataset is assessed using two techniques.

3.13.1. Discriminant ability.

Receiver operator characteristic (ROC) curves are used to test the ability of the score to discriminate between patients with the primary outcome and those without. ROC curves will be used as they provide a graphic representation of the trade-off at different score values between sensitivity and specificity.

3.13.2. Calibration.

The accuracy of model fit is assessed using Nagelkerke's pseudo R^2, which is a measure of the improvement of the model over a null model including no predictor variables. It is analogous to the R^2 measure used in linear regression but designed specifically for the binary outcome of logistic regression. For the purposes of this study, model calibration is of less importance than discriminant ability as it is envisaged that the relative risk of two patients (accuracy of which is measured by discriminant ability) will be of more clinical use in prioritisation than absolute risk (accuracy of which is measured by calibration).

3.14 Temporal validation

Validation of the score will occur using the data collection techniques as above in a temporally separate data set collected at the Northern General Hospital from

October to December 2008. This will address the issue of over-fitting as discussed in 1.10 above.

The calculation of sample size for validation studies is not well established. However, a simulation study by Vergouwe et al (122) suggested a minimum sample size of 265, including 106 events, to detect a difference in c-statistic from 0.83 to 0.73.

Vergouwe further calculated that the power for a given sample size is:

$$Z_{\beta} = \sqrt{(N\delta^2/\sigma^2)} - Z_{1/2\alpha}$$

where N = sample size, δ = difference in model performance, Z_{β} = value of the standard normal distribution corresponding to β, with β = type II error rate, $Z_{1/2\alpha}$ = value of the standard normal distribution corresponding to 1/2α, with α = type I error rate, σ = standard error of performance measure*√n.

3.15 External validation

External validation of the score at different sites and in different health care settings is outwith the scope of this thesis.

3.16 Comparison with other scoring systems

The data set collected for validation as in 3.14 above will also be used to compare the performance of the developed score with others identified in the literature review to determine which is likely to prove most useful in the patient population in question.

The statistical comparison of two ROC curves derived from the same data developed by Hanley and McNeil (123) and the technique subsequently promulgated by Obuchowski for calculating sample sizes was (124) confirm the adequacy of the size of the data set for this purpose.

3.17 Summary

- A prospective cohort design has been used to address the lack of prognostic decision aids in unselected medical emergency patients.
- Data on potential predictor variables have been collected in a standardised manner.
- Outcome measures have been developed to identify patients at high risk of death, and those where emergency care has the potential to have affected survival.

Chapter 4
Derivation of a score to predict death at 7 days

4.1 Introduction

This chapter will describe the data set assembled for the derivation of a score predicting death at 7 days, and the process of deriving that score.

4.2 Data set

Patients presenting by emergency ambulance to the Emergency Department of the Northern General Hospital, Sheffield, from February to May 2008 were identified prospectively by DAVROS researchers. Patients with obstetric, traumatic or purely psychiatric presentations were excluded. The rationale underlying the selection criteria for the data set is discussed in 3.4 above. Data on demographics, presenting complaint, past medical history (as recorded by the treating clinician) and physiological variables were abstracted by trained research assistants. Data from blood tests and ICD-10 coding were later added to the main DAVROS dataset but this will not be considered here as it would be unavailable to the clinician at the time of presentation to the Emergency Department. 2437 patients were included in this derivation cohort.

4.3 Study population

Characteristics of the study population are in table 10.

Table 10: Characteristics of the derivation study population

Variable	Mean (standard deviation)	Range
Age	69 (19)	18-103
Pulse	88 (24)	21-215
Respiratory rate	19 (6)	6-60
Systolic BP	136 (29)	24-266
Diastolic BP	75 (15)	30-153
Temperature	36.6 (1.2)	26.0-41.0
Variable	Median (interquartile range)	Range
GCS	15 (15-15)	3-15
SaO2 breathing air	97 (95-98)	50-100
SaO2 breathing oxygen	98 (95-100)	24-100
Variable	Number	Percentage
Male	1131	46.4
Active malignancy	110	4.5
Chronic respiratory disease	292	12.0
Heart disease	829	34.0
Asthma	269	11.0
Diabetes	382	15.7
Epilepsy	97	4.0
Warfarin usage	154	6.3
Steroid usage	149	6.1
Outcome		
Alive at 7 days	2235	91.6
Dead at 7 days	128	5.2
Unknown status at 7 days	77	3.2

The 77 patients whose outcome was unknown were excluded from further analysis, leaving 2363.

4.4 Univariate analysis of potential predictor variables

4.4a Demographic and historical variables

Categorical demographic and historical values were crosstabulated against outcome at 7 days and significant associations identified using chi-squared testing. Full details are in table 11 below. Categorical variables significant at $p<0.1$ were active malignancy ($p<0.001$) and history of chronic respiratory disease (not asthma) ($p=0.07$).

Table 11: Univariate analysis of categorical variables

Variable	Number (%age) of deaths	Total
Gender p=0.630		
Female	72 (5.7)	1260
Male	56 (5.1)	1094
Missing	0	9
Active malignancy p<0.001		
No active malignancy	110 (4.9)	2252
Active malignancy	17 (15.6)	109
Chronic respiratory disease p=0.07		
No history	105 (5.1)	2073
History present	22 (7.6)	288
Heart disease p=0.885		
No history	83 (5.3)	1557
History present	44 (5.5)	804
Asthma p=0.154		
No history	118 (5.6)	2103
History present	9 (3.6)	249
Diabetes p=0.618		
No history	105 (5.3)	1989
History present	22 (5.9)	372
Epilepsy p=0.308		
No history	124 (5.5)	2264
History present	3 (3.1)	97
Warfarin therapy p=0.244		
No warfarin	122 (5.5)	2210
Warfarin	5 (3.3)	151
Steroid therapy p=0.664		
No steroids	118 (5.3)	2215
Steroids	9 (6.2)	146

4.4b Continuous variables

Continuous physiological variables and age were assessed in three ways: quintiles of the study population, quintiles of the variable range and (where included in DAVROS) groups defined *a priori* within the DAVROS project to reflect generally accepted clinical definitions of normal. Full details are in appendix 6. GCS and oxygen saturations (both breathing air and on oxygen) were assessed in tertiles and quartiles to better reflect their population distributions. GCS was also assessed in *a priori* groups broadly consistent with currently accepted clinical risk categories.

Continuous variables significant at p<0.1 were pulse, respiratory rate, systolic BP, diastolic BP, temperature, pulse pressure, GCS, oxygen saturations and age. After visual inspection of the histograms in appendix 6 population quintiles were chosen

for all variables except GCS and oxygen saturations where *a priori* groups were used, as shown in table 12.

Table 12: Univariate analysis of continuous variables

Variable	**Dead (percentage)**	**Total**
Pulse population quintiles p<0.001		
<70	18 (3.8)	475
70-79	11 (2.7)	415
80-89	23 (5.1)	447
90-105	23 (5.1)	451
>105	44 (9.2)	479
Missing	9 (9.4)	96
Respiratory rate population quintiles p<0.001		
<15	11 (3.1)	357
15-16	11 (2.4)	467
17-18	11 (2.6)	421
19-22	19 (6.3)	301
>22	52 (15.6)	333
Missing	24 (5.0)	484
Systolic BP population quintiles p<0.001		
<110	46 (11.6)	397
110-129	25 (4.6)	545
130-139	12 (3.5)	346
140-160	9 (1.6)	573
>160	22 (5.5)	398
Missing	14 (13.5)	104
Diastolic BP population quintiles p<0.001		
<65	47 (10.1)	466
65-69	14 (3.6)	391
70-75	13 (3.4)	381
76-90	22 (3.1)	713
>90	16 (5.3)	300
Missing	16 (14.3)	112
Temperature population quintiles p=0.012		
<36	42 (9.0)	469
36-36.3	19 (5.1)	373
36.4-36.8	22 (4.6)	477
36.9-37.3	15 (4.2)	358
>37.3	17 (4.5)	377
Missing	13 (4.2)	309
Pulse pressure population quintiles p<0.001		
<40	35 (9.4)	373
40-52	25 (5.0)	505
53-64	14 (3.1)	456
65-80	16 (3.2)	498
>80	21 (5.0)	417
Missing	17 (14.9)	114
GCS a priori groups p<0.001		

3-5	13 (54.2)	24
6-8	7 (26.9)	26
9-12	26 (23.4)	111
13-14	18 (5.5)	327
15	40 (2.6)	1564
Missing	24 (7.7)	311
Oxygen saturations breathing air *a priori* groups p<0.001		
Very low (<90)	22 (16.8)	131
Low (90-3)	7 (4.0)	174
Normal (94-100)	28 (2.0)	1417
Missing	71 (11.1)	641
Oxygen saturations breathing supplemental oxygen population tertiles p<0.001		
<95	17 (15.2)	112
95-8	17 (7.7)	220
99-100	15 (6.9)	219
Missing	79 (4.4)	1812
Age population quintiles p<0.001		
<50	4 (1)	435
50-69	17 (3.4)	504
70-78	27 (5.4)	498
79-85	30 (6.4)	466
>85	50 (10.9)	460

Where visual inspection of the groups suggested similar risk profiles across different groups, these were collapsed. Thus the variables in table 13 remained for further analysis.

Table 13: Simplified variables for further analysis

Variable	Simplified groups
Pulse	<80, 80-105, >105
Respiratory rate	<19, 19-22, >22
Systolic blood pressure	<110, 110-139, 140-160, >160
Diastolic blood pressure	<65, 65-90, >90
Temperature	<36, >35.9
Pulse pressure	<40, 40-52, 53-80, >80
GCS	3-5, 6-12, 13-14, 15
Oxygen saturations	Low risk: >95 breathing air Moderate risk: 90-95 breathing air or >94 breathing oxygen High risk: <90 breathing air or <95 breathing oxygen
Age	<50, 50-69, 70-85, >85

4.5 Handling of missing data

In 79 cases with known outcomes all physiological variables were missing. The number of deaths was 6 (8%) among those with missing data compared with 122 (6%) without. This gives a X^2 value of 0.757 (df=1, p=0.384) for missingness of all variables being associated with poor outcome and thus it was not assumed that missingness was systematically associated either with "obvious wellness" or severe illness as discussed in 3.9.2 above.

Each variable was therefore assessed visually using the histograms in appendix 6 and missing data collapsed into the most closely comparable risk group. Thus missing pulse, systolic blood pressure, diastolic blood pressure, oxygen saturations and pulse pressure were treated as the highest risk groups (>105, <110, <65, high risk and <40 respectively). Missing respiratory rate and GCS were treated as moderate risk (19-23 and 13-14), while missing temperature was treated as low risk (>36). It is plausible that this reflects pragmatic practice, where pulse, blood pressure and oxygen saturations are measured using an automated machine and therefore their absence indicates "unmeasurability" (see 3.9 above), while respiratory rate, GCS and temperature need to be measured manually and often require a specific patient-based trigger to be recorded and therefore their absence indicates "unmeasuredness". High rates of missingness in respiratory rate recording have also been noted in the TARN trauma dataset, leading that research group to abandon a prognostic model that requires respiratory rate (125).

4.6 Univariate logistic regression analysis of predictor variables

Binary logistic regression was used with the simplified groups as developed above to generate estimated odds ratios (expressed as exp(B)) for each of the risk groups. Where multiple groups existed for a variable, the odds ratio was calculated in comparison with the lowest risk group. Variables reaching significance at the 90% level are highlighted in red in table 14.

Table 14: Univariate logistic regression analysis

Variable	**exp(B)**	**p**	**95% CI for exp(B)**	
Active malignancy present	3.59	<0.001	2.072	6.248
Age (ref <50)		<0.001		
50-69	3.76		1.256	11.264
70-85	6.771		2.441	18.783
>85	13.140		4.704	36.710
Pulse (ref <80)		<0.001		
80-105	1.603		0.998	2.576
>105	3.014		1.892	4.802
Respiratory rate (ref <19)		<0.001		
19-23	2.221		1.405	3.513
>23	6.582		4.162	10.410
Systolic BP (ref 140-60)		<0.001		
<110	8.526		4.185	17.370
110-139	2.715		1.300	5.669
>160	3.667		1.670	8.050
Diastolic BP (ref 65-90)		<0.001		
<65	3.585		2.435	5.277
>90	1.651		0.926	2.944
GCS (ref 15)		<0.001		
3-5	45.027		19.013	106.634
6-12	12.089		7.319	19.970
13-14	2.685		1.724	4.182
SaO2 (ref low >95 breathing air)		<0.001		
High (<90 air/<95 O2)	9.721		5.903	16.008
Moderate (90-5 air/ >94 O2)	2.490		1.472	4.213
Temperature <36	2.133	<0.001	7.019	10.157
Pulse pressure (ref 53-80)		<0.001		
<40	3.682		2.316	5.853
40-52	1.604		0.933	2.759
>80	1.633		0.924	2.888
Respiratory disease present	1.550	0.072	0.962	2.498
Heart disease present	1.028	0.885	0.706	1.497
Asthma present	0.608	0.158	0.305	1.213
Diabetes present	1.128	0.619	0.702	1.811
Epilepsy present	0.551	0.315	0.172	1.763
Warfarin therapy	0.586	0.250	0.236	1.456
Steroid therapy	1.167	0.664	0.580	2.350

4.7 Interactions between significant variables

Logistic regression was used to examine first order interactions between variables significant at the 10% level by assessing the predictive significance of both variables and their interaction term. Wald and p-values for the non-significant interactions are in appendix 7. Significant interactions are in table 15.

Table 15: Significant interactions

Interaction term	Wald	Degrees of freedom	P
Active malignancy by age	182.419	3	<0.001
Active malignancy by GCS	51.059	3	<0.001
Active malignancy by temperature	144.752	1	<0.001
Active malignancy by pulse pressure	25.060	3	<0.001
Age by temperature	90.820	3	<0.001
Pulse by temperature	319.717	2	<0.001
Respiratory rate by temperature	186.342	2	<0.001
SBP by temperature	29.071	3	<0.001
DBP by temperature	156.644	2	<0.001
GCS by temperature	15.375	3	0.002
SaO2 by temperature	216.794	2	<0.001
Temperature by pulse pressure	19.101	3	<0.001
Temperature by respiratory disease	158.828	1	<0.001
Pulse pressure by respiratory disease	46.431	3	<0.001

Full details are in appendix 7, but briefly it appears that active malignancy is more predictive of poor outcome in younger patients, and those with a low GCS and temperature or normal pulse pressure; that tachycardia, hypoxia and extremes of SBP and pulse pressure are more predictive of poor outcome in patients with a low temperature; that low temperature is less predictive of poor outcome in patients with a very low GCS; and that widened pulse pressure is less predictive of poor outcome in patients with respiratory disease.

4.8 Multivariate analysis of predictor variables

Variables and interactions significant at the 10% level were block entered into a multivariate logistic regression analysis. This produced estimated odds ratios (Exp(B)) and p values as in table 16. Variables significant at the 95% level are highlighted in red.

Table 16: Multivariate analysis of predictor variables

	Exp(B)	p
Active malignancy	230.974	<0.001

Age (ref <50)		0.906
50-69	*20253*	*0.998*
70-85	*29027*	*0.998*
>85	*24519*	*0.998*
Pulse (ref <80)		0.295
80-105	*1.406*	*0.293*
>105	*1.735*	*0.123*
Respiratory rate (ref <19)		0.001
19-23	*1.482*	*0.190*
>23	*3.731*	*<0.001*
Systolic BP (ref 140-60)		0.040
<110	*9.543*	*0.007*
110-139	*5.682*	*0.010*
>160	*4.243*	*0.053*
Diastolic BP (ref 65-90)		0.380
<65	*1.399*	*0.301*
>90	*1.514*	*0.340*
GCS (ref 15)		0.116
3-5	*7.01*e8*	*0.997*
6-12	*4.190*	*0.022*
13-14	*0.920*	*0.881*
SaO2 (ref low risk >95 breathing air)		<0.001
High risk (<90 air/<95 O2)	*3.639*	*<0.001*
Moderate risk (90-5 air/ >94 O2)	*1.400*	*0.327*
Temperature <36	2.417	0.218
Pulse pressure (ref 53-80)		0.790
<40	*1.368*	*0.702*
40-52	*1.886*	*0.399*
>80	*0.825*	*0.867*
Respiratory disease present	0.821	0.640
Active malignancy by age interaction		0.258
Active malignancy and age 50-69	*59232852*	*0.999*
Active malignancy and age 70-85	*21081320*	*0.999*
Active malignancy and age >85	*4100951*	*0.999*
Active malignancy by GCS interaction		0.545
Active malignancy and GCS 3-5	*5.189*e14*	*0.997*
Active malignancy and GCS 6-12	*0.369*	*0.424*
Active malignancy ad GCS 13-14	*0.222*	*0.166*
Active malignancy by temperature interaction	5.306	0.083
Active malignancy by pulse pressure interaction		0.139
Active malignancy and pulse pressure <40	*4.472*	*0.276*
Active malignancy and pulse pressure 40-52	*13.782*	*0.045*
Active malignancy and pulse pressure >80	*1.633*	*0.777*
Age by temperature interaction		0.582
Age 50-69 and temperature <36	*0.761*	*0.841*
Age 70-85 and temperature <36	*0.296*	*0.338*
Age >85 and temperature <36	*0.433*	*0.506*

Pulse by temperature interaction		0.076
Pulse 80-105 and temperature <36	*0.234*	*0.025*
Pulse >105 and temperature <36	*0.535*	*0.382*
Respiratory rate by temperature interaction		0.956
Respiratory rate 19-23 and temperature <36	*1.09*	*0.885*
Respiratory rate >23 and temperature <36	*0.893*	*869*
SBP by temperature interaction		0.307
SBP <110 and temperature <36	*6.723*	*0.253*
SBP 110-139 and temperature <36	*5.514*	*0.207*
SBP >160 and temperature <36	*15.385*	*0.067*
DBP by temperature interaction		0.841
DBP <65 and temperature <36	*1.185*	*0.794*
DBP >90 and temperature <36	*0.635*	*0.603*
GCS by temperature interaction		0.470
GCS 3-5 and temperature <36	*0.146*	*0.154*
GCS 6-12 and temperature <36	*1.216*	*0.776*
GCS 13-14 and temperature <36	*1.16*	*0.804*
SaO2 by temperature interaction		0.134
High risk (<90 air/<95 O2) and temperature <36	*4.171*	*0.046*
Moderate risk (90-5 air/ >94 O2) and temperature <36	*2.270*	*0.228*
Pulse pressure by temperature interaction		0.799
Pulse pressure <40 and temperature <36	*0.884*	*0.9*
Pulse pressure 40-52 and temperature <36	*0.475*	*0.398*
Pulse pressure >80 and temperature <36	*1.210*	*0.881*
Respiratory disease by temperature interaction	1.264	0.761
Pulse pressure by respiratory disease interaction		0.061
Pulse pressure <40 and respiratory disease	0.255	0.061
Pulse pressure 40-52 and respiratory disease	0.292	0.159
Pulse pressure >80 and respiratory disease	0.112	0.077

Taking the model at face value, some combinations of factors lead to inevitable death (see, for example, the coefficients for malignancy, age 50 and above, or GCS between 3 to 5). However, these very large odds ratios are suggestive of a model which has failed to converge; in other words, there are combinations of these factors with few and/or no events, causing the model to ascribe non-finite event probabilities. The likeliest explanation is in the interaction terms, of which five produced at least one category containing one or no events. Thus after taking statistical advice the following interaction terms were removed:

Active malignancy by age

Active malignancy by GCS

Active malignancy by pulse pressure

SBP by temperature

Pulse pressure by respiratory disease

The rerun multivariate logistic regression analysis produced results as in table 17

Table 17: Rerun multivariate analysis

	Exp(B)	**p**
Active malignancy	0.141	0.001
Age (ref <50)		0.062
50-69	*0.125*	*0.021*
70-85	*0.461*	*0.172*
>85	*0.472*	*0.081*
Pulse (ref <80)		0.797
80-105	*1.021*	*0.968*
>105	*0.764*	*0.622*
Respiratory rate (ref <19)		0.134
19-23	*0.328*	*0.045*
>23	*0.485*	*0.178*
Systolic BP (ref 140-60)		0.053
<110	*0.425*	*0.132*
110-139	*1.799*	*0.469*
>160	*1.324*	*0.666*
Diastolic BP (ref 65-90)		0.535
<65	*1.203*	*0.754*
>90	*1.887*	*0.347*
GCS (ref 15)		<0.001
3-5	*0.669*	*0.389*
6-12	*9.75*	*0.013*
13-14	*4.656*	*0.002*
SaO2 (ref low risk >95 breathing air)		<0.001
High risk (<90 air/<95 O2)	*0.461*	*0.148*
Moderate risk (90-5 air/ >94 O2)	*3.684*	*0.002*
Temperature <36	0.556	0.475
Pulse pressure (ref 53-80)		0.778
<40	*0.538*	*0.399*
40-52	*0.790*	*0.779*
>80	*0.587*	*0.505*
Respiratory disease present	1.407	0.566
Active malignancy by temperature interaction	2.526	0.197
Age by temperature interaction		0.477
Age 50-69 and temperature <36	*0.407*	*0.462*
Age 70-85 and temperature <36	*0.494*	*0.335*
Age >85 and temperature <36	*1.37*	*0.546*
Pulse by temperature interaction		0.088
Pulse 80-105 and temperature <36	*0.374*	*0.158*
Pulse >105 and temperature <36	*1.47*	*0.54*
Respiratory rate by temperature interaction		0.890
Respiratory rate 19-23 and temperature <36	*0.795*	*0.728*
Respiratory rate >23 and temperature <36	*0.739*	*0.633*
DBP by temperature interaction		0.536

DBP <65 and temperature <36	*0.494*	*0.335*
DBP >90 and temperature <36	*0.422*	*0.271*
GCS by temperature interaction		0.531
GCS 3-5 and temperature <36	*0.935*	*0.905*
GCS 6-12 and temperature <36	*5.017*	*0.199*
GCS 13-14 and temperature <36	*0.736*	*0.628*
SaO2 by temperature interaction		0.134
High risk (<90 air/<95 O2) and temperature <36	*2.252*	*0.206*
Moderate risk (90-5 air/ >94 O2) and temperature <36	*0.573*	*0.295*
Pulse pressure by temperature interaction		0.529
Pulse pressure <40 and temperature <36	*2.153*	*0.288*
Pulse pressure 40-52 and temperature <36	*1.517*	*0.573*
Pulse pressure >80 and temperature <36	*2.785*	*0.175*
Respiratory disease by temperature interaction	0.671	0.565

4.8b Recalculated linear coefficients

Variables significant at the 95% level in the multivariate analysis were block re-entered into a multivariate analysis to develop linear coefficients for the final model, with results as in table 18.

Table 18: Multivariate analysis for linear coefficients

	Exp(B)	**p**	**95% CI for exp (B)**	
Active malignancy	13.320	<0.001	8.36	21.22
GCS (ref 15)		<0.001		
3-5	342.163	<0.001	143.854	813.851
6-12	13.433	<0.001	8.122	22.217
13-14	2.200	<0.001	1.421	3.405
SaO2 (ref low risk >95 breathing air)		<0.001		
High risk (<90 air/<95 O2)	5.815	<0.001	3.579	9.449
Moderate risk (90-5 air/ >94 O2)	1.455	0.134	0.891	2.374

4.8c Rerun model

Given the very large coefficient for GCS 3-5 in the previous model and the small numbers in the GCS 3-5 group (24, with 13 deaths), the multivariate regression was rerun with the GCS categories 3-5 and 6-12 combined, with results as in table 19.

Table 19: Model with GCS categories collapsed

	p	**Exp(B)**	**95% CI for Exp(B)**	
Age (ref <50)	<0.001			
50-69	0.124	0.608	0.322	1.147
70-85	<0.001	0.348	0.201	0.603
>85	0.787	0.924	0.520	1.642
Respiratory rate (ref <19)	<0.001			

19-23	0.012	1.753	1.129	2.721
>23	<0.001	2.791	1.730	4.5
Diastolic BP (ref 65-90)	<0.001			
<65	<0.001	3.172	2.090	4.815
>90	<0.001	5.204	3.250	8.333
SaO2 (ref low risk >95 breathing air)	<0.001			
High risk (<90 air/<95 O2)	<0.001	3.202	1.956	5.243
Moderate risk (90-5 air/ >94 O2)	0.754	0.932	0.598	1.452
Temperature <36	<0.001	4.714	3.104	7.158
GCS (ref 15)	<0.001			
3-12	<0.001	29.372	18.674	46.197
13-14	<0.001	5.311	3.403	8.291
Respiratory disease present	<0.001	12.355	8.202	18.609
Respiratory disease and temperature <36	<0.001	7.466	3.134	17.784

4.9 Equation to predict death at 7 days

The coefficients generated in stage 4.8c were entered into an equation to generate a scaled probability of death:

$P(death) = e^a/1+e^a$

where a = (age50-69*0.608) + (age70-85*0.348) + (age>85*0.924) + (respiratoryrate19-23*1.753) + (respiratoryrate>23*2.791) + (DBP<65*3.172) + (DBP>90*5.204) + (sao2high*3.202) + (sao2mod*0.932) + (temperature<36*4.714) + (GCS3-12*29.372) + (GCS13-14*5.311) + (respiratorydisease*12.355) + (respiratorydisease^temperature * 7.466).

Entering the value generated by the equation as the sole variable in a logistic regression analysis generated a Nagelkerke R^2 value of 0.778, indicating adequate calibration.

4.10 Generating a bedside score to predict death at 7 days

In order to generate a score useable at the bedside, the coefficients generated in 4.8c were approximated to integers as discussed in 3.12 above, resulting in the score in table 20.

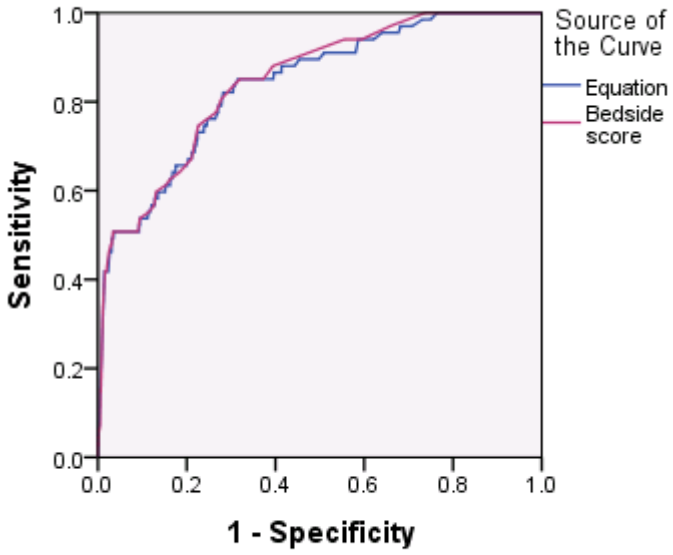
Table 20: Bedside score to predict death at 7 days

	Exp (B)	**Rounded**
Age (ref <50)		
50-69	0.608	1
70-85	0.348	1
>85	0.924	1
Respiratory rate (ref <19)		
19-23	1.753	2
>23	2.791	3
Diastolic BP (ref 65-90)		
<65	3.172	3
>90	5.204	5
SaO2 (ref low risk >95 breathing air)		
High risk (<90 air/<95 O2)	3.202	3
Moderate risk (90-5 air/ >94 O2)	0.932	1
Temperature <36	4.714	5
GCS (ref 15)		
3-12	29.372	30
13-14	5.311	5
Respiratory disease present	12.355	12
Respiratory disease and temperature <36	7.466	7

4.10a Performance of bedside score in predicting death

The performance of the bedside score was assessed using a ROC curve, which had an area under the curve (c-statistic) of 0.847 (95% confidence interval 0.8-0.894), indicating no loss of discrimination during the rounding process. The ROC curve is shown in figure 15 superimposed on that of the equation for comparison. It is possible that the slight improvement in performance from the equation to the bedside score highlights a small degree of overfitting in the equation, particularly in the age categories.

Figure 15: ROC curves for equation and bedside score to predict death in 7 days



Diagonal segments are produced by ties.

4.11 Limitations

The major limitation of this section of the study is its single-site nature. Ideally data from multiple hospitals would have been examined, with site of care included in logistic regression analysis, to explore whether predictors of death were consistent across all sites and address the constant risk fallacy (72). The full DAVROS study in fact found that some variables (particularly oxygen saturations) were not consistently predictive of death in all sites (76). However the Northern General dataset was selected in order that the subset of patients used for chapter 5 would be as close to this dataset as possible.

4.12 Summary

- Data from 2437 patients including 128 who died within 7 days of ED presentation was collected.
- Multivariate analysis identified 7 readily available factors and one interaction (age, respiratory rate, diastolic blood pressure, oxygen saturation, temperature, GCS, pre-existing respiratory disease and respiratory disease by temperature interaction) predicting death within 7 days.
- These factors were used to generate an equation with adequate discrimination and calibration: $P(death) = e^a/1+e^a$

  where a = (age50-69*0.608) + (age70-85*0.348) + (age>85*0.924) + (respiratoryrate19-23*1.753) + (respiratoryrate>23*2.791) + (DBP<65*3.172) + (DBP>90*5.204) + (sao2high*3.202) + (sao2mod*0.932) + (temperature<36*4.714) + (GCS3-12*29.372) + (GCS13-14*5.311) + (respiratorydisease*12.355) + (respiratorydisease^temperature * 7.466).
- The equation was simplified to integer values to enable use at the bedside with no loss of discriminatory ability.

Chapter 5
Derivation of a score to predict potentially preventable or potentially prevented death at 7 days

5.1 Introduction

This chapter will describe the data set assembled for derivation of a score predicting potentially prevented and preventable deaths ("potentially preventable" deaths), or solely potentially prevented deaths, and the process of deriving those scores.

5.2 Data set

Sets of casenotes relating to patients presenting to the Northern General Hospital by ambulance from February to May 2008, as described in 4.2 above, were screened in date order until 99 outcomes of interest (death or potentially prevented death) were included, for a sample size as discussed in 3.10 above. In total 697 cases were assessed, presenting between 11/2/08 and 4/03/08. 242 sets of notes were inaccessible due to changes in the casenote numbering system. 57 sets of notes were scrutinised and the case then excluded as the patient had sustained trauma (18), or a fractured neck of femur (22) or there was no entry in the notes for the relevant attendance (17). 398 cases were therefore included in the derivation set.

5.3 Study population

Characteristics of the study population are in table 21. To explore the representativeness of the data set, comparison was made of the recorded physiological variables between included patients and those with missing outcomes. Characteristics of the population with missing outcomes are also in table 21. The only notable difference is in the proportion of patients with active malignancy, which may reflect that patients with malignancy are more likely to be followed up, and their casenotes therefore to be active and accessible.

Table 21: Characteristics of the derivation study population

	Study population		Missing	
Variable	Mean (sd)	Range	Mean (sd)	Range
Age	66.5 (20.2)	18-102	67.2 (18.5)	18-97
Pulse	92 (24)	35-188	87 (24)	37-196
Respiratory rate	20 (7)	6-45	19 (6)	12-46
Systolic BP	133 (29)	45-243	132 (27)	62-208
Diastolic BP	75 (15)	36-130	75 (15)	35-124
Temperature	36.5 (1.2)	26.0-40.0	36.5 (1.1)	30.5-39.7
Variable	Median (IQR)	Range	Median (IQR)	Range
GCS	15 (15-15)	3-15	15 (15-15)	3-15
SaO2 breathing air	97 (94-98)	66-100	97 (95-98)	71-100
SaO2 breathing oxygen	97 (96-99)	71-100	97 (95-99)	73-100
Variable	Number	Percentage	Number	Percentage
Male	183	46	117	48.3
Active malignancy	23	5.8	7	2.9

5.4 Patients considered to have outcomes of interest

The patient group of interest in this study is those where emergency care had the potential to affect survival (or non-survival), ie where a patient death was potentially preventable or potentially prevented as discussed in 3.7.2 above. 4 outcomes have been recorded:

- inevitable death, where the patient died within 7 days of presentation and a decision was made within 24 hours of admission not to attempt CPR in the event of cardiac arrest;
- potentially preventable death, where the patient died within 7 days of presentation, but no decision was made to withdraw or limit care;
- potentially prevented death, where the patient survived for 7 days following presentation, but within those 7 days received a potentially life-saving intervention (as defined in 3.7.3b);
- non-critical illness, where the patient survived for 7 days following presentation without receiving a potentially live-saving intervention.

Full details of the patients are given in appendix 8. 15 patients were judged to have sustained inevitable death, 5 patients were judged to have sustained potentially preventable death and in 79 patients death was judged to have been potentially prevented.

5.5 Univariate analysis of potential predictor variables

Each analysis was carried out for two outcomes: potentially prevented death alone, and potentially prevented death plus potentially preventable death. It was hypothesised at the outset that there might be some physiological indicators of reversibility of the patient's condition such that factors which predicted potentially prevented death would not be the same as those which predicted potentially preventable death.

5.5a Demographic and historical variables

Categorical demographic and historical values were crosstabulated against outcome at 7 days and significant associations identified using chi-squared testing. Full details are in table 22 below. No categorical variables were significant at the 90% level.

Table 22: Univariate analysis of categorical variables

Variable	**Potentially preventable and potentially prevented death**	**Total**
Gender p=0.194		
Female	39 (18.5)	211
Male	45 (24.6)	183
Missing	0	4
Active malignancy p=0.653		
No active malignancy	80 (21.3)	375
Active malignancy	4 (17.4)	23
Variable	**Potentially prevented death**	**Total**
Gender p=0.334		
Female	38 (18)	211
Male	41 (22.4)	183
Missing	0	4
Active malignancy p=0.761		
No active malignancy	75 (20)	375
Active malignancy	4 (17.4)	23

5.5b Continuous variables

Continuous physiological variables and age were assessed in three ways: quintiles of the study population, quintiles of the variable range and (where included in DAVROS) groups defined *a priori* within the DAVROS project to reflect generally accepted clinical definitions of normal. Full details are in appendix 9. GCS and oxygen saturations (both breathing air and on oxygen) were assessed in tertiles and quartiles to better reflect their population distributions. GCS was also assessed in *a priori* groups broadly consistent with currently accepted clinical risk categories.

After visual inspection of the histograms in appendix 9 population quintiles were chosen for all variables except systolic blood pressure, GCS and oxygen saturations where *a priori* groups were used.

Continuous variables significant at the 90% level for potentially prevented and potentially preventable death were pulse, respiratory rate, systolic blood pressure, pulse pressure, GCS, oxygen saturations and age, as shown in table 23.

Table 23: Univariate analysis of continuous variables to predict potentially preventable and potentially prevented death

Variable	**Potentially preventable and potentially prevented death**	**Total**
Pulse population quintiles p<0.001		
<71	6 (8)	76
71-82	15 (20)	76
83-95	16 (20)	80
96-110	11 (14)	81
>110	34 (44)	77
Missing	2 (25)	8
Respiratory rate population quintiles p=0.016		
<16	7 (12)	59
16-17	17 (21)	80
18	13 (17)	76
19-23	15 (25)	60
>23	25 (35)	72
<16	7 (14)	51
Systolic BP a priori groups p=0.003		
Very low (<100)	15 (36)	42
Low (100-120)	22 (26)	84
Normal (120-180)	37 (15)	244
High (>180)	7 (39)	18
Missing	3 (30)	10

Pulse pressure population quintiles p=0.005		
<40	21 (28)	76
40-50	26 (30)	86
51-62	7 (10)	72
63-76	10 (13)	75
>76	16 (21)	77
Missing	4 (33)	12
GCS a priori groups p=0.015		
3-5	2 (66)	3
6-8	4 (57)	7
9-12	7 (33)	21
13-14	10 (20)	49
15	49 (18)	275
Missing	12 (28)	43
Oxygen saturations breathing air *a priori* groups p=0.056		
Very low (<90)	10 (32)	31
Low (90-3)	8 (33)	24
Normal (94-100)	39 (17)	231
Missing	27 (24)	112
Oxygen saturations breathing supplemental oxygen population tertiles p=0.439		
<96	7 (32)	22
96-99	16 (25)	63
100	4 (19)	21
Missing	57 (20)	292
Age population quintiles p=0.039		
<45	22 (28)	78
45-64	15 (19)	79
65-75	21 (29)	72
76-85	18 (19)	97
>85	8 (11)	72

To predict potentially prevented death alone, pulse, respiratory rate, systolic blood pressure, pulse pressure, oxygen saturations and age were significant at the 90% level (table 24).

Table 24: Univariate analysis of continuous variables to predict potentially prevented death

Variable	Potentially prevented	Total
Pulse population quintiles p<0.001		
<71	6 (8)	76
71-82	13 (17)	76
83-95	15 (19)	80
96-110	10 (12)	81
>110	33 (43)	77
Missing	2 (25)	8
Respiratory rate population quintiles p=0.056		
<16	6 (10)	59
16-17	17 (21)	80
18	13 (17)	76
19-23	14 (23)	60
>23	22 (31)	72
Missing	7 (14)	51
Systolic BP a priori groups p=0.005		
Very low (<100)	14 (33)	42
Low (100-120)	20 (24)	84
Normal (120-180)	35 (14)	244
High (>180)	7 (39)	18
Missing	3 (30)	10
Pulse pressure population quintiles p=0.003		
<40	21 (28)	76
40-50	24 (28)	86
51-62	6 (8)	72
63-76	8 (11)	75
>76	16 (21)	77
Missing	4 (33)	12
Oxygen saturations breathing air population quartiles p=0.05		
<95	20 (27)	73
95-6	13 (19)	67
97-8	13 (16)	79
99-100	6 (9)	67
Missing	27 (24)	112
Oxygen saturations breathing supplemental oxygen population tertiles p=0.264		
<96	7 (32)	22
96-99	16 (25)	63
100	4 (19)	21
Missing	52 (18)	292
Age population quintiles p=0.047		
<45	22 (28)	78
45-64	15 (19)	79
65-75	18 (25)	72
76-85	17 (18)	97
>85	7 (10)	72

Where visual inspection of the groups suggested similar risk profiles across different groups, these were collapsed. Thus the variables as in table 25 remained for further analysis.

Table 25: Simplified variables for further analysis

Potentially preventable and potentially prevented death	
Age	<45, 45-64, 65-75, 75-85, >85
Pulse	<71, 71-110, >110
Respiratory rate	<16, 16-18, 19-23, >23
SBP	<100, 100-120, 121-180, >180
Pulse pressure	<51, 51-76, >76
GCS	3-8, 9-12, 13-15
Oxygen saturations	High (<95 air or <96 O2), moderate (95-98 air or >95 O2), low (99-100 air)
Potentially prevented death	
Age	<45, 45-64, 65-75, 75-85, >85
Pulse	<71, 71-110, >110
Respiratory rate	<16, 16-18, 19-23, >23
SBP	<100, 100-120, 121-180, >180
Pulse pressure	<51, 51-76, >76
Oxygen saturations	High (<95 air or <96 O2), moderate (95-98 air or >95 O2), low (99-100 air)

5.6 Handling of missing data

Each variable was assessed visually using the histograms in appendix 9 and missing data collapsed into the most closely comparable risk group.

Thus missing pulse, systolic blood pressure, and pulse pressure were treated as the highest risk groups (>110, <110, and <51 respectively). Missing GCS was treated as moderate risk (9-12), while missing respiratory rate was treated as low risk (<16). If both oxygen saturations breathing air and oxygen were missing the variable was left as missing.

As discussed in chapter 4, it is plausible that this reflects pragmatic practice, where the absence of pulse and blood pressure indicates “unmeasurability”, while respiratory rate and GCS need to be measured manually and therefore their absence indicates “unmeasuredness”.

5.7 Univariate logistic regression analysis of predictor variables

Binary logistic regression was used with the simplified groups as developed above to generate estimated odds ratios (expressed as exp(B)) for each of the risk groups. Where multiple groups existed for a variable, the odds ratio was calculated in comparison with the lowest risk group. Variables reaching significance at the 90% level are highlighted in red in tables 26 and 27.

Table 26:Univariate logistic regression analysis of predictor variables for potentially preventable and potentially prevented death

Variable	**exp(B)**	**p**	**95% CI for exp(B)**	
Age (ref <45)		0.104		
45-64	*0.733*	*0.341*	*0.387*	*1.390*
65-75	*1.031*	*0.925*	*0.545*	*1.950*
75-85	*0.538*	*0.044*	*0.295*	*0.983*
>85	*0.554*	*0.074*	*0.290*	*1.058*
Pulse (ref <71)		<0.001		
71-110	*0.522*	*0.011*	*0.316*	*0.862*
>110	*3.756*	*<0.001*	*1.930*	*7.309*
Respiratory rate (ref <16)		<0.001		
16-18	*1.009*	*0.973*	*0.605*	*1.683*
19-23	*3.463*	*<0.001*	*1.819*	*6.594*
>23	*3.991*	*<0.001*	*2.140*	*7.443*
Systolic BP (ref 121-180)		<0.001		
<100	*4.461*	*<0.001*	*2.397*	*8.302*
100-120	*2.140*	*0.006*	*1.250*	*3.665*
>180	*31.475*	*<0.001*	*13.479*	*73.496*
Pulse pressure (ref 51-76)		<0.001		
<51	*1.791*	*0.011*	*1.143*	*2.805*
>76	*3.540*	*<0.001*	*2.030*	*6.174*
GCS (ref 13-15)		<0.001		
3-8	*34.446*	*<0.001*	*16.607*	*71.447*
9-12	*2.203*	*0.007*	*1.235*	*3.930*
SaO2 (ref low risk)		<0.001		
High	*1.583*	*0.156*	*0.839*	*2.987*
Moderate	*0.470*	*0.004*	*0.283*	*0.781*

Table 27: Univariate logistic regression analysis of predictor variables for potentially prevented death

Variable	**exp(B)**	**p**	**95% CI for exp(B)**	
Age (ref <45)		0.123		
45-64	*0.734*	*0.345*	*0.387*	*1.394*
65-75	*0.880*	*0.693*	*0.466*	*1.663*
75-85	*0.511*	*0.029*	*0.279*	*0.933*
>85	*0.526*	*0.052*	*0.275*	*1.005*
Pulse (ref <71)		<0.001		
71-110	*0.495*	*0.007*	*0.297*	*0.825*
>110	*4.583*	*<0.001*	*2.289*	*9.180*
Respiratory rate (ref <16)		<0.001		
16-18	*0.560*	*0.053*	*0.311*	*1.008*
19-23	*1.660*	*0.171*	*0.803*	*3.431*
>23	*1.744*	*0.117*	*0.870*	*3.497*
Systolic BP (ref 121-180)		<0.001		
<100	*5.294*	*<0.001*	*2.703*	*10.368*
100-120	*2.050*	*0.010*	*1.186*	*3.542*
>180	*34.284*	*<0.001*	*14.209*	*82.723*
Pulse pressure (ref 51-76)		<0.001		
<51	*1.947*	*0.004*	*1.230*	*3.081*
>76	*3.966*	*<0.001*	*2.262*	*6.953*
SaO2 (ref low risk)		<0.001		
High	*1.352*	*0.350*	*0.718*	*2.545*
Moderate	*0.427*	*0.001*	*0.257*	*0.710*

5.8 Interactions between significant variables

Logistic regression was used to examine first order interactions between variables significant at the 10% level by assessing the predictive significance of both variables and their interaction term. Wald and p-values for the non-significant interactions are in appendix 10 (potentially preventable and potentially prevented death) and appendix 11 (potentially prevented death alone). Significant interactions are in tables 28 and 29.

Table 28: Significant interactions: potentially preventable and potentially prevented death

Interaction term	Wald	Degrees of freedom	P
Pulse by respiratory rate	19.290	6	0.004
Pulse by SBP	19.200	6	0.004
Pulse by SaO2	20.088	4	<0.001
Respiratory rate by pulse pressure	16.431	6	0.012
Respiratory rate by SaO2	14.071	6	0.029
SBP by pulse pressure	32.734	4	<0.001
SBP by SaO2	20.303	6	0.002
Pulse pressure by SaO2	17.221	4	0.002

Full details are in appendix 10, but briefly it appears that hypoxia confers increased risk if tachycardia is absent or if tachypnoea is present, and that normal pulse pressure confers increased risk in the presence of systolic hypotension.

Table 29: Significant interactions: potentially prevented death

Interaction term	Wald	Degrees of freedom	p
Pulse by respiratory rate	16.646	6	0.011
Pulse by SBP	21.388	6	0.002
Respiratory rate by pulse pressure	14.386	6	0.026
Respiratory rate by SaO2	16.403	6	0.012
SBP by pulse pressure	36.601	4	<0.001
SBP by SaO2	22.282	6	0.001
Pulse pressure by SaO2	19.012	4	0.001

Again, full details are in appendix 11. It seems that hypoxia becomes of increasingly higher risk at higher respiratory rates.

5.9 Multivariate analysis of predictor variables

Variables significant at the 90% level were block entered into a multivariate logistic regression analysis. Based on a discernible pattern, interactions of oxygen saturations with pulse and respiratory rate were also entered for potentially prevented and preventable death, and interaction of oxygen saturation with respiratory rate for potentially prevented death alone. This produced estimated odds ratios (Exp(B)) and p values as in tables 30 and 31. Variables significant at the 95% level are highlighted in red.

Table 30: Multivariate analysis of predictor variables for potentially preventable and potentially prevented deaths

Variable	**exp(B)**	**p**	**95% CI for exp(B)**	
Pulse (ref <71)		0.02		
71-110	*0.371*	*0.419*	*0.034*	*4.101*
>110	*5.176*	*0.035*	*1.125*	*23.816*
Respiratory rate (ref <16)		0.228		
16-18	4.814	0.128	0.636	36.443
19-23	19.144	0.047	1.035	354.172
>23	6.664	0.094	0.723	61.405
Systolic BP (ref 121-180)		0.061		
<100	*3.687*	*0.036*	*1.089*	*12.481*
100-120	*1.112*	*0.836*	*0.407*	*3.043*
>180	*2.468*	*0.203*	*0.615*	*9.905*
Pulse pressure (ref 51-76)		0.536		
<51	*1.479*	*0.440*	*0.548*	*3.992*
>76	*1.729*	*0.302*	*0.611*	*4.894*
GCS (ref 13-15)		<0.001		
3-8	*23.013*	*0.001*	*3.806*	*139.163*
9-12	*4.060*	*0.034*	*1.114*	*14.790*
SaO2 (ref low risk)		0.971		
High	*0.946*	*0.954*	*0.146*	*6.123*
Moderate	*0.840*	*0.834*	*0.164*	*4.300*
Pulse/SaO2 interaction		0.210		
Respiratory rate/SaO2 interaction		0.210		

Table 31: Multivariate analysis of predictor variables for potentially prevented deaths

Variable	Exp(B)	p	95% CI Exp(B)	
Pulse (ref <71)		<0.001		
71-100	*0.809*	*0.587*	*0.376*	*1.739*
>110	*3.607*	*0.006*	*1.446*	*9.0*
Respiratory rate (ref <16)		0.342		
16-18	*0.864*	*0.778*	*0.314*	*2.379*
19-23	*1.861*	*0.276*	*0.609*	*5.686*
>23	*1.552*	*0.449*	*0.497*	*4.845*
Systolic BP (ref 121-180)		<0.001		
<100	*3.432*	*0.014*	*1.287*	*9.151*
100-120	*1.489*	*0.335*	*0.662*	*3.349*
>180	*15.342*	*<0.001*	*4.410*	*53.374*
Pulse pressure (ref 51-76)		0.554		
<51	1.458	0.371	0.638	3.332
>76	1.495	0.378	0.611	3.654
SaO2 (ref low risk)		0.257		
High	0.954	0.926	0.353	2.580
Moderate	0.572	0.201	0.243	1.346
Respiratory rate/SaO2 interaction		0.434		

5.9b Recalculated linear coefficients

Variables significant at the 95% level in the multivariate analysis were block re-entered into a multivariate analysis to develop linear coefficients for the final model, with results as in tables 32 and 33.

Table 32: Multivariate analysis for linear coefficients to predict potentially preventable and potentially prevented death

	Exp(B)	p	95% CI for exp (B)	
Pulse (ref <71)		<0.001		
71-110	*2.373*	*0.068*	*0.937*	*6.010*
>110	*9.533*	*<0.001*	*3.440*	*26.418*
Systolic BP (ref 120-180)		0.015		
<100	*3.614*	*0.004*	*1.521*	*8.586*
100-120	*1.488*	*0.252*	*0.754*	*2.934*
>180	*2.888*	*0.062*	*0.947*	*8.810*
GCS (ref 13-15)		<0.001		
3-8	*13.280*	*<0.001*	*4.543*	*38.819*
9-12	*2.770*	*0.045*	*1.022*	*7.508*

Table 33: Multivariate analysis for linear coefficients to predict potentially prevented death

	Exp(B)	p	95% CI for exp (B)	
Pulse (ref <71)		<0.001		
71-110	*0.883*	*0.707*	*0.462*	*1.690*
>110	*4.595*	*<0.001*	*2.158*	*9.784*
Systolic BP (ref 120-180)		<0.001		
<100	*4.700*	*<0.001*	*2.306*	*9.580*
100-120	*1.928*	*0.028*	*1.075*	*3.459*
>180	*21.700*	*<0.001*	*8.636*	*54.531*

5.10 Equation to predict potentially preventable or potentially prevented death

The coefficients generated in stage 5.9b were entered into an equation to generate a scaled probability of potentially preventable or potentially prevented death:

P(potentially preventable or potentially prevented death) = $e^a/1+e^a$

where a = (pulse[71-110]*2.373) + (pulse[>110]*9.533) + (sbp[<100]*3.614) + (sbp[100-120]*1.488) + (sbp[>180]*2.888) + (gcs[3-8]*13.28) + (gcs[9-12]*2.770)

Entering the value generated by the equation as the sole variable in a logistic regression analysis generated a Nagelkerke R^2 value of 0.383, indicating poor calibration.

5.11 Equation to predict potentially prevented death

The coefficients generated in stage 5.9b were entered into an equation to generate a scaled probability of potentially prevented death:

P(potentially prevented death) = $e^a/1+e^a$

where a = (pulse[71-110]*0.883) + (pulse[>110]*4.595) + (sbp[<100]*4.7) + (sbp[100-120]*1.928) + (sbp>[180]*21.7)

Entering the value generated by the equation as the sole variable in a logistic regression analysis generated a Nagelkerke R^2 value of 0.368, indicating poor calibration.

5.12 Generation of bedside scores

The coefficients generated in 5.9b above were approximated to integers (see 3.12 above) for ease of use in a bedside score as in tables 34 and 35.

Table 34: Bedside score to predict potentially preventable and potentially prevented death

	Exp(B)	**Rounded**
Pulse (ref <71)		
71-110	2.373	2
>110	9.533	10
Systolic BP (ref 120-180)		
<100	3.614	4
100-120	1.488	1
>180	2.888	3
GCS (ref 13-15)		
3-8	13.280	13
9-12	2.770	3

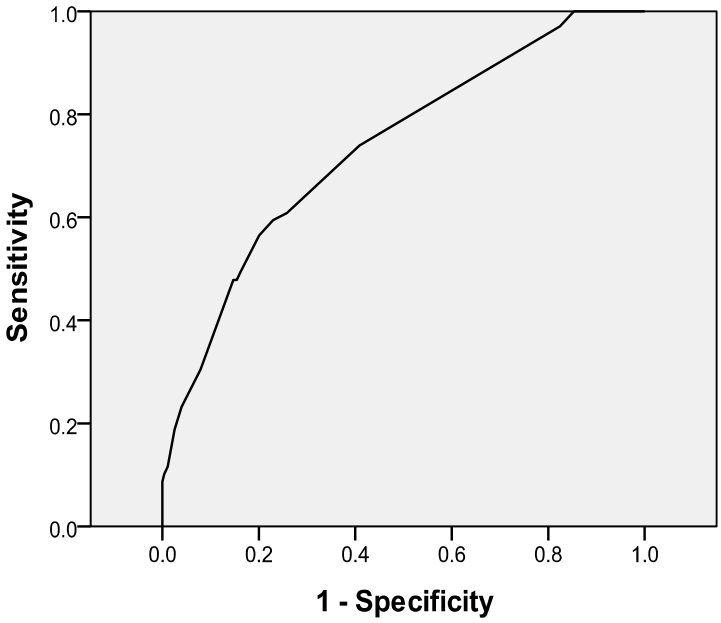
Table 35: Bedside score to predict potentially prevented death

	Exp(B)	**Rounded**
Pulse (ref <71)		
71-110	0.883	1
>110	4.595	5
Systolic BP (ref 120-180)		
<100	4.700	5
100-120	1.928	2
>180	21.700	22

5.12b Performance of score in predicting potentially preventable and potentially prevented death

The performance of the score was assessed using a ROC curve (figure 16), which had an area under the curve (c-statistic) of 0.737 (95% confidence interval 0.671-0.804), indicating no loss of discrimination during the rounding process.

Figure 16: ROC curve for bedside score to predict potentially preventable and potentially prevented death

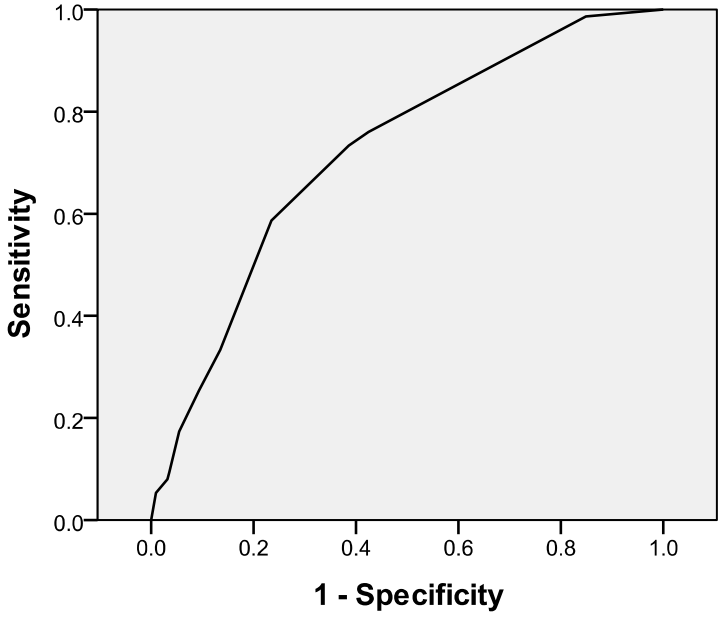


Diagonal segments are produced by ties.

5.12c Performance of score in predicting potentially prevented death

The performance of the score was assessed using a ROC curve (figure 17), which had an area under the curve (c-statistic) of 0.720 (95% confidence interval 0.658-0.782), indicating no loss of discrimination during the rounding process.

Figure 17: ROC curve for bedside score to predict potentially prevented death



Diagonal segments are produced by ties.

5.13 Limitations

This part of the study is limited by the exclusion of patients who were not admitted to hospital. This limits this study in terms of the development of a clinical score as there is no analysis of patients who were discharged from the ED, which may restrict generalisability. Ideally a full cohort of presenting patients would be studied and those discharged from the ED followed up as outpatients to ensure the absence of post-discharge adverse events. However within the scope of this study that was logistically unfeasible.

We did not analyse whether this standardized score added value in the detection of the at-risk patient to clinician gestalt. The original DAVROS protocol planned to collect paramedic impressions of severity of illness but this had to be abandoned due to poor rates of completion.

5.14 Summary

- Data from 398 patients including 5 who suffered potentially preventable death and 79 in whom death was potentially prevented within 7 days of ED presentation was collected.
- Multivariate analysis identified 3 readily available factors predicting potentially preventable or potentially prevented death within 7 days.
- These factors were used to generate an equation with reasonable discrimination but poor calibration: P(potentially preventable or potentially prevented death) = $e^a/1+e^a$

  where a = (pulse[71-110]*2.373) + (pulse[>110]*9.533) + (sbp[<100]*3.614) + (sbp[100-120]*1.488) + (sbp[>180]*2.888) + (gcs[3-8]*13.28) + (gcs[9-12]*2.770).
- This was noted to differ from the equation predicting all deaths: P(death) = $e^a/1+e^a$

  where a = (age50-69*0.608) + (age70-85*0.348) + (age>85*0.924) + (respiratoryrate19-23*1.753) + (respiratoryrate>23*2.791) + (DBP<65*3.172) + (DBP>90*5.204) + (sao2high*3.202) + (sao2mod*0.932) + (temperature<36*4.714) + (GCS3-12*29.372) + (GCS13-14*5.311) + (respiratorydisease*12.355) + (respiratorydisease^temperature * 7.466).

- The equation was simplified to integer values to enable use at the bedside with minimal loss of discriminatory ability.

Chapter 6
Validation of scores to predict death at 7 days

6.1 Introduction

This chapter will describe the data set assembled for validation of a score predicting death at 7 days, and the process of validating that score. As discussed in 1.10 and 3.14 above, the process of validation addresses the issue of overfitting of a model, whereby a model fits the sampled data closely but misrepresents the true relationship between predictor and outcome variables.

This chapter will also describe the assessment of other published scores in predicting death at 7 days. As discussed in 2.3.2, a number of scoring systems have been developed and proposed for risk stratification in the undifferentiated emergency population. However as noted earlier there is little external validation of these systems and minimal head-to-head comparison to enable an informed choice between the scores in a clinical setting.

6.2 Data set

Patients presenting by emergency ambulance to the Emergency Department of the Northern General Hospital, Sheffield, from October to December 2008 were identified prospectively by DAVROS researchers, and data abstracted and recorded as in section 4.2 above. 2322 patients were included. This cohort was collected using the same setting and patient selection processes as the derivation cohort, but was temporally separate.

Characteristics of the study population are presented in table 36 (with the values of the derivation set for comparison):

Table 36: Characteristics of validation and derivation cohorts

Variable	Mean (sd; range in validation cohort)	Mean (sd; range) in derivation cohort	Difference between means (p-value)
Age	69.5 (19.1; 18-103)	69 (19; 18-103)	0.5 (0.51)
Respiratory rate	20 (6.6; 8-80)	19 (6; 6-60)	1 (0.564)
Systolic BP	139 (28.4; 44-261)	136 (29; 24-266)	3 (0.572)
Diastolic BP	76 (15; 11-151)	75 (15; 30-153)	1 (0.532)
Pulse rate	88 (23; 20-180)	88 (24; 21-215)	0
Temperature	36.5 (1.14; 25.2-40.5)	36.6 (1.2; 26.0-41.0)	0.1 (0.495)
Mean arterial pressure	97 (17.8; 32-186)		
Variable	Median (IQR; range) in validation cohort	Median (IQR;range) in derivation cohort	
SaO2 breathing air	97 (95-98; 45-100)	97 (95-98; 50-100)	
SaO2 breathing oxygen	98 (95-100; 60-100)	98 (95-100; 24-100)	
GCS	15 (15-15; 3-15)	15 (15-15; 3-15)	
Variable	Number (percentage) in validation cohort	Percentage in derivation cohort	p-value for difference in proportions
Male	1093 (46.5)	46.4	0.944
Active malignancy	96 (4.1)	4.5	0.503
Chronic respiratory disease	247 (10.5)	12.0	0.105
Heart disease	694 (29.5)	34.0	<0.001
Asthma	236 (10.0)	11.0	0.268
Diabetes	349 (14.9)	15.7	0.447
Epilepsy	63 (2.7)	4.0	0.014
Warfarin usage	139 (5.9)	6.3	0.569
Steroid usage	127 (5.4)	6.1	0.303
Outcome			
Alive at 7 days	2210 (94)	91.6	
Dead at 7 days	141 (6.0)	5.2	

The differences in characteristics between the derivation and validation sets are interesting; physiologically the groups were very similar, but the validation set had a slightly lower recorded incidence of pre-existing cardiac morbidity. This might reflect a different patient demographic in presentations, whereby previously well patients present early in the winter with seasonal problems, while by the spring many presentations are of patients with acute decompensation of chronic disease.

6.3 Statistical considerations

As discussed in 3.13 above, the performance of the various scores will be assessed using a ROC curve, as this provides information about the trade-off between sensitivity and specificity across all values of the score.

6.3a Sample size for validation

Using the Vergouwe calculation (122) discussed in 3.14 above:

$Z_\beta = \sqrt{(N\delta^2/\sigma^2)} - Z_{1/2\alpha}$

where N = sample size, δ = difference in model performance, Z_β = value of the standard normal distribution corresponding to β, with β = type II error rate, $Z_{1/2\alpha}$ = value of the standard normal distribution corresponding to 1/2α, with α = type I error rate and σ = standard error of performance measure*√n.

Solving this for N, for a 10% difference in model performance (δ=0.1), and using α=0.05, β=0.2, σ=0.019*√2361=0.9232 would require a total sample size of 53. The standard error of the performance measure was approximated from the ROC curve at 4.9. The apparently small sample size is related to the narrow confidence intervals of the original ROC and the statistical assumption that the rate of death would be the same in derivation and validation sets.

6.3b Sample size for comparison of scores

Using the Obuchowski tables (124) as discussed in 3.14 above and her closest assumptions to this data (moderate accuracy of test, six observers, small observer variability and 4:1 ratio of control:event patients), 103 patients would be required to detect a 0.1 difference in AUROC, and 424 to detect a 0.05 difference in AUROC.

6.4 Performance of equation in predicting death at 7 days

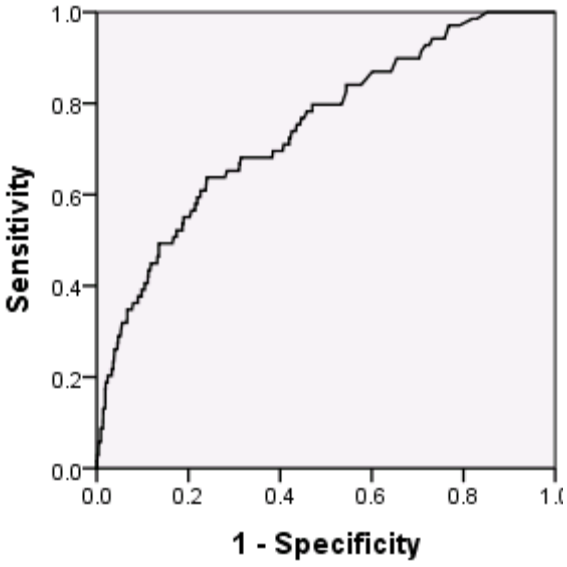
The equation generated in stage 4.9:

$P(death) = e^a/1+e^a$

where a = (age50-69*0.608) + (age70-85*0.348) + (age>85*0.924) + (respiratoryrate19-23*1.753) + (respiratoryrate>23*2.791) + (DBP<65*3.172) + (DBP>90*5.204) + (sao2high*3.202) + (sao2mod*0.932) + (temperature<36*4.714) + (GCS3-12*29.372) + (GCS13-14*5.311) + (respiratorydisease*12.355) + (respiratorydisease^temperature * 7.466).

was applied to this data set. The performance of the equation was assessed using a ROC curve (figure 18), which had an area under the curve (c-statistic) of 0.741 (95% confidence interval 0.685-0.806), indicating moderate discrimination.

Figure 18: ROC curve for equation to predict death at 7 days in validation cohort



Diagonal segments are produced by ties.

Entering the value generated by the equation as the sole variable in a logistic regression analysis generated a Nagelkerke R^2 value of 0.821, indicating good calibration.

6.5 Performance of bedside score in predicting death at 7 days

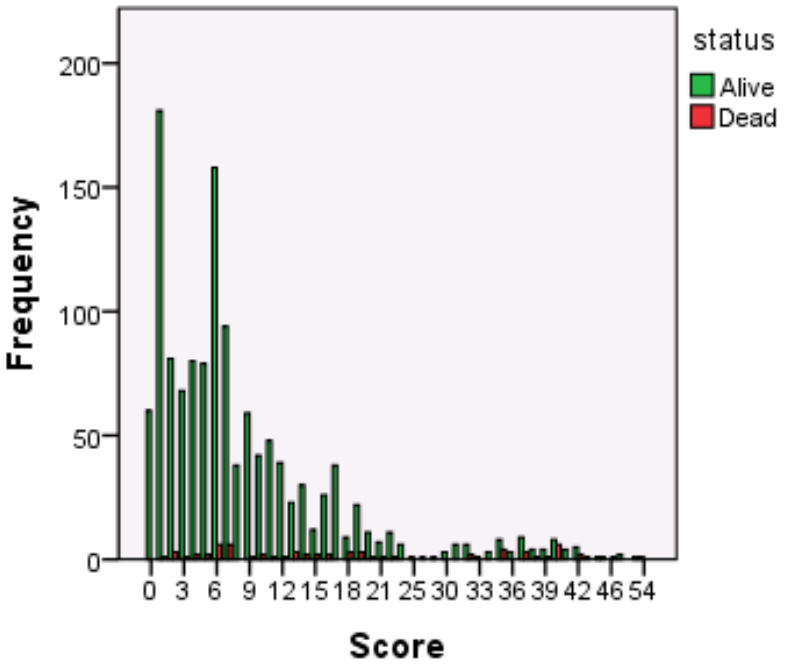
The bedside score generated in 4.10 (table 37) was applied to the data set.

Table 37: Bedside score to predict death at 7 days

Age	
50 or above	1
Respiratory rate	
19-23	2
>23	3
Diastolic BP	
<65	3
>90	5
SaO2	
High risk (<90 breathing air/<95 with supplemental O2)	3
Moderate risk (90-5 breathing air/ >94 with supplemental O2)	1
Temperature <36	5
GCS	
3-12	30
13-14	5
Respiratory disease present	12
Respiratory disease and temperature <36	7

The frequency of occurrence of various scores is shown in figure 19, subdivided by outcome. The distribution is heavily skewed to the lower scores.

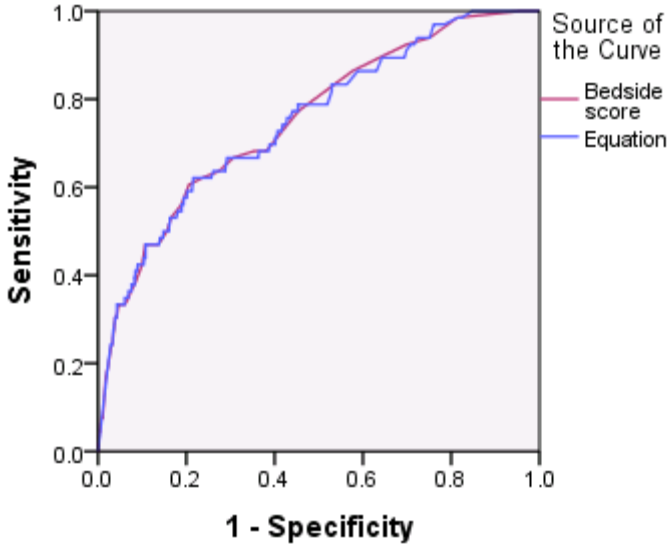
Figure 19: Frequency distribution of scores in the validation cohort



The performance of the score was assessed using a ROC curve which is shown in figure 20 superimposed on that of the equation for comparison. The area under the

curve (c-statistic) is 0.753 (95% confidence interval 0.691-0.815), indicating minimal loss of discrimination during the rounding process.

Figure 20: ROC curves for equation and bedside score in predicting death at 7 days in the validation cohort

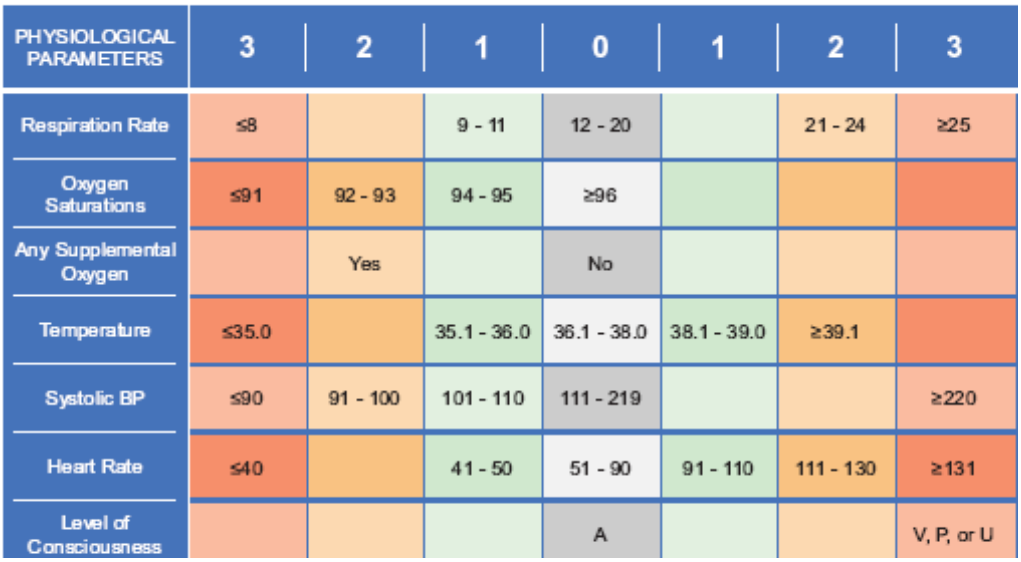


Diagonal segments are produced by ties.

6.6 Selection of other scores for analysis of performance

The scores identified in 2.3.2 and described in appendix 1 were examined for suitability for this section. They were identified in the literature review as having been developed or advocated for use in the setting of undifferentiated patients, the population of interest in this study and represented by this cohort. The RCP NEWS score (figure 21) was also assessed as its use has been advocated in all acute care settings.

Figure 21: RCP National Early Warning Score



PHYSIOLOGICAL PARAMETERS	3	2	1	0	1	2	3
Respiration Rate	≤8		9 - 11	12 - 20		21 - 24	≥25
Oxygen Saturations	≤91	92 - 93	94 - 95	≥96			
Any Supplemental Oxygen		Yes		No			
Temperature	≤35.0		35.1 - 36.0	36.1 - 38.0	38.1 - 39.0	≥39.1	
Systolic BP	≤90	91 - 100	101 - 110	111 - 219			≥220
Heart Rate	≤40		41 - 50	51 - 90	91 - 110	111 - 130	≥131
Level of Consciousness				A			V, P, or U

Not all scores identified could be calculated as some variables required for the scores were not available in the data set. These were HOTEL, PRE-EMPT2, the simple clinical score and the Sun score and the missing variables are shown in table 38.

Table 38: Scores which could not be calculated

HOTEL	ECG, mobility status
PRE-EMPT 2	PaCO2, H^+
Simple clinical score	Clinical features, ECG, premorbid functional status
Sun score	Racial group, recent hospitalisation, unstructured triage

However, adequate information was available to calculate the majority of scores, which are shown with their constituent variables in table 39. Throughout, where an AVPU was required, GCS was converted to AVPU as 15=A, 13-14=V, 9-12=P, 3-8=U.

Table 39: Scores analysed for prediction of death at 7 days

	Structure	**Variables included**
Bispebjerg (87)	Weighted summative score	Consciousness (AVPU) Heart rate Respiratory rate Systolic blood pressure Temperature
Hillerød (88) (Vital signs section)	Multiple level all-or-nothing	Consciousness (GCS) Heart rate Oxygen saturation Respiratory rate Systolic blood pressure Temperature
Rapid Emergency Medicine Score (91)	Weighted summative score	Consciousness (GCS) Heart rate Mean arterial pressure Oxygen saturation Respiratory rate Temperature
South African Triage Score (93) and Cape Triage Score (94) (TEWS component excluding mobility)	Weighted summative score	Consciousness (AVPU) Heart rate Respiratory rate Systolic blood pressure Temperature
Vital Signs Score (96) (excluding airway and fitting)	Unweighted summative score	Consciousness (GCS) Heart rate Oxygen saturation Respiratory rate Systolic blood pressure
VitalPAC™ EWS (ViEWS) (97)	Weighted summative score	Consciousness (AVPU) Heart rate Inspired oxygen Oxygen saturation Respiratory rate Systolic blood pressure Temperature

6.7 Frequency distribution of other scores

Full details of the distribution of scores calculated from the data set are shown in appendix 10. Briefly, for all scores except the Hillerod score, higher scores are expected to be associated with a higher risk of poor outcome. Generally the majority of patients were scored as low risk on all the scores, although the higher weighting of any impaired consciousness in NEWS reduced the number of very low risk patients.

6.8 Performance of scores in predicting death at 7 days

The performance of the scores was assessed using ROC curves (figure 22 and table 40), which had areas under the curve (c-statistics) of 0.658 to 0.762, indicating poor to moderate discrimination.

Figure 22: ROC curves for scores to predict death at 7 days

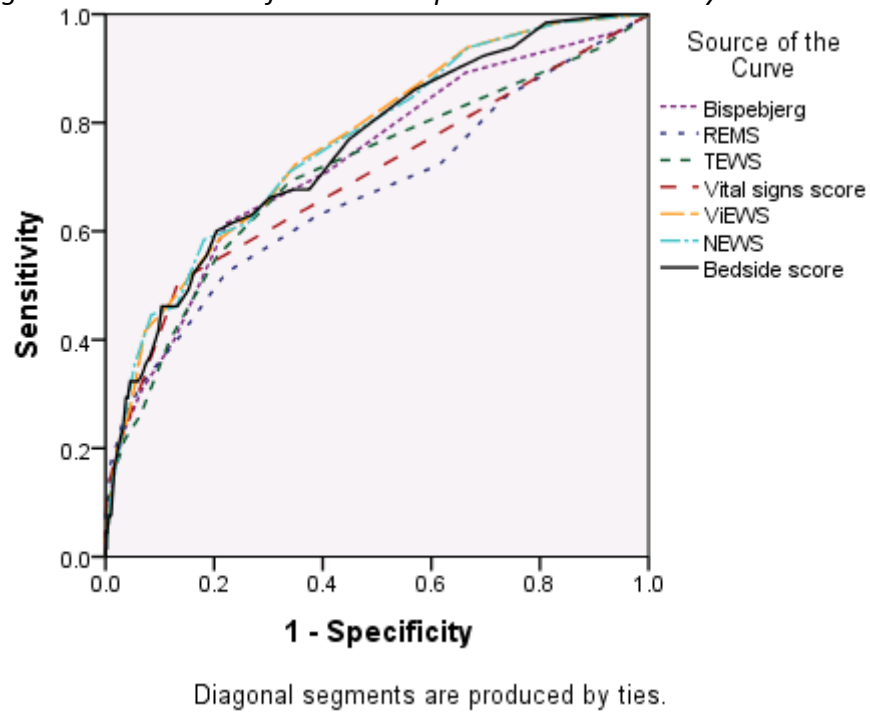


Table 40: Performance of scores in predicting death at 7 days

Score	AUROC	95% confidence interval
Bispebjerg	0.728	0.662-0.794
Hillerod	0.658	0.605-0.710
REMS	0.663	0.586-0.739
TEWS	0.711	0.640-0.782
Vital signs score	0.692	0.616-0.767
ViEWS	0.762	0.703-0.821
NEWS	0.761	0.702-0.821
New bedside score	0.753	0.691-0.815

6.9 Limitations

Some scores could not be calculated as not all variables were included in the data set. Of these only one of the four was designed for non-selective use in the emergency department.

6.10 Summary

- Data from 2322 patients including 383 who died within 7 days of ED presentation was collected.
- The equation derived earlier was validated in this data set and demonstrated moderate discrimination and calibration (AUROC 0.731).
- The simplified integer score derived earlier demonstrated minimal loss of discriminatory ability (AUROC 0.719).
- Other scores identified earlier were validated in this data set and demonstrated moderate discrimination, with AUROC from 0.658 – 0.762.

Chapter 7
Validation of score to predict potentially prevented and potentially preventable death

7.1 Introduction

This chapter will describe the data set assembled for validation of a score predicting potentially prevented and potentially preventable deaths, or solely potentially prevented deaths, and the process of validating those scores. As discussed in chapter 6, the process of validation addresses the issue of overfitting of a model.

It will also describe the assessment of other published scores in predicting potentially prevented and potentially preventable death at 7 days. As discussed previously, a number of scoring systems have been developed and proposed for risk stratification in the undifferentiated emergency population. However as noted there is even less data regarding their validation and performance in predicting non-death outcomes.

7.2 Data set

Sets of casenotes relating to patients presenting by emergency ambulance to the Emergency Department of the Northern General Hospital from October to December 2008 (those described in 6.2 above) were screened in date order. In total 498 cases were assessed. 252 sets of notes were inaccessible due to changes in the casenote numbering system. 19 sets of notes were scrutinised and the case then excluded as the patient had sustained trauma (3) or a fractured neck of femur (6) or there was no entry in the notes for the relevant attendance (10). 227 cases were therefore included in the validation set.

7.3 Study population

Characteristics of the study population are in table 41, with characteristics of the derivation set for comparison. The striking difference is the drop in numbers of patients receiving supplemental oxygen, given that the distributions of oxygen saturations breathing air are similar; it may be that this represents a change in practice following the publication of the British Thoracic Society guideline for emergency use of oxygen in 2008 (126).

Table 41: Characteristics of validation and derivation cohorts

Variable	**Mean (standard deviation)**	**Range**	**Mean (sd; range) in derivation set**
Age	71.3 (18.3)	19-96	66.5 (20.2; 18-102)
Respiratory rate	20 (5.8)	12-40	20 (7; 6-45)
Systolic BP	138 (28.6)	60-249	133 (29; 45-243)
Diastolic BP	74 (15)	36-142	75 (15; 36-130)
Pulse rate	89 (21.9)	35-152	92 (24; 35-188)
Temperature	36.6 (1.40)	26.4-39.6	36.5 (1.2; 26.0-40.0)
Mean arterial pressure	95 (18.3)	46-178	94 (17; 39-158)
Variable	**Median (interquartile range)**	**Range**	**Median (IQR; range) derivation set**
SaO2 breathing air	96 (94-98)	45-100	97 (94-98; 66-100)
SaO2 breathing oxygen	98 (96-100)	84-100	97 (96-99; 71-100)
GCS	15 (15-15)	7-15	15 (15-15; 3-15)
Variable	**Number**	**Percentage**	**Percentage in derivation set**
Male	94	42	46
Active malignancy	6	2.7	5.8
Receiving supplemental oxygen	34	14.9	26.6

7.3a Comparison of missing with included patients

To explore the representativeness of the data set, comparison was made of the recorded physiological variables between included patients and those with missing outcomes (table 42). There do not appear to be any notable differences, probably reflecting that casenotes tended to be missing relating to a casenote renumbering issue rather than due to any clinical matters.

Table 42: Comparison of validation population with those with missing outcomes

Variable	Mean (sd) included	Mean (sd) missing
Age	71.3 (18.3)	70.1 (17.8)
Respiratory rate	20 (5.8)	21 (6.9)
Systolic BP	138 (28.6)	138 (28.5)
Diastolic BP	74 (15)	75 (15)
Pulse rate	89 (21.9)	89 (23.6)
Temperature	36.6 (1.40)	36.5 (1.38)
Mean arterial pressure	95 (18.3)	96 (17.9)
Variable	Median (IQR) included	Median (IQR) missing
SaO2 breathing air	96 (94-98)	96 (94-98)
SaO2 breathing oxygen	98 (96-100)	98 (94-100)
GCS	15 (15-15)	15 (14-15)
Variable	Number (%) included	Number (%) missing
Male	94 (42)	121 (45.8)
Active malignancy	6 (2.7)	13 (4.9)
Receiving supplemental oxygen	34 (14.9)	52 (19.7)

7.3b Missing data for individual variables

Amongst included cases, data was missing for the following variables: pulse (23), respiratory rate (61), systolic blood pressure (24), diastolic blood pressure (24), GCS (53), oxygen saturations breathing air (62), oxygen saturations breathing supplemental oxygen (192) and temperature (39).

7.4 Patients considered to have outcomes of interest

The same four outcomes as in chapter 5 above have been recorded:

- inevitable death, where the patient died within 7 days of presentation and a decision was made within 24 hours of admission not to attempt CPR in the event of cardiac arrest;
- potentially preventable death, where the patient died within 7 days of presentation, but no decision was made to withdraw or limit care;
- potentially prevented death, where the patient survived for 7 days following presentation, but within those 7 days received a life-saving intervention (as defined in 3.7.1);
- non-critical illness, where the patient survived for 7 days following presentation without a requirement for live-saving intervention.

2 patients were judged to have sustained inevitable death, 1 patient was judged to have sustained potentially preventable death and in 35 patients death was judged to have been potentially prevented. Full details of the patients are given in appendix 13.

7.5 Statistical considerations

As discussed in 3.11 above, the performance of the various scores will be assessed using a ROC curve, as this provides information about the trade-off between sensitivity and specificity across all values of the score.

7.5a Sample size for validation

Using the calculation discussed in 3.14 above:

$Z_\beta = \sqrt{(N\delta^2/\sigma^2)} - Z_{1/2\alpha}$

where N = sample size, δ = difference in model performance, Z_β = value of the standard normal distribution corresponding to β, with β = type II error rate, $Z_{1/2\alpha}$ = value of the standard normal distribution corresponding to 1/2α, with α = type I error rate and σ = standard error of performance measure*√n.

Solving this for N, for a 10% difference in model performance (δ=0.1), and using α=0.05, β=0.2, σ=0.032*√386=0.629 would require a sample size of 255.

7.5b Sample size for comparison of scores

Using the Obuchowski tables (124) as discussed in 3.14 above and her closest assumptions to this data (moderate accuracy of test, six observers, small observer variability and 4:1 ratio of control:event patients), 103 patients would be required to detect a 0.1 difference in AUROC, and 424 to detect a 0.05 difference in AUROC.

7.6 Performance of equation for predicting potentially preventable and death

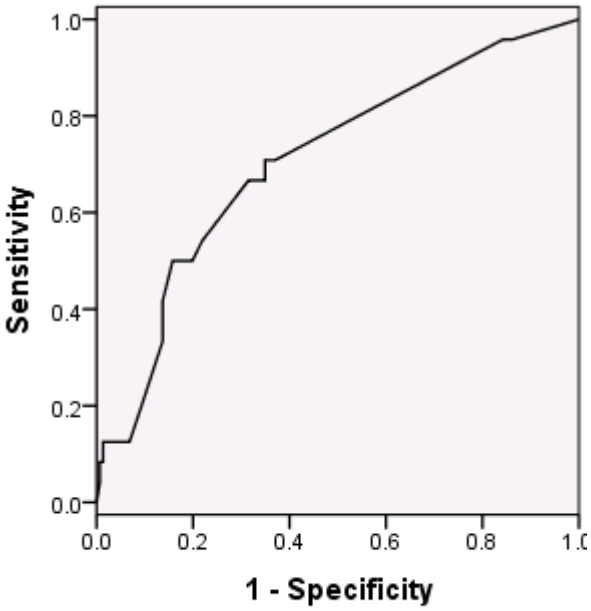
The equation generated in stage 5.8c above:

P(potentially preventable or potentially prevented death) = $e^a/1+e^a$

where a = (pulse71-110*2.373) + (pulse>110*9.533) + (sbp<100*3.614) + (sbp100-120*1.488) + (sbp>180*2.888) + (gcs3-8*13.28) + (gcs9-12*2.770)

was applied to this data set. The performance of the equation was assessed using a ROC curve (figure 23), which had an area under the curve (c-statistic) of 0.707 (95% confidence interval 0.594 - 0.820), indicating moderate discrimination.

Figure 23: ROC curve for equation in predicting potentially preventable or potentially prevented death at 7 days



Entering the value generated by the equation as the sole variable in a logistic regression analysis generated a Nagelkerke R^2 value of 0.545, indicating reasonable calibration.

7.7 Performance of bedside score in predicting potentially preventable and potentially prevented death at 7 days

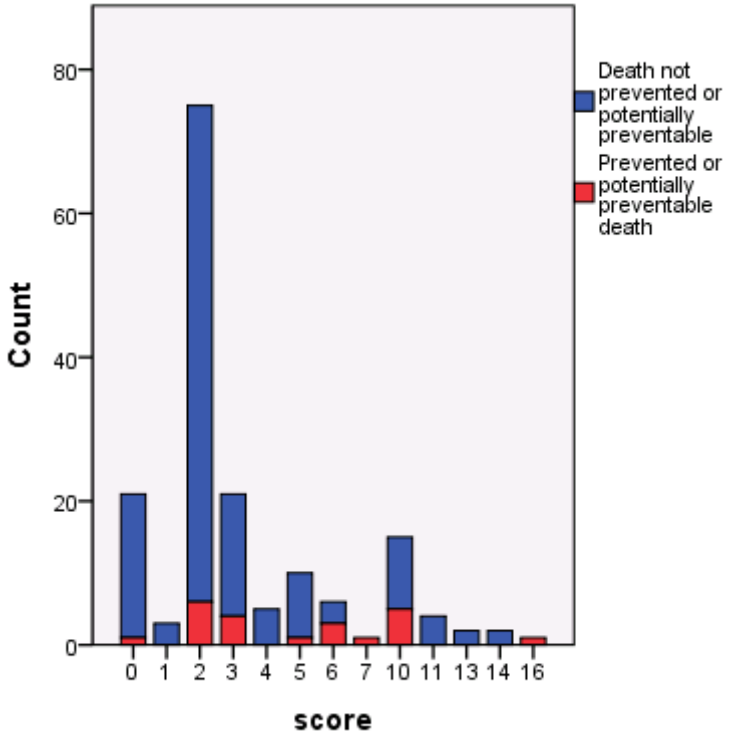
The bedside score generated in 5.12 (table 43) was applied to the data set.

Table 43: Bedside score to predict potentially preventable and potentially prevented death at 7 days

Pulse	
71-110	2
>110	10
Systolic BP	
<100	4
100-120	1
>180	3
GCS	
3-8	13
9-12	3

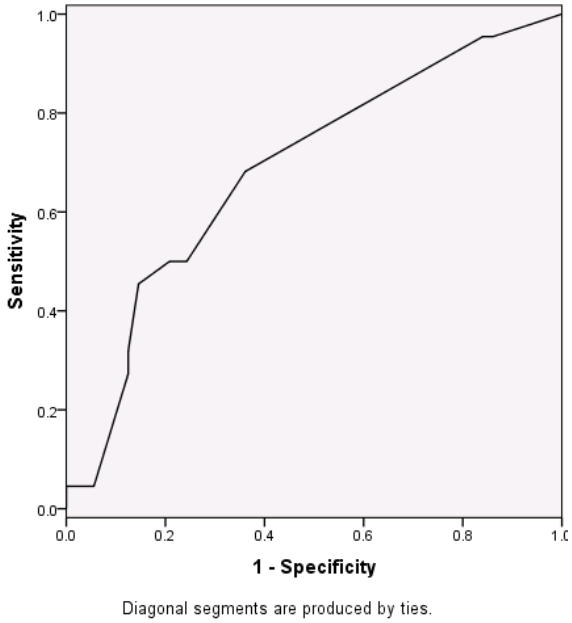
The frequency of occurrence of various scores is shown in figure 24, subdivided by outcome. The distribution is skewed to the lower scores.

Figure 24: Frequency distribution of bedside scores in the validation cohort



The performance of the score was assessed using a ROC curve (figure 25), which had an area under the curve (c-statistic) of 0.686 (95% confidence interval 0.568-0.804), indicating some loss of discrimination during the rounding process.

Figure 25: ROC curve for performance of the bedside score in predicting potentially preventable and potentially prevented death at 7 days



7.8 Selection of other scores for analysis of performance

The scores identified in 2.3.2 and described in appendix 1 were examined for suitability for this section. They were identified in the literature review as having been developed or advocated for use in the setting of undifferentiated patients. The RCP NEWS score (see section 6.6 above) was also assessed as its use has been advocated in all acute care settings.

7.9 Frequency distribution of other scores

Full details of the distribution of scores within the data set are shown in appendix 14. Generally one third to one half of patients were scored as low risk on all the scores, although as before the higher weighting of any alteration in consciousness in NEWS reduced the number of very low risk patients.

7.10 Performance of scores in predicting potentially preventable or potentially prevented death at 7 days

The performance of the scores was assessed using ROC curves (table 44), which had areas under the curve (c-statistics) of 0.559 to 0.684, indicating poor to moderate discrimination.

Table 44: Performance of scores in predicting potentially preventable and potentially prevented death at 7 days

Score	AUROC (95% CI)	AUROC in previous publications
Bispebjerg	0.684 (0.521-0.848)	Not previously assessed
Hillerod	0.613 (0.506-0.721)	Not previously assessed
REMS	0.615 (0.442-0.787)	0.852 (inhospital death from ED) (91) 0.74 (inhospital death in ED patients brought by ambulance) (127) 0.723 (14-day mortality in ED resus patients) (128) 0.771 (30-day mortality in ED resus patients) (129) 0.696 (7-day mortality or ICU admission in ED resus patients) (129)
TEWS	0.681 (0.523-0.839)	Not previously assessed
Vital signs score	0.589 (0.382-0.797)	0.72 (inhospital death) (96)
ViEWS	0.559 (0.357-0.762)	0.88 (death in 24hours on MAU) (97)
NEWS	0.637 (0.457-0.817)	Not previously assessed
New bedside score	0.686 (0.568-0.804)	

7.11 Limitations

This section of the study was greatly hampered by the non-availability of both ED records and complete patient notes, despite repeated attempts and hand-searching of the casenote library and archive at Northern General. Fortunately it appears (see 7.3a above) that this should not have skewed the results, although their power will have been affected.

7.12 Summary

- Data were collected from 227 patients including 36 patients where death was potentially prevented or potentially preventable.
- The equation derived earlier was validated in this data set and demonstrated moderate discrimination and calibration.
- The simplified integer score derived earlier demonstrated some loss of discriminatory ability.

- Other scores identified earlier were validated in this data set and demonstrated poor to moderate discrimination, with AUROC from 0.559 – 0.684.

Chapter 8
Operationalisation and implications

8.1 Introduction

This chapter will describe the process of using two expert groups to develop thresholds and responses for the implementation of the score developed in chapter 5. A formal cost-benefit analysis to develop explicit thresholds is outwith the scope of this thesis, both in terms of complexity and availability of data. The process will therefore involve implicit threshold setting by clinically credible active professionals. The operational implications of these thresholds will then be explored.

The score as developed might conceivably be used in a number of settings or for different purposes (at triage to stream patients to a co-located primary care provider, within the majors area of the ED to structure the order in which patients are assessed by a clinician, after treatment to determine whether patients should be admitted to a monitored area within an assessment unit etc). The process described in this chapter aims to explore how it might be used in a typical ED to prioritise patients and determine their location within the ED. This relates to the concepts of over- and under-triage discussed in 1.2.3 above; setting thresholds for various responses to a score requires consideration of the relative risks of over- and under-triaging to individual patients and across the ED. It is possible, indeed likely, that thresholds may not be absolute but may vary dependent on circumstances; most ED clinicians will recognise the scenario of having a patient they would rather be treating in the resuscitation area if only all the resuscitation beds were not already full. This is a necessary part of the research as realistically if implementation of the score resulted in more patients being located in the resuscitation area than in the waiting room, adoption would be unpopular and limited.

8.2 Settings

Two groups of emergency medicine consultants were assembled, one at the Northern General Hospital in Sheffield, as described in chapter 3. A second group was assembled at Lancashire Teaching Hospitals Trust. This is a trust which

operates Emergency Departments at two sites (Preston and Chorley), seeing 120,000 patients per year. Chorley is a District General Hospital with a limited range of services, while Preston is a teaching hospital of the University of Manchester with the regional neurosurgical service, an adult major trauma centre and a hyperacute stroke centre. Interventional cardiology is provided outside the Trust by transfer to Blackpool. Consultants and foundation trainees rotate between the two Emergency Departments, while middle grades and nursing staff are generally site-specific.

8.3 Methods

Eventually it would be ideal to validate the score as developed empirically in a functioning ED. This would, however, be well beyond the scope of this thesis. It was therefore decided to use the consensus opinion of experienced clinicians as a partial proxy for this.

Both processes as described below received ethical approval via the University of Sheffield. A number of techniques for achieving group consensus have been described; the first choice method for this study was a nominal group technique, as it allows face to face discussion but the structured process aims to reduce domination of the group by more senior or vocal individuals (130). The silent generation of responses (see 8.3a below) may also be useful in encouraging contributions from participants who prefer to consider then articulate their opinions. It is also logistically feasible with a group of clinicians, as it requires only a single group meeting. Initially it was hoped to carry out a nominal group process at both sites. However, due to service pressures at the Northern General Hospital it was not possible to gather all the consultants at the same time. The process was therefore amended to be performed electronically.

Both groups considered the same briefing document (appendix 15) including background information on the role of severity assessment and resources currently available for assessing patients in the ED, together with the score developed in chapter 5. The document included a deliberately non-specific description of an older patient presenting with lightheadedness and “weakness” and generating 10 different scores at presentation. The presenting complaint was intentionally non-

specific to reduce the chance of participants presuming a diagnosis which might affect their responses; it was hoped that use of a consistent scenario with varying scores would ensure that threshold decisions were based as much as possible on the score alone. It was also deliberately chosen to be potentially compatible with the full range of scores on the triage tool. The sets of observations were selected from the original data sets to include the minimum and maximum scores occurring in the data set (0 and 23) and various representative values across the possible range. It was felt to be impractical to ask for responses to all 36 possible scores, so a limit of 10 was set.

Participants were asked to consider each clinical vignette, addressing four specific points:

- Where in the ED should this patient be accommodated (resuscitation, majors, waiting room etc). This may include streaming to primary care if appropriate.
- How often should this patient have observations repeated?
- Which staff should be informed about the patient (nurse in charge, doctor in charge, any nurse, any doctor, no specific need to inform anyone)?
- How soon should the patient be assessed by a doctor or nurse practitioner?

8.3a Nominal group technique, Lancashire Teaching Hospitals

Consultants were recruited by email before a monthly meeting. No formal prespecified sample size or selection criteria were set, but the date of the meeting was selected to maximise the number of attendees.

At the meeting, the participants took part in a nominal group technique as described by Jones (131), facilitated by KC. After a welcome participants were asked to write their responses to each clinical vignette silently. Responses were entirely open, with no suggested options provided with the vignettes. Participants were then asked to share in turn their responses, without debate at this stage. Following this the group discussed their responses to reach consensus.

8.3b Electronic consensus process, Northern General Hospital

Consultants were recruited by email via a link consultant (Dr Shammi Ramlakhan). Participants were sent the same briefing document and vignettes and asked to respond to the vignettes via email.

Initially it was intended to return anonymised responses to participants for further consideration to reach consensus but this did not prove necessary.

8.3c Analysis

Formal statistical analysis of the nominal group data was not undertaken. Responses from the two sites was tabulated and compared visually to generate thresholds for various systematic and clinical responses to the score.

8.4 Results

8.4a Lancashire Teaching Hospitals

7 consultants (of 14) attended the nominal group session, with time in post ranging from just over 1 year to over 15 years. The group's initial responses are in table 45. The table shows the range of responses given to each request for opinion. As the aim was to reach consensus by discussion rather than weighting, frequencies of each response are not shown.

Table 45: Initial responses to clinical vignettes from Lancashire Teaching Hospitals

Area of ED	**Routine repeat observations**	**Staff to be informed**	**Timeliness of assessment**
Case 1; P70, BP 130/72, GCS 15; score 0			
Waiting room Primary care	3 hourly 2 hourly Hourly/when assessed None	None Nurse	1 hour 1-2 hour 2 hours 4 hours No urgency
Case 2; P61, BP 120/65, GCS 15; score 1			
Waiting room Primary care	3 hourly 1-2 hourly Hourly/when assessed None	None Any nurse	1 hour 1-2 hour 4 hours No urgency
Case 3; P84, BP 111/64, GCS 15; score 3			
Waiting room Waiting room/majors	Hourly 2-hourly 3-hourly	None Any nurse Nurse in charge	1 hour 1-2 hours 2 hours 4 hours
Case 4; P63, BP 93/69, GCS 14; score 4			
Majors Majors telemetry Resus	5 mins 15 mins 30 mins	Nurse in charge + any doctor Nurse in charge + doctor in charge Doctor in charge Any doctor Any doctor + any nurse	10 mins 15-30 mins 30 mins 1 hour
Case 5; P102, BP 99/56, GCS 14; score 6			
Majors Resus	5 mins 15 mins 30 mins	Nurse in charge Nurse in charge + any doctor Nurse in charge + doctor in charge Any doctor Any doctor + any nurse	10 mins 15 mins 30 mins 1 hour
Case 6; P35, BP 70/40, GCS 12; score 7			
Resus	5 mins 10 mins 15 mins On resus monitor	Nurse in charge + doctor in charge Doctor in charge	Immediate 10 mins 15 mins
Case 7; P134, BP 149/80, GCS 15; score 10			
Resus Majors	10 mins 15 mins 20 mins 30 mins	Nurse in charge Nurse in charge + any doctor Any doctor Any nurse	10 mins 15 mins 30 mins 1 hour
Case 8; P123, BP 199/94, GCS 15; score 13			
Resus Majors	10 mins 15 mins 20 mins	Nurse in charge Nurse in charge + senior doctor Nurse in charge + any doctor	10 mins 15 mins 30 mins

	30 mins	Any doctor Any nurse	1 hour
Case 9; P129, BP 82/57, GCS 12; score 17			
Resus	5 mins On resus monitor	Nurse in charge + doctor in charge Doctor in charge	Immediate 10 mins
Case 10; P114, BP 133/92, GCS 8; score 23			
Resus	5 mins 10 mins On resus monitor	Nurse in charge + doctor in charge Doctor in charge	Immediate

P: pulse; BP: blood pressure; GCS: Glasgow Coma Scale

Consensus was reached easily with brief discussion; issues recurrent in the discussion were:

- mismatches between responses that the respondents would have considered ideal and those achievable in terms of resources and external constraints (eg targets)
- respondents' perceptions of patient risk notwithstanding the score (particularly in cases 5 and 6).

The consensus achieved is shown in table 46. Notably there appear to be clear thresholds at case 2 (suitability for primary care streaming), case 3 (safety of the waiting room) and case 5 (need for resuscitation care). Interestingly case six, with a score of 7 (generated by a bradycardia of 35 and hypotension at 70/40), provoked more concern than cases seven (tachycardic at 134) and eight (tachycardic 123 and hypertensive 199/94). One respondent explicitly quoted the risk of death in undifferentiated hypotension whilst discussing case six. However, given that all respondents felt that cases seven and eight could only be managed in majors with continuous telemetry, this was not felt to be grossly inconsistent.

Table 46: Summary consensus at Lancashire Teaching Hospitals

Case	Area	Routine repeat observations	Personnel to be informed	Time to be seen
1	Primary care/waiting room	None	None	1-2 hours
2	Primary care/waiting room	None	None	1-2 hours
3	Waiting room	1-2 hours	Nurse	1-2 hours
4	Majors telemetry	15-30 mins	Nurse in charge + any doctor	15-30 mins
5	Majors telemetry	15-30 mins	Nurse in charge + any doctor	15-30 mins
6	Resus	Continuous	Nurse in charge + doctor in charge	Up to 10 mins
7	Resus/majors	15-30 mins	Nurse in charge + any doctor	15-30 mins
8	Resus/majors	15-30 mins	Nurse in charge + any doctor	15-30 mins
9	Resus	Continuous	Nurse in charge + doctor in charge	Immediately
10	Resus	Continuous	Nurse in charge + doctor in charge	Immediately

8.4b Northern General Hospital

3 consultants (of 17) responded to the questionnaire. The group's initial responses are in table 47.

Table 47: Initial responses to clinical vignettes from Northern General Hospital

Area of ED	**Routine repeat observations**	**Staff to be informed**	**Timeliness of assessment**
Case 1; P70, BP 130/72, GCS 15; score 0			
Waiting room Majors	None 2 hours	None Any nurse	2 hours
Case 2; P61, BP 120/65, GCS 15; score 1			
Waiting room Majors	None 2 hours	None Any nurse	2 hours
Case 3; P84, BP 111/64, GCS 15; score 3			
Waiting room Majors	None 1 hour 2 hours	None Any nurse	2 hours
Case 4; P63, BP 93/69, GCS 14; score 4			
Majors	15 mins 30 mins 1 hour	Any doctor Any doctor or nurse in charge Any nurse in majors	30 mins 1 hour
Case 5; P102, BP 99/56, GCS 14; score 6			
Majors	15 mins 30 mins 1 hour	Any doctor Any doctor or nurse in charge Any nurse in majors	30 mins 1 hour
Case 6; P35, BP 70/40, GCS 12; score 7			
Resus	5 mins Continuous	Doctor in charge	Immediate 1 min 5 mins
Case 7; P134, BP 149/80, GCS 15; score 10			
Majors Majors/resus	30 mins Continuous pulse monitoring	Any nurse Any doctor Doctor in charge	30 mins 30-60 mins
Case 8; P123, BP 199/94, GCS 15; score 13			
Majors Majors/resus	30 mins Continuous pulse monitoring	Any nurse Any doctor Doctor in charge	30 mins 30-60 mins
Case 9; P129, BP 82/57, GCS 12; score 17			
Resus	5 mins 15 mins Continuous	Doctor in charge	Immediate 1 min 5 mins
Case 10; P114, BP 133/92, GCS 8; score 23			
Resus	10 mins 15 mins Continuous	Doctor in charge	Immediate 5 mins

Given the low levels of response and the broad agreement with results from Lancashire Teaching Hospitals (apart from a single respondent who felt that the presenting complaint merited a majors assessment irrespective of physiological

findings), a decision was taken not to undertake a second round of consensus development.

8.4c Summary thresholds

After examination of the consensus responses in table 45, it was apparent that several broader groups of patient risk were perceived by clinicians. The categories and responses were therefore collapsed to:

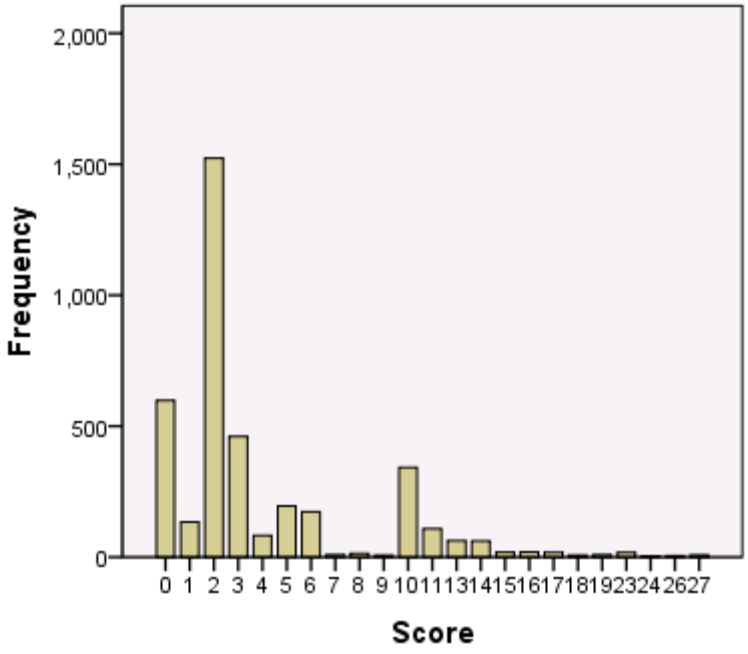
- Score 0 or 1: suitable for primary care streaming if available; no routine repeat observations required.
- Score 2 or 3: to wait in waiting room; 1-2 hourly repeat observations; a nurse to be aware of the patient.
- Score 4-6: to be allocated to majors area, preferably with telemetry; observations every 15-30 minutes; the nurse in charge and a doctor to be aware of the patient.
- Score 7 or above: to be allocated to the resuscitation area with continuous monitoring; the nurse and doctor in charge to be aware of the patient.

8.5 Implications for the running of the Emergency Department

In order to explore the potential implications of using the thresholds as developed above, the entire data set from the Northern General Hospital, as described in 4.2 and 6.2, was pooled. This totalled 4758 patients, with adequate data to calculate the score in 3880.

Frequencies of scores are as shown in figure 26. As noted in chapter 7, the distribution is generally skewed towards lower scores, although the spike with a value of 10 presumably represents a group of patients with an isolated tachycardia.

Figure 26: Distribution of bedside scores in the pooled data set

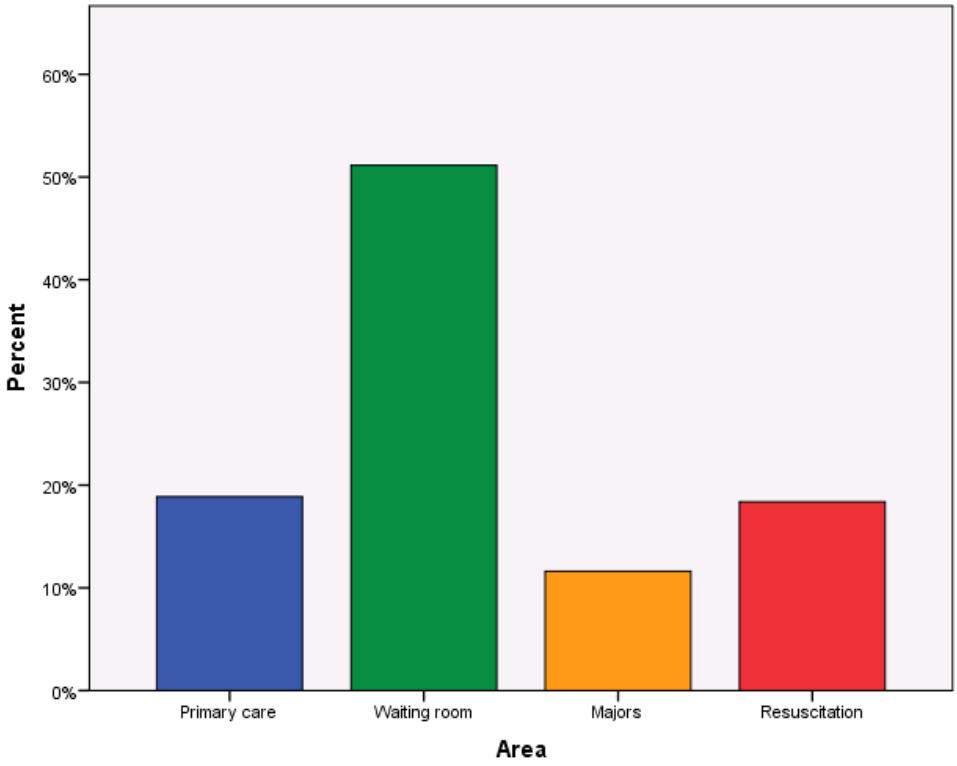


Applying the thresholds described in 8.4 above would result in 713 patients (18.4%) being allocated to the resuscitation area, 451 (11.6%) to majors, 1984 (51.1%) to the waiting room and 732 (18.9%) being streamed to primary care if available (figure 27). Over the 163 days of the study this approximates to 4.4 patients per day allocated to the resuscitation area, 2.8 to majors, 12.2 to the waiting room and 4.5 to primary care. This is of course a broad approximation as of course local policies may be presentation-dependent; for example that all patients with a focal neurological deficit must be seen in the resuscitation area (for potential stroke thrombolysis) or that no patients complaining of chest pain can be streamed to primary care.

Without formal assessment it seems that the score directs a much lower proportion of ambulance patients to the resuscitation or majors areas than currently happens in standard practice. This may represent an artefact from the deliberately non-specific nature of the case description, whereby non-physiologic moderating factors such as pain or vomiting were not included. It might also be a feature of the consensus process used, whereby responding clinicians were inclined to demonstrate their robust nature while when relatively isolated in actual clinical practice they might be more conservative. It must also be remembered that all patients in the data set were admitted to hospital; this provides further support for

the premise that potential to benefit in the longer term is not the same as acuity in the Emergency Department, as discussed in 1.4.3.

Figure 27: Potential allocation of patients using bedside score



8.5a Temporal distribution of severity

It is possible that patterns of patient acuity may differ by times of day, due to multiple factors (availability or lack of it of other services, changing patient thresholds for calling for help, reluctance of ambulance crews not to transport apparently vulnerable patients at antisocial hours). Crosstabulation was therefore carried out of acuity of presentation by time of presentation as recorded in the dataset. Results of this are in table 48. Rates of presentation per hour were calculated by dividing the total by the 163 days of data collection.

Table 48: Presentations crosstabulated by acuity and time of presentation

Time	Primary care (score 0-1)	Waiting room (score 2-3)	Majors (score 4-6)	Resuscitation (score >6)	Total
00-08	148 0.1/hr	454 0.3/hr	95 0.1/hr	148 0.1/hr	845
08-12	162 0.25/hr	413 0.6/hr	99 0.15/hr	167 0.25/hr	841
12-17	196 0.25/hr	490 0.6/hr	137 0.2/hr	162 0.2/hr	985
17-21	127 0.2/hr	398 0.6/hr	74 0.1/hr	135 0.2/hr	734
21-00	97 0.2/hr	224 0.5/hr	46 0.1/hr	100 0.2/hr	467
Total	730	1979	451	712	

It appears that presentations are fairly consistent across the day, apart from in the early hours where they decrease. Obviously this does not take into account self-presenting patients, those who were brought by ambulance but not admitted to hospital and trauma patients, all of whom were excluded from the dataset.

8.6 Limitations

This section of the study was limited by the non-inclusion of senior nursing staff. This decision was based on two factors: firstly that we wished to gather the opinions of senior clinicians with significant experience and not all the nursing staff in band 6 or 7 (sister) roles have this experience. Secondly it was a pragmatic decision based on their availability which was low due to clinical commitments.

The non-specific nature of the case vignette was deliberately structured to remove non-physiological cues (such as abdominal pain which might have made respondents concerned to exclude a rupturing aneurysm). However it is possible that the non-reporting of features such as pain, pallor or sweating may have made the respondents assume their absence and therefore be more reassured than they might have been in day-to-day practice.

8.7 Summary

- Consensus methodology was used in two hospital sites to develop threshold responses to the bedside score.
- Four levels of response were constructed, with linked levels of monitoring and urgency of clinical review: waiting room or primary care if available, waiting room, majors, resuscitation.
- Just under one-third of patients in the DAVROS dataset would be allocated to one of the two highest acuity categories (score 4 or above), representing around one patient every 2-3 hours presenting by ambulance and being admitted to hospital.

Chapter 9
Discussion

9.1 Introduction

Chapters 1 to 8 have described the current state of the literature regarding assessment of risk in unselected non-trauma Emergency Department patients, the derivation and validation of a bedside score to identify those patients at high risk of needing life-saving treatment within 7 days of presentation, and the thresholds set by a group of clinicians in terms of responses to that score. This chapter will explore the meaning and implications of this research for the NHS, potential limitations in the research and priorities for future research.

9.2 Existing evidence

Despite the significant and increasing logistical challenges presented by acute illness and acute severe illness to the NHS, proportionately minimal research has addressed the issue of identifying the acutely ill patient at high risk of adverse outcome and the acutely ill patient who has the potential to benefit from time-sensitive emergency care.

9.2.1 Clinical tools for the identification of the patient at risk

Of 136 clinical tools identified which had been developed for, or whose use had been advocated, in the emergency setting, only ten could reasonably be applied to the undifferentiated population which arrives in the Emergency Department and only two of these had been prospectively externally validated (table 7). The clinical utility of an instrument which requires a firm diagnosis to the general emergency physician is highly questionable, except in certain very limited circumstances such as ST-elevation myocardial infarction. The use of these tools also risks miscalculation of risk where an aberrant diagnosis is made initially.

The quality of the literature is alarmingly poor, with inadequately described data sets and statistical analysis. Where statistical analysis is described, it appears that there are generally unsafe assumptions of a monotonic linear relationship between most predictor variables and poor outcome. Many studies appear to have been

conducted *de novo* without reference to existing literature and hence there is minimal external validation of scores other than in specific diagnostic groups, notably community-acquired pneumonia and acute coronary syndromes.

Even less information exists regarding the impact on patients of these scores; the value of a score with perfect prediction might still be questioned if the identification of high-risk patients does not result in an intervention to reduce their risk. It often seems to be assumed in the literature that the risk of poor outcome (usually death) is the same as potential to benefit from treatment; as discussed in 1.4.3 above this is unlikely to be true in all patient populations. It may be that this will become more of a limitation as the emergency population becomes older with more chronic illness and the prediction of death becomes less helpful.

No well-validated clinical decision aid could be identified for the assessment of patients with medical emergencies.

9.2.2 Individual physiological variables in the prediction of poor outcome

Although vast numbers of publications report relationships between various physiological variables and assorted poor outcomes (again, usually death), accurate assessment of the true impact of alterations in physiology is impossible for a number of reasons: a) incomplete reporting; the majority of studies report only those variables found to be predictive in that study, with no description of other variables which were assessed; b) poor statistical reporting; as can be seen in table 8, means and p values are often reported as demonstrating a “significant” relationship between variables and outcome without fully exploring the nature of that relationship and whether it might be clinically useful; c) inadequate analysis; as above, odds ratios for risk are often constructed around either an arbitrary cut-off value or one derived using an assumption of linearity in the data.

Cautious analysis of the literature as reported does not seem to indicate any consistently reported effect of physiological derangement on the risk of adverse outcome, although such an effect would be biologically plausible.

9.3 New findings from this research

9.3.1 Objective 1: to generate a list of interventions which could be used to define potential to benefit from emergency (time-critical) care: section 3.7.3b.

We have, in this study, defined *a priori* a group of interventions which appear on best available evidence to be potentially life-saving and time-sensitive. However, the evidence supporting these is incomplete and our beliefs underlying many frequently-used interventions would bear further scrutiny; patients who could benefit could then be more reliably identified. Although our definitions of time-sensitive interventions might be criticized for being defined arbitrarily, the concept of an "Emergency Care Sensitive Condition" is being more widely embraced and supports our methodology (132).

9.3.2 Objective 2: to identify variables predicting potential to benefit from emergency care in a broader population than disease-specific scores: chapter 5.

This study adds the findings from a prospectively collected cohort where a bedside score including only readily-available data (pulse, systolic blood pressure and Glasgow Coma Score) could predict potentially prevented or potentially preventable death in the next seven days with moderate discrimination (AUROC 0.737), which fell to AUROC 0.686 in a validation set. Potential presentations of the score for clinical practice are shown in appendix 16.

9.3.3 Objective 3: to identify variables predicting death in emergency patients and explore whether these are the same as those predicting potential to benefit: chapter 4.

A bedside score to predict death within seven days was more complex, requiring age, respiratory rate, diastolic blood pressure, oxygen saturations, temperature, Glasgow Coma Score and a pre-existing diagnosis of respiratory disease. However, it had good discrimination (AUROC 0.847) although this fell to 0.753 in the validation set.

It is apparent in other published work that patterns of physiological derangement differ between patients who die and those who receive life-saving intervention (133). The most significant finding of this study is the discrepancy between

prediction of all deaths and prediction of potentially preventable or potentially prevented death. The two scores developed differed both in included variables (age, respiratory rate, diastolic blood pressure, oxygen saturation, temperature, GCS and pre-existing respiratory disease versus the simpler pulse, systolic blood pressure and GCS) and weightings (although included in both scores, the lowest values of GCS indicated death far more strongly than potential to benefit).

9.3.4 Objective 4: to develop near-patient scoring systems reflecting the variables identified in 2 and 3: sections 4.10 and 5.12 and objective 5: to explore whether the scoring systems generated in 4 as well as those already identified perform adequately in the prediction of death and potential to benefit in a broader patient population: sections 6.8 and 7.10.

The score to predict potential to benefit outperformed all others identified in the validation set (AUROCs 0.559-0.684) and the single existing published analysis (table 44). The score to predict death outperformed all other decision aids except ViEWS and NEWS (AUROCs 0.658-0.762) in the validation set.

9.4 Strengths of the study

This is the first study specifically to address the identification of at-risk patients in a multi-diagnosis emergency department population who had not already undergone triage, and to use the provision of a life-saving intervention as an outcome measure. It therefore addresses two of the major flaws inherent in applying the existing literature in emergency care; firstly, the lack of data from an emergency setting and failure to recognise the prognostic significance of emergency care and (non-)response to that care and secondly, the danger of assuming that variables which predict death will predict successful response to treatment.

9.5 Limitations of the study

Specific limitations are discussed in each chapter. There are a number of features of the study which might, in retrospect, have been structured differently.

It was felt at the outset that use of the existing DAVROS dataset of physiological variables would strengthen the study by easing the data collection load (effectively

restricting it to identification of outcomes) and therefore enabling a larger, and more statistically powerful, study cohort. The study was greatly hampered by the non-availability of both ED records and complete patient notes. This was multifactorial but the major underlying problem, that of casenote renumbering and subsequent loss of notes at Sheffield Teaching Hospitals, was outwith the control of the study team and not predictable at the inception of the project. In retrospect it might have been as effective to collect a data set *de novo,* particularly if one or more sites could have been identified with a robust electronic patient record such that the issue of missing data and sets of notes could have been minimised. It is possible that this might have affected the interpretation of outcomes in that automatically extracted data from an electronic record might have missed some subtleties of treatment obvious to a clinician reading the full notes. This is unlikely, however, as specific treatments of interest were defined *a priori* in this study and specifically sought in the paper records.

Patients who were not admitted to hospital were excluded from the initial DAVROS dataset as the aim was to develop a risk-adjustment tool for emergency admissions to hospital. This limits this study in terms of developing a clinical score as there is no analysis of patients who were discharged from the ED, which may restrict generalisability. Ideally a full cohort of presenting patients would be studied and those discharged from the ED followed up as outpatients to ensure the absence of post-discharge adverse events. However within the scope of this study that was logistically unfeasible. Rates of short-term death after discharge from the ED have been reported as 30/100000 (113) to 50/10000 (114), so the required cohort size to investigate this fully would have been impractical. No data was collected on potential non-death adverse outcomes in patients who were discharged from the ED; the international literature would suggest a 7-day bounce-back readmission rate of around 2.5% (134) but it is unclear how many of those patients receive a potentially life-saving intervention; it is likely to be low.

As a single-site study it may be that these results are not generalisable; interpretation of vital sign derangement is not only affected by patient factors but also by the health care system, staffing levels and types and time available for

patient care (135). Thus a process of external validation might find that life-saving interventions are provided differently in other settings. It also appears from the parent DAVROS study that, in the prediction of death, even physiological measurements are subject to the constant risk fallacy (76), which is likely also to limit generalisability of any decision aid developed in a single-site population.

It was not possible to analyse whether this standardized score added value in the detection of the at-risk patient to clinician gestalt; initial attempts were made in the DAVROS study to collect paramedic impressions of severity of illness but had to be abandoned due to poor rates of completion. A purpose-designed data set would ideally collect real-time impressions of patient acuity from nursing and medical staff to assess whether a decision aid contributed added value. Where this has been done in the existing literature (for assessing risk of acute coronary syndrome and pulmonary embolus), clinicians overestimated risk compared with a machine-based attribute matching system, but their discrimination (assessed using ROC curves) was the same (136). Interestingly, physician performance did not improve with experience, raising queries about the need for formalised risk assessment "to support junior staff".

The consensus methods used in chapter 8 to derive thresholds for responding to the score only involved doctors at two institutions and did not include nursing staff. This clearly has potential to be unrealistic in terms of non-physiological factors which may affect interpretation of risk in day-to-day clinical practice, particularly the holistic "end of the bed" assessment.

9.6 Implications for clinical practice and policy

This study has added to the volume of literature a purpose-derived score which identifies patients with the potential to benefit from time-sensitive care. It has potential to be valuable in clinical practice by enabling the prioritisation of patients for whom urgency of treatment will be beneficial. There are, however, a number of reasons why it should not yet be widely applied in standard practice. Green et al have recently updated the standards for the application of clinical decision rules originally developed by Stiell and Wells and discussed at 1.3.6 (137) and this study

has not fully addressed all their requirements. The score clearly needs wider validation, ideally including comparison with unstructured clinician (doctor or triage nurse) gestalt and with NEWS as the currently mandated standard of care. There must also be careful consideration as the score is applied of whether the outcomes used to develop definitions of urgency are still valid; the interventions listed at 3.7.3b were acknowledged at the time to be based on incomplete evidence; it is to be hoped that as the evidence base for emergency care is developed these can be refined (for example, intravenous magnesium in acute severe asthma would no longer be considered a potentially life-saving intervention (138)).

This leaves the working emergency clinician in the situation of not having a score developed for and demonstrated to work in the emergency setting. He or she has a number of options: firstly to use one of the existing scores (of which NEWS is probably the most appropriate) but to recognise its limitations in the emergency department; secondly not to use a score but to rely on the unstructured judgement of clinical staff. In the current politicomanagerial climate of high regard for standardised paperwork easily amenable to retrospective audit this is unlikely to be managerially palatable, despite the lack of supporting evidence. Given the current state of equipoise over the utility of standardised scores in terms of patient benefit in the ED clinicians should also be encouraged to participate in formal research to address the issue.

These results highlight the potential flaws in applying clinical scores to predict outcomes other than those for which they were originally derived. It is clear that in this data set variables which predict death are not the same as those which predict life-saving intervention, similar to the findings of Churpek et al, who found that amongst 291 deaths and 2638 emergency admissions to critical care from general wards, patterns of vital sign derangement differed (133). As the health economist Tony Culyer said “capacity to benefit is not identical to need” (139), and clinicians should be clear about the reasons for which a score is being used.

The importance of appropriate choice of outcomes is also relevant where acuity scores are used at an organisational level. If a scoring system is used during a

process of casemix adjustment in an attempt to assess or improve quality, it should be clear that it identifies conditions and outcomes which are amenable to alteration with good care (132). The small proportion of deaths deemed to be potentially preventable in this data set calls into question the practice of using mortality rates (even if standardised) to identify apparent extremes of organisational performance, due to an inappropriate signal to noise ratio, which has been modelled elsewhere (140).

The logistical issues encountered during this study highlight the need for the development of an agreed minimum data set for emergency care to enable more indepth analysis of care provided and its usefulness. Currently it is labour-intensive, if not impossible, to identify fundamental aspects of care provided (such as fluids and drugs administered), and diagnoses made and excluded. The College of Emergency Medicine Informatics Group has recognised this and is working on a minimum data set to demonstrate the "added value" of ED care.

Far more patients were deemed to have benefited from life-saving treatment than to have died, with 114 patients identified where death was potentially prevented versus 23 who died. Despite our liberal definition of potentially preventable death (any death where end of life care or "do not attempt resuscitation" was not specified within 24 hours of admission), only around 1 in 4 deaths (6 of 23) met our criteria for potential preventability. Although higher than previously reported "preventable deaths", deemed to be around 5% of all deaths in a larger UK cohort (141), this is still low and supports concerns about the use of mortality rates, however well standardised, as indicators of quality of care.

9.7 Implications for research

The variables we have identified as predictive of a need for life-saving intervention are not the same as those in use in many standardized early warning scores; this may well reflect the difficulty of using data sets which have been collected using death as an outcome. We used a single snapshot set of physiological data; it has been demonstrated in the acute medicine setting that changes in physiological scores are of value in prognostication (142); ideally ongoing research would

examine the prognostic value of response (or non-response) to treatment provided prehospitally or in the Emergency Department. Even a score developed to predict response to life-saving intervention does not perform more than adequately; it is likely that this outcome represents a heterogenous group of patients and this complexity should be recognised and explored in the ED setting.

Clearly identification of and response to the patient at risk requires more than a reliable scoring system; complex psychosocial and cognitive factors are at play in decision-making, particularly in the pressured environment of the ED (143), and examination of the interaction of these factors should be a priority for future research (144) .

The first issue requires a larger multisite cohort study with prospective data collection, including clinician gestalt from prehospital, nursing and medical staff. This would currently be very resource-intensive in terms of collecting outcome data due to the lack of an integrated robust patient record system in many hospitals. However it is likely to become more feasible as health information technology continues to develop.

The second and third issues could be addressed with a programme of research conceptualising assessment and response to risk in the ED patient as a complex intervention (145) and described in more detail below.

9.7.1 Development of a consensus regarding factors which should affect prioritisation in emergency care.

Patient-only focus groups would enable patients to contribute factors they feel are relevant to patient prioritisation without contamination from professional participants.

Discrete choice experimental technique formally addresses the issue of multiple potentially conflicting criteria (146) and is well established as a means of eliciting patient preferences (147). It has been used successfully to assess features of an emergency primary care service of importance to patients (148). For example, the

respondent might be asked “Of 2 patients, one with severe pain but no risk of death, and the other with mild pain but a high risk of death, who should be treated first?”. A pair of experiments, one recruiting members of the public and the other recruiting healthcare professionals working in UK EDs could explore similarities and differences between public and professionals in the prioritisation of patients in the ED.

9.7.2 Literature review to identify time-critical interventions

Once data on care provided in UK EDs can be collected from evolving IT systems, the most-frequently provided interventions could be the subject of a systematic literature review to examine the strength of the evidence base for the effectiveness and time-dependent benefit of these interventions.

9.7.3 Qualitative study of staff attitudes and behaviour

Face-to-face interviews and observational study along ethnographic principles could explore particularly attitudes to and constructs of patient prioritisation and how this relates to time-dependence of treatments. They would need to include staff across all professional groups in the ED (including medical and nursing staff from other specialities providing care and assessment within the ED). These staff would need be selected to represent professional attributes which might feasibly influence attitudes to patient risk (profession, specialty, grade and length of time in post) and EDs studied would need to be diverse in terms of size, training status and specialist services (such as trauma centre status, hyperacute stroke services).

9.7.4 Design and piloting of a toolkit

A toolkit might potentially include a scoring system, presentation-specific "red flag" reminders, or flowcharts. Based on findings from the earlier stages of the programme, it might include paper-based, website or app-based components, or parts may be integrated into an electronic patient record. It is likely to be designed to be used multiple times during the patient's ED stay, unlike current triage tools which are only used at initial presentation.

Piloting could involve toolkit implementation at several sites at varying times in a stepped wedge design to mitigate against external confounders such as central targets or initiatives. Outcomes should include:

- Observational study as conducted before the intervention
- Process measures of time-critical interventions identified earlier (for example antibiotics in sepsis).
- Other markers of healthcare resource use, such as the avoidable admission rate as developed by O'Cathain et al (149), which reflects the proportion of patients who present to acute hospitals with specific conditions who are then admitted to inpatient care. The conditions have previously been identified as those being amenable to ambulatory care (ie where admission might potentially be avoided).
- Critical incident reports generated within and relating to the ED which reflect issues around patient risk assessment or communication (for example a patient arriving on a general ward then requiring early transfer to a high dependency area), or prioritisation (for example delays in transfer to an operating theatre). Incident reporting can provide insight into both organizational culture and staff perceptions of risk and causation (150).

9.7.5 Formal trial of toolkit

Should the pilot trial demonstrate feasibility of implementing the toolkit, further research would require multicentre analysis of its effectiveness in improving patient-centred outcomes.

9.8 Summary

- No previously published tool has been developed and validated which addresses identification of the patient at risk of adverse outcome from the unselected ED population.
- Mixed, often conflicting, evidence exists for the prognostic value of multiple physiological variables in different patient sets.
- It cannot be assumed that the implementation of early warning scores suitable for inpatient areas of the hospital will benefit ED patients.

- Further research is required, ideally with a standardised minimum data set, to address the limitations of this study, particularly which patients will benefit from time-critical interventions.

Appendix 1
Decision aids for undifferentiated patient populations

The Bispebjerg Early Warning Score (BEWS) was developed in Denmark and involves a two-step nursing assessment; initially immediately life-threatening conditions such as uncontrolled haemorrhage or respiratory arrest are identified. If the patient does not have any of these a vital signs assessment is used to calculate the BEWS and a score of over 4 triggers an emergency call (87).

Bispebjerg Early Warning Score

	3	2	1	0	1	2	3
Respiratory rate	-	≤8	-	9-14	15-20	21-30	>30
Heart rate	-	≤40	41-50	51-100	101-110	111-130	>130
SBP	≤70	71-80	81-100	101-199	-	>199	-
Temperature	-	≤35	35.1-36	36.1-38	38.1-39	>39	-
Level of consciousness	-	-	-	Awake	Responds to voice	Responds to pain	Unresponsive

The Hillerød Acute Triage (HAPT) system (88) identifies high risk patients according to standardised vital signs monitoring and presentation-specific high risk criteria. Where the two are discordant, the patient is considered to be in the higher risk category.

Hillerød Acute Triage (HAPT) system

	1 Red Resuscitation	2 Orange Urgent	3 Yellow Less urgent	4 Green Not urgent
Vital signs				
A	Obstructed airway Stridor	Threatened airway		
B	SpO2 <80 RR >35 or <8	SpO2 80-89 RR 31-35	SpO2 90-94 RR 26-30	SpO2 ≥95 RR 8-25
C	HR >130 SBP <80	HR 121-30 or <40 SBP 80-89	HR 111-120 or 40-49	HR 50-110
D	GCS ≤8	GCS 9-13	GCS 14	GCS 15
E		Temp >40 or <32	Temp 38.1-40 or 32-34	Temp 34-38
Presentation-specific				
ECG changes	Lifethreatening	High risk	Low risk	Minor/normal
Ongoing chest pain	Very severe (VAS 10)	Severe (VAS 6-9)	Moderate (VAS 1-5)	No pain
Chest pain in last 24 hours		Ischaemic type		Non-ischaemic type
Dyspnoea on exertion		Very severe	Severe	None
Severe comorbidity			Yes	No

The HOTEL score was developed in Ireland to predict early mortality amongst acute medical patients. It allocates a point for each of Hypotension (SBP<100mmHg), low Oxygen saturation (<90%), Temperature (<35C), abnormal ECG and Loss of ability to stand unaided (89).

PREEMPT-2 was developed in Scotland as a response to the loss of ICU provision at one hospital site and the subsequent need to triage patients for transfer (90).

PREEMPT-2 = $0.39713 + f_1(x_1) + f_2(x_2) + f_3(x_3) + f_4(x_4) + f_5(x_5) + f_6(x_6) + f_7(x_7)$, where:

x_1 = age	$f_1(x_1) = 15.03253 \times (x_1/100)^2 - 19.37260 \times (x_1/100)^3$
x_2 = respiratory rate	$f_2(x_2) = 123.31922 \times (1/x_2)^2 + 19.29794 \times (x_2/100)^3$
x_3 = SBP	$f_3(x_3) = - 0.335839 \times x_3^{1/2}$
x_4 = AVPU	$f_4(x_4) = 1.60142 \times (x_4=V) + 1.90037 \times (x_4=P) + 3.15080 \times (x_4=U)$
$x_5 = SpO_2$	$f_5(x_5) = 1.90957 \times (100/x_5)^2$
$x_6 = PaCO_2$	$f_6(x_6) = 0.76629 \times (x_6/10)^3$
$x_7 = H^+$	$f_7(x_7) = - 95.22595 \times (1/x_7)$

The Rapid Emergency Medicine Score is a modification of the Intensive Care APACHE II score (64) designed to be calculable in the Emergency Department (91).

Rapid Emergency Medicine Score

	0	1	2	3	4
Temperature	36-38.4	38.5-38.9 34-35.9	32-33.9	39-40.9 30-31.9	>40.9 <30
MAP	70-109	-	110-129 50-69	130-159	>159 <49
Heart rate	70-109	-	110-139 55-69	140-179 40-54	>179 <39
Respiratory rate	12-24	25-34 10-11	6-9	35-49	>49 <6
Peripheral oxygen saturation	>89	86-89		75-85	<75
GCS	>13	11-13	8-10	5-7	3-5

The Simple Clinical Score was based on a large cohort from a small rural Irish hospital and designed specifically to include only variables immediately available or discernible at presentation (92).

Simple Clinical Score

Age	
≥50 (men) ≥55 (women)	2
>75 (men and women)	4
SBP	
81-100	2
71-80	3
<70	4
Pulse rate > SBP	2
Temperature <35 or ≥39	2
Respiratory rate	
21-30	1
>30	2
Oxygen saturation	
90-94	1
<90	2
Breathless on presentation	1
Abnormal ECG	2
Diabetes (type I or II)	1
Coma without intoxication or overdose	4
Altered mental status without coma, intoxication or overdose and aged >49y	2
New stroke	3
Unable to stand unaided, or nursing home resident	2
Prior to current illness, spent some part of daytime in bed	2

The South African triage score was developed to address the combined issues of ED overcrowding and a high population burden of chronic disease. It combines a physiological score (TEWS) with specific discriminator features (93).

South African Triage Score physiological score (TEWS)

	0	1	2	3
Mobility	Walking	With help	Stretcher /immobile	
RR	9-14	15-20	<9 21-29	>30
HR	51-100	41-50 101-110	<41 111-129	>129
SBP	101-199	81-100	71-80 >199	<71
Temperature	35-38.4		COLD or <35 HOT or >38.4	
AVPU	Alert	Voice	Pain Confused	Unresponsive
Trauma	No	Yes		

South African Triage Score discriminator features

	Red	Orange	Yellow	Green
TEWS	7 or more	5-6	3-4	<3
Target time to treat	Immediate	<10 min	<60 min	<4 hours
Mechanism of injury		High energy transfer		All other patients
Presentation		Acute shortness of breath		
		Coughing blood		
		Chest pain		
		Uncontrolled haemorrhage	Controlled haemorrhage	
	Current seizure	Postictal		
		Acute focal neurology		
		Reduced consciousness		
		Psychosis/aggression		
		Threatened limb		
		Dislocation – other	Dislocation – finger or toe	
		Fracture – compound	Fracture – closed	
	Burn – face / inhalational	Burn over 20%		
		Burn – electrical		
		Burn – circumferential		
		Burn – chemical		
		Poisoning/overdose		
	Glucose <3	Glucose >11 and ketonuria	Glucose >17 (no ketonuria)	
		Vomiting – fresh blood	Vomiting – persistent	
		Pregnancy – abdominal trauma/pain	Pregnancy – trauma	
			Pregnancy – PV bleed	
Pain		Severe	Moderate	Mild

The Cape Triage Score also uses TEWS, with slightly different discriminators (94).

Cape Triage Score discriminators

	Red	Orange	Yellow	Green
TEWS	8 or more	6-7	3-5	2 or less
Mechanism of injury	Entrapment	Impact – high	Impact – low	
Pain		Severe	Moderate	Mild
Respiratory	Asthma – status	Asthma		
Cardiac		Chest pain		
Vascular		Arterial haemorrhage		
Neurological	Current seizure	Postictal		
	Unresponsive	Responds to pain	Responds to voice	Alert
Psychiatric		Psychosis/aggression		
Orthopaedic		Threatened limb		
		Dislocation – major joint	Dislocation – minor joint	
		Fracture – open	Fracture – closed	
Burn	Face/inhalation	>20%	Minor	
Metabolic	Glucose <2.2	Overdose/poisoning		
Intestinal		Haematemesis	Abdominal pain	
Obstetric		Pregnancy – trauma	Pregnancy – PV bleed	
Anatomy	Airway trauma	Head/neck/torso/ evisceration trauma	Limb trauma	

The Sun score (95) predicts the likelihood of hospital admission at the point of ED triage, using an equation including age, racial group, acuity as allocated in an unstructured manner by the triage nurse (PAC), and past history or comorbidities. The natural log of p/1-p (where p is probability of admission) is calculated using a constant of -2.903 and scores as below. Of note it was developed in Singapore, so the risks related to racial group are probably not generalisable to UK practice.

Sun score

Age group	
<15	-1.698
15-24	-0.453
35-44	0.233
45-54	0.423
55-64	0.580
65-74	0.921
75-84	1.280
>84	1.662
Racial group	
Malay	-0.127
Indian	0.091
Other	0.028
Arrival by ambulance	0.537
PAC group	
1	3.007
2	1.488
Prior ED visit 3 months	0.220
Prior hospital admission 3 months	0.360
Chronic conditions	
Diabetes only	0.760
Hypertension only	0.383
Dyslipidaemia only	0.633
Diabetes and hypertension	0.979
Diabetes, hypertension and dyslipidaemia	0.965
Diabetes and dyslipidaemia	0.719
Dyslipidaemia and hypertension	0.642

The Vital Signs Score was developed in Bern to be applied at ED presentation and involves summation of seven parameters derived from Medical Emergency Team calling criteria (96).

Vital Signs Score

Airway - threatened	Need for intratracheal suctioning, oro or nasopharyngeal tubes, intubation, bronchoscopy.
Breathing – respiratory rate	<6/min or >36/min
Breathing – SaO_2	<90% despite supplemental oxygen
Circulation – SBP	<90mmHg
Circulation – heart rate	<40/min or >140/min
Neurology – GCS	<13
Neurology – seizures	Repeated or prolonged (>5 min)

The VitalPAC™ EWS (ViEWS) was developed in Portsmouth from a large database of computerised vital signs recordings on a medical assessment unit and has since formed the basis for the RCP-approved National Early Warning Score (97).

VitalPAC™ EWS (ViEWS)

	0	1	2	3
Pulse	51-90	41-50 91-110	<41 111-120	>120
Breathing rate	12-20	9-11	21-24	<9 >24
Temperature	36.1-38	35.1-36 38.1-39	>39	<35.1
SBP	111-249	101-110 >249	91-100	<91
SaO_2	>95	94-95	92-93	<92
Inspired O_2	Air			Any O_2
AVPU	Alert			Any other

Appendix 2 Existing clinical decision aids for specific patient populations

Abdominal aortic aneurysm

APACHE II				
Lazarides 1997 (151)	External validation	40 patients in theatre for repair of ruptured infrarenal AAA	Hospital mortality	Mean APACHE II for mortality 14.5 (SD 5.1) vs 11.3 (SD 3.6) p=0.02
Edinburgh aneurysm score				
Tambyraja 2007 (152)	Derivation	105 patients undergoing repair of ruptured AAA	Hospital mortality	Score 0-1 29%, score 2 50%, score 3 80%
Tambyraja 2008 (153)	External validation	84 patients undergoing repair of ruptured AAA	Hospital or 30-day mortality	AUROC 0.72 (0.61-0.83)
Glasgow aneurysm score				
Tambyraja 2005 (154)	External validation	82 patients undergoing repair of ruptured AAA	Hospital mortality	AUROC 0.606 (0.483-0.729)
Leo 2006 (155)	External validation	114 patients undergoing repair of ruptured AAA	"Immediate" post-operative mortality	AUROC 0.906 (0.85-0.962)
Tambyraja 2008 (153)	External validation	84 patients undergoing repair of ruptured AAA	Hospital or 30-day mortality	AUROC 0.64 (0.52-0.76)
Hardman score				
Prance 1999 (156)	External validation	69 patients undergoing repair of ruptured AAA	Hospital mortality	Score 0 18%, score 1 28%, score 2 48%, score 3 100%
Neary 2003 (157)	External validation	188 patients undergoing repair of ruptured AAA	Hospital mortality	Score 0 35%, score 1 55%, score 2 74%, score 3 90%
Calderwood 2004 (158)	External validation	137 patients undergoing repair of ruptured AAA	Hospital mortality	Score 0 40%, score 1 46%, score 2 77%, score 3 92%, score 4 100%
Tambyraja 2005 (154)	External validation	82 patients undergoing repair of ruptured AAA	Hospital mortality	Score 0 15%, score 1 55%, score 2 38%, score >2 33% p=0.211
Leo 2006 (155)	External validation	114 patients undergoing repair of ruptured AAA	"Immediate" post-operative mortality	AUROC 0.834 (0.759-0.91)

Sharif 2007 (159)	External validation	52 patients undergoing EVAR for ruptured AAA	Hospital mortality	Score 0 15%, score 1 27%, score 2 50%, score 3/4 66%
		74 patients undergoing open repair of ruptured AAA		Score 0 41%, score 1 46%, score 2 61%, score 3/4 70%
Karkos 2008 (160)	External validation	41 patients undergoing EVAR for ruptured AAA	30-day mortality	X2 for trend p=0.02
Tambyraja 2008 (153)	External validation	84 patients undergoing repair of ruptured AAA	Hospital or 30-day mortality	AUROC 0.69 (0.57-0.8)
Hardman score (modified)				
Calderwood 2004 (158)	Derivation	137 patients undergoing repair of ruptured AAA	Hospital mortality	Score 0 22%, score 1 47%, score 2 67%, score 3 84%, score 4 100%
POSSUM				
Lazarides 1997 (151)	External validation	40 patients in theatre for repair of ruptured infrarenal AAA	Hospital mortality	Mean POSSUM for mortality 66.5 (SD 7.2) vs 63.3 (SD 7.6) p=0.18
RAAA-POSSUM				
Neary 2003 (157)	External validation	188 patients undergoing repair of ruptured AAA	Hospital mortality	X^2 for trend p<0.001

AAA: Abdominal aortic aneurysm; EVAR: endovascular aneurysm repair

Acute coronary syndrome (suspected or confirmed)

Acute physiology score (APS)				
Moreau 1989 (161)	External validation	76 patients admitted to hospital with AMI	Hospital mortality	AUROC 0.749 +/- 0.075
APACHE II				
Moreau 1989 (161)	External validation	76 patients admitted to hospital with AMI	Hospital mortality	AUROC 0.823 +/- 0.067
Alemi 1990 (162)	External validation	775 patients hospitalised with AMI	Hospital mortality	AUROC 0.7
Australia/New Zealand				
Macdonald 2011 (163)	External validation	1714 ED patients with potential ACS having troponin sampling	30-day MACE	AUROC 0.82 (0.8-0.84)
Bazzino				
Bazzino 1999 (164)	Derivation	1038 patients hospitalised with unstable angina	Hospital mortality or AMI (pre-troponin definition)	AUROC 0.59 +/- 0.03
Chang score				
Chang 2006 (165)	Derivation Secondary analysis of RCT	6066 patients hospitalised with STEMI	30-day mortality	AUROC 0.8 (0.77-0.82)
Cheshire Merseyside and North Wales score				
Rawlings 2012 (166)	External validation	104 patients hospitalised with NSTEACS	30-day mortality	AUROC 0.845
Coronary prognostic index				
Moreau 1989 (161)	External validation	76 patients admitted to hospital with AMI	Hospital mortality	AUROC 0.816 +/- 0.068
EMMACE				
Gale 2009 (167)	External validation National registry	100686 patients hospitalised with ACS (including STEMI)	30-day mortality	AUROC 0.78 (0.77-0.78)
Freedom-from-event score				
Brieger 2009 (168)	Derivation International registry	16127 patients admitted for 24h with ECG change, biomarker change or previous coronary artery disease	Hospital mortality or heart failure, shock, AF, VF, cardiac arrest, VT, MI, stroke, major bleed	AUROC 0.77

	Validation International registry	6820 patients admitted for 24h with ECG change, biomarker change or previous coronary artery disease		AUROC 0.77
Soderholm 2012 (169)	External validation	559 ED patients admitted with suspicion of ACS	Inpatient complications	AUROC 0.69 (0.6-0.79)
Get with the guidelines score				
Chin 2011 (170)	Derivation National registry	65668 patients admitted with STEMI or STEMI	Hospital mortality	AUROC 0.85 (0.8-0.9)
	Internal validation National registry	16336 patients admitted with STEMI or STEMI		AUROC 0.84
Goldman index				
Goldman 1996 (171)	Derivation	10682 patients presenting to ED with chest pain unexplained by trauma or CXR	VF, cardiac arrest, complete heart block, PPM, emergency cardioversion, shock, IABP, ETT, CABG/PTCA predischarge	AUROC 0.82
	Temporal validation	4676 patients presenting to ED with chest pain unexplained by trauma or CXR		AUROC 0.80
Durairaj 2001 (172)	External validation	1061 patients with ACS admitted to inpatient telemetry	Major complication (VF, cardiac arrest, new complete AV block, temporary PPM, emergency cardioversion, shock, IABP, ETT, CABG or PCI)	NPV for very low risk with chest pain 1 (0.988-1) and without chest pain 0.995 (0.974-0.999)
Limkakeng 2001 (173)	External validation	998 ED patients with chest pain prompting ECG	30-day MACE	Goldman >4% sens 0.54 (0.47-0.6) spec 0.74 (0.72-0.76) PPV 0.16 (0.13-0.19) NPV 0.95 (0.94-0.96)
Hollander 2004 (174)	External validation	1029 patients with ACS admitted to telemetry beds	Mortality preventable by monitoring or VF/VT requiring treatment	Goldman risk <8%, no outcomes of interest (CI 0-3%)

Manini 2009 (175)	External validation Secondary analysis of diagnostic study	148 ED patients with chest pain and non-diagnostic ECG	Hospital diagnosis of ACS	High or intermediate risk sensitivity 0.53 (0.29-0.77) specificity 0.72 (0.64-0.79) PPV 0.2 (0.08-0.31) NPV 0.92 (0.87-0.97)
Soderholm 2012 (169)	External validation	559 ED patients admitted with suspicion of ACS	Inpatient complications	AUROC 0.6 (0.49-0.72)
GRACE				
de Araujo Goncales 2002 (176)	External validation	460 hospitalised patients with NSTEACS	30-day mortality or AMI	AUROC 0.672 (0.627-0.714)
Yan 2004 (177)	External validation National registry	2925 patients hospitalised with NSTEACS	Hospital mortality	AUROC 0.83 (0.77-0.89)
Rahimi 2006 (178)	External validation	558 patients hospitalised with NSTEMI	Hospital mortality	AUROC 0.578 (0.457-0.699)
			Malignant arrhythmia	AUROC 0.573 (0.444–0.701)
Lyon 2007 (179)	External validation	1000 ED patients with NSTEACS	30-day MACE	AUROC 0.8 (0.75-0.85)
Sinclair 2007 (180)	External validation	149 patients hospitalised with chest pain and biomarker or ECG change or history of coronary artery disease	Hospital complications	Recurrent ACS p<0.05, arrhythmia p<0.02, CCF p<0.02, PCI p<0.05
Yan 2007 (181)	External validation National registry	1728 patients hospitalised with NSTEACS	Hospital mortality	AUROC 0.81 (0.73-0.89)
Lev 2008 (182)	External validation	855 patients with STEMI undergoing PCI	30-day MACE	AUROC 0.544
			30-day mortality	AUROC 0.471
Elbarouni 2009 (183)	Temporal validation International registry	11118 patients hospitalised with NSTEACS	Hospital mortality	AUROC 0.842 (0.823-0.861)
Gale 2009 (167)	External validation National registry	85771 patients hospitalised with ACS (including STEMI)	Hospital mortality	AUROC 0.8 (0.8-0.81)
Goodacre 2011 (184)	External validation	1772 ED patients with chest pain in previous 12 hours without ECG	30-day MACE	AUROC 0.722

		change or high-risk coronary artery disease		
Goodacre 2012 (185)	External validation	2243 ED patients with chest pain in previous 12 hours without ECG change or high-risk coronary artery disease	30-day MACE	AUROC 0.717 (0.698-0.735)
Soderholm 2012 (169)	External validation	559 ED patients admitted with suspicion of ACS	Inpatient complications	AUROC 0.76 (0.65-0.85)
Hasdai				
Hasdai 2000 (68)	Derivation Secondary analysis of RCT	38942 patients thrombolysed for STEMI	Development of cardiogenic shock	AUROC 0.761
	Validation Secondary analysis of RCT	14960 patients thrombolysed for STEMI		AUROC 0.796
IDHI				
Alemi 1990 (162)	External validation	775 patients hospitalised with AMI	Hospital mortality	AUROC 0.69
Mayo score				
Williams 2006 (67)	Derivation Hospital registry	809 patients hospitalised with AMI (including STEMI)	30-day mortality	AUROC 0.81
	Internal validation Hospital registry	403 patients hospitalised with AMI (including STEMI)		AUROC 0.79
MINAP				
Gale 2008 (186)	Derivation National registry	34722 patients hospitalised with STEMI	Hospital mortality	AUROC 0.8 (0.79-0.8)
Normand				
Normand 1996 (187)	Derivation	10936 patients hospitalised with ACS (including STEMI)	30-day mortality	AUROC 0.79
	Internal validation	3645 patients hospitalised with ACS (including STEMI)		AUROC 0.78
North American Chest Pain Rule				
Hess 2012 (188)	Derivation	2718 ED patients with chest pain and possible ACS	30-day MACE	Sensitivity 100 (97.2-100), specificity 20.9 (16.9-24.9) if

				using a cutoff of 50y
PAMI				
Lev 2008 (182)	External validation	855 patients with STEMI undergoing PCI	30-day MACE	AUROC 0.65
			30-day mortality	AUROC 0.742
PIMI				
Alemi 1990 (162)	External validation	775 patients hospitalised with AMI	Hospital mortality	AUROC 0.44
PREDICT				
Rahimi 2006 (178)	External validation	558 patients hospitalised with NSTEMI	Hospital mortality	AUROC 0.829 (0.744-0.914)
			Malignant arrhythmia	AUROC 0.531 (0.366–0.697)
PURSUIT				
Boersma 2000 (189)	Derivation Secondary analysis of RCT	9461 hospitalised patients with NSTEACS	30-day mortality	AUROC 0.814
			30-day mortality or reinfarct	AUROC 0.669
de Araujo Goncales 2002 (176)	External validation	460 hospitalised patients with NSTEACS	30-day mortality or AMI	AUROC 0.615 (0.569-0.660)
Brilakis 2003 (190)	External validation	337 patients admitted to coronary care with NSTEMI	30-day mortality	AUROC 0.78
Yan 2004 (177)	External validation National registry	2925 patients hospitalised with NSTEACS	Hospital mortality	AUROC 0.84 (0.79-0.89)
Rahimi 2006 (178)	External validation	558 patients hospitalised with NSTEMI	Hospital mortality	AUROC 0.86 (0.778-0.942)
			Malignant arrhythmia	AUROC 0.523 (0.358–0.688)
Yan 2007 (181)	External validation National registry	1728 patients hospitalised with NSTEACS	Hospital mortality	AUROC 0.8 (0.71-0.88)
Gale 2009 (167)	External validation National registry	49995 hospitalised patients with NSTEACS	30-day mortality	AUROC 0.79 (0.78-0.8)
Sanchis				
Manini 2009 (175)	External validation	148 ED patients with chest pain	Hospital diagnosis of ACS	Sanchis >1

	Secondary analysis of diagnostic study	and non-diagnostic ECG		sensitivity 0.41 (0.18-0.65) specificity 0.86 (0.8-0.92) PPV 0.28 (0.1-0.46) NPV 0.92 (0.87-0.97)
SAPS				
Moreau 1989 (161)	External validation	76 patients admitted to hospital with AMI	Hospital mortality	AUROC 0.869 +/- 0.059
Selker				
Selker 1991 (191)	Derivation	4099 patients aged >29 (men) or >39 (women) hospitalised with potential ACS	Hospital mortality	AUROC 0.82
	Split sample validation	1387 patients aged >29 (men) or >39 (women) hospitalised with potential ACS		AUROC 0.85
Simple risk index				
Rathore 2003 (192)	External validation	49711 patients age >65 with STEMI	30-day mortality	AUROC 0.62
Das 2006 (193)	External validation Registry data	2153 patients hospitalised with AMI (NSTEMI or STEMI)	30-day mortality	AUROC 0.77 (0.74-0.79) AUROC in STEMI subset 0.76 (0.72-0.8)
Gale 2009 (167)	External validation National registry	100686 patients hospitalised with ACS (including STEMI)	30-day mortality	AUROC 0.79 (0.78-0.8)
TIMI risk index				
Wiviott 2004 (194)	External validation National registry	153486 patients hospitalised with STEMI	Hospital mortality	AUROC 0.79
Ilkhanoff 2005 (195)	External validation	719 patients hospitalised with suspected ACS	30-day mortality	AUROC 0.79
Bradshaw 2007 (196)	External validation Secondary analysis of RCT	10487 hospitalised patients with AMI (NSTEMI or STEMI) excluding patients with overt shock	30-day mortality	AUROC 0.81 (0.79-0.82)
Garcia-Almagro 2008	External validation	661 ED patients with potential	1-month MACE	HR per unit increase 1.3 (1.0-

(197)		ACS		1.6) p=0.007
TIMI score NSTEMI				
Antman 2000 (66)	Derivation Secondary analysis of RCT	1953 patients with ACS including ST changes or raised biomarkers	14-day mortality	AUROC 0.74
			14-day mortality or AMI	AUROC 0.63
			AMI in 14 days	AUROC 0.66
			Urgent revascularisation in 14 days	AUROC 0.68
Samaha 2002 (198)	External validation Secondary analysis of RCT	919 patients with NSTEMI	30-day MACE	AUROC 0.59 (0.63 for RIP, 0.61 for AMI)
de Araujo Goncales 2002 (176)	External validation	460 hospitalised patients with NSTEACS	30-day mortality or AMI	AUROC 0.551 (0.504-0.597)
Foussas 2005 (199)	External validation	985 hospitalised patients with NSTEACS	14-day mortality, AMI or recurrent ischaemia	Score 1-2 9%, score 3 16.5%, score 4 21.5%, score 5 26.7%, score 6-7 36.6% p for trend <0.001
Conway Morris 2006 (200)	External validation	954 ED patients with ACS	30-day MACE	AUROC 0.79 (0.75-0.84)
Pollack 2006 (201)	External validation	3929 ED patients with chest pain prompting ECG	30-day MACE	Chi-square p<0.001
Rahimi 2006 (178)	External validation	558 patients hospitalised with NSTEMI	Hospital mortality	AUROC 0.638 (0.515-0.760)
			Malignant arrhythmia	AUROC 0.486 (0.328–0.645)
Soiza 2006 (202)	External validation	834 patients hospitalised with ACS (including STEMI) or potential ACS	Hospital mortality, AMI or revascularisation	Chi-square for trend p<0.001
Chase 2007 (203)	External validation	238 patients with chest pain within 1 week of cocaine use	30-day mortality, AMI or revascularisation	TIMI 0, 3.7% (1-8.3), TIMI I 13.2% (5.7-20.7), TIMI 2 17.1% (4.3-29.8), TIMI 3 21.4% (4.4-38.4), TIMI 4 20.0% (0.1-43.6), TIMI 5/6

				50.0% (0.1-100)
Jaffery 2007 (204)	External validation	947 ED patients with potentially cardiac chest pain	30-day MACE	TIMI >2 sens 0.537 (0.449-0.623) spec 0.752 (0.721-0.782) PPV 0.266 (0.215-0.323) NPV 0.906 (0.882-0.927)
Karounos 2007 (205)	External validation	2022 ED patients age >30 with chest pain prompting ECG	30-day MACE	p<0.001 for trend
Lyon 2007 (179)	External validation	1000 ED patients with NSTEACS	30-day MACE	AUROC 0.79 (0.74-0.85)
Yan 2007 (181)	External validation National registry	1728 patients hospitalised with NSTEACS	Hospital mortality	AUROC 0.68 (0.59-0.77)
Body 2009 (206)	External validation	796 ED patients with NSTEACS	30-day MACE	AUROC 0.77 (0.73-0.8)
Campbell 2009 (207)	External validation	3169 ED patients with chest pain prompting ECG	30-day MACE	TIMI <2 RR 0.21 (0.18-0.31)
Manini 2009 (175)	External validation Secondary analysis of diagnostic study	148 ED patients with chest pain and non-diagnostic ECG	Hospital diagnosis of ACS	TIMI >2 sensitivity 0.35 (0.13-0.58) specificity 0.85 (0.79-0.91) PPV 0.23 (0.07-0.39) NPV 0.91 (0.86-0.96)
Goodacre 2011 (184)	External validation	1772 ED patients with chest pain in previous 12 hours without ECG change or high-risk coronary artery disease	30-day MACE	AUROC 0.67
Macdonald 2011 (163)	External validation	1714 ED patients with potential ACS having troponin sampling	30-day MACE	AUROC 0.76 (0.73-0.79)
Goodacre 2012 (185)	External validation	2243 ED patients with chest pain in previous 12 hours without ECG change or high-risk coronary artery disease	30-day MACE	AUROC 0.682 (0.662-0.701)
Rawlings 2012 (166)	External validation	104 patients hospitalised with	30-day mortality	AUROC 0.67

		NSTEACS		
TIMI score STEMI				
Morrow 2000 (208)	Derivation Secondary analysis of RCT	14114 patients with STEMI	30-day mortality	AUROC 0.779
	Validation Secondary analysis of RCT	3687 patients with STEMI		AUROC 0.746
Foussas 2005 (199)	External validation	861 patients with STEMI	30-day mortality	Score 0-1 3.3%, score 2-3 7.4%, score 4-5 12%, score >5 25.6% p for trend <0.001
Lev 2008 (182)	External validation	855 patients with STEMI undergoing PCI	30-day MACE	AUROC 0.635
			30-day mortality	AUROC 0.724
Troponin prediction score				
Januzzi 2006 (209)	Derivation Secondary analysis of RCT	769 patients with NSTEACS and negative troponin on presentation	Troponin rise at 12h	AUROC 0.76
	External validation Secondary analysis of cohort	493 patients with NSTEACS and negative troponin on presentation		AUROC 0.73

AF: atrial fibrillation; AMI: acute myocardial infarction; CABG: coronary artery bypass grafting; ETT: endotracheal tube; IABP: intra-aortic balloon pump; MACE: major adverse cardiac event; NSTEACS: non-ST-elevation acute coronary syndrome; NSTEMI: non-ST-elevation acute myocardial infarction; PCI: percutanous coronary intervention; PPM: pacemaker; PTCA: percutaneous transluminal coronary angioplasty; STEMI: ST-elevation myocardial infarction; VF: ventricular fibrillation; VT: ventricular tachycardia

Asthma and COPD (chronic obstructive pulmonary disease)

BAP-65				
Tabak 2009 (210)	Derivation	43893 patients hospitalised with acute exacerbation of COPD	Hospital mortality	AUROC 0.72 (0.7-0.74)
	Secondary endpoint in derivation cohort		Need for mechanical ventilation	AUROC 0.77 (0.75-0.79)
	Internal validation	44181 patients hospitalised with acute exacerbation of COPD	Hospital mortality	AUROC 0.71 (0.7-0.73)
			Need for mechanical ventilation	AUROC 0.77 (0.75-0.79)
Shorr 2011 (211)	External validation	34669 patients admitted with acute exacerbation of COPD	Inpatient mortality	AUROC 0.77 (0.76-0.78)
			Need for invasive ventilation	AUROC 0.78 (0.78-0.79)
CHOP				
Tsai 2010 (212)	Derivation Observational study	1824 ED patients age 18-54 with acute asthma	Hospitalisation	AUROC 0.72
	External validation Observational study	1335 ED patients age 18-54 with acute asthma		AUROC 0.65
CURB-65				
Chang 2011 (213)	External validation	249 patients admitted with exacerbation of COPD	30-day mortality	AUROC 0.7334
Steer 2012 (214)	External validation	920 patients admitted with exacerbation of COPD	Hospital mortality	AUROC 0.717 (0.66-0.77)
Extended MRC dyspnoea scale				
Steer 2012 (214)	External validation	920 patients admitted with exacerbation of COPD	Hospital mortality	AUROC 0.794 (0.75-0.84)
National Asthma Guidelines				
Kelly 2004 (215)	External validation	831 ED patients with acute asthma	Hospitalisation	At presentation 13% mild, 57% moderate, 89% severe, At 1 hour 18% mild, 84% moderate, 86% severe
Rodrigo Index				
Rodrigo 1997 (216)	Derivation	184 ED patients age 18-50 with acute asthma	Hospitalisation	AUROC 0.91

	Temporal validation	91 ED patients age 18-50 with acute asthma		AUROC 0.9

GI bleed

Blatchford				
Gralnek 2004 (217)	External validation	175 inpatients undergoing OGD	Rebleed requiring OGD, surgery or readmission	Blatchford 0 - no recurrence or mortality
Chen 2007 (218)	External validation	354 patients undergoing OGD for non-variceal UGIB	Mortality	Sensitivity 1 (0.44-1) specificity 0.08 (0.06-0.13) PPV 0.01 (0.003-0.03) NPV 1 (0.88-1)
			Rebleed requiring OGD, surgery or readmission	Sensitivity 1 (0.86-1) specificity 0.09 (0.06-0.12) PPV 0.07 (0.05-0.1) NPV 1 (0.88-1)
Masaoka 2007 (219)	External validation	93 patients undergoing emergency OGD	Transfusion, operative or endoscopic intervention	AUROC 0.628
Stanley 2009 (220)	External validation	647 patients admitted for UGIB	Mortality or need for intervention	AUROC 0.92 (0.9-0.94)
Chandra 2012 (221)	External validation	171 ED patients with UGIB	30-day mortality or intervention	AUROC 0.79
Farooq 2012 (222)	External validation	195 ED patients with UGIB	Requirement for endoscopic therapy	Cutoff >0 sensitivity 100, specificity 4 Cutoff >5 sensitivity 87, specificity 33
Blatchford (modified)				
Romagnuolo 2007 (223)	External validation Registry data	1869 patients undergoing OGD for non-variceal UGIB	High-risk stigmata at OGD	mBRS <2 OR 0.4 (0.3-0.6)
			Rebleed or death before OGD	mBRS <2 5% vs 19% p<0.001
BLEED				
Kolleff 1997 (224)	External validation	465 patients admitted via ED with GI bleed	Mortality, recurrent bleed or need for surgery	AUROC 0.72
Bordley				
Bordley 1985 (225)	Derivation	110 inpatients with UGIB	Mortality, urgent	Low risk poor outcome 0/52,

			surgery, rebleed or complication	high risk 26/58
	Retrospective validation	52 inpatients with UGIB		Low risk poor outcome 0/22, high risk 8/30
	Prospective validation	52 inpatients with UGIB		Low risk poor outcome 2/52, high risk 29/59
Rockall				
Church 2006 (226)	External validation Secondary analysis of trial	247 undergoing OGD for peptic ulcer with age>60, shock, comorbidity or haemoglobin<10	30-day rebleed	AUROC 0.634
			30-day mortality	AUROC 0.843
Sarwar 2007 (227)	External validation	402 patients admitted for UGIB	Mortality	AUROC 0.834
			Rebleeding	AUROC 0.798
Rockall (clinical element)				
Sanders 2002 (228)	External validation	162 inpatients undergoing OGD for bleeding peptic ulcer	Mortality	0/51 score <3 vs 5/30 score 3 vs 7/41 score 4 vs 5/26 score 5 vs 5/14 score 6 p=0.01
		196 inpatients undergoing OGD for bleeding varices		7/131 score <4 vs 7/39 score 4 vs 9/26 score 5 p<0.0005
Gralnek 2004 (217)	External validation	175 inpatients undergoing OGD	Rebleed requiring OGD, surgery or readmission	Clinical Rockall 0 - no recurrence or mortality
Chen 2007 (218)	External validation	354 patients undergoing OGD for non-variceal UGIB	Mortality	Sensitivity 1 (0.44-1) specificity 0.19 (0.15-0.23) PPV 0.01 (0.004-0.03) NPV 1 (0.94-1)
			Rebleed requiring OGD, surgery or readmission	Sensitivity 0.69 (0.49-0.84) specificity 0.18 (0.14-0.22) PPV 0.06 (0.03-0.09) NPV 0.89 (0.79-0.94)
Das 2008 (229)	External validation	200 patients admitted for UGIB	Need for OGD to control bleeding	AUROC 0.65 (0.56-0.74)
		194 patients admitted for UGIB		AUROC 0.53 (0.43-0.62)
Stanley 2009 (220)	External validation	647 patients admitted for UGIB	Mortality or need for intervention	AUROC 0.72 (0.68-0.76)
Chandra 2012 (221)	External validation	171 ED patients with UGIB	30-day mortality or	AUROC 0.62

			intervention	
Farooq 2012 (222)	External validation	195 ED patients with UGIB	Requirement for endoscopic therapy	Cutoff >0 sensitivity 95, specificity 9 Cutoff >2 sensitivity 84, specificity 29
Strate				
Strate 2005 (230)	Derivation	254 patients admitted for LGIB	>2 unit blood transfusion, fall >20% haematocrit in 24h, recurrence after 24h	AUROC 0.761
	Temporal validation	275 patients admitted for LGIB		AUROC 0.754

LGIB: lower gastrointestinal bleed; OGD: oesophagogastroduodenoscopy; UGIB: upper gastrointestinal bleed

Heart failure

3CPO				
Gray 2010 (121)	Derivation Secondary analysis of RCT	1069 patients admitted with acute cardiogenic pulmonary oedema	1-week mortality	AUROC 0.794 (0.745-0.843)
3CPO (simplified)				
Gray 2010 (121)	Derivation Secondary analysis of RCT	1069 patients admitted with acute cardiogenic pulmonary oedema	1-week mortality	AUROC 0.754 (0.701-0.807)
ADHERE decision tree				
Auble 2007 (231)	External validation State database	33533 patients admitted for heart failure	Inpatient mortality	AUROC 0.68 (0.67-0.7)
			Inpatient mortality or lifethreatening event	AUROC 0.58 (0.57-0.59)
ADHERE logistic regression				
Auble 2007 (231)	External validation State database	33533 patients admitted for heart failure	Inpatient mortality	AUROC 0.73 (0.72-0.75)
			Inpatient mortality or lifethreatening event	AUROC 0.61 (0.6-0.62)
Brigham				
Auble 2007 (231)	External validation State database	33533 patients admitted for heart failure	Inpatient mortality	AUROC 0.61 (0.59-0.62)
			Inpatient mortality or lifethreatening event	AUROC 0.61 (0.6-0.62)
EFFECT				
Lee 2003 (232)	Derivation Secondary analysis of trial	2624 patients admitted for acute heart failure	30-day mortality	AUROC 0.8
	Internal validation Secondary analysis of trial	1407 patients admitted for acute heart failure		AUROC 0.79
Auble 2007 (231)	External validation State database	33533 patients admitted for heart failure	Inpatient mortality	AUROC 0.74 (0.72-0.75)
			Inpatient mortality or lifethreatening event	AUROC 0.62 (0.61-0.63)
Get with the guidelines score				
Peterson 2010 (233)	Combined derivation and validation	39783 patients admitted for LVF	Hospital mortality	AUROC 0.75

	Registry data			
Le Conte				
Le Conte 1999 (234)	Derivation	186 patients admitted with acute pulmonary oedema	Hospital mortality	Correct classification 89.8% NPV 98% PPV 41%
Pulmonary edema prognostic score				
Fiutowski 2008 (235)	Derivation	276 patients admitted with pulmonary oedema	Hospital mortality	AUROC 0.78

Influenza

CAP				
Muller 2010 (236)	External validation	607 patients admitted with influenza A or B	Hospital mortality	AUROC 0.65 (0.58–0.71)
			ICU admission	AUROC 0.65 (0.59–0.71)
CURB-65				
Muller 2010 (236)	External validation	607 patients admitted with influenza A or B	Hospital mortality	AUROC 0.66 (0.60–0.72)
			ICU admission	AUROC 0.58 (0.52–0.64)
Mulrennan 2010 (237)	External validation	35 ED patients with H1N1	ICU admission	82% ICU adm (9/11) had score 0 or 1
MEDS				
Muller 2010 (236)	External validation	607 patients admitted with influenza A or B	Hospital mortality	AUROC 0.77 (0.71–0.83)
			ICU admission	AUROC 0.67 (0.61–0.73)
NHAP				
Muller 2010 (236)	External validation	607 patients admitted with influenza A or B	Hospital mortality	AUROC 0.68 (0.62–0.74)
			ICU admission	AUROC 0.62 (0.57–0.68)
PMEWS				
Muller 2010 (236)	External validation	607 patients admitted with influenza A or B	Hospital mortality	AUROC 0.68 (0.61–0.74)
			ICU admission	AUROC 0.63 (0.57–0.69)
Pneumonia Severity Index				
Muller 2010 (236)	External validation	607 patients admitted with influenza A or B	Hospital mortality	AUROC 0.78 (0.72-0.83)
			ICU admission	AUROC 0.67 (0.61–0.73)
SMART-COP				
Muller 2010 (236)	External validation	607 patients admitted with influenza A or B	Hospital mortality	AUROC 0.69 (0.62–0.75)
			ICU admission	AUROC 0.73 (0.67–0.79)
SOFA				
Adenji 2011 (238)	External validation	62 inpatients with H1N1	ICU admission	AUROC 0.77 (.65-.89)
			IPPV	AUROC 0.87 (.72-1)
STSS				
Muller 2010 (236)	External validation	607 patients admitted with influenza A or B	Hospital mortality	AUROC 0.71 (0.66–0.77)
			ICU admission	AUROC 0.63 (0.57–0.69)
Adenji 2011 (238)	External validation	62 inpatients with H1N1	ICU admission	AUROC 0.88 (.78-.98)

			IPPV	AUROC 0.91 (.83-.99)

Pancreatitis

APACHE II				
Meek 2000 (239)	External validation	92 patients admitted with gallstone pancreatitis	>24h ICU care	APACHE II>4 sensitivity 0.67 specificity 0.74 PPV 0.29 NPV 0.94
Gan 2003 (240)	External validation Retrospective review	53 HIV +ve patients admitted with pancreatitis	Mortality, ICU, local complications or surgery	Sensitivity 1, specificity 0.7, PPV 0.41, NPV 1
Halonen 2003 (241)	External validation	60 patients admitted with "severe" pancreatitis	Inpatient mortality	AUROC 0.817 (0.702-0.931)
Gurleyik 2005 (242)	External validation	55 patients admitted with pancreatitis	Atlanta severity	APACHE II >=7 Sensitivity 0.615 Specificity 0.857 PPV 0.571 NPV 0.878
Taylor 2005 (243)	External validation Retrospective chart review	49 patients admitted with pancreatitis	Over 5/7 ICU	AUROC 0.55
Papachristou 2006 (244)	External validation	102 patients admitted with pancreatitis	ICU, necrosis or mortality	AUROC 0.893
Yeung 2006 (245)	External validation	101 patients admitted with pancreatitis	Atlanta severity	AUROC 0.904
Ueda 2007 (246)	External validation	137 patients admitted with pancreatitis	Bacteraemia or infected pancreatic necrosis	AUROC 0.73
			Renal/hepatic/lung dysfunction	AUROC 0.88
			Mortality	AUROC 0.81
Garcea 2008 (247)	External validation	181 patients admitted with pancreatitis	Mortality	AUROC 0.875
			Atlanta severity	AUROC 0.861
APACHE-O				
Papachristou 2006 (244)	External validation	102 patients admitted with pancreatitis	ICU, necrosis or mortality	AUROC 0.895
Yeung 2006 (245)	External validation	101 patients admitted with pancreatitis	Atlanta severity	AUROC 0.904
BISAP				

Singh 2009 (248)	External validation	397 patients admitted with pancreatitis	Hospital mortality	AUROC 0.82
Early Warning Score				
Garcea 2008 (247)	External validation	181 patients admitted with pancreatitis	Mortality	AUROC 0.827
			Atlanta severity	AUROC 0.853
Glasgow				
Gan 2003 (240)	External validation Retrospective review	39 HIV +ve patients admitted with pancreatitis	Mortality, ICU, local complications or surgery	Sensitivity 0.67, specificity 0.7, PPV 0.29, NPV 0.92
Taylor 2005 (243)	External validation Retrospective chart review	49 patients admitted with pancreatitis	Over 5/7 ICU	AUROC 0.67
Ueda 2007 (246)	External validation	137 patients admitted with pancreatitis	Bacteraemia or infected pancreatic necrosis	AUROC 0.73
			Renal/hepatic/lung dysfunction	AUROC 0.74
			Mortality	AUROC 0.73
Glasgow (modified)				
De Beaux 1995 (249)	External validation	279 patients admitted to tertiary centre with pancreatitis	Mortality	Score 0: 0/74, score 1 3/77, score 5 4/7, score 6 2/4.
Meek 2000 (239)	External validation	92 patients admitted with gallstone pancreatitis	>24h ICU care	Score >2 sensitivity 0.67 specificity 0.85 PPV 0.4 NPV 0.93
Imrie score				
Meek 2000 (250)	External validation	66 patients admitted with pancreatitis	Severe complications	Imrie >2 sensitivity 0.86 specificity 0.87 PPV 0.5 NPV 0.98
Halonen 2003 (241)	External validation	60 patients admitted with “severe” pancreatitis	Inpatient mortality	AUROC 0.536 (0.364-0.708)
Garcea 2008 (247)	External validation	181 patients admitted with pancreatitis	Mortality	AUROC 0.794
			Atlanta severity	AUROC 0.747
MODS				
Garcea 2008 (247)	External validation	181 patients admitted with pancreatitis	Mortality	AUROC 0.783
			Atlanta severity	AUROC 0.793

Ranson				
Meek 2000 (250)	External validation	66 patients admitted with pancreatitis	Severe complications	Ranson >2 sensitivity 0.75 specificity 0.67 PPV 0.29 NPV 0.94
Gan 2003 (240)	External validation Retrospective review	31 HIV +ve patients admitted with pancreatitis	Mortality, ICU, local complications or surgery	Sensitivity 1, specificity 0.33, PPV 0.3, NPV 1
Halonen 2003 (241)	External validation	60 patients admitted with “severe” pancreatitis	Inpatient mortality	AUROC 0.655 (0.503-0.808)
Taylor 2005 (243)	External validation Retrospective chart review	49 patients admitted with pancreatitis	Over 5/7 ICU	AUROC 0.54
Ueda 2007 (246)	External validation	137 patients admitted with pancreatitis	Bacteraemia or infected pancreatic necrosis	AUROC 0.82
			Renal/hepatic/lung dysfunction	AUROC 0.84
			Mortality	AUROC 0.83
Ranson (biliary)				
Meek 2000 (239)	External validation	92 patients admitted with gallstone pancreatitis	>24h ICU care	Score >2 sensitivity 0.69 specificity 0.67 PPV 0.31 NPV 0.91
SAPS				
van den Biezenbos 1998 (251)	External validation Secondary analysis of observational study	78 patients undergoing CT for pancreatitis	Mortality	AUROC 0.747 (SE 0.085)

Pneumonia

A-DROP				
Shindo 2008 (252)	External validation	329 patients hospitalised with pneumonia	30-day mortality	AUROC 0.846 (0.79-0.903)
APACHE II				
Jeong 2011 (253)	External validation	502 ED patients with pneumonia	30-day mortality	AUROC 0.847 (0.804-890)
ATS				
Angus 2002 (254)	External validation	1339 patients admitted with pneumonia	30-day mortality	AUROC 0.6 (0.54-0.65)
			ICU admission	AUROC 0.61 (0.57-0.65)
Buising 2006 (255)	External validation	392 patients admitted via ED with pneumonia	Inpatient mortality	AUROC 0.63
			ICU admission	AUROC 0.9
Kontou 2009 (256)	External validation	158 patients hospitalised with pneumonia	ICU admission	Sensitivity 0.9 (0.75-0.97), specificity 0.8 (0.73-0.86), PPV 0.53 (0.4-0.66), NPV 0.97 (0.92-0.99)
			Mortality	Sensitivity 0.65 (0.43-0.82), specificity 0.71 (0.63-0.78), PPV 0.25 (0.15-0.38), NPV 0.93 (0.87-0.97)
ATS (modified)				
Valencia 2007 (257)	External validation	457 patients admitted with PSI class V pneumonia	Inpatient mortality	Sensitivity 0.75 specificity 0.8 PPV 0.53 NPV 0.91
			ICU admission	Sensitivity 0.72 specificity 0.77 PPV 0.44 NPV 0.91
Feldman 2009 (258)	External validation	766 patients admitted with pneumonia	14-day mortality	AUROC 0.7491
ATS 2007				
Feldman 2009 (258)	External validation	766 patients admitted with pneumonia	14-day mortality	AUROC 0.7099
Kontou 2009 (256)	External validation	158 patients hospitalised with pneumonia	ICU admission	Sensitivity 0.9 (0.75-0.97), specificity 0.72 (0.64-0.80), PPV 0.44 (0.33-0.57), NPV

				0.97 (0.91-0.99)
			Mortality	Sensitivity 0.75 (0.53-0.89), specificity 0.65 (0.57-0.73), PPV 0.24 (0.15-0.36), NPV 0.95 (0.88-0.98)
Phua 2009 (259)	External validation	1242 patients admitted with pneumonia	Inpatient mortality	AUROC 0.88 (0.86-0.91)
			ICU admission	AUROC 0.85 (0.81-0.88)
ATS 2007 (modified)				
Kontou 2009 (256)	External validation	158 patients hospitalised with pneumonia	ICU admission	For 1+1 criteria; Sensitivity 0.77 (0.62-0.89), specificity 0.77 (0.69-0.83), PPV 0.45 (0.33-0.59), NPV 0.93 (0.87-0.97) For 2+1 criteria: Sensitivity 0.77 (0.62-0.89), specificity 0.84 (0.77-0.9), PPV 0.55 (0.4-0.68), NPV 0.94 (0.88-0.97)
			Mortality	For 1+1 criteria: Sensitivity 0.45 (0.26-0.66), specificity 0.68 (0.6-0.75), PPV 0.17 (0.09-0.29), NPV 0.9 (0.82-0.94) For 2+1 criteria: Sensitivity 0.45 (0.26-0.66), specificity 0.75 (0.67-0.81), PPV 0.2 (0.11-0.35), NPV 0.9 (0.84-0.95)
Man 2011 (260)	External validation	767 ED patients with nursing-home acquired pneumonia	30-day mortality or ICU admission	AUROC 0.627 (0.562-0.692)
BTS				
Angus 2002 (254)	External validation	1339 patients admitted with	30-day mortality	AUROC 0.62 (0.57-0.69)

		pneumonia	ICU admission	AUROC 0.58 (0.53-0.63)
BTS (modified)				
Loh 2004 (261)	External validation	108 patients admitted with pneumonia, including TB	Inpatient mortality	Sensitivity 0.15, specificity 0.93, PPV 0.71, NPV 0.52
CORB				
Buising 2007 (262)	Internal validation	330 patients presenting to ED with pneumonia	Mortality or ICU admission	AUROC 0.72 (0.63-0.76)
CRB				
Bauer 2006 (263)	External validation	1343 ED and primary care patients with pneumonia	30-day mortality	AUROC 0.72 (0.654-0.787)
Buising 2007 (262)	External validation	330 patients presenting to ED with pneumonia	Mortality or ICU admission	AUROC 0.74 (0.66-0.83)
CRB-65				
Bauer 2006 (263)	External validation	1343 ED and primary care patients with pneumonia	30-day mortality	AUROC 0.785 (0.736-0.833)
Capelastegui 2006 (264)	External validation	1776 ED patients with pneumonia	30-day mortality	AUROC 0.864 (0.835-0.892)
Barlow 2007 (265)	External validation Secondary analysis quality improvement study	419 patients admitted with pneumonia	30-day mortality	AUROC 0.73 (0.67-0.79)
Buising 2007 (262)	External validation	330 patients presenting to ED with pneumonia	Mortality or ICU admission	AUROC 0.67 (0.61-0.73)
Man 2007 (266)	External validation	1016 patients admitted with pneumonia	30-day mortality	AUROC 0.694 (0.634-0.753)
			ICU admission	Score 0 3.9%, score 1-2 3.3%, score 3-4 9.5% p=0.01
Schaaf 2007 (267)	External validation	105 patients hospitalised with pneumococcal pneumonia	Inpatient mortality	AUROC 0.845 (0.739-0.951)
Chalmers 2008 (268)	External validation	1007 patients hospitalised with pneumonia, excluding those with malignancy or immunosuppression	30-day mortality	AUROC 0.74 (0.73-0.79)
			IPPV or inotropes	AUROC 0.77 (0.74-0.8)
Kruger 2008 (269)	External validation	1671 in CAPNETZ cohort (primary	28-day mortality	AUROC 0.79

	Secondary registry analysis	and secondary care)		
Schuetz 2008 (270)	External validation Secondary analysis of RCT	373 ED patients with pneumonia, multiple exclusions	30-day mortality	AUROC 0.66 (0.58-0.73)
Zuberi 2008 (271)	External validation	137 patients admitted with pneumonia excluding nursing home residents	30-day mortality	AUROC 0.84
Chalmers 2009 (272)	External validation	1269 patients admitted with pneumonia	Complicated parapneumonic effusion or empyema	AUROC 0.52 (0.49-0.55)
Feldman 2009 (258)	External validation	744 patients admitted with pneumonia	14-day mortality	AUROC 0.7365
Menendez 2009 (273)	External validation	453 patients admitted with pneumonia	30-day mortality	AUROC 0.79 (0.72-0.87)
El-Solh 2010 (274)	External validation	457 patients hospitalised with nursing-home acquired pneumonia	30-day mortality	AUROC 0.605 (0.559-0.650)
			ICU admission	AUROC 0.62 (0.574-0.655)
CURB				
Ewig 2004 (275)	External validation	696 patients hospitalised with pneumonia	ICU admission	AUROC 0.732 (0.676-0.787)
Aujesky 2005 (276)	External validation	3181 ED patients with pneumonia	30-day mortality	AUROC 0.73 (0.68-0.76)
Bauer 2006 (263)	External validation	1343 ED and primary care patients with pneumonia	30-day mortality	AUROC 0.793 (0.745-0.841)
Buising 2006 (255)	External validation	392 patients admitted via ED with pneumonia	Inpatient mortality	AUROC 0.74
			ICU admission	AUROC 0.7
Buising 2007 (262)	External validation	330 patients presenting to ED with pneumonia	Mortality or ICU admission	AUROC 0.76 (0.71-0.82)
Valencia 2007 (257)	External validation	457 patients admitted with PSI class V pneumonia	Inpatient mortality	Sensitivity 0.78 specificity 0.45 PPV 0.3 NPV 0.87
			ICU admission	Sensitivity 0.72 specificity 0.42 PPV 0.24 NPV 0.86
Kontou 2009 (256)	External validation	158 patients hospitalised with	ICU admission	Score >1 sensitivity 0.58

		pneumonia		(0.41-0.74), specificity 0.79 (0.71-0.85), PPV 0.4 (0.27-0.55), NPV 0.89 (0.81-0.93)
			Mortality	Score >1 sensitivity 0.5 (0.3-0.7), specificity 0.75 (0.67-0.81), PPV 0.22 (0.13-0.36), NPV 0.91 (0.85-0.95)
El-Solh 2010 (274)	External validation	457 patients hospitalised with nursing-home acquired pneumonia	30-day mortality	AUROC 0.592 (0.543-0.638)
			ICU admission	AUROC 0.651 (0.605-0.695)
CURB-65				
Lim 2003 (65)	Combined derivation and validation	1068 patients admitted with pneumonia	30-day mortality	Score 0/1 1.5%, score 2 9.2%, score >2 22%.
Aujesky 2005 (276)	External validation	3181 ED patients with pneumonia	30-day mortality	AUROC 0.76 (0.73-0.8)
Buising 2006 (255)	External validation	392 patients admitted via ED with pneumonia	Inpatient mortality	AUROC 0.74
			ICU admission	AUROC 0.61
Capelastegui 2006 (264)	External validation	1776 ED patients with pneumonia	30-day mortality	AUROC 0.87 (0.844-0.895)
España 2006 (277)	External validation	1057 ED patients with pneumonia	Mortality, IPPV or septic shock	AUROC 0.78
		719 ED patients with pneumonia		AUROC 0.79
		1121 ED patients with pneumonia		AUROC 0.71
Barlow 2007 (265)	External validation Secondary analysis quality improvement study	419 patients admitted with pneumonia	30-day mortality	AUROC 0.78 (0.73-0.83)
Buising 2007 (262)	External validation	330 patients presenting to ED with pneumonia	Mortality or ICU admission	AUROC 0.68 (0.6-0.76)
Man 2007 (266)	External validation	1016 patients admitted with pneumonia	30-day mortality	AUROC 0.733 (0.689-0.787)
			ICU admission	Score 0-1 2.3%, score 2 4.4%, score >2 6.5% p=0.02

Valencia 2007 (257)	External validation	457 patients admitted with PSI class V pneumonia	Inpatient mortality	Sensitivity 0.73 specificity 0.8 PPV 0.53 NPV 0.85
			ICU admission	Sensitivity 0.6 specificity 0.44 PPV 0.21 NPV 0.81
Ananda-Rajah 2008 (278)	External validation	408 patients hospitalised with pneumonia	30-day mortality	AUROC 0.69
			ICU admission	AUROC 0.63
Chalmers 2008 (268)	External validation	1007 patients hospitalised with pneumonia, excluding those with malignancy or immunosuppression	30-day mortality	AUROC 0.76 (0.74-0.79)
			IPPV or inotropes	AUROC 0.78 (0.75-0.81)
Charles 2008 (279)	External validation	882 patients admitted with pneumonia	ETT, NIV or vasopressor	AUROC 0.67
Schuetz 2008 (280)	External validation	281 patients admitted with pneumonia	Mortality or intensive care admission	AUROC 0.66 (0.58-0.73)
Schuetz 2008 (270)	External validation Secondary analysis of RCT	373 ED patients with pneumonia, multiple exclusions	30-day mortality	AUROC 0.69 (0.61-0.77)
Shindo 2008 (252)	External validation	329 patients hospitalised with pneumonia	30-day mortality	AUROC 0.835 (0.763-0.908)
Zuberi 2008 (271)	External validation	137 patients admitted with pneumonia excluding nursing home residents	30-day mortality	AUROC 0.86
Chalmers 2009 (281)	External validation Secondary analysis diagnostic study	314 ED patients with pneumonia	30-day mortality	AUROC 0.79 (0.74-0.85)
			IPPV or vasopressor use	AUROC 0.77 (0.72-0.83)
Chalmers 2009 (272)	External validation	1269 patients admitted with pneumonia	Complicated parapneumonic effusion or empyema	AUROC 0.54 (0.51-0.57)
Feldman 2009 (258)	External validation	744 patients admitted with pneumonia	14-day mortality	AUROC 0.7361
Huang 2009 (282)	External validation	1653 ED patients with pneumonia	30-day mortality	AUROC 0.78
Menendez 2009 (273)	External validation	453 patients admitted with	30-day mortality	AUROC 0.82 (0.76-0.89)

		pneumonia		
Parsonage 2009 (283)	External validation	132 patients age 16-64 admitted with pneumonia	30-day mortality	AUROC 0.93
		92 patients age 65-74 admitted with pneumonia		AUROC 0.74
		128 patients age 75-84 admitted with pneumonia		AUROC 0.74
		76 patients age >84 admitted with pneumonia		AUROC 0.59
Phua 2009 (259)	External validation	1242 patients admitted with pneumonia	Inpatient mortality	AUROC 0.82 (0.78-0.85)
			ICU admission	AUROC 0.68 (0.63-0.72)
Renaud 2009 (69)	External validation	6560 patients hospitalised with pneumonia	ICU admission in 1-3 days	AUROC 0.69 (0.66-0.72)
Yandiola 2009 (284)	External validation	671 patients hospitalised with pneumonia	ICU admission	AUROC 0.61
			Mechanical ventilation	AUROC 0.61
			Severe sepsis	AUROC 0.66
Chen 2010 (285)	External validation	348 ED patients age 18-64 with pneumonia	30-day mortality	AUROC 0.8 (0.67-0.93)
		438 ED patients age 65-84 with pneumonia		AUROC 0.73 (0.65-0.82)
		201 ED patients age >84 with pneumonia		AUROC 0.6 (0.48-0.73)
El-Solh 2010 (274)	External validation	457 patients hospitalised with nursing-home acquired pneumonia	30-day mortality	AUROC 0.593 (0.543-0.638)
			ICU admission	AUROC 0.657 (0.611-0.7)
Schuetz 2010 (286)	External validation	925 patients hospitalised with pneumonia	30-day mortality	AUROC 0.74
			Serious complication	AUROC 0.66
Albrich 2011 (287)	External validation	1359 patients hospitalised with pneumonia or LRTI	30-day mortality	AUROC 0.73 (.68-.75)
			30-day mortality, ICU admission or complication	AUROC 0.65 (.61-.69)
Jeong 2011 (253)	External validation	502 ED patients with pneumonia	30-day mortality	AUROC 0.764 (0.703-0.825)

Jones 2011 (288)	External validation	2069 ED patients with pneumonia, multiple exclusions	30-day mortality	AUROC 0.82
Man 2011 (260)	External validation	767 ED patients with nursing-home acquired pneumonia	30-day mortality or ICU admission	AUROC 0.702 (0.649-0.755)
Labarère 2012 (289)	External validation	850 ED patients with pneumonia, excluding nursing home residents	30-day mortality	AUROC 0.74 (0.67-0.82)
			30-day ICU admission	AUROC 0.65 (0.57-0.72)
			3-day ICU admission	AUROC 0.63 (0.55-0.71)
Park 2012 (290)	External validation	126 patients hospitalised with pneumonia	28-day mortality	AUROC 0.864
CURXO-80				
Labarère 2012 (289)	External validation	850 ED patients with pneumonia, excluding nursing home residents	30-day mortality	AUROC 0.78 (0.72-0.84)
			30-day ICU admission	AUROC 0.74 (0.68-0.81)
			3-day ICU admission	AUROC 0.74 (0.67-0.8)
eCURB				
Jones 2011 (288)	Derivation	2069 ED patients with pneumonia, multiple exclusions	30-day mortality	AUROC 0.86
	External validation	1048 ED patients with pneumonia, multiple exclusions		AUROC 0.845
IDSA/ATS				
Man 2011 (260)	External validation	767 ED patients with nursing-home acquired pneumonia	30-day mortality or ICU admission	AUROC 0.712 (0.654-0.770)
Labarère 2012 (289)	External validation	850 ED patients with pneumonia, excluding nursing home residents	30-day mortality	AUROC 0.75 (0.68-0.82)
			30-day ICU admission	AUROC 0.74 (0.67-0.81)
			3-day ICU admission	AUROC 0.72 (0.64-0.8)
Park 2012 (290)	External validation	126 patients hospitalised with pneumonia	28-day mortality	AUROC 0.844
Pitt bacteremia score				
Feldman 2009 (258)	External validation	766 patients admitted with pneumonia	14-day mortality	AUROC 0.8397
Pneumonia severity score				
Carusone 2007 (291)	External validation Control arm of RCT	353 nursing home residents with LRTI or pneumonia	30-day mortality	Score 0 8%, score 1 8%, score 2 10%, score 3 33%, score 4

				0%
			Hospitalisation	Score 0 19%, score 1 14%, score 2 45%, score 3 58%, score 4 100%
PSI				
Flanders 1999 (292)	External validation Secondary analysis of quality improvement data	1024 patients admitted with pneumonia	Hospital mortality	AUROC 0.847
Feagan 2000 (293)	External validation	858 patients hospitalised with pneumonia	30-day mortality	Class I-II 0, class III 3.3%, class IV-V 22.4%
			ICU admission	Class I-II 7.5%, class III 9.4%, class IV-V 17%
Dedier 2001 (294)	External validation	1062 patients hospitalised with pneumonia	Hospital mortality	class I 0, class II 1.6%, class III 3.5%, class IV 3.6%, class V 22.3% p<0.001
Angus 2002 (254)	External validation	1339 patients admitted with pneumonia	30-day mortality	AUROC 0.75 (0.71-0.78)
			ICU admission	AUROC 0.6 (0.56-0.65)
Mody 2002 (295)	External validation	101 patients >60y admitted with pneumonia	30-day mortality	Class 2 0, class 3 0.5%, class 4 10.8%, class 5 25%
Ewig 2004 (275)	External validation	696 patients hospitalised with pneumonia	ICU admission	AUROC 0.607 (0.607-0.727)
Querol-Ribelles 2004 (296)	External validation	302 ED patients with pneumonia	30-day mortality	AUROC 0.91 (0.88-0.95)
Aujesky 2005 (276)	External validation	3181 ED patients with pneumonia	30-day mortality	AUROC 0.81 (0.78-0.84)
Buising 2006 (255)	External validation	392 patients admitted via ED with pneumonia	Inpatient mortality	AUROC 0.73
			ICU admission	AUROC 0.65
Capelastegui 2006 (264)	External validation	1776 ED patients with pneumonia	30-day mortality	AUROC 0.888 (0.864-0.912)
España 2006 (277)	External validation	1057 ED patients with pneumonia	Mortality, IPPV or septic shock	AUROC 0.81
		719 ED patients with pneumonia		AUROC 0.79

		1121 ED patients with pneumonia		AUROC 0.71
Migliorati 2006 (297)	External validation	148 patients hospitalised with pneumonia	30-day mortality	Class I-II 0, class III 6.2%, class IV 1.8%, class V 30.2%
			ICU admission	Class I-II 0, class III 6.2%, class IV 0, class V 13.2%
Sanders 2006 (298)	External validation	284 immunocompromised patients admitted with pneumonia	Inpatient mortality	AUROC 0.7 (0.6-0.79) overall; 0.77 (0.52-1) low-risk (HIV, solid organ xplant); 0.6 (0.46-0.74) high-risk (marrow xplant, haem malignancy, post-chemo)
Buising 2007 (262)	External validation	330 patients presenting to ED with pneumonia	Mortality or ICU admission	AUROC 0.68 (0.62-0.75)
Etzion 2007 (299)	External validation	591 patients hospitalised with pneumonia	30-day mortality	AUROC 0.86 (0.8-0.92)
Man 2007 (266)	External validation	1016 patients admitted with pneumonia	30-day mortality	AUROC 0.736 (0.687-0.786)
			ICU admission	Group II-III 2.7%, group IV 4.5%, group V 6.6 % p=0.063
Renaud 2007 (300)	External validation	925 ED patients with pneumonia	28-day mortality	AUROC 0.85 (0.81-0.88)
			ICU admission	Class I 4.3%, class II 7.6%, class III 6.5%, class IV 11.1%, class V 17.9%
		853 ED patients with pneumonia	28-day mortality	AUROC 0.89 (0.85-0.93)
			ICU admission	Class I 2.1%, class II 4.9%, class III 3.8%, class IV 1.4%, class V 7.1%
Ananda-Rajah 2008 (278)	External validation	408 patients hospitalised with pneumonia	30-day mortality	AUROC 0.72
			ICU admission	AUROC 0.58
Charles 2008 (279)	External validation	882 patients admitted with pneumonia	ETT, NIV or vasopressor	AUROC 0.69
Chen 2008 (301)	External validation	250 patients hospitalised with	Hospital mortality	Class I-II 0, class III 2.9%,

		pneumonia		class IV 7.8%, class V 25.3% p<0.001
Garau 2008 (302)	External validation	3233 patients admitted with pneumonia	2-day mortality	Class I: 0, II 0.2%, III 0.3%, IV 1.3%, V 7.5%
			ICU admission	Class I: 2.5%, II 3.7%, III 3.9%, IV 5%, V 10.2%
Restrepo 2008 (303)	External validation	730 patients hospitalised with pneumonia	30-day mortality	Class I 2.7%, class II 2.6% class III 5.8%, class IV 8.6%, class V 27%
			ICU admission	Class I 9%, class II 7.7%, class III 14.4%, class IV 25%, class V 50%
Schuetz 2008 (280)	External validation	281 patients admitted with pneumonia	Mortality or intensive care admission	AUROC 0.71 (0.65-0.78)
Schuetz 2008 (270)	External validation Secondary analysis of RCT	373 ED patients with pneumonia, multiple exclusions	30-day mortality	AUROC 0.72 (0.65-0.78)
Chalmers 2009 (281)	External validation Secondary analysis diagnostic study	314 ED patients with pneumonia	30-day mortality	AUROC 0.79 (0.73-0.84)
			IPPV or vasopressor use	AUROC 0.73 (0.67-0.78)
Chalmers 2009 (272)	External validation	1269 patients admitted with pneumonia	Complicated parapneumonic effusion or empyema	AUROC 0.55 (0.52-0.58)
Feldman 2009 (258)	External validation	742 patients admitted with pneumonia	14-day mortality	AUROC 0.721
Huang 2009 (282)	External validation	1653 ED patients with pneumonia	30-day mortality	AUROC 0.83
Kontou 2009 (256)	External validation	158 patients hospitalised with pneumonia	ICU admission	PSI V: sensitivity 0.45 (0.29-0.62), specificity 0.83 (0.76-0.89), PPV 0.4 (0.26-0.56), NPV 0.9 (0.83-0.94) PSI IV and V: Sensitivity 0.81 (0.64-0.91),

				specificity 0.5 (0.41-0.58), PPV 0.28 (0.2-0.38), NPV 0.91 (0.82-0.96)
			Mortality	PSI V: sensitivity 0.5 (0.30-0.7), specificity 0.82 (0.75-0.87), PPV 0.29 (0.16-0.45), NPV 0.92 (0.86-0.96) PSI IV and V: sensitivity 0.95 (0.76-0.99), specificity 0.49 (0.41-0.58), PPV 0.21 (0.14-0.31), NPV 0.99 (0.92-0.99)
Menendez 2009 (273)	External validation	453 patients admitted with pneumonia	30-day mortality	AUROC 0.81 (0.75-0.87)
Pilotto 2009 (304)	External validation	134 patients age >65 admitted with pneumonia	30-day mortality	AUROC 0.71 (0.62-0.78)
Phua 2009 (259)	External validation	1242 patients admitted with pneumonia	Inpatient mortality	AUROC 0.86 (0.83-0.88)
			ICU admission	AUROC 0.75 (0.71-0.79)
Renaud 2009 (69)	External validation	6560 patients hospitalised with pneumonia	ICU admission in 1-3 days	AUROC 0.75 (0.73-0.78)
Yandiola 2009 (284)	External validation	671 patients hospitalised with pneumonia	ICU admission	AUROC 0.63
			Mechanical ventilation	AUROC 0.69
			Severe sepsis	AUROC 0.72
Chen 2010 (285)	External validation	348 ED patients age 18-64 with pneumonia	30-day mortality	AUROC 0.87 (0.77-0.97)
		438 ED patients age 65-84 with pneumonia		AUROC 0.85 (0.8-0.9)
		201 ED patients age >84 with pneumonia		AUROC 0.69 (0.6-0.79)
Schuetz 2010 (286)	External validation	925 patients hospitalised with pneumonia	30-day mortality	AUROC 0.84
			Serious complication	AUROC 0.69
Jeong 2011 (253)	External validation	502 ED patients with pneumonia	30-day mortality	AUROC 0.795 (0.742-0.848)
Man 2011 (260)	External validation	767 ED patients with nursing-	30-day mortality or ICU	AUROC 0.688 (0.633-0.744)

		home acquired pneumonia	admission	
Labarère 2012 (289)	External validation	850 ED patients with pneumonia, excluding nursing home residents	30-day mortality	AUROC 0.82 (0.77-0.88)
			30-day ICU admission	AUROC 0.68 (0.61-0.75)
			3-day ICU admission	AUROC 0.66 (0.59-0.74)
Park 2012 (290)	External validation	126 patients hospitalised with pneumonia	28-day mortality	AUROC 0.875
REA-ICU				
Renaud 2009 (69)	Derivation	4593 patients hospitalised with pneumonia	ICU admission in 1-3 days	AUROC 0.8 (0.77-0.83)
	Internal validation	1967 patients hospitalised with pneumonia		AUROC 0.8 (0.76-0.84)
Labarère 2012 (289)	External validation	850 ED patients with pneumonia, excluding nursing home residents	30-day mortality	AUROC 0.74 (0.66-0.82)
			30-day ICU admission	AUROC 0.78 (0.72-0.84)
			3-day ICU admission	AUROC 0.76 (0.7-0.83)
SCAP				
España 2006 (277)	Derivation	1057 ED patients with pneumonia	Mortality, IPPV or septic shock	AUROC 0.83
	Internal validation	719 ED patients with pneumonia		AUROC 0.86
	External validation	1121 ED patients with pneumonia		AUROC 0.72
Renaud 2009 (69)	External validation	6560 patients hospitalised with pneumonia	ICU admission in 1-3 days	AUROC 0.74 (0.71-0.76)
Yandiola 2009 (284)	External validation	671 patients hospitalised with pneumonia	ICU admission	AUROC 0.746
			Mechanical ventilation	AUROC 0.76
			Severe sepsis	AUROC 0.79
Man 2011 (260)	External validation	767 ED patients with nursing-home acquired pneumonia	30-day mortality or ICU admission	AUROC 0.650 (0.596-0.703)
SEWS				
Barlow 2007 (265)	External validation Secondary analysis quality improvement study	419 patients admitted with pneumonia	30-day mortality	AUROC 0.64 (0.57-0.7)
Chalmers 2009 (272)	External validation	1269 patients admitted with	Complicated	AUROC 0.53 (0.50-0.56)

		pneumonia	parapneumonic effusion or empyema	
SIRS				
Barlow 2007 (265)	External validation Secondary analysis quality improvement study	419 patients admitted with pneumonia	30-day mortality	AUROC 0.68 (0.61-0.75)
Schaaf 2007 (267)	External validation	105 patients hospitalised with pneumococcal pneumonia	Inpatient mortality	Non-SIRS 0%, SIRS 0%, severe/shock 30.6% ($p<0.0001$)
SMART-COP				
Charles 2008 (279)	Derivation	882 patients admitted with pneumonia	ETT, NIV or vasopressor	AUROC 0.87
	External validation	2067 patients admitted with pneumonia		AUROC 0.72 (0.68-0.77)
		1307 patients admitted with pneumonia		AUROC 0.78 (0.72-0.83)
		408 patients admitted with pneumonia		AUROC 0.81 (0.74-0.88)
		608 patients admitted with pneumonia		AUROC 0.82 (0.77-0.86)
		3074 patients admitted with pneumonia		AUROC 0.87 (0.83-0.91)
Labarère 2012 (289)	External validation	850 ED patients with pneumonia, excluding nursing home residents	30-day mortality	AUROC 0.71 (0.6-0.81)
			30-day ICU admission	AUROC 0.76 (0.79-0.83)
			3-day ICU admission	AUROC 0.75 (0.67-0.83)
SMRT-CO				
Charles 2008 (279)	Derivation Secondary rule	882 patients admitted with pneumonia	ETT, NIV or vasopressor	AUROC 0.8
	External validation	2067 patients admitted with pneumonia		AUROC 0.69 (0.65-0.73)
		1307 patients admitted with pneumonia		AUROC 0.74 (0.69-0.79)

		408 patients admitted with pneumonia		AUROC 0.78 (0.7-0.85)
		608 patients admitted with pneumonia		AUROC 0.76 (0.7-0.81)
		3074 patients admitted with pneumonia		AUROC 0.8 (0.76-0.84)
SOAR				
El-Solh 2010 (274)	External validation	457 patients hospitalised with nursing-home acquired pneumonia	30-day mortality	AUROC 0.765 (0.724-0.803)
			ICU admission	AUROC 0.734 (0.691-0.774)

ETT: endotracheal intubation; IPPV: intermittent positive pressure ventilation; NIV: non-invasive ventilation

Pulmonary embolism

Aujesky				
Aujesky 2006 (305)	Derivation	10354 patients with discharge diagnosis of PE	30-day mortality	Sensitivity 0.99 (0.98-0.99) specificity 0.24 (0.23-0.25) PPV 0.12 (0.11-0.12) NPV 0.99 (0.99-1)
	Internal validation	5177 patients with discharge diagnosis of PE		Sensitivity 0.97 (0.95-0.98) specificity 0.23 (0.22-0.25) PPV 0.12 (0.11-0.13) NPV 0.98 (0.98-0.99)
	External validation Secondary analysis of diagnostic study	221 patients undergoing CT for potential PE		Sensitivity 1 (0.54-1) specificity 0.35 (0.29-0.42) PPV 0.04 (0.02-0.09) NPV 1 (0.95-1)
Palmieri 2008 (306)	External validation	89 ED patients with non-massive PE	Hospital mortality	Score <65 0, 65-85 0, 86-105 11%, 106-25 23%, >125 22%
			Haemodynamic instability	Score <65 0, 65-85 20%, 86-105 56%, 106-125 39%, >125 56%
PESI				
Choi 2009 (307)	External validation	90 patients with PE diagnosed by CT	30-day mortality	Class I 0, class II 10.3%, class III 9.1%, class IV 0, class V 50%.
Nordenholz 2011 (308)	External validation Retrospective review	168 patients with PE	Mortality, respiratory failure, need for ventilation or thrombolysis	2.1% class I, 2.3% class II, 15.6% class III, 8% class IV, 15% class V
PESI (simplified)				
Jimenez 2010 (71)	Derivation Prospective registry data	995 patients with symptomatic PE	30-day PE-specific mortality	AUROC 0.75 (0.69-0.8)
	External validation Prospective registry data	7106 patients with PE	30-day mortality	NPV 98.9% (98.5-99.3)

Utrecht score				
Agterof 2011 (309)	Derivation	210 normotensive patients admitted with PE	10-day adverse events	AUROC 0.82

Sepsis

APACHE II				
Nguyen 2008 (310)	External validation Secondary analysis of RCT	246 patients meeting EGDT criteria	Hospital mortality	AUROC 0.73 (0.67-0.8)
Chou 2010 (311)	External validation	90 patients with Vibrio vulnificus infection	Hospital mortality	AUROC 0.928 (0.854-0.972)
CURB-65				
Howell 2007 (312)	External validation	2132 ED patients with infection	28-day mortality	AUROC 0.788 (0.744-0.833)
Crowe 2010 (313)	External validation	216 ED patients receiving EGDT	Hospital mortality	AUROC 0.59 (0.51-0.67)
MEDS				
Shapiro 2003 (74)	Derivation	2070 ED patients having blood cultures taken	Hospital mortality	AUROC 0.82 HL 0.21
	Internal validation	1109 ED patients having blood cultures taken		AUROC 0.76 HL 0.39
Howell 2007 (312)	External validation	2132 ED patients with infection	28-day mortality	AUROC 0.849 (0.812-0.877)
Jones 2008 (314)	External validation Secondary analysis quality improvement study	143 patients with severe sepsis or septic shock	Hospital mortality	AUROC 0.61 (0.5-0.72)
Lee 2008 (315)	External validation	525 ED patients with SIRS	5-day mortality	AUROC 0.89
			5 to 30-day mortality	AUROC 0.78
Nguyen 2008 (310)	External validation Secondary analysis of RCT	246 patients meeting EGDT criteria	Hospital mortality	AUROC 0.6 (0.53-0.67)
Sankoff 2008 (316)	External validation	385 ED patients with SIRS	28-day mortality	AUROC 0.88 (0.83-0.92)
Vorwerk 2009 (317)	External validation	307 ED patients with 2 SIRS criteria + infection	28-day mortality	AUROC 0.82 (0.78-0.87)
Chou 2010 (311)	External validation	90 patients with Vibrio vulnificus infection	Hospital mortality	AUROC 0.830 (0.736-0.901)
Crowe 2010 (313)	External validation	216 ED patients receiving EGDT	Hospital mortality	AUROC 0.74 (0.67-0.81)
Ghanem-Zoubi 2011 (318)	External validation	1072 inpatients with sepsis	1-day mortality	AUROC 0.79 (0.73-0.85)
			5-day mortality	AUROC 0.77 (0.73-0.81)
			30-day mortality	AUROC 0.73 (0.71-0.78)
Hermans 2012 (319)	External validation	331 patients admitted via ED	28-day mortality	AUROC 0.81 (.73-.88)

		with sepsis		
MEWS				
Vorwerk 2009 (317)	External validation	307 ED patients with 2 SIRS criteria + infection	28-day mortality	AUROC 0.72 (0.67-0.77)
Ghanem-Zoubi 2011 (318)	External validation	1072 inpatients with sepsis	1-day mortality	AUROC 0.83 (0.77-0.88)
			5-day mortality	AUROC 0.73 (0.68-0.78)
			30-day mortality	AUROC 0.67 (0.63-0.71)
MPM0				
Nguyen 2008 (310)	External validation Secondary analysis of RCT	246 patients meeting EGDT criteria	Hospital mortality	AUROC 0.72 (0.65-0.79)
REMS				
Howell 2007 (312)	External validation	2132 ED patients with infection	28-day mortality	AUROC 0.802 (0.752-0.852)
Crowe 2010 (313)	External validation	216 ED patients receiving EGDT	Hospital mortality	AUROC 0.62 (0.54-0.69)
Ghanem-Zoubi 2011 (318)	External validation	1072 inpatients with sepsis	1-day mortality	AUROC 0.87 (0.83-0.92)
			5-day mortality	AUROC 0.8 (0.76-0.84)
			30-day mortality	AUROC 0.76 (0.72-0.79)
SAPS II				
Kofoed 2008 (320)	External validation	151 patients hospitalised with 2 SIRS criteria + infection	30-day mortality	AUROC 0.89 (0.8-0.98)
Nguyen 2008 (310)	External validation Secondary analysis of RCT	246 patients meeting EGDT criteria	Hospital mortality	AUROC 0.71 (0.64-0.78)
Chen 2009 (321)	External validation	298 patients hospitalised with pyogenic liver abscess	Hospital mortality	AUROC 0.857 (0.812-0.895)
Sepsis algorithm				
Thiel 2010 (322)	Derivation	13785 hospitalised patients	ICU admission in next 24h	Sensitivity 56.9%, specificity 93.7%
	Temporal validation	13737 hospitalised patients		Sensitivity 54.7, specificity 93.4
	Temporal validation	13937 hospitalised patients		Sensitivity 55, specificity 93
Simple Clinical Score				
Ghanem-Zoubi 2011 (318)	External validation	1072 inpatients with sepsis	1-day mortality	AUROC 0.85 (0.8-0.9)
			5-day mortality	AUROC 0.79 (0.76-0.83)

			30-day mortality	AUROC 0.77 (0.74-0.81)
SOFA				
Kofoed 2008 (320)	External validation	151 patients hospitalised with 2 SIRS criteria + infection	30-day mortality	AUROC 0.8 (0.65-0.94)

EGDT: early goal-directed therapy; SIRS: systemic inflammatory response syndrome

Surgical

APACHE II				
Kulkarni 2007 (323)	External validation	50 patients with peritonitis secondary to perforated hollow viscus (including from blunt trauma)	Hospital mortality	AUROC 0.984
Ertan 2008 (324)	External validation	102 patients undergoing emergency surgery for colorectal cancer	Hospital mortality	AUROC 0.78
APACHE II (modified)				
Mishra 2003 (325)	External validation	140 patients with perforated peptic ulcer	Hospital mortality	AUROC 0.84 +/- 0.06
APACHE III				
Ertan 2008 (324)	External validation	102 patients undergoing emergency surgery for colorectal cancer	Hospital mortality	AUROC 0.773
Boey				
Mishra 2003 (325)	External validation	140 patients with perforated peptic ulcer	Hospital mortality	AUROC 0.85 +/- 0.06
CR-POSSUM				
Ertan 2008 (324)	External validation	102 patients undergoing emergency surgery for colorectal cancer	Hospital mortality	AUROC 0.718
Hacettepe				
Mishra 2003 (325)	External validation	140 patients with perforated peptic ulcer	Hospital mortality	AUROC 0.72 +/- 0.08
Jabalpur				
Mishra 2003 (325)	Derivation	140 patients with perforated peptic ulcer	Hospital mortality	AUROC 0.92 +/- 0.03
Mannheim Peritonitis Index				
Mäkelä 2005 (326)	External validation	172 patients undergoing surgery for perforated diverticular	Mortality	Age <70 score I 1% score II 21% p=0.002

		disease		Age >70 score 1 2% score II 50% p=0.001
Notash 2005 (327)	External validation	80 patients undergoing surgery for peritonitis	Mortality	AUROC 0.972
Biondo 2006 (328)	External validation	156 patients undergoing surgery for distal colonic peritonitis	Hospital mortality	AUROC 0.725 (0.648-0.794)
Mannheim (modified)				
Mishra 2003 (325)	External validation	140 patients with perforated peptic ulcer	Hospital mortality	AUROC 0.85 +/- 0.06
MPM II				
Ertan 2008 (324)	External validation	102 patients undergoing emergency surgery for colorectal cancer	Hospital mortality	AUROC 0.714
Peritonitis severity score				
Biondo 2006 (328)	External validation	156 patients undergoing surgery for distal colonic peritonitis	Hospital mortality	AUROC 0.793 (0.721-0.854)
P-POSSUM				
Poon 2005 (329)	External validation	160 patients undergoing surgery for obstructing colorectal cancer	Mortality	AUROC 0.75
SAPS II				
Ertan 2008 (324)	External validation	102 patients undergoing emergency surgery for colorectal cancer	Hospital mortality	AUROC 0.83

Syncope

EGSYS				
Del Rosso 2008 (330)	Derivation	260 ED patients with syncope	Cardiac cause found	AUROC 0.904 (0.864-0.943)
	Internal validation	256 ED patients with syncope		AUROC 0.849 (0.778-0.921)
OESIL				
Hing 2005 (331)	External validation	100 ED patients with syncope	Adverse cardiac outcome	AUROC 0.73 (0.63-0.84)
Dipaola 2010 (332)	External validation	492 ED patients with syncope	10-day mortality, major intervention or readmission	Sensitivity 0.88 (0.7-0.98), specificity 0.59 (0.55-0.64), PPV 0.11 (0.07-0.15), NPV 0.99 (0.98-1)
Numeroso 2010 (333)	External validation	200 patients with syncope on ED observation ward	Cardiogenic cause found	OESIL >1 sensitivity 0.98, specificity 0.28, PPV 0.284, NPV 0.978
San Francisco Syncope Rule				
Quinn 2004 (83)	Derivation	684 ED patients with syncope or near syncope	7-day serious outcome	Sensitivity 0.96 (0.92-1), specificity 0.62 (0.58-0.66)
Quinn 2006 (334)	External validation	791 ED patients with syncope or near syncope	30-day serious outcome	Sensitivity 0.98 (0.89-1), specificity 0.56 (0.52-0.6)
Cosgriff 2007 (335)	External validation	89 ED patients with syncope	7-day serious outcome	Sensitivity 0.9 (0.6-0.98) specificity 0.57 (0.46-0.67) PPV 0.21 (0.11-0.35) NPV 0.98 (0.89-0.99)
Sun 2007 (336)	External validation	351 ED patients with syncope or near syncope	7-day serious outcome	Sensitivity 0.89 (0.81-0.97) specificity 0.42 (0.37-0.48) PPV 0.22 (0.17-0.28) NPV 0.95 (0.92-0.99)
Birnbaum 2008 (337)	External validation	713 ED patients with syncope or near syncope	7-day serious outcome	Sensitivity 0.74 (0.61-0.84) specificity 0.57 (0.53-0.61) PPV 0.14 (0.1-0.18) NPV 0.96 (0.93-0.98)
Dipaola 2010 (332)	External validation	492 ED patients with syncope	10-day mortality, major intervention or	Sensitivity 0.81 (0.61-0.92), specificity 0.63 (0.58-0.67),

			readmission	PPV 0.11 (0.06-0.15), NPV 0.98 (0.97-1)

Transient ischaemic attack

ABCD				
Bray 2007 (338)	External validation	98 ED patients with TIA	7-day CVA incidence	ABCD>4 sensitivity 1 (0.4-1) specificity 0.53 (0.43-0.63) PPV 0.08 (0.03-0.21) NPV 1 (0.91-1)
Tsivgoulis 2007 (339)	External validation	226 patients hospitalised with TIA	7-day CVA incidence	ABCD>4 HR 8.74 (2.72-29.32)
			30-day CVA incidence	AUROC 0.78 (0.69-0.87)
Sciolla 2008 (340)	External validation	274 ED patients with TIA	7-day CVA incidence	AUROC 0.75 (0.63-0.88)
			30-day CVA incidence	AUROC 0.76 (0.66-0.86)
ABCD2				
Johnstone 2007 (63)	External validation	1069 ED patients with TIA	2-day CVA incidence	AUROC 0.72 (0.6-0.84)
			7-day CVA incidence	AUROC 0.63 (0.57-0.69)
Ay 2009 (341)	External validation	479 ED patients with TIA	7-day CVA incidence	AUROC 0.66 (0.57-0.76)
Giles 2011 (342)	External validation	3363 patients with TIA, tissue negative on MRI	7-day CVA incidence	AUROC 0.73 (0.67-0.8)
		1211 patients with TIA, tissue positive on MRI		AUROC 0.68 (0.63-0.73)
Stead 2011 (343)	External validation	637 ED patients with TIA	7-day CVA incidence	Low risk 1.1%, intermediate 0.3%, high 2.7%

CVA: cerebrovascular accident; MRI: magnetic resonance imaging

Multi-diagnosis groups

APACHE II				
Man 2007 (344)	External validation	867 patients in ED resus area	14-day mortality	AUROC 0.743 (0.705-0.798)
Cattermole 2009 (129)	External validation	330 patients in ED resus area	30-day mortality	AUROC 0.838 (0.793-0.876)
			7-day mortality or ICU admission	AUROC 0.733 (0.681-0.780)
Emergency Severity Index				
Platts-Mills 2010 (120)	External validation	782 ED patients aged >65y	Immediate life-saving intervention	ESI 1 sensitivity 42% (26%-61%) specificity 99.1% (98.1-99.5%)
HOTEL				
Kellett 2008 (89)	Derivation	6947 patients admitted to MAU	Mortality in 15min-24h	AUROC 0.865 (0.793-0.937)
	Temporal validation	3343 patients admitted to MAU		AUROC 0.854 (0.746-0.962)
MARS				
Silke 2010 (345)	Derivation	10712 patients admitted to MAU	5-day mortality	AUROC 0.93 (0.92-0.94)
	External validation	13182 patients admitted to MAU		AUROC 0.92 (0.9-0.94)
MEWS				
Kellett 2006 (92)	External validation	3228 patients admitted to MAU	30-day mortality	AUROC 0.649
Cattermole 2009 (129)	External validation	330 patients in ED resus area	30-day mortality	AUROC 0.754 (0.703-0.799)
			7-day mortality or ICU admission	AUROC 0.761 (0.711-0.806)
PEDS				
Cattermole 2009 (129)	Secondary endpoint	330 patients in ED resus area	30-day mortality	AUROC 0.898 (0.860-0.928)
	Derivation		7-day mortality or ICU admission	AUROC 0.909 (0.872-0.938)
RAPS				
Olsson 2003 (91)	External validation	11751 patients at non-surgical Emergency Department	Hospital mortality	AUROC 0.652 +/- 0.019
Goodacre 2006 (127)	External validation Secondary analysis quality data	2215 patients transported to ED by emergency ambulance	Hospital mortality	AUROC 0.64 (0.59-0.69)
Man 2007 (344)	External validation	867 patients in ED resus area	14-day mortality	AUROC 0.654 (0.596-0.712)
REMS				

Olsson 2003 (91)	External validation	11751 patients at non-surgical Emergency Department	Hospital mortality	AUROC 0.852 +/- 0.019
Goodacre 2006 (127)	External validation Secondary analysis quality data	2215 patients transported to ED by emergency ambulance	Hospital mortality	AUROC 0.74 (0.7-0.78)
Man 2007 (344)	External validation	867 patients in ED resus area	14-day mortality	AUROC 0.723 (0.674-0.771)
Cattermole 2009 (129)	External validation	330 patients in ED resus area	30-day mortality	AUROC 0.771 (0.722-0.816)
			7-day mortality or ICU admission	AUROC 0.696 (0.643-0.745)
RTS				
Cattermole 2009 (129)	External validation	330 patients in ED resus area	30-day mortality	AUROC 0.766 (0.717-0.811)
			7-day mortality or ICU admission	AUROC 0.748 (0.698-0.794)
SAPS II				
Cosentini 2009 (346)	External validation	211 patients admitted to MAU	Hospital mortality	AUROC 0.84 (0.77-0.91)
SEWS				
Paterson 2006 (347)	External validation Secondary analysis of clinical audit	848 patients admitted to MAU or SAU	Hospital mortality	Score 0-1 0.5%, 2-3 6%, 4-5 15%, >5 20%
Simple Clinical Score				
Kellett 2006 (92)	Derivation	6736 patients admitted to MAU	30-day mortality	AUROC 0.858 +/- 0.009
	Internal validation	3228 patients admitted to MAU		AUROC 0.856 +/- 0.013
Emmanuel 2010 (348)	External validation	207 patients admitted to MAU	Hospital mortality	AUROC 0.94
Worthing				
Duckitt 2007 (349)	Derivation	3184 patients admitted to MAU	Hospital mortality	AUROC 0.74 (0.71-0.77)
	Internal validation	1102 patients admitted to MAU		AUROC 0.72 (0.66-0.79)

MAU: medical assessment unit; SAU: surgical assessment unit

Other

AF nomogram				
Barrett 2011 (350)	Derivation	832 ED patients with symptomatic AF	30-day mortality, cardiovascular event or return to ED	AUROC 0.67 (0.63-0.71)
APACHE II (modified)				
Eizadi-Mood 2007 (351)	External validation	131 patients hospitalised with organophosphate poisoning	Need for ET intubation	AUROC 0.892 (0.826-0.940)
Mood 2011 (352)	External validation	92 patients hospitalised with mixed drug overdose	Hospital mortality	AUROC 0.81 (0.69-0.89)
Aronin				
Aronin 1998 (353)	Derivation	176 patients hospitalised with bacterial meningitis on LP	Hospital mortality	AUROC 0.73 (0.62-0.81)
	Internal validation	93 patients hospitalised with bacterial meningitis on LP		AUROC 0.81 (0.71-0.92)
Elbaz				
Elbaz 2008 (354)	Derivation	169 patients hospitalised with core temperature <35	Hospital mortality	AUROC 0.81 (0.75-0.87)
Fournier's Gangrene Severity Index				
Corcoran 2008 (355)	External validation	68 patients hospitalised with Fournier's gangrene	Hospital mortality	For cutoff 9 sensitivity 71.4% specificity 90%
GCS				
Davies 2008 (356)	External validation Secondary analysis of RCT	990 patients hospitalised with organophosphate poisoning	Hospital mortality	AUROC 0.84 (0.8-0.87)
Mood 2011 (352)	External validation	92 patients hospitalised with mixed drug overdose	Hospital mortality	AUROC 0.77 (.64-.86)
Get with the Guidelines Stroke Score				
Smith 2010 (357)	Derivation	164993 ED patients with ischaemic CVA	Hospital mortality	AUROC 0.72
	Internal validation	109995 ED patients with ischaemic CVA		AUROC 0.72
Get with the Guidelines Stroke Score with NIHSS				

Smith 2010 (357)	Derivation	164993 ED patients with ischaemic CVA	Hospital mortality	AUROC 0.84
	Internal validation	109995 ED patients with ischaemic CVA		AUROC 0.85
NIHSS				
Fonarow 2012 (358)	External validation	33102 patients hospitalised with ischaemic CVA	30-day mortality	AUROC 0.82 (0.81-0.83)
Poison severity score				
Davies 2008 (356)	External validation Secondary analysis of RCT	990 patients hospitalised with organophosphate poisoning	Hospital mortality	AUROC 0.81 (0.77-0.85)
SOFA				
Dutta 2008 (359)	External validation	23 patients hospitalised with myxoedema coma	Hospital mortality	AUROC 0.629 (0.38-0.878)

CVA: cerebrovascular accident; LP: lumbar puncture

Appendix 3 Variables in existing scoring systems

Age

Single cut-point		
Mannheim	Peritonitis	50
PSI	Pneumonia	50
Modified Glasgow	Pancreatitis	55
Ranson	Pancreatitis	55
ABCD	CVA/TIA	60
ABCD2	CVA/TIA	60
Altona	Peritonitis	60
BISAP	Asthma/COPD	60
BALI	Pancreatitis	65
BAP-65	COPD	65
Bazzino	ACS	65
CRB-65	Pneumonia	65
CURB-65	Pneumonia	65
Glasgow (Imrie) pancreatitis	Pancreatitis	65
MEDS	Sepsis	65
Modified TIMI	ACS	65
OESIL	Syncope	65
PMEWS	Unselected	65
TIMI	ACS	65
Sanchis	ACS	67
Elbaz	Hypothermia	70
Peritonitis severity score	Peritonitis	70
Ranson (Biliary)	Pancreatitis	70
Bordley		75
Hardman	AAA	76
Modified Hardman	AAA	76
Mayo	ACS	80
REA-ICU	Pneumonia	80
SCAP	Pneumonia	80
Multiple cut-points		
GRACE	ACS	30, 40, 50, 60, 70, 80, 90
Hasdai	ACS	30, 40, 50, 60, 70, 80, 90
Weisfelt	Meningitis	30, 40, 50, 60, 70, 80
SAPS II	Unselected	40, 60, 70, 75, 80
IHDI	ACS	40, 70
REMS	Unselected	45, 55, 65, 75
SAPS	Unselected	45, 55, 65, 75
Kellett	Unselected	50(m)/55(w), 75
PURSUIT	ACS	50, 60, 70, 80
Freedom-from-event score	ACS	40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90
POSSUM	Surgical	60, 70

PREDICT	ACS	60, 70
RAAA-POSSUM	AAA	60, 70
V-POSSUM	AAA	60, 70
CR POSSUM	Surgical	60, 70, 80
Rockall	GI bleed	60, 80
Chang	ACS	65, 75
PAMI	ACS	65, 75
TIMI STEMI	ACS	65, 75
Continuous variable		
Coronary prognostic index (Norris)	ACS	
EFFECT	Heart failure	
EMMACE	ACS	
Glasgow aneurysm score	AAA	
MINAP	ACS	
MPM II	Unselected	
Normand	ACS	
PESI	PE	
Selker	ACS	Truncated at 50 and 80
ADHERE logistic regression	ACS	x0.0288
Non-linear correlation		
TIMI risk index	ACS	(Age/10)sq

AVPU consciousness level

Single cut-point		
National asthma guidelines	Asthma/COPD	Not A
Kellett	Unselected	"Coma"
Worthing	Unselected	Not A
Multiple cut-points		
Pitt Bacteremia score	Sepsis	AVPU
MEWS	Unselected	AVPU
PMEWS	Unselected	AVPU
SEWS	Unselected	AVPU

Diastolic blood pressure

Single cut-point		
ABCD	CVA/TIA	90
ABCD2	CVA/TIA	90
CRB-65	Pneumonia	60
CURB-65	Pneumonia	60
American Thoracic Society 2001	Pneumonia	60
British Thoracic Society	Pneumonia	60
CORB	Pneumonia	60
CRB	Pneumonia	60
CURB	Pneumonia	60
Modified BTS	Pneumonia	60
Multiple cut-points		
Hasdai	ACS	20mmHg intervals 40-200
Non-linear		
Pitt Bacteremia score	Sepsis	Drop >20mmHg

Glasgow Coma Scale

Single cut-point		
Edinburgh aneurysm score	AAA	14
BISAP	Pancreatitis	15
MPM II	Unselected	5
Multiple cut-points		
V-POSSUM	AAA	9, 12, 15
Poison severity score	Poisoning	9, 14
POSSUM	Surgical	9, 12, 15
APACHE-O	Unselected	8, 11, 14
PEDS	Unselected	9, 13
RAPS	Unselected	5, 8, 11, 14
REMS	Unselected	5, 8, 11, 14
RTS	Unselected	3, 4, 6, 9, 13
SAPS	Unselected	4, 7, 10, 13
SAPS II	Unselected	6, 9, 11, 14
SOFA	Unselected	6, 10, 13, 15
Continuous variable		
Weisfelt	Meningitis	
APACHE II Acute physiology	Unselected	
APACHE II	Unselected	

Mean arterial pressure

Single cut-point		
Elbaz	Hypothermia	90
SOFA	Unselected	70
Multiple cut-points		
CAPS	Asthma/COPD	40, 50, 60, 70, 90, 100
APACHE II Acute physiology	Unselected	49, 69, 109, 129, 159
APACHE II	Unselected	49, 69, 109, 129, 159
REMS	Unselected	50, 60, 110, 130, 160
RAPS	Unselected	50, 60, 110, 130, 160
Non-linear correlation		
Normand	ACS	Log
MODS	Unselected	1/MAP

Oxygen saturation

Single cut-off		
Pulmonary embolism severity index	PE	90 (No FiO2 analysis)
REA-ICU	Pneumonia	90 (No FiO2 analysis)
CORB	Pneumonia	90 (No FiO2 analysis)
MEDS	Sepsis	90 (or requiring supplemental O2)
HOTEL	Unselected	90 (No FiO2 analysis)
Multiple cut-offs		
National asthma guidelines	Asthma	92, 95 (No FiO2 analysis)
SMART-COP	Pneumonia	93 (<50y), 90 (>50y) (Could use PaO2/FiO2)
Kellett	Unselected	90, 95 (No FiO2 analysis)
PMEWS	Unselected	90, 94, 96 (No FiO2analysis)
REMS	Unselected	75, 85, 89 (No FiO2 analysis)
SEWS	Unselected	85, 90, 93 (No FiO2 analysis)
Worthing	Unselected	96, 94, 92 (Breathing room air)

Pulse rate

Single cut-point		
BISAP	Pancreatitis	90
SIRS	Unselected	90
Normand	ACS	100
Rockall	GI bleed	100
Blatchford	GI bleed	100
Strate	GI bleed	100
TIMI STEMI	ACS	100
Worthing	Unselected	100
Modified Blatchford	GI bleed	100
PAMI	ACS	100
BAP-65	COPD	110
Pulmonary embolism severity index	PE	110
Pulmonary edema prognostic score	LVF	115
Weisfelt	Meningitis	120
PSI	Pneumonia	124
SMART-COP	Pneumonia	125
SMRT-CO	Pneumonia	125
REA-ICU	Pneumonia	125
Multiple cut-points		
Chang	ACS	63, 85
Freedom-from-event score	ACS	10bpm intervals 30-150
GRACE	ACS	50, 70, 90, 110, 150, 200
PREDICT	ACS	100, 120
Hasdai	ACS	20bpm intervals 40-260
CAPS	Asthma/COPD	80, 110, 130, 150, 170
National asthma guidelines	Asthma	100, 120
Poison severity score	Poisoning	40, 50, 140, 180
CR POSSUM	Surgical	40, 100, 120
POSSUM	Surgical	40, 50, 80, 100, 120
APACHE II Acute physiology	Unselected	39, 54, 69, 109, 139, 179
APACHE II	Unselected	39, 54, 69, 109, 139, 179
APACHE-O	Unselected	39, 54, 69, 109, 139, 179
LODS	Unselected	30, 140
MEWS	Unselected	40, 50, 100, 110, 130
PMEWS	Unselected	40, 50, 100, 110, 130
REMS	Unselected	40, 55, 70, 110, 140, 180
RAPS	Unselected	40, 55, 70, 110, 140, 180
SAPS	Unselected	40, 55, 70, 110, 140, 180
SAPS II	Unselected	40, 70, 120, 160

SEWS	Unselected	30, 40, 50, 100, 110, 130
Continuous variable		
TIMI risk index	ACS	
EMMACE	ACS	
MINAP	ACS	
ADHERE logistic regression	LVF	
MODS	Unselected	
MPM II	Unselected	
Non-linear correlation		
Glasgow aneurysm score	AAA	"Shock"
Selker	ACS	Dependent on SBP
Kellett	Unselected	>SBP

Respiratory rate

Single cut-point		
BISAP	Pancreatitis	20
SIRS	Unselected	20
MEDS	Sepsis	20
PSI	Pneumonia	30
REA-ICU	Pneumonia	30
CRB-65	Pneumonia	30
CURB-65	Pneumonia	30
CORB	Pneumonia	30
CRB	Pneumonia	30
CURB	Pneumonia	30
Le Conte	LVF	30
Brigham	LVF	30
British Thoracic Society	Pneumonia	30
American Thoracic Society 2007	Pneumonia	30
Modified BTS	Pneumonia	30
SCAP	Pneumonia	30
Multiple cut-points		
SMART-COP	Pneumonia	25 (<50), 30 (>50y)
SMRT-CO	Pneumonia	25 (<50), 30 (>50y)
APACHE II Acute physiology	Unselected	5, 9, 11, 24, 34, 49
APACHE II	Unselected	5, 9, 11, 24, 34, 49
APACHE-O	Unselected	5, 9, 11, 24, 34, 49
MEWS	Unselected	9, 15, 21, 30
PMEWS	Unselected	8, 18, 25, 30
RAPS	Unselected	6, 10, 12, 25, 35, 50
REMS	Unselected	6, 10, 12, 25, 35, 50
RTS	Unselected	0, 1, 6, 10, 30
SAPS	Unselected	6, 10, 12, 25, 35, 50
SEWS	Unselected	8, 20, 30, 35
Worthing	Unselected	20, 22
Continuous variable		
Normand	ACS	Truncated <12
EFFECT	LVF	

Systolic blood pressure

Single cut-point		
CRB-65	Pneumonia	90
CURB-65	Pneumonia	90
CORB	Pneumonia	90
CRB	Pneumonia	90
CURB	Pneumonia	90
Edinburgh aneurysm score	AAA	90
Brigham	LVF	90
MEDS	Sepsis	90 after fluids
PSI	Pneumonia	90
San Francisco	Syncope	90
SMART-COP	Pneumonia	90
SMRT-CO	Pneumonia	90
Peritonitis severity score	Peritonitis	90
SCAP	Pneumonia	90
Modified Hardman	AAA	100
BLEED	GI bleed	100
Bordley	GI bleed	100
Rockall	GI bleed	100
TIMI STEMI	ACS	100
Pulmonary embolism severity index	PE	100
Worthing	Unselected	100
HOTEL	Unselected	100
Goldman	ACS	110
Strate	GI bleed	115
Pulmonary edema prognostic score	LVF	130
ABCD	TIA	140
ABCD2	TIA	140
Mayo	ACS	140
Multiple cut-points		
V-POSSUM	AAA	90, 100, 110, 130, 170
Chang	ACS	120, 132
Freedom-from-event score	ACS	100, 120, 140, 160, 180, 200
GRACE	ACS	80, 100, 120, 140, 160, 200
Hasdai	ACS	20mmHg intervals 80-280
PREDICT	ACS	60, 100
Blatchford	GI bleed	90, 100
Modified Blatchford	GI bleed	90, 100, 110
EFFECT	LVF	90, 100, 120, 140, 160,

		180
Aronin	Meningitis	90 or 40 drop
Poison severity score	Poisoning	80, 100
American Thoracic Society 2001	Pneumonia	90, shock
Modified ATS	Pneumonia	90, shock
American Thoracic Society 2007	Pneumonia	"hypotension", "shock"
Pitt Bacteremia score	Sepsis	90 or drop 30
CR POSSUM	Surgical	90, 100, 170
POSSUM	Surgical	90, 100, 110, 130, 170
Kellett	Unselected	70, 80, 100
LODS	Unselected	40, 70, 90, 240, 270
MEWS	Unselected	70, 80, 100, 200
PEDS	Unselected	100, 140
PMEWS	Unselected	70, 90, 100
RTS	Unselected	0, 1, 50, 76, 90
SAPS	Unselected	55, 80, 150, 190
SAPS II	Unselected	70, 100, 200
SEWS	Unselected	70, 80, 100, 200
Continuous variable		
Coronary prognostic index (Norris)	ACS	Cont
EMMACE	ACS	Cont
MINAP	ACS	Cont
ADHERE logistic regression	LVF	Cont
Non-linear correlation		
Glasgow aneurysm score	AAA	"Shock"
Selker	ACS	Continuous (truncated 175) also squared
TIMI risk index	ACS	1/SBP
Mannheim	Peritonitis	"Shock"
MPM II	Unselected	Continuous and squared

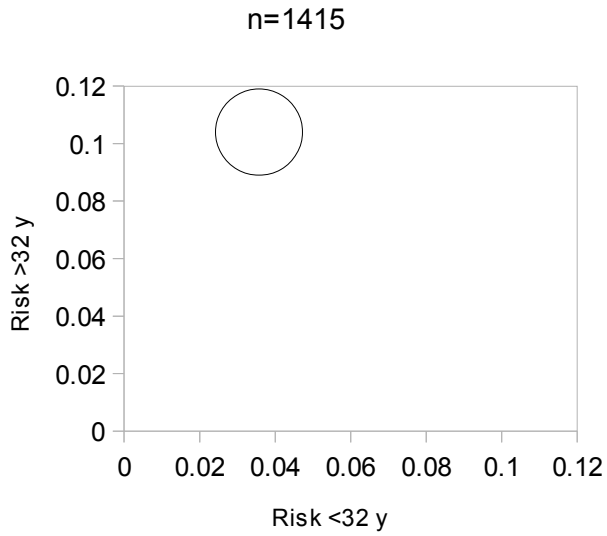
Temperature

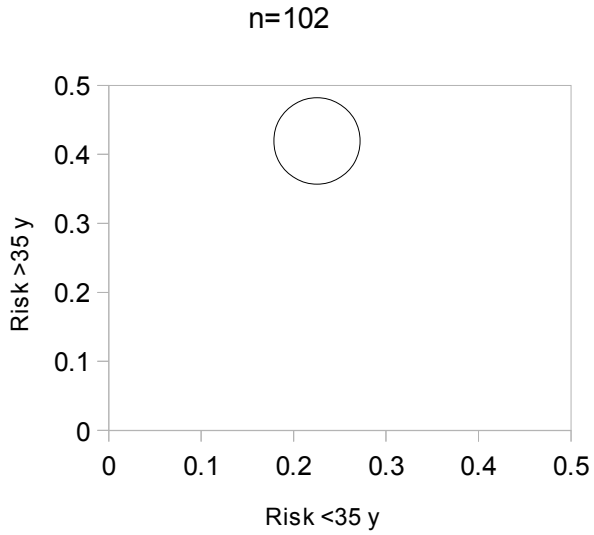
Single cut-point		
HOTEL	Unselected	35
Worthing	Unselected	35.3
American Thoracic Society 2007	Pneumonia	36
SAPS II	Unselected	39
Multiple cut-points		
BISAP	Pancreatitis	36, 38
PSI	Pneumonia	35, 39.9
Pitt Bacteremia score	Sepsis	35, 36, 39, 40
APACHE II Acute physiology	Unselected	30, 32, 34, 36, 38.5, 39, 41
APACHE II	Unselected	30, 32, 34, 36, 38.5, 39, 41
APACHE-O	Unselected	30, 32, 34, 36, 38.5, 39, 41
Kellett	Unselected	35, 39
MEWS	Unselected	35, 38.5
PMEWS	Unselected	35, 36, 38, 39
REMS	Unselected	30, 32, 34, 36, 38.5, 39, 41
SAPS	Unselected	30, 32, 34, 36, 38.5, 39, 41
SEWS	Unselected	34, 35, 36, 38, 39
SIRS	Unselected	36, 38

Appendix 4: L'Abbé plots of risk for individual physiological variables

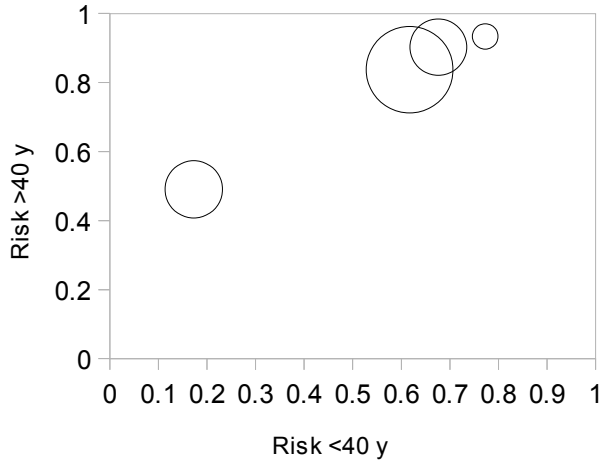
Each circle represents one study.

Age

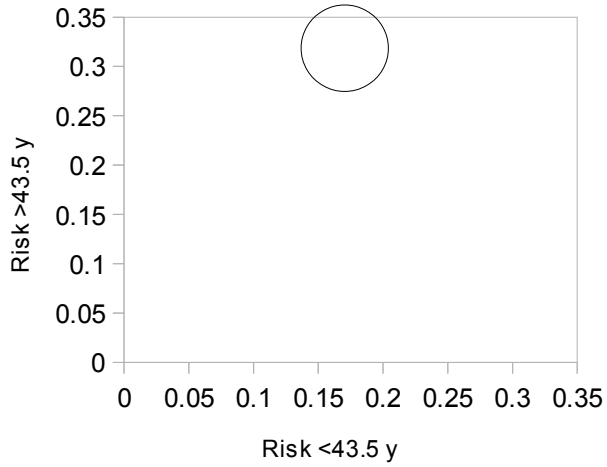




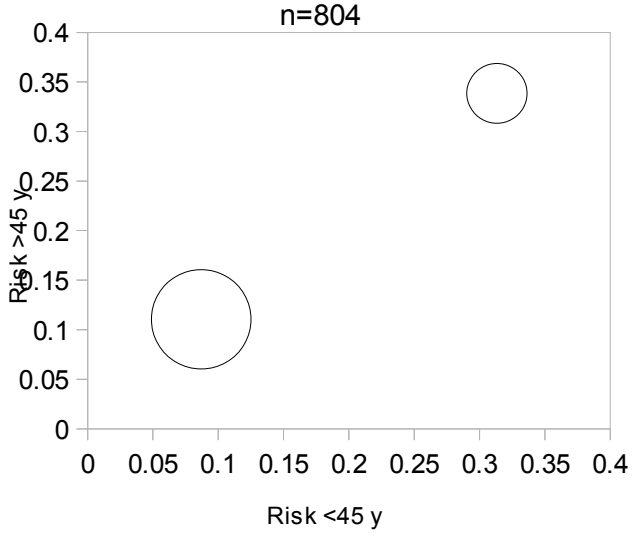
n=826



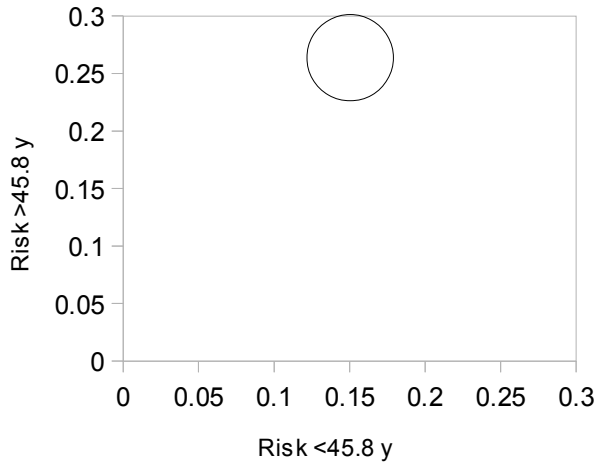
n=270



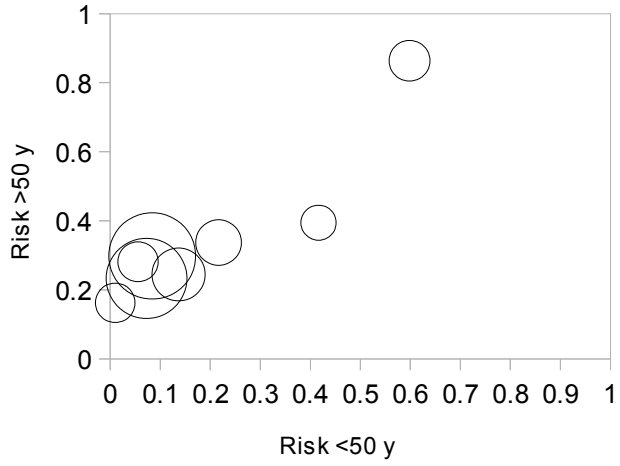
n=804



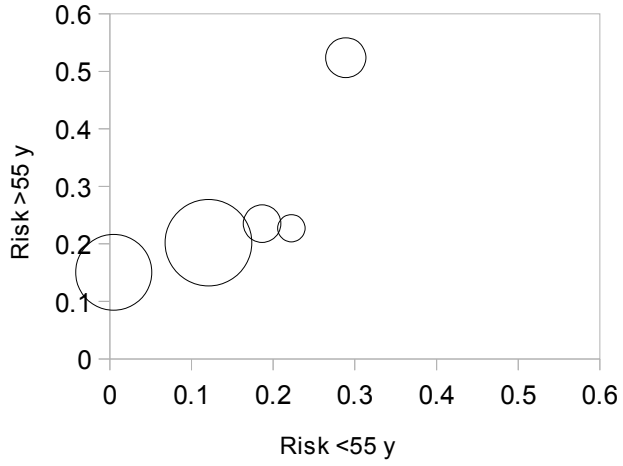
n=1317



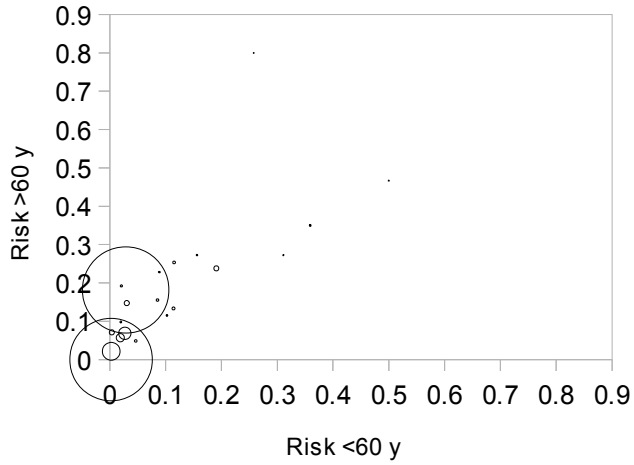
n=2163



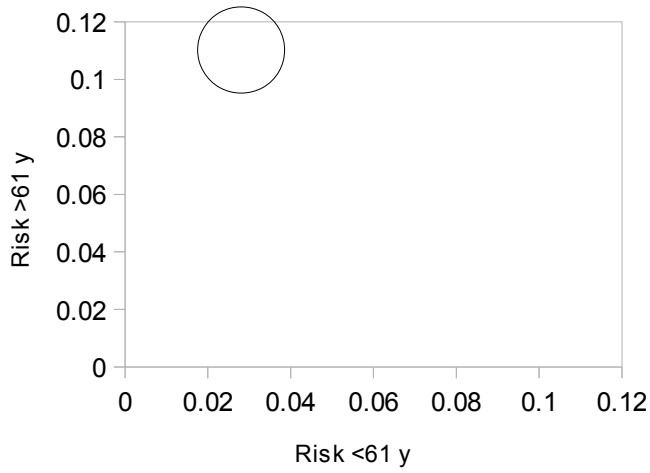
n=920



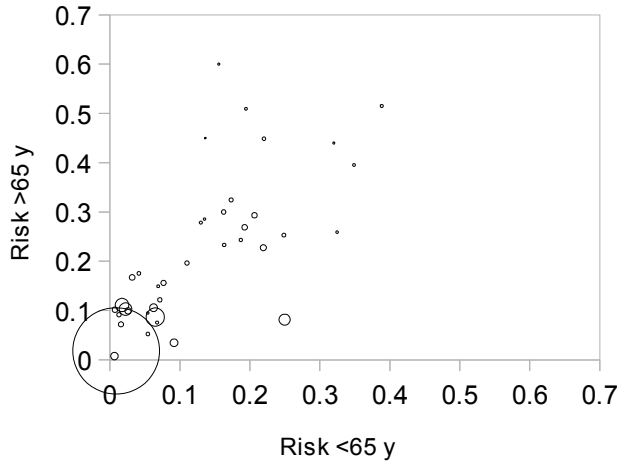
n=522009



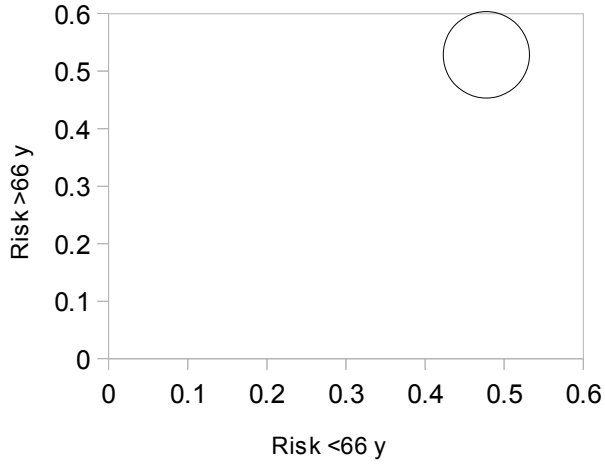
n=1493



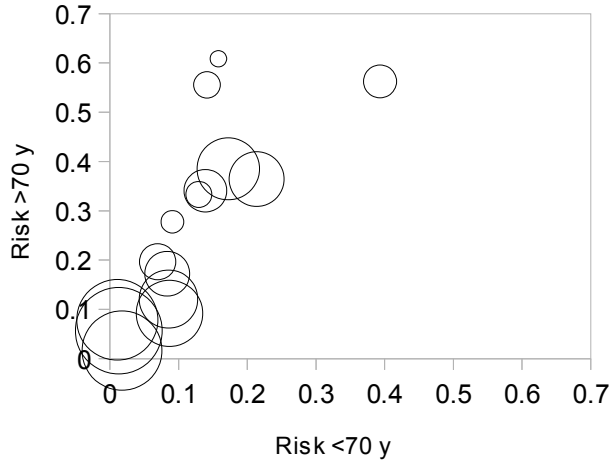
n=151800



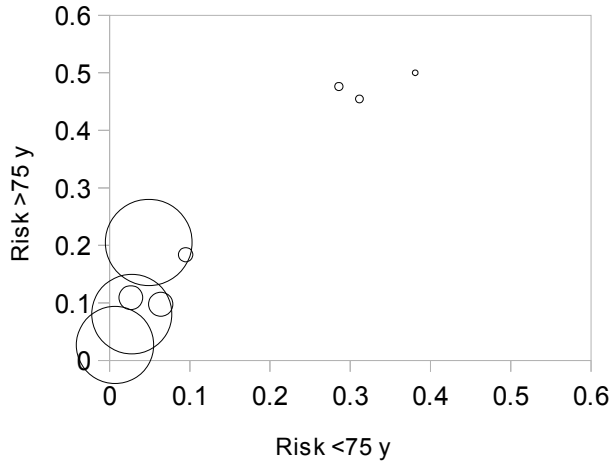
n=252

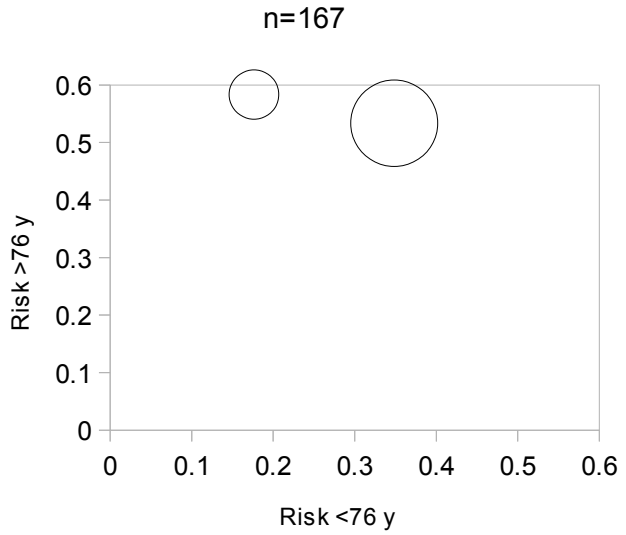


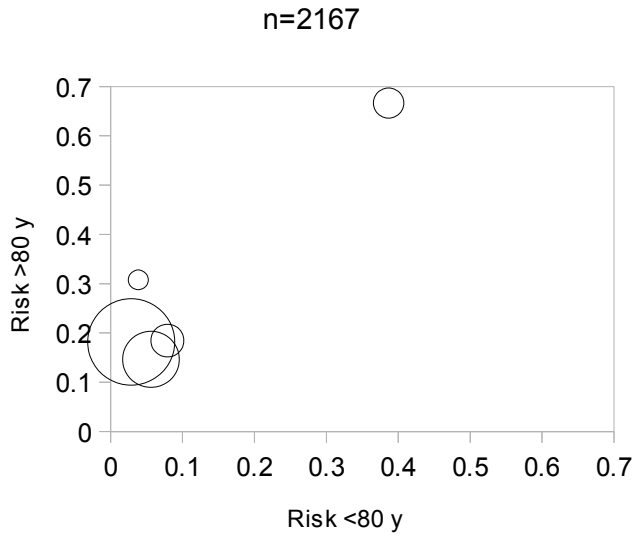
n=6795

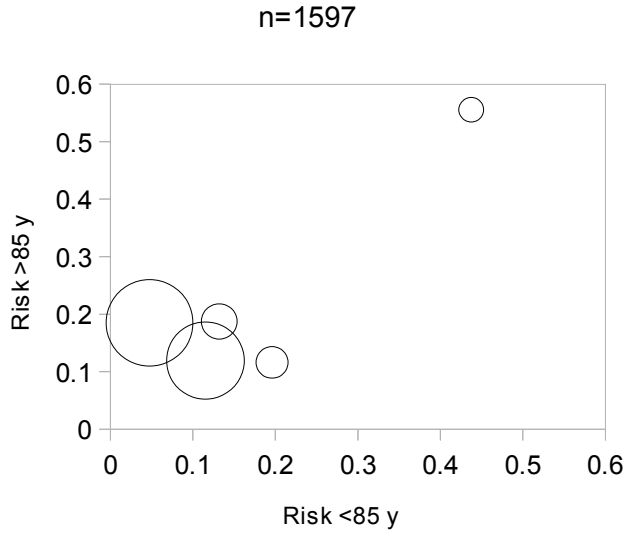


n=36172

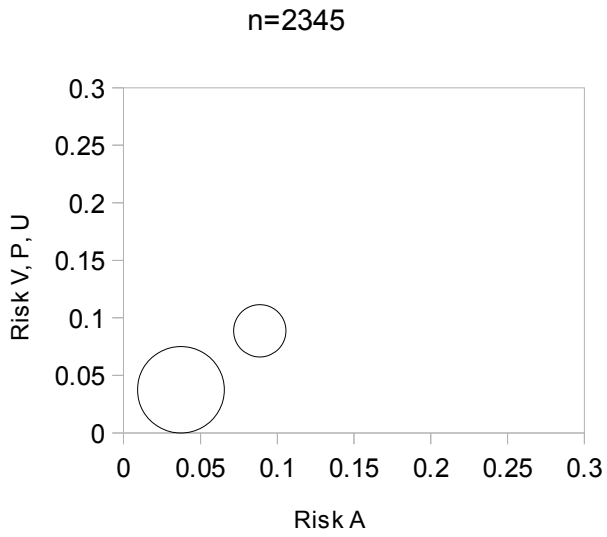


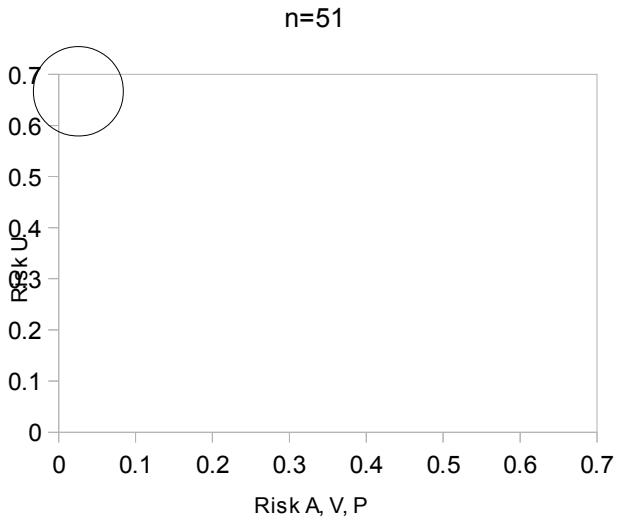

n=167
Risk >76 y
Risk <76 y
0
0.1
0.2
0.3
0.4
0.5
0.6


n=2167
Risk >80 y
Risk <80 y
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7

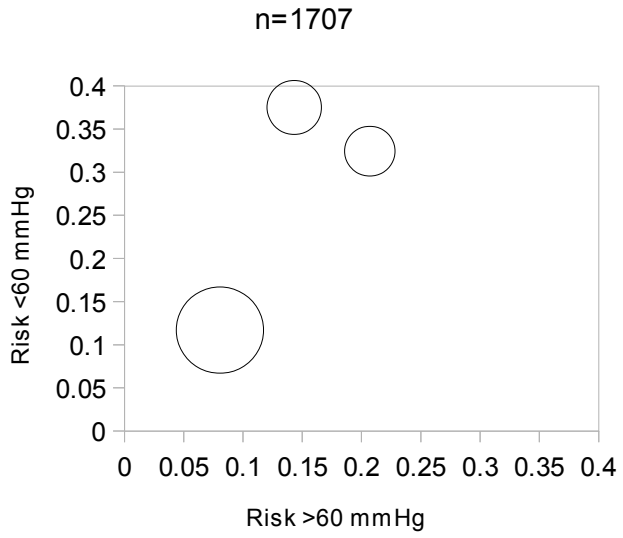

n=1597
Risk >85 y
Risk <85 y
0
0.1
0.2
0.3
0.4
0.5
0.6

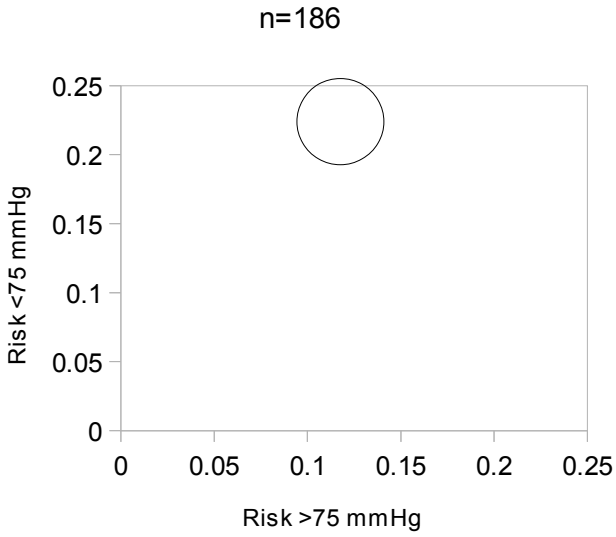
AVPU level of consciousness

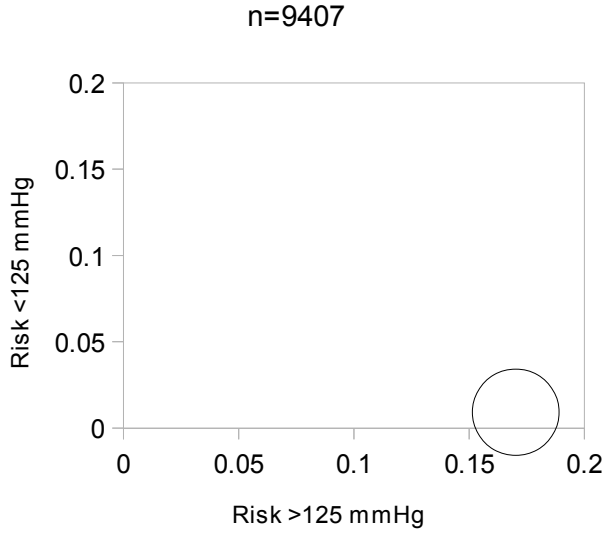




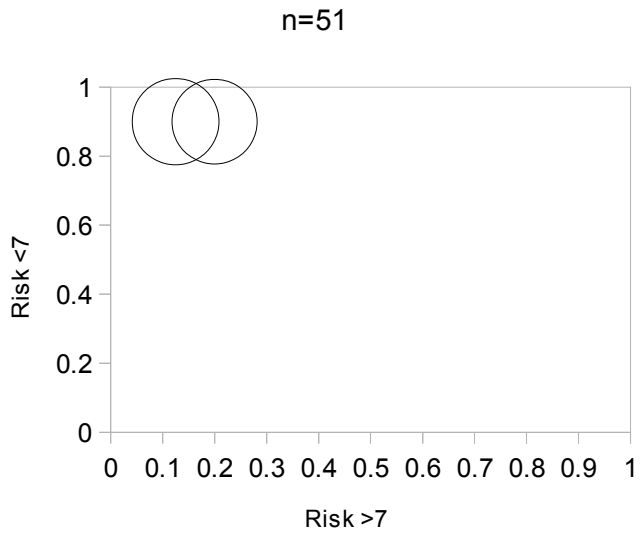
Diastolic Blood Pressure

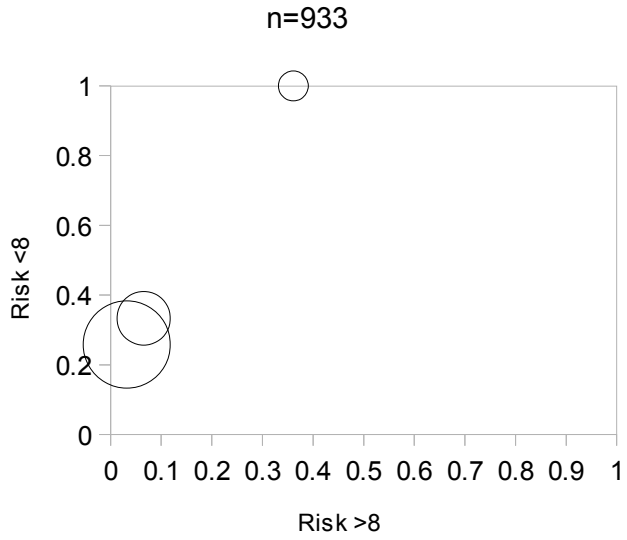


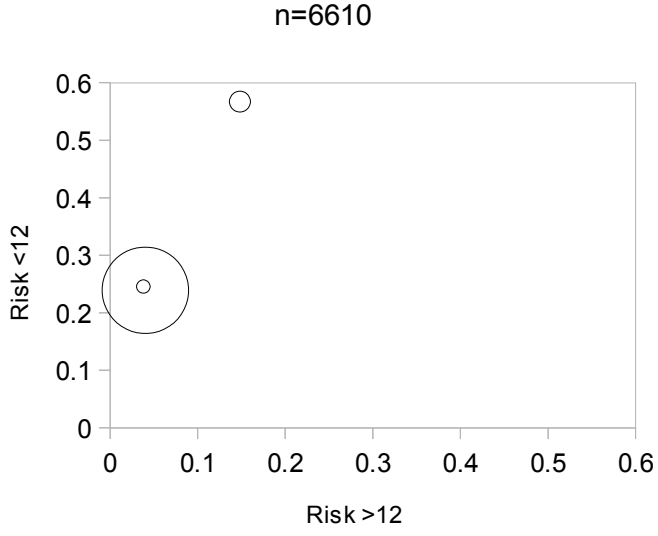


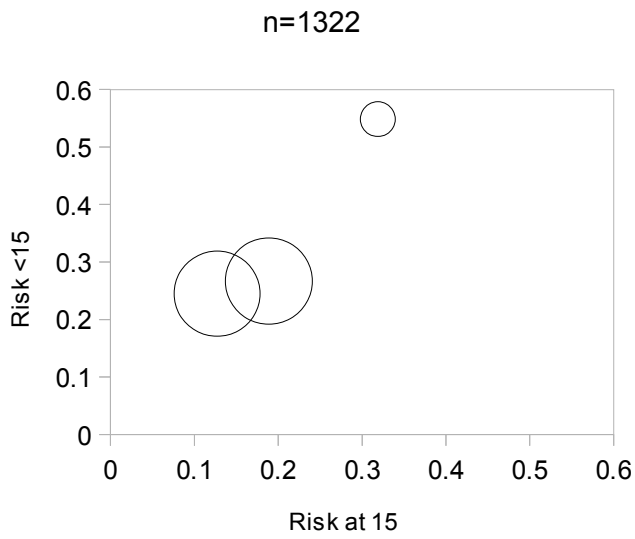


Glasgow Coma Scale

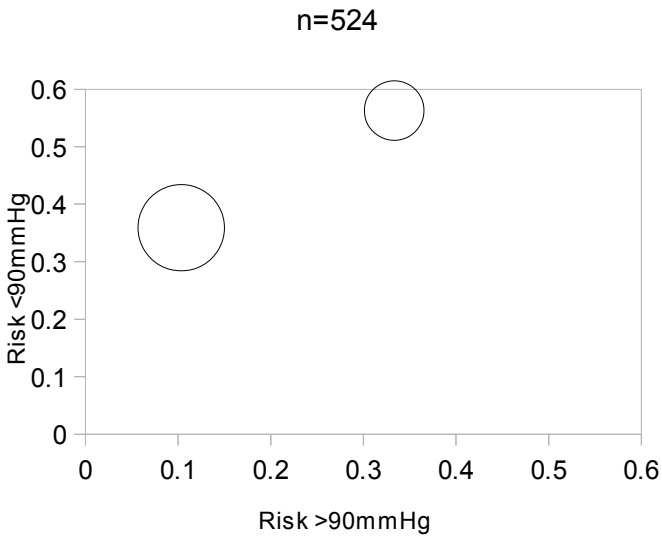




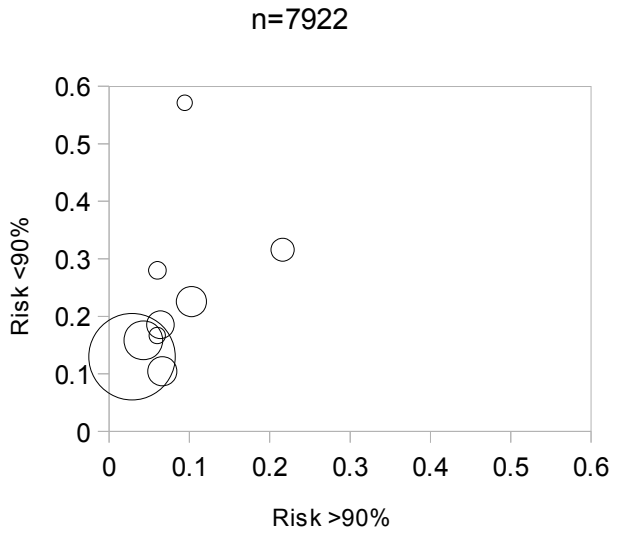


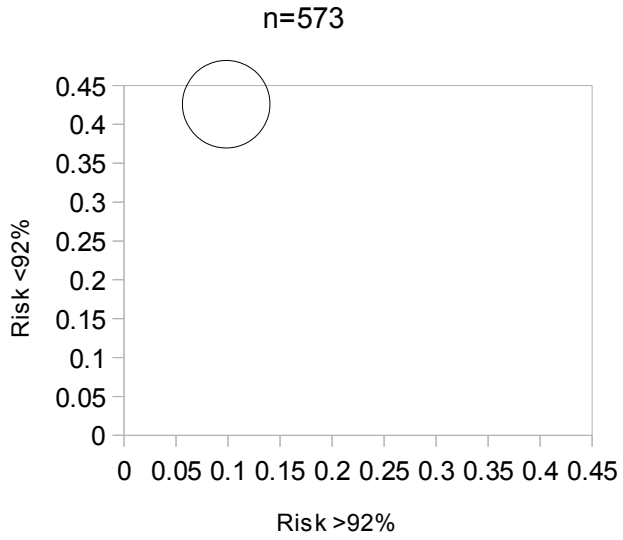

n=1322
0.6
0.5
0.4
0.3
0.2
0.1
0
Risk <15
0
0.1
0.2
0.3
0.4
0.5
0.6
Risk at 15

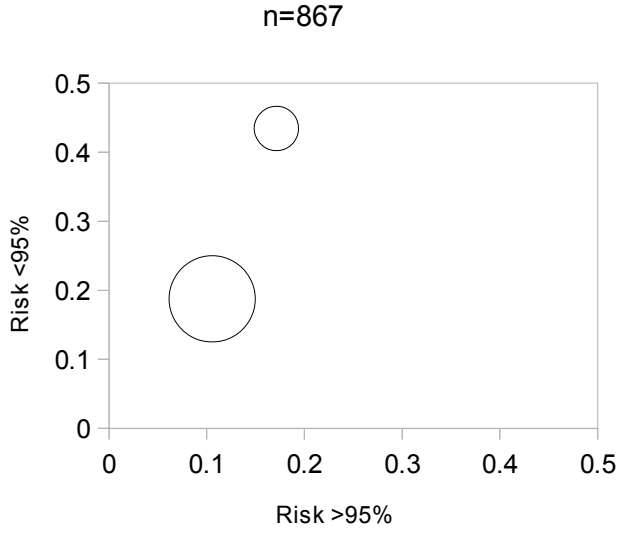
Mean arterial pressure



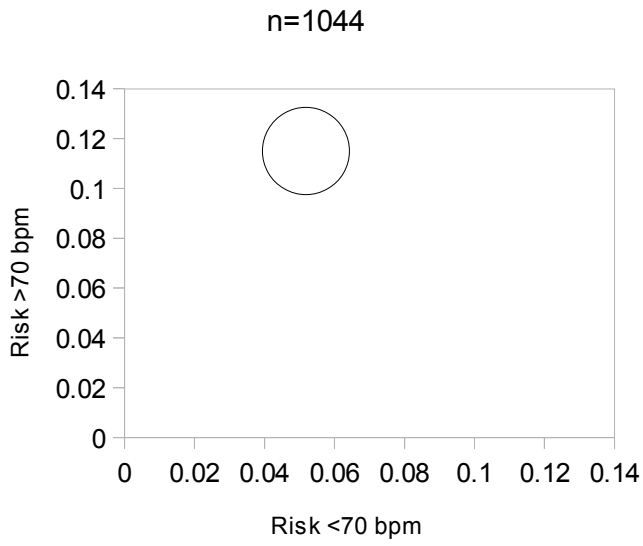
Oxygen saturations (SaO2)

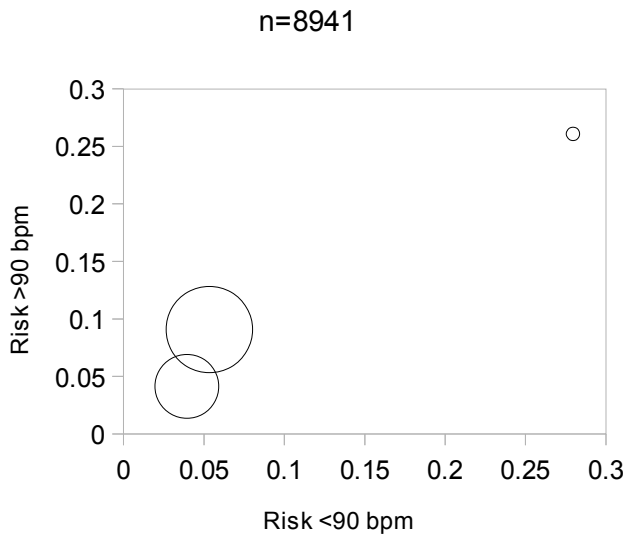


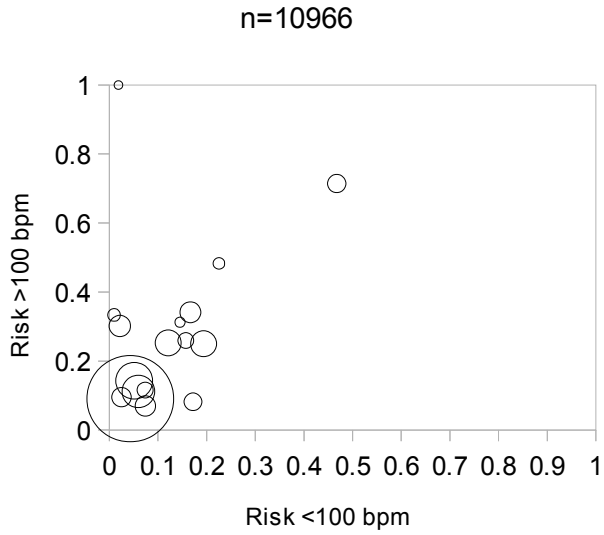




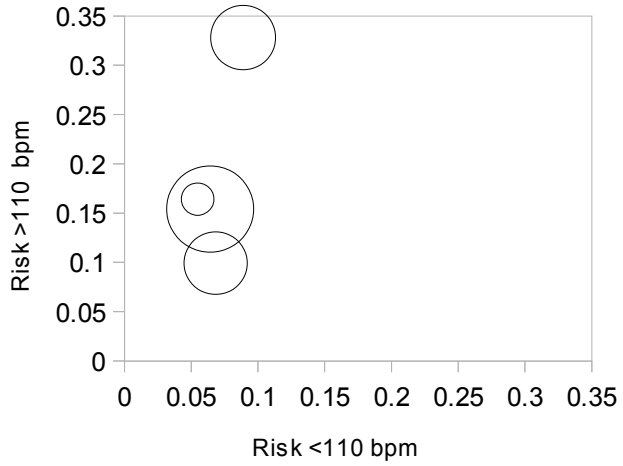
Pulse


n=1044
Risk >70 bpm
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
0 0.02 0.04 0.06 0.08 0.1 0.12 0.14
Risk <70 bpm

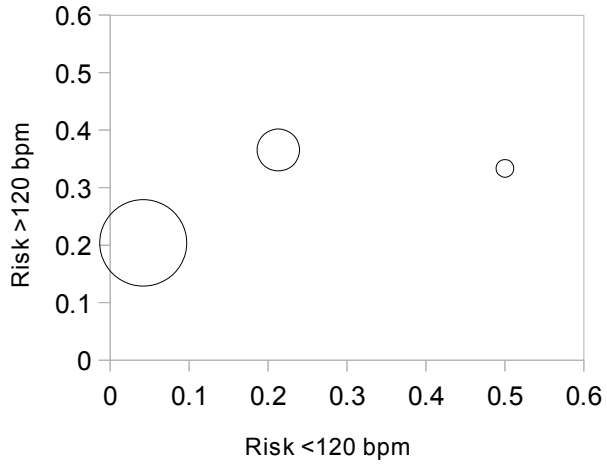

n=8941
Risk >90 bpm
0.3
0.25
0.2
0.15
0.1
0.05
0
0 0.05 0.1 0.15 0.2 0.25 0.3
Risk <90 bpm


n=10966
Risk >100 bpm
1
0.8
0.6
0.4
0.2
0
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
Risk <100 bpm

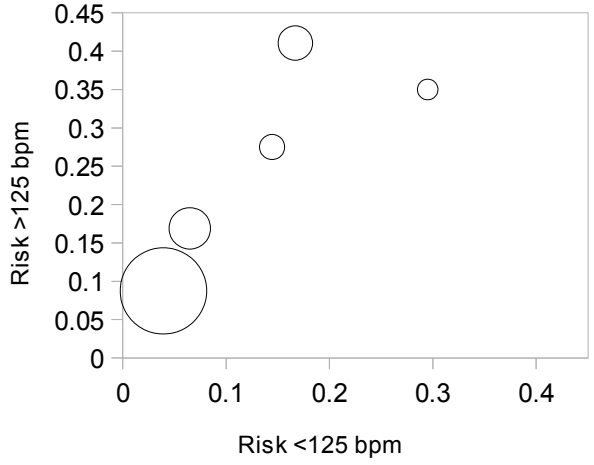
n=2225

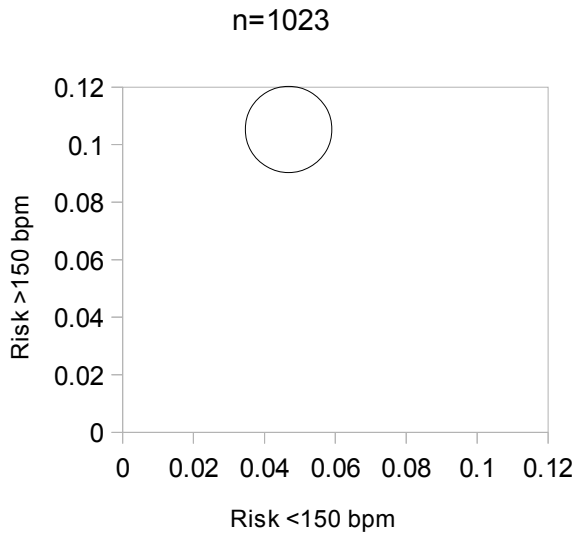


n=1195

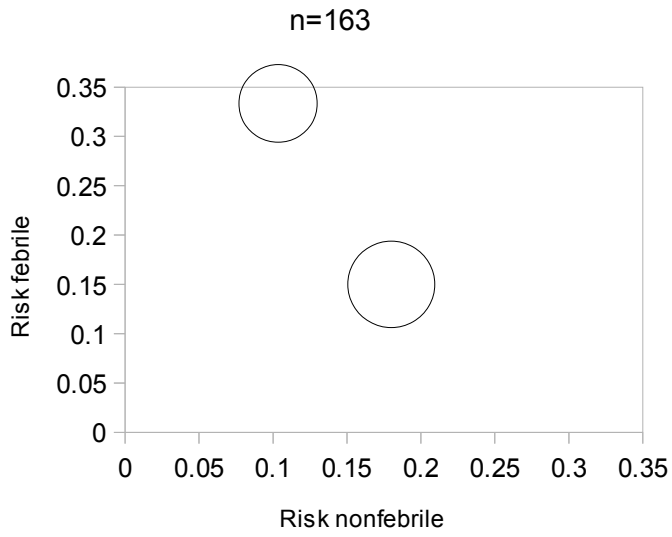


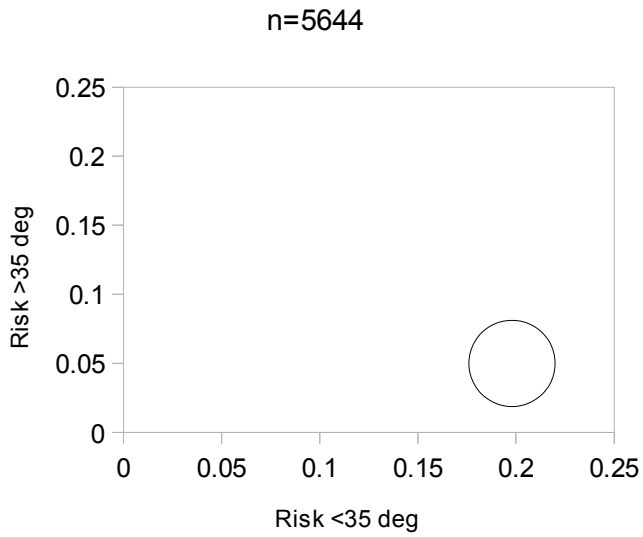
n=7026

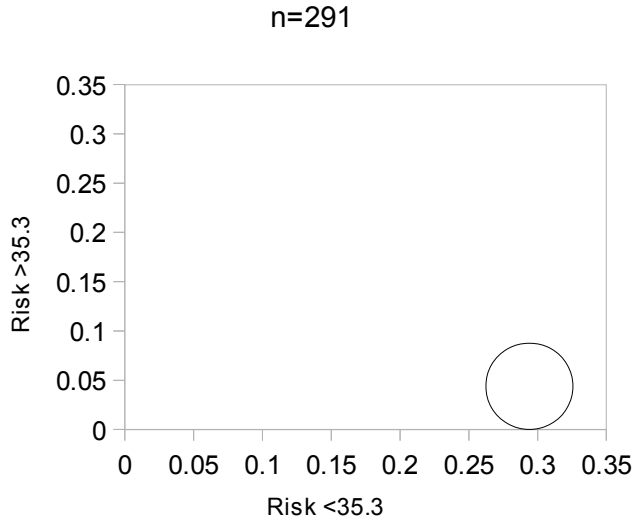



n=1023
0.12
0.1
0.08
0.06
0.04
0.02
0
Risk >150 bpm
0
0.02
0.04
0.06
0.08
0.1
0.12
Risk <150 bpm

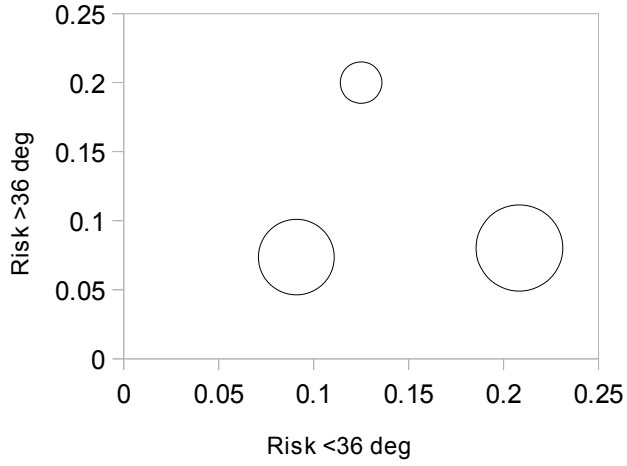
Temperature



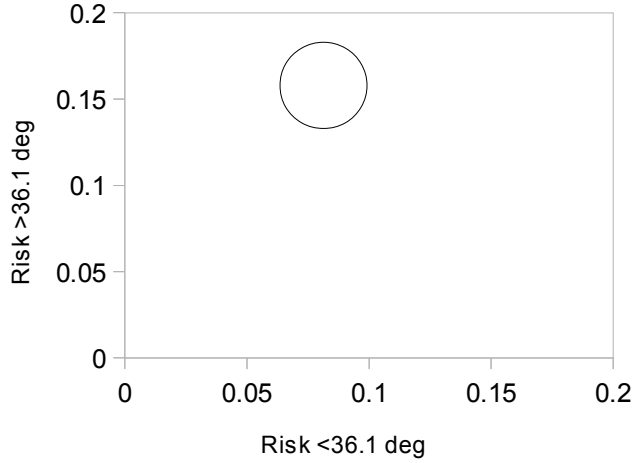




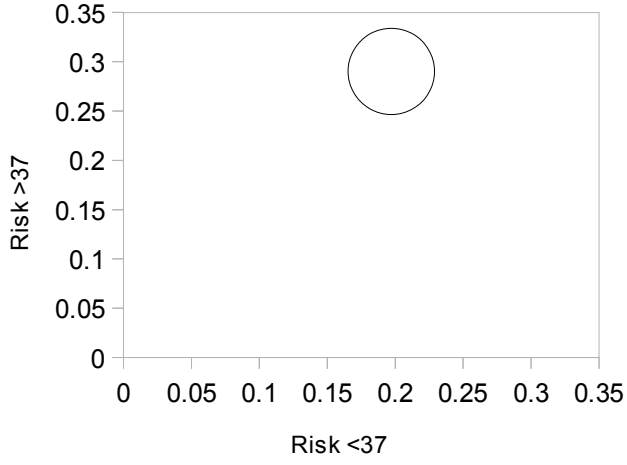
n=1369

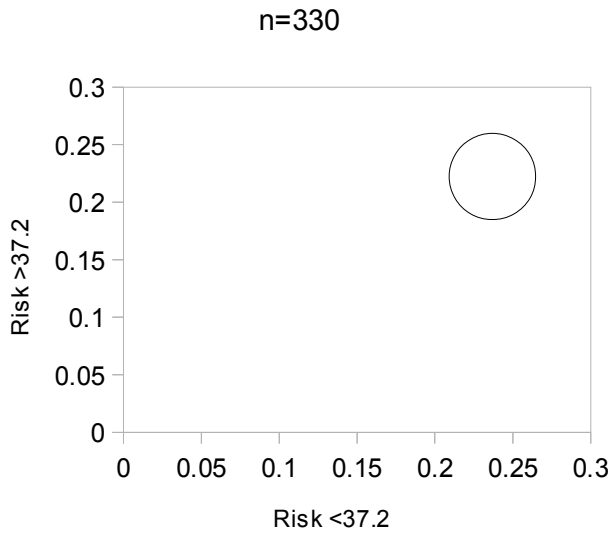


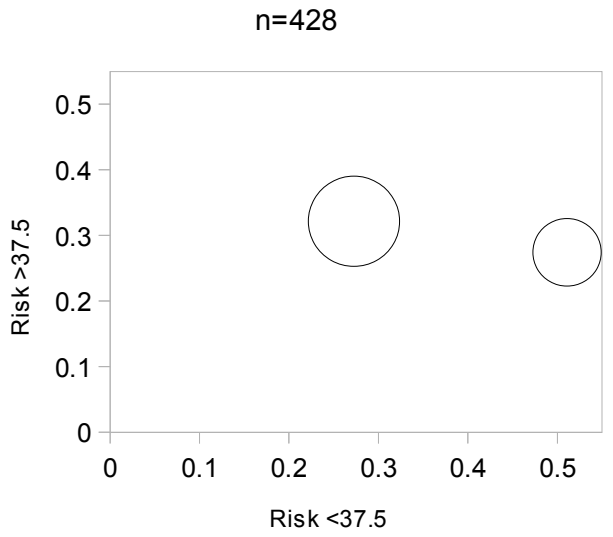
n=998

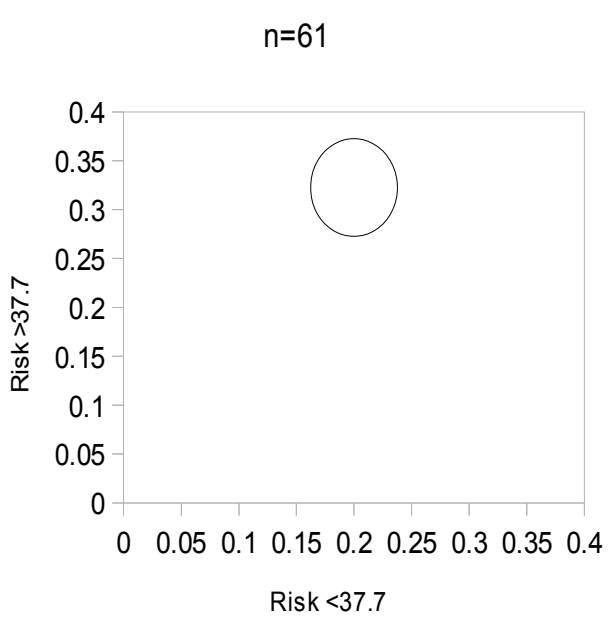


n=619

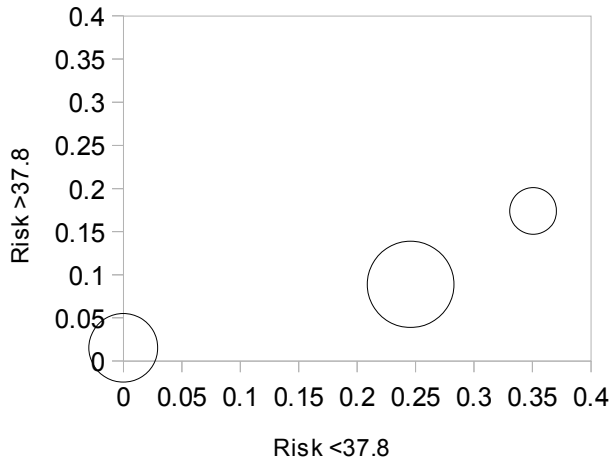



n=330
Risk >37.2
0.3
0.25
0.2
0.15
0.1
0.05
0
0 0.05 0.1 0.15 0.2 0.25 0.3
Risk <37.2

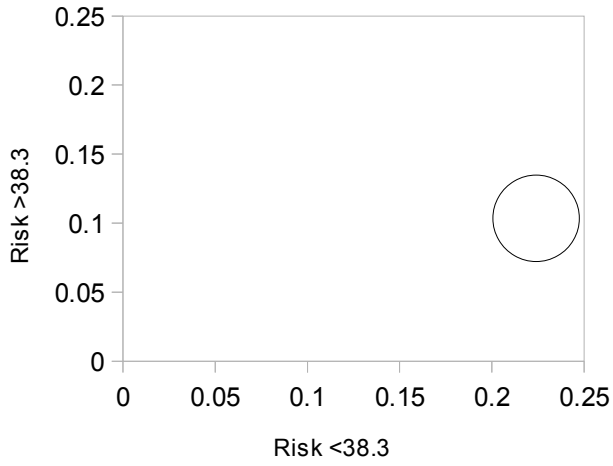

n=428
Risk >37.5
0.5
0.4
0.3
0.2
0.1
0
0 0.1 0.2 0.3 0.4 0.5
Risk <37.5


n=61
Risk >37.7
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0 0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4
Risk <37.7

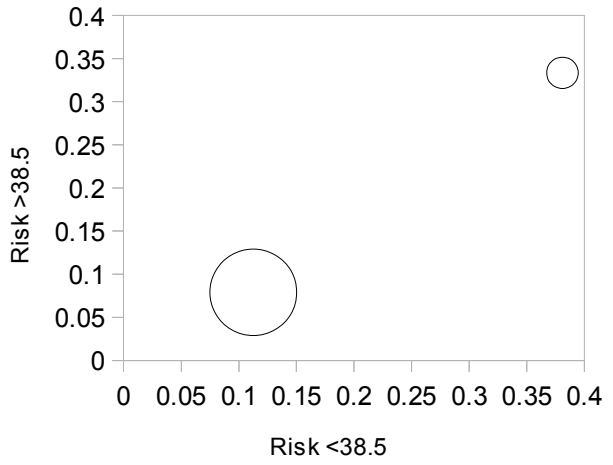
n=660

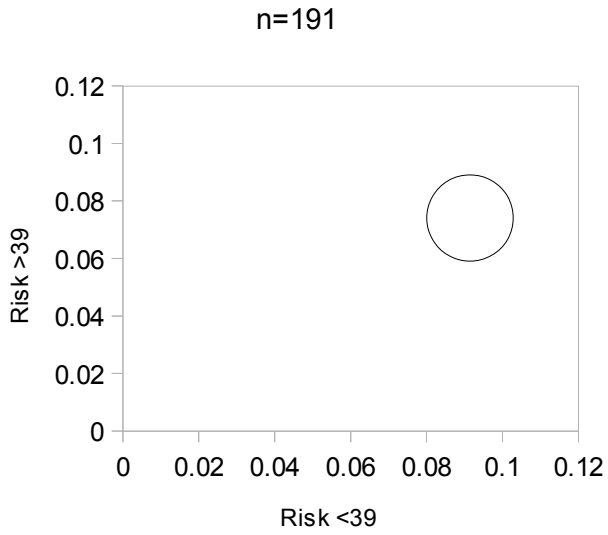


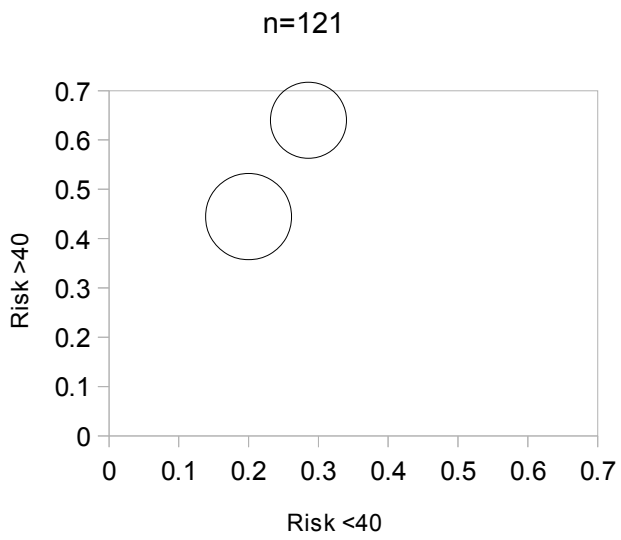
n=395



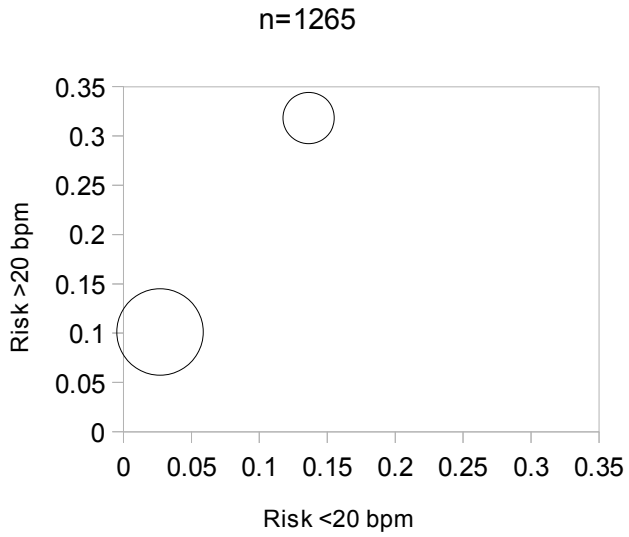
n=257

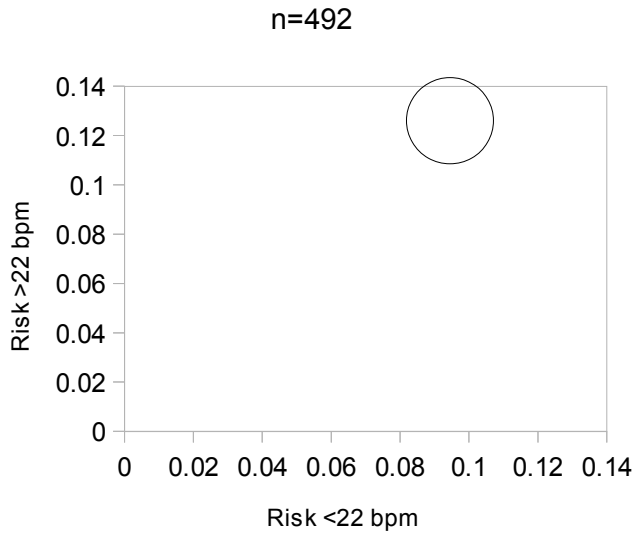


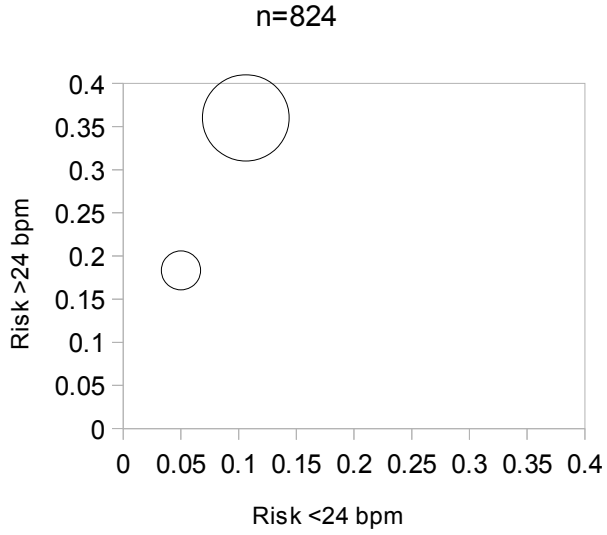

n=191
0.12
0.1
0.08
0.06
0.04
0.02
0
Risk >39
0
0.02
0.04
0.06
0.08
0.1
0.12
Risk <39


n=121
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Risk >40
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Risk <40

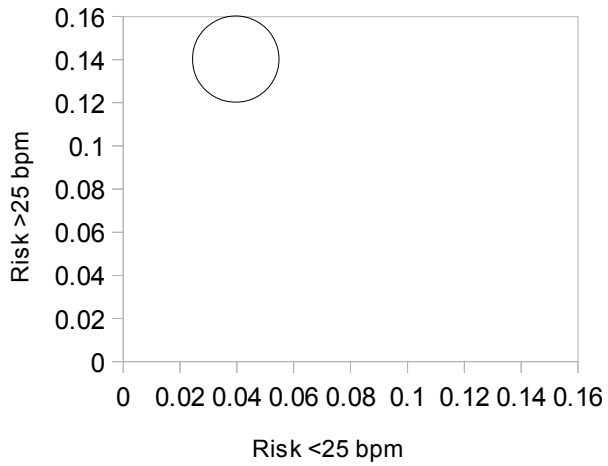
Respiratory rate



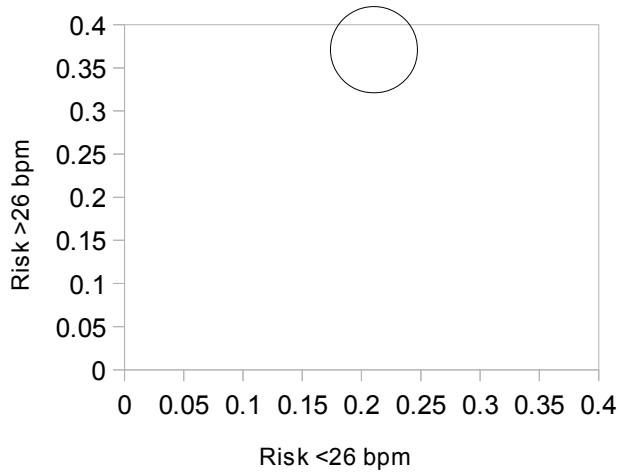




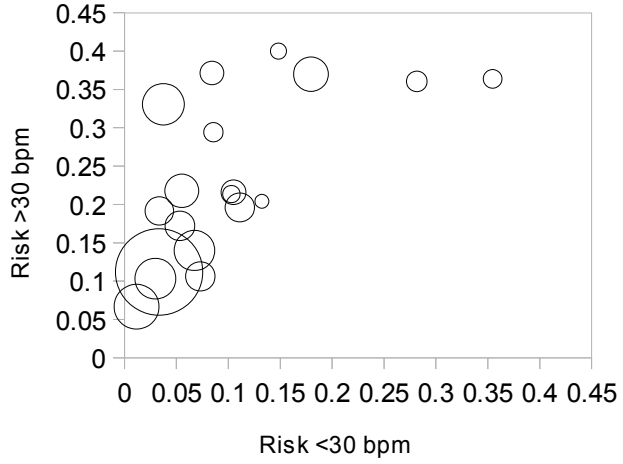
n=5644



n=100

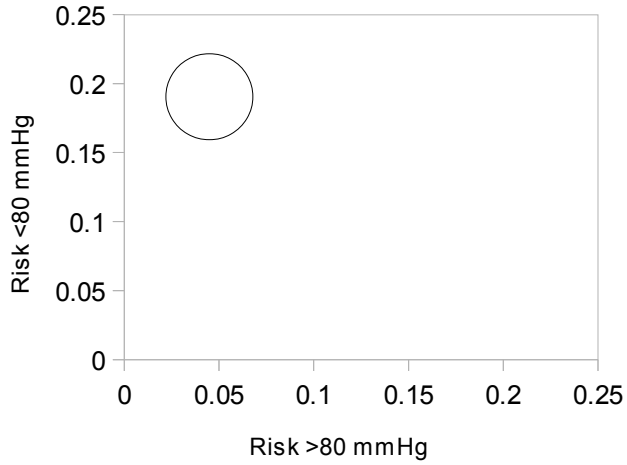


n=14262

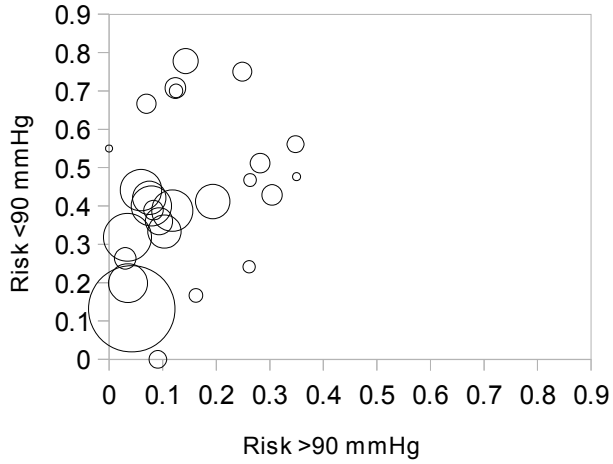


Systolic blood pressure

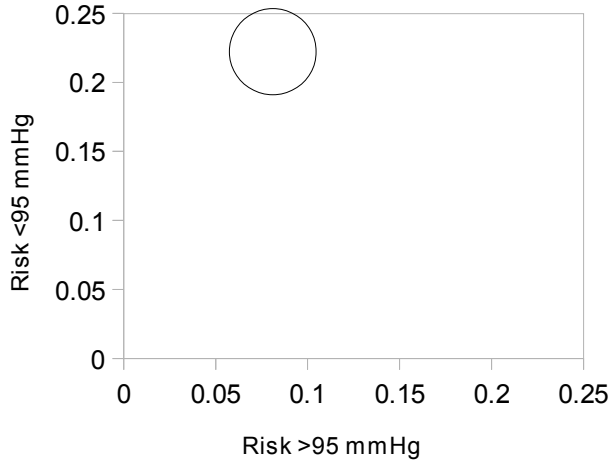
n=1022



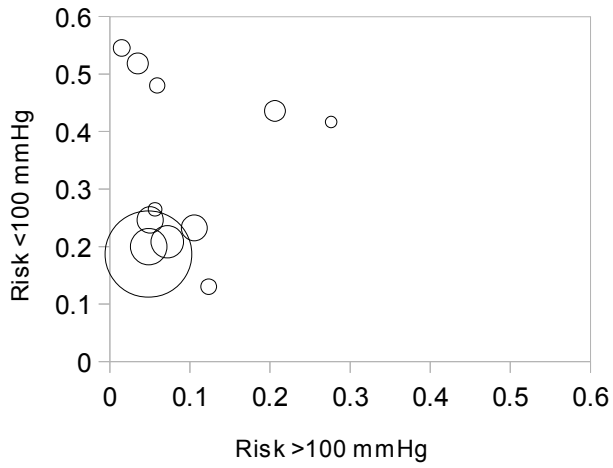
n=15587



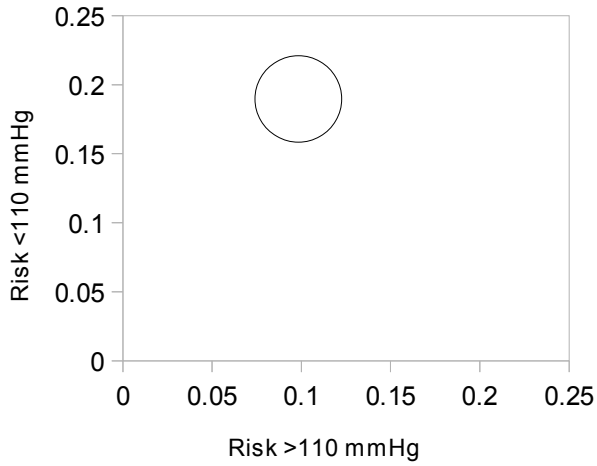
n=778



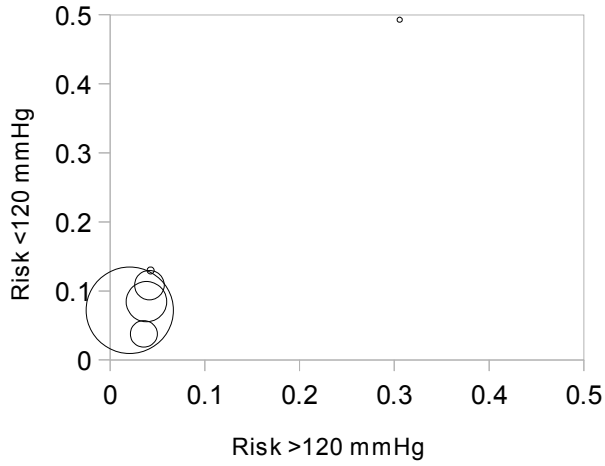
n=10035

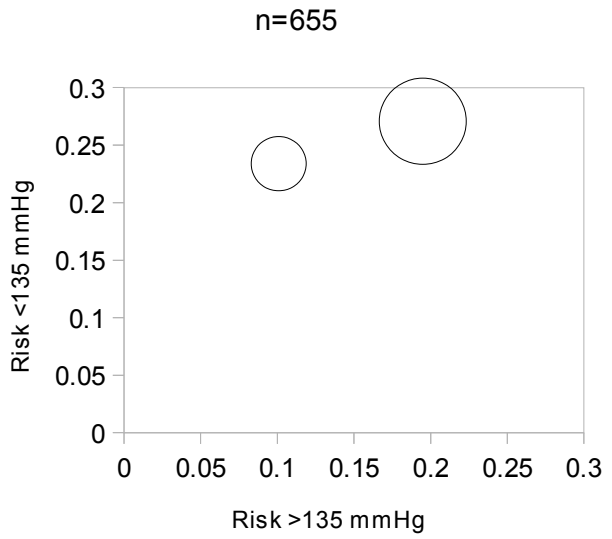


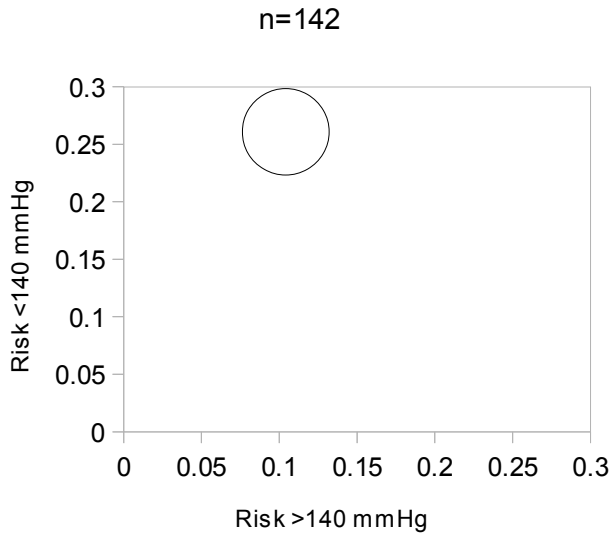
n=414



n=69840



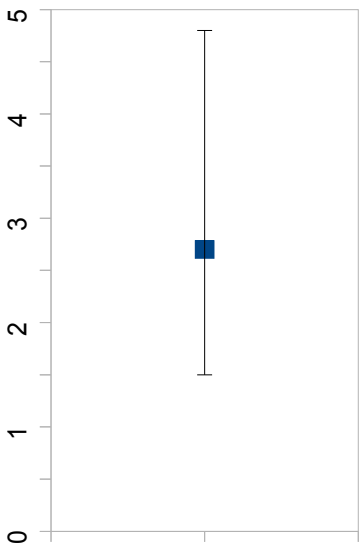




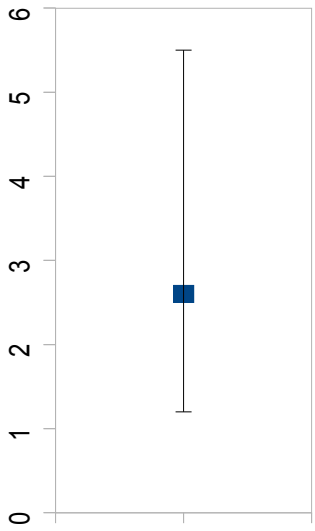
Appendix 5: Odds ratios for individual physiological variables

Age

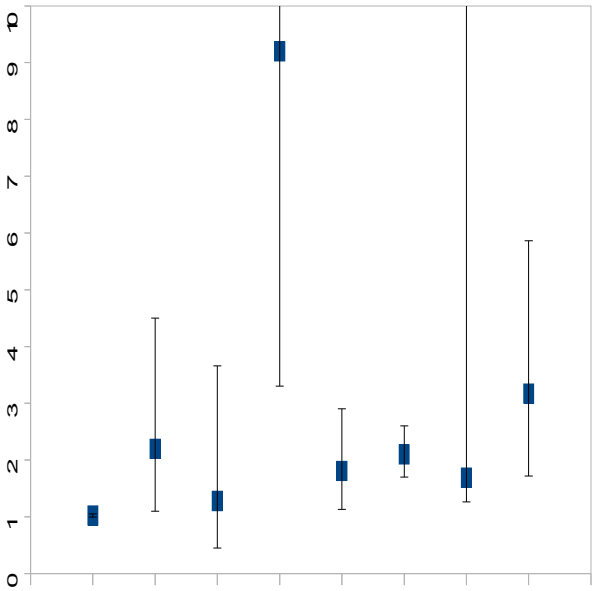
Dichotomised at 48y



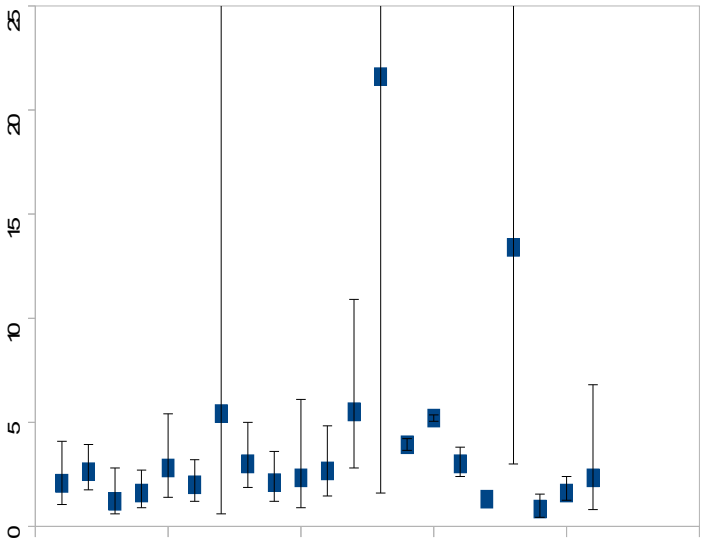
Dichotomised at 50y



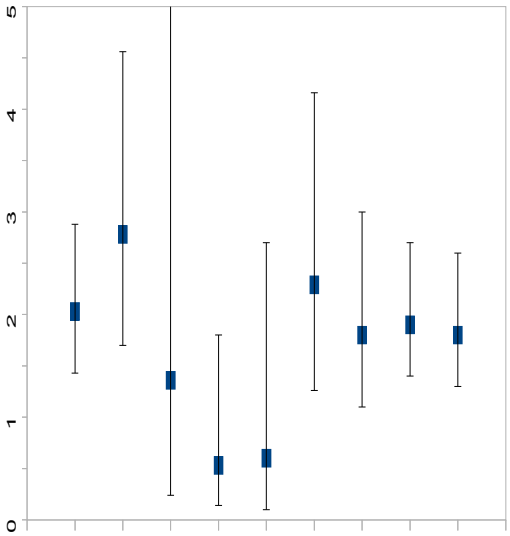
Dichotomised at 60y



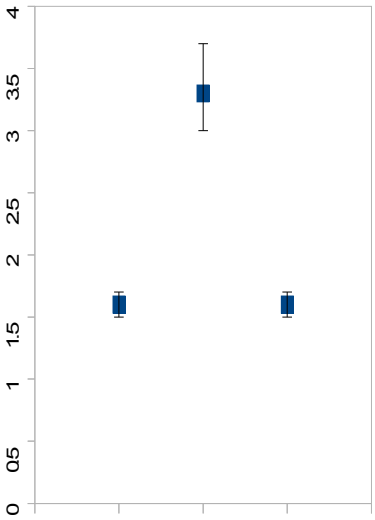
Dichotomised at 65y



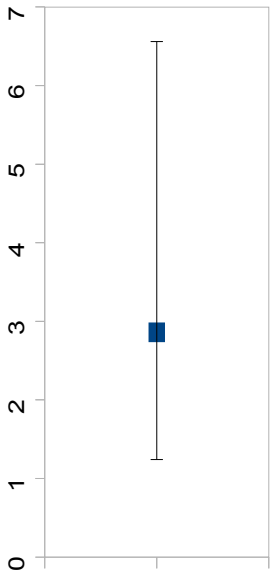
Dichotomised at 70y



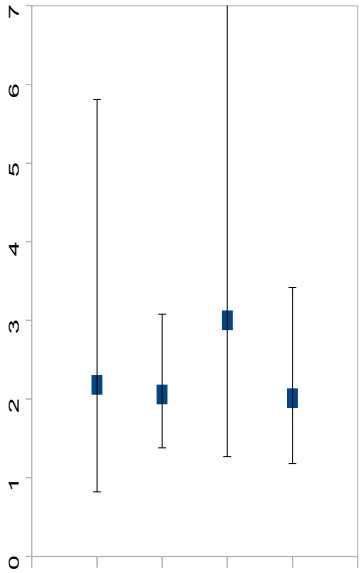
Dichotomised at 75y



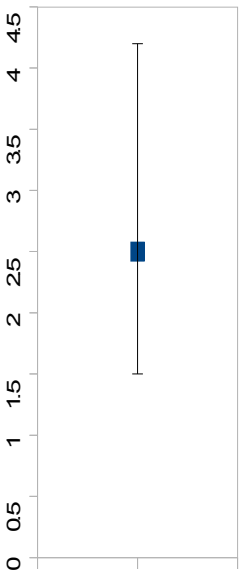
Dichotomised at 76y



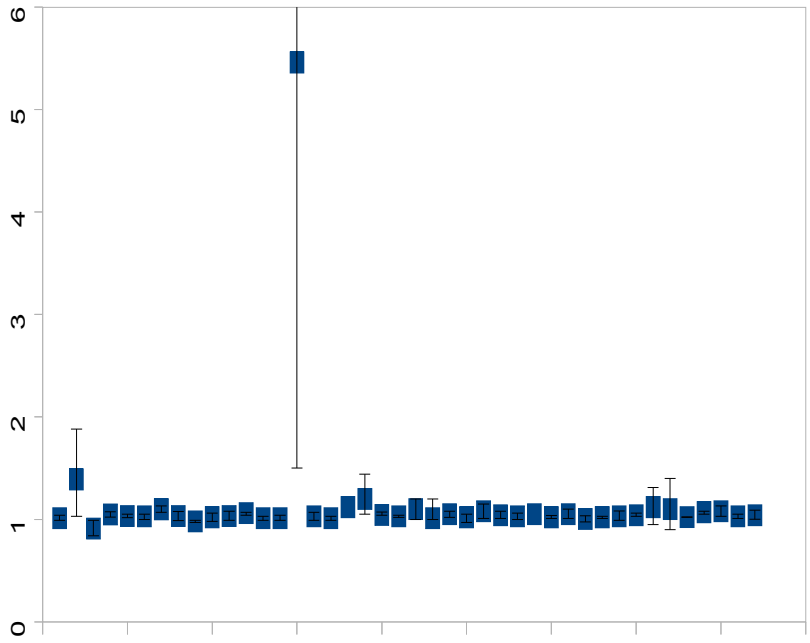
Dichotomised at 80y



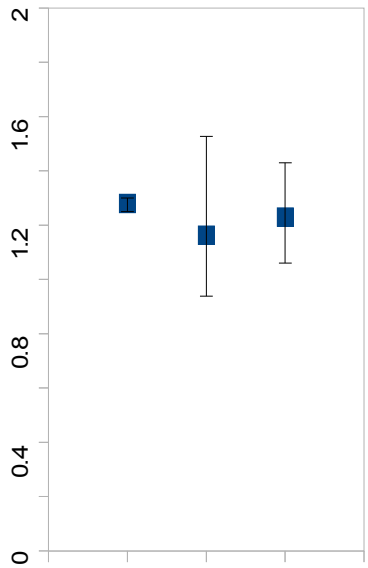
Dichotomised at 85y



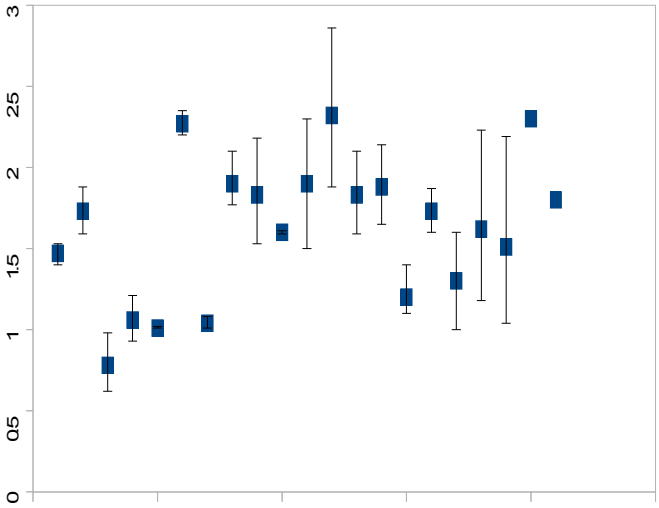
Continuous variable per year



Continuous variable per 5 years

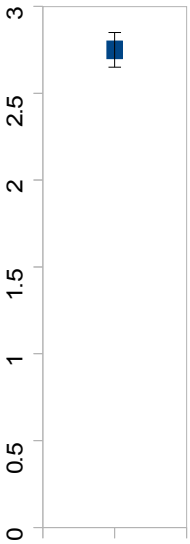


Continuous variable per 10 years

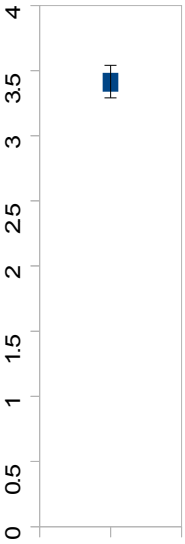


AVPU

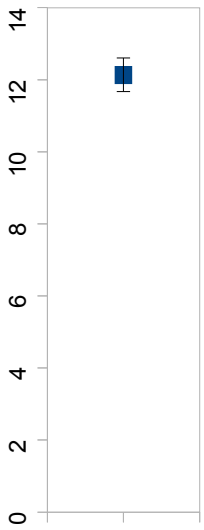
Dichotomised at V



Dichotomised at P

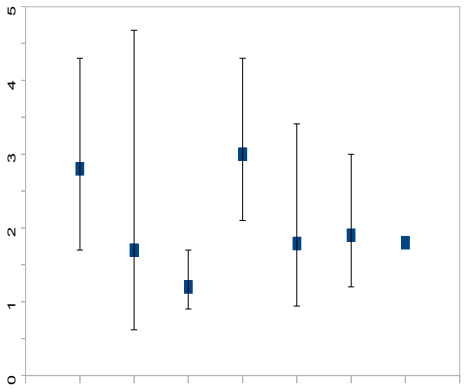


Dichotomised at U

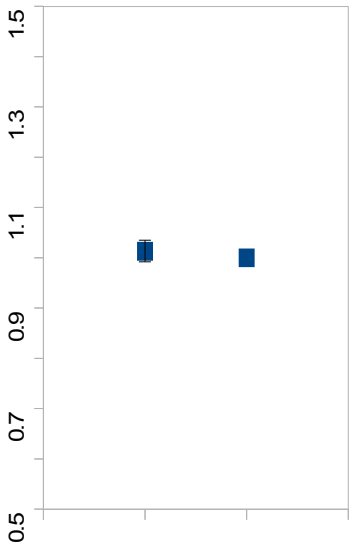


Diastolic BP

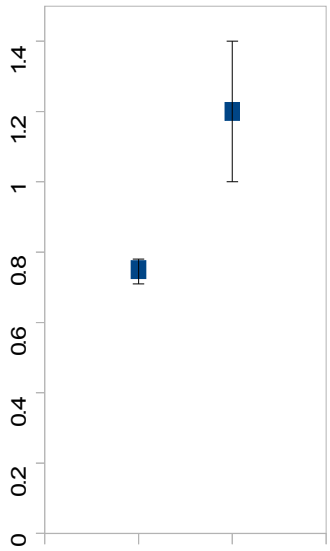
Dichotomised at 60mmHg



Continuous variable per mmHg

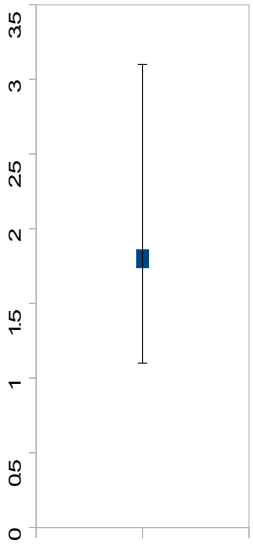


Continuous variable per 10mmHg

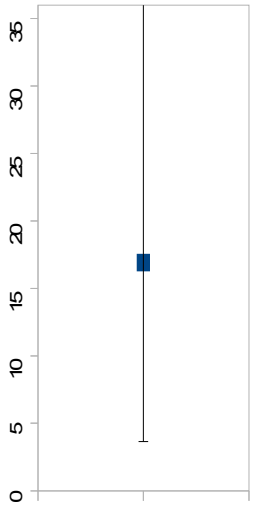


GCS

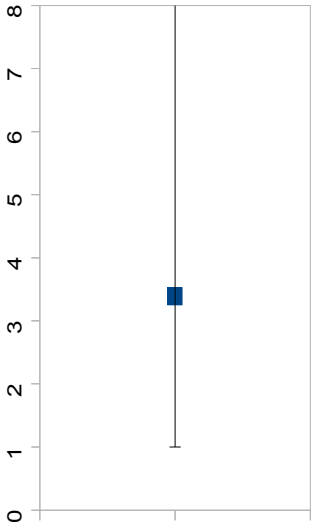
Dichotomised at 14



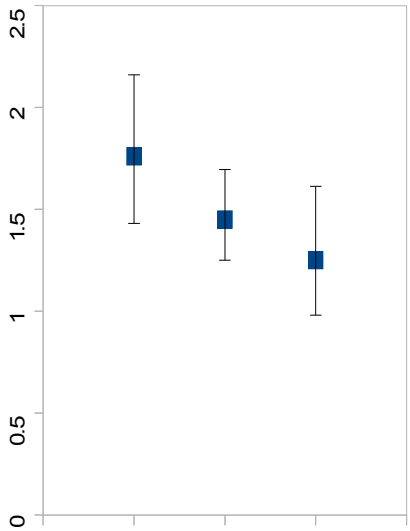
Dichotomised at 12



Dichotomised at 8

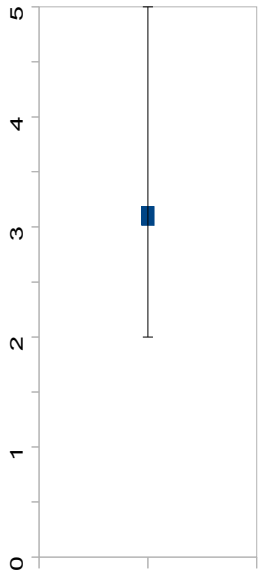


Continuous variable

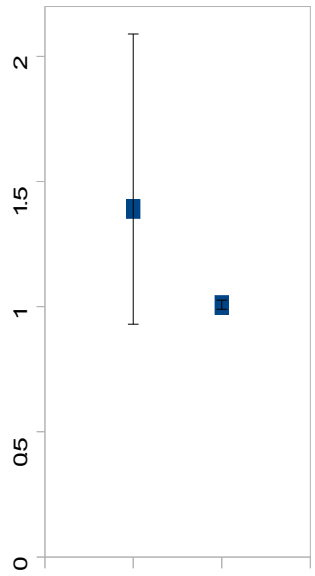


MAP

Dichotomised at 70mmHg

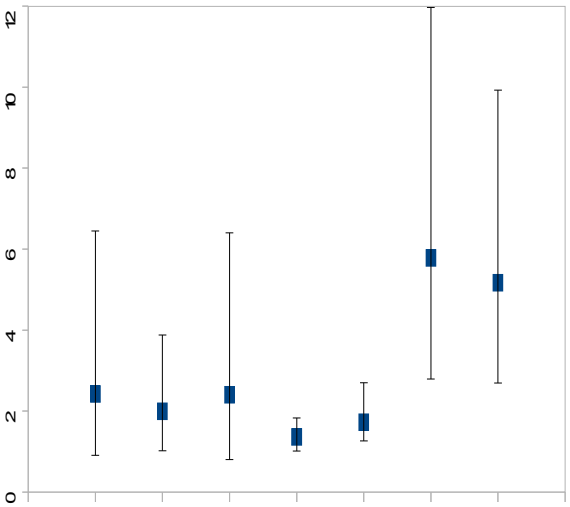


Continuous variable

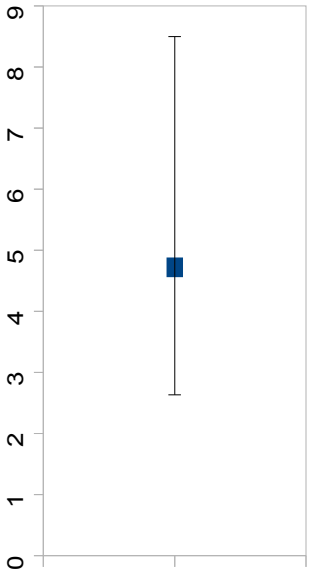


SaO2

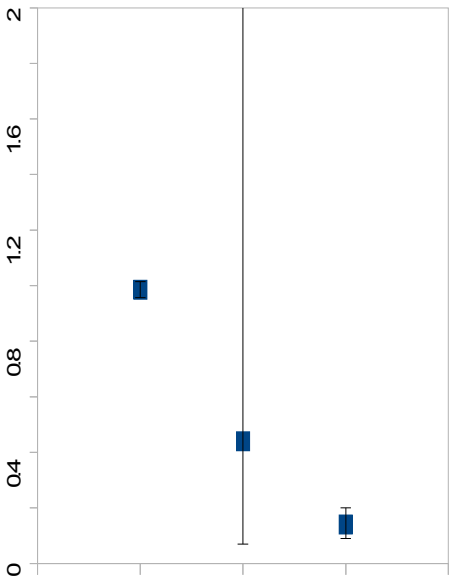
Dichotomised at 90%



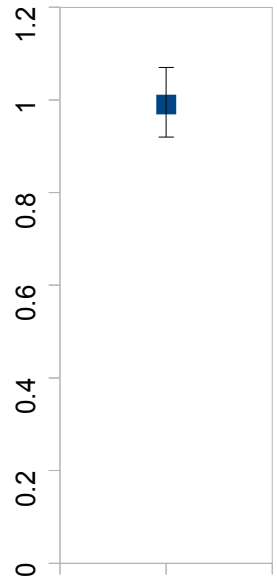
Dichotomised at 95%



Continuous variable per %

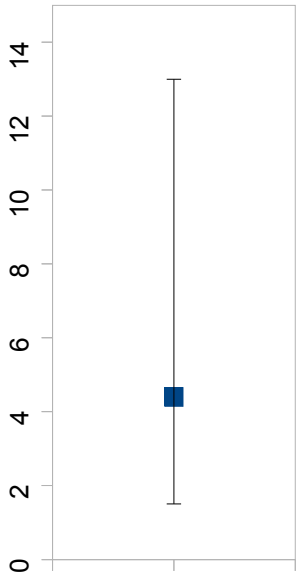


Continuous variable per 5%

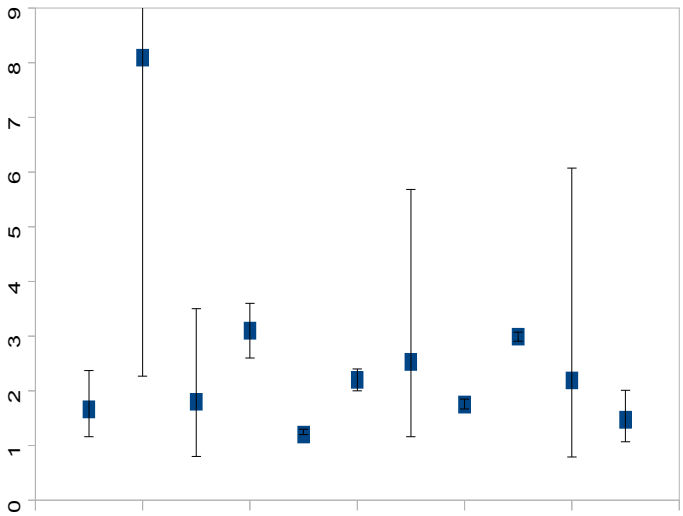


Pulse

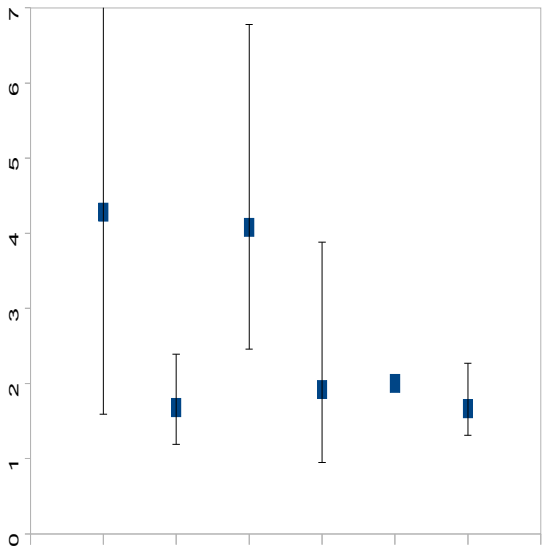
Dichotomised at 93 bpm



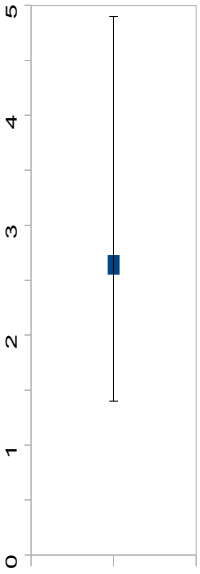
Dichotomised at 100 bpm



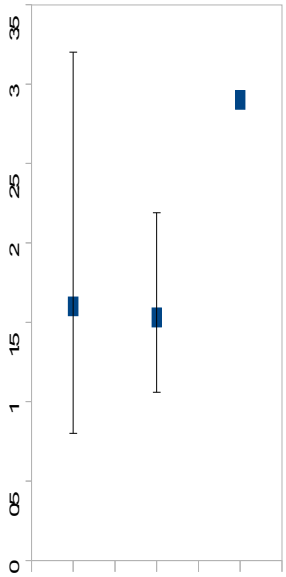
Dichotomised at 110 bpm



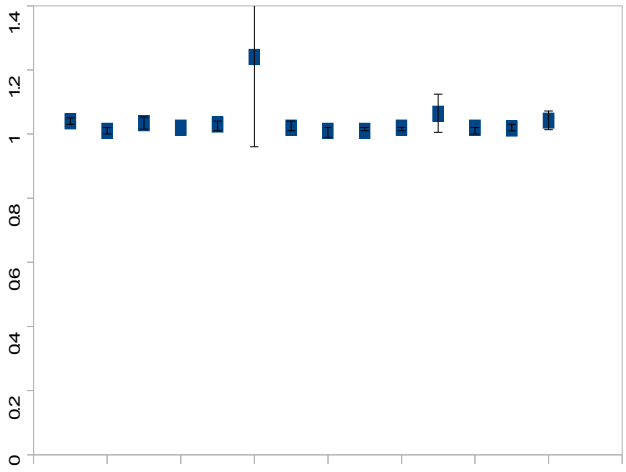
Dichotomised at 120 bpm



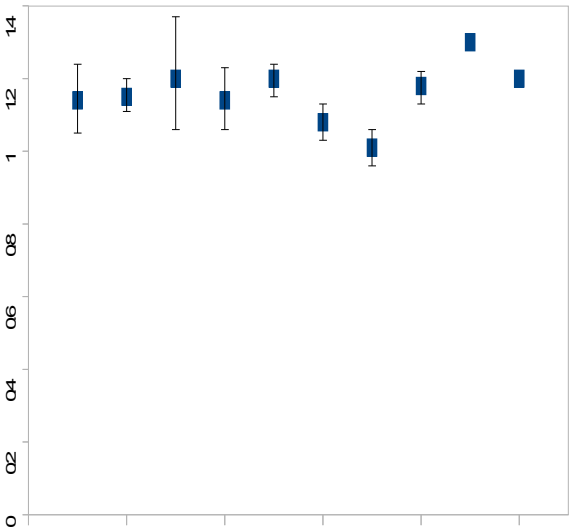
Dichotomised at 125 bpm



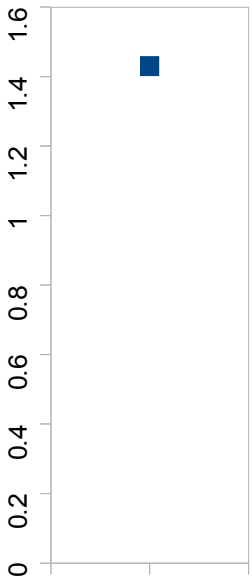
Continuous per bpm



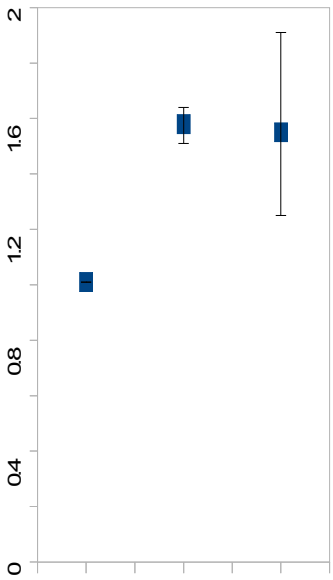
Continuous per 10 bpm



Continuous per 20 bpm

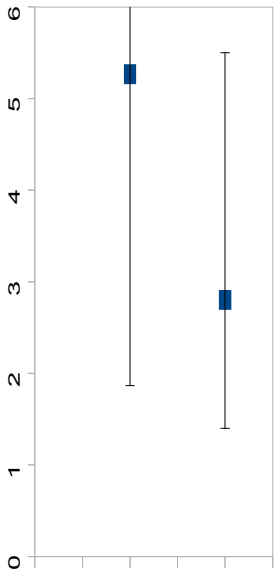


Continuous per 30 bpm

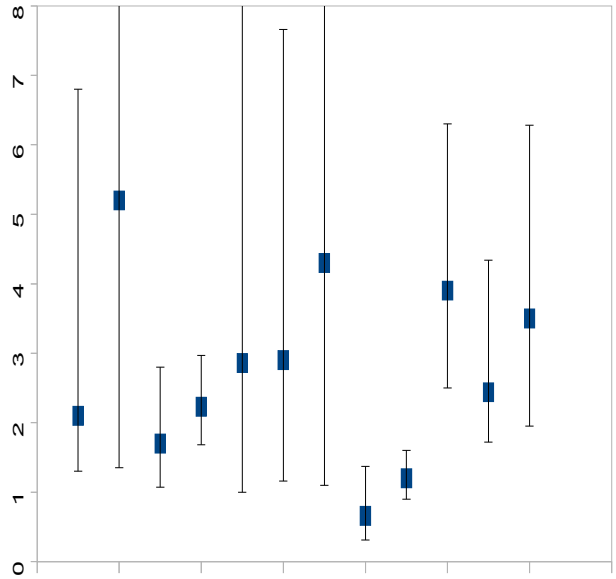


Respiratory rate

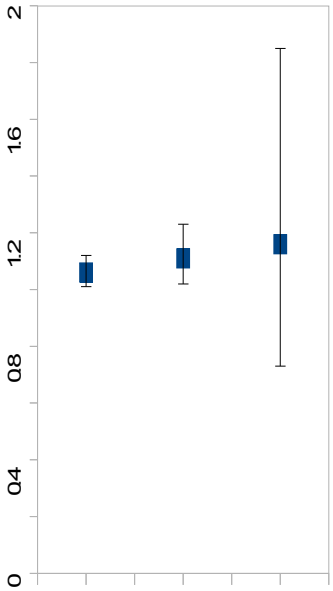
Dichotomised at 25bpm



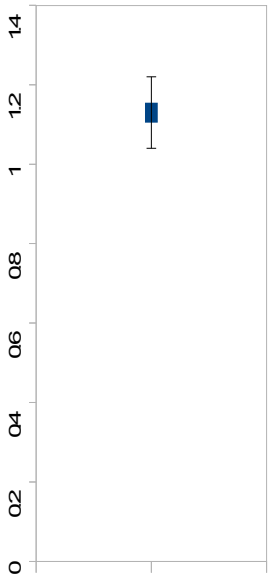
Dichotomised at 30bpm



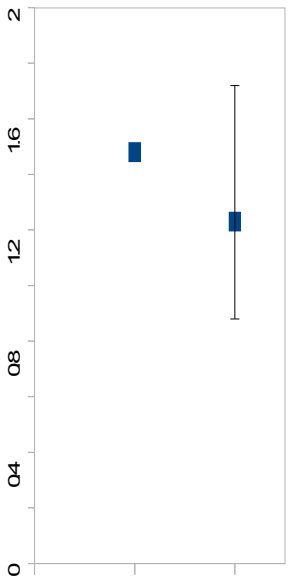
Continuous per bpm



Continuous per 5bpm

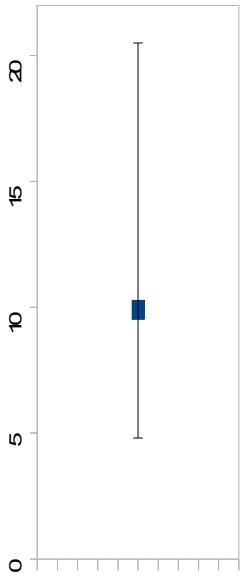


Continuous per 10bpm

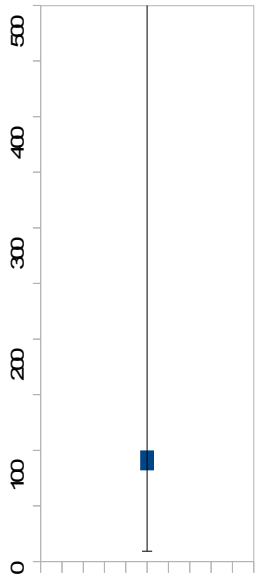


Systolic blood pressure

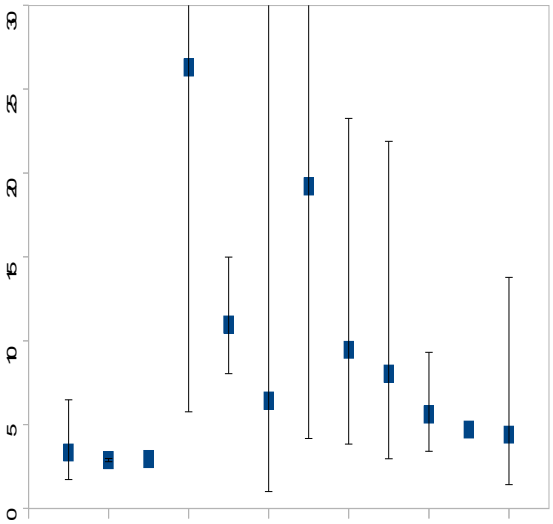
Dichotomised at 70mmHg



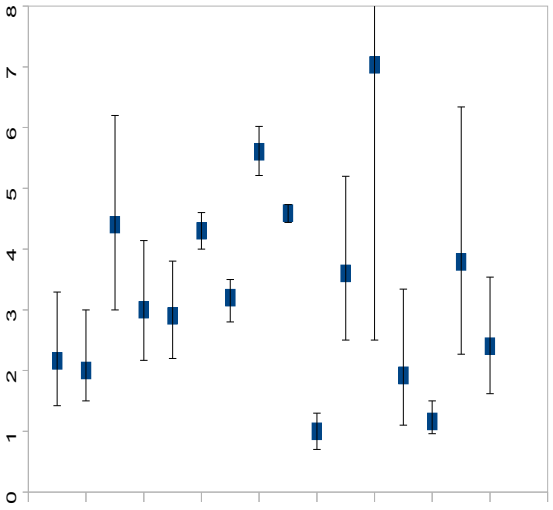
Dichotomised at 80mmHg



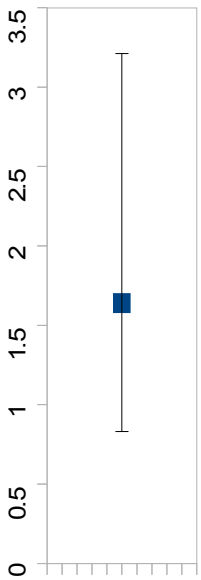
Dichotomised at 90mmHg



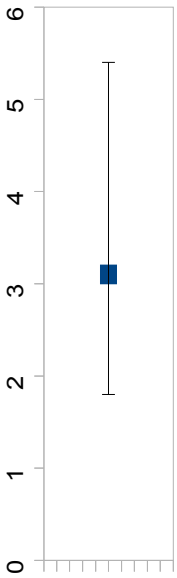
Dichotomised at 100mmHg



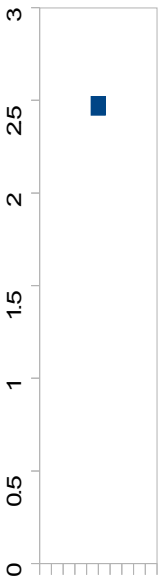
Dichotomised at 120mmHg



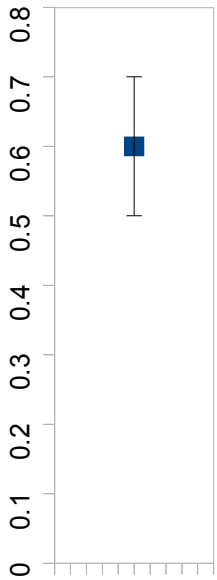
Dichotomised at 125mmHg



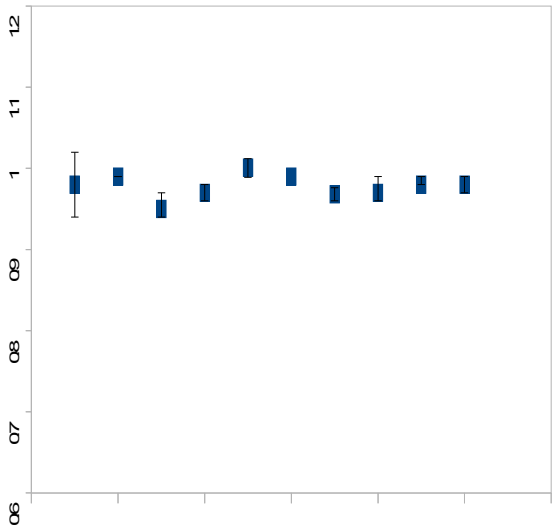
Dichotomised over 130mmHg



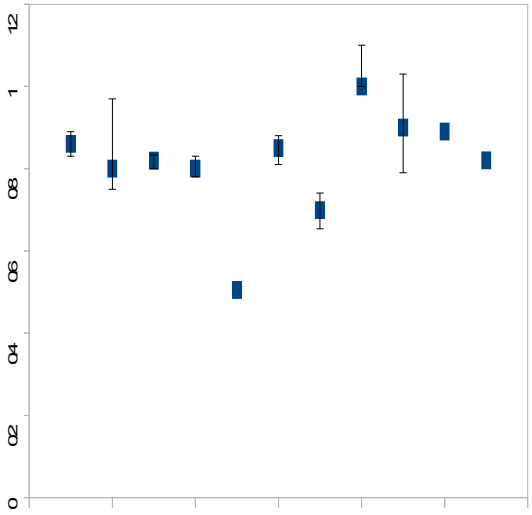
Dichotomised over 180mmHg



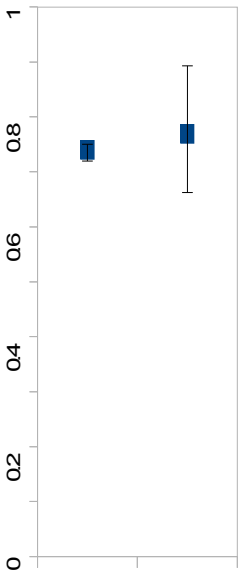
Continuous per mmHg



Continuous per 10mmHg

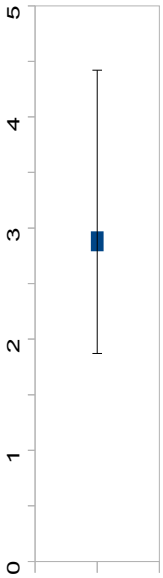


Continuous per 20mmHg

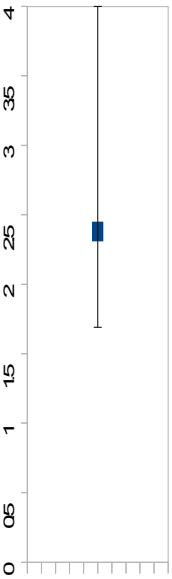


Temperature

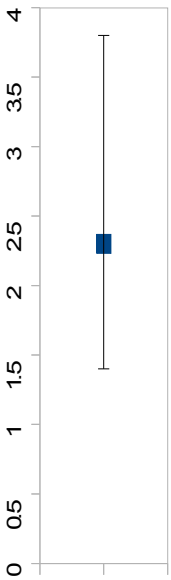
Dichotomised below 35.3



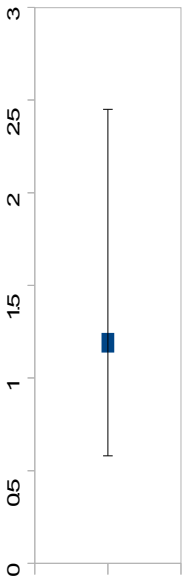
Dichotomised below 36



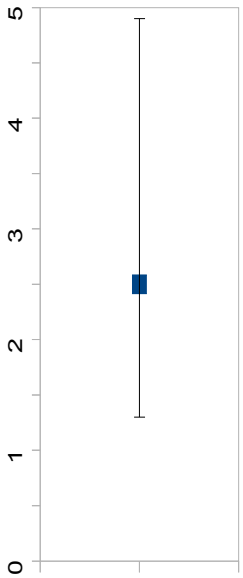
Dichotomised below 37



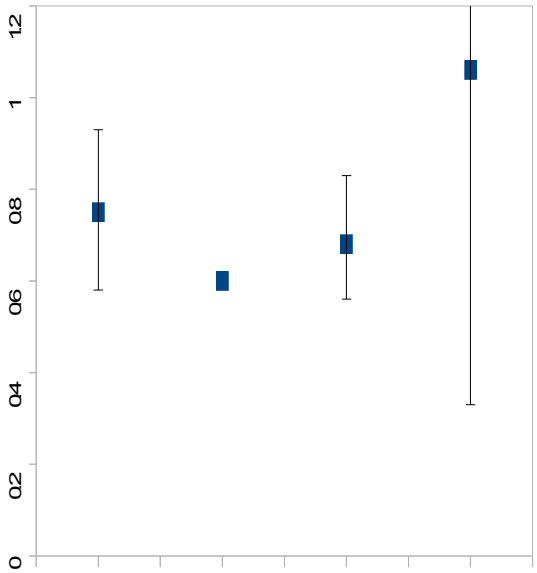
Dichotomised above 37.5



Dichotomised above 38.3



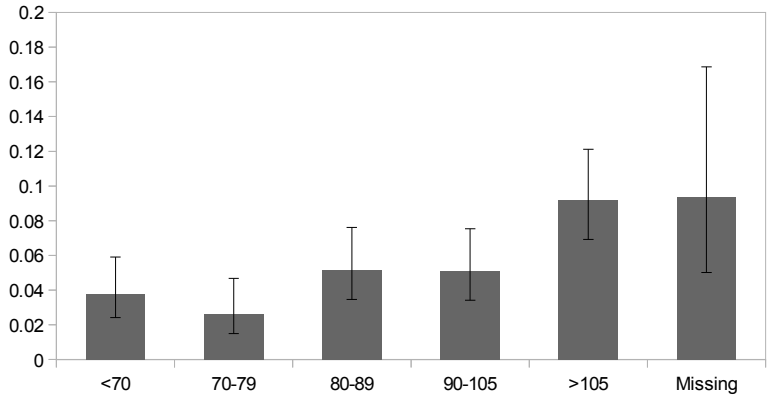
Continuous per degree C



Appendix 6: Univariate analysis of continuous variables in the prediction of death at 7 days

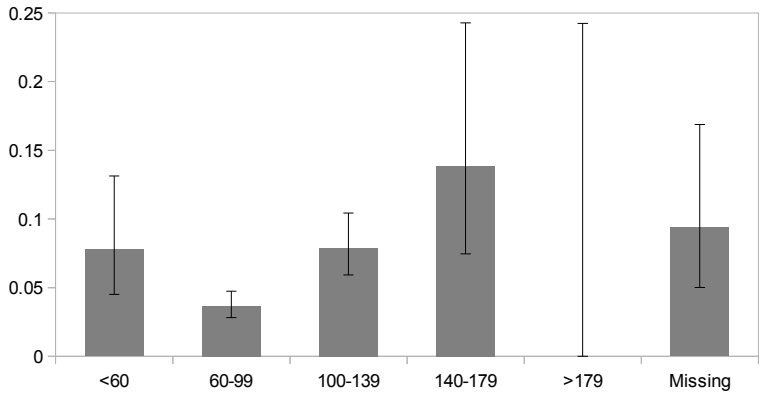
Graphs present the proportion of patients dying, with 95% confidence intervals.

	Dead	Total
Pulse population quintiles		
<70	18 (3.8)	475
70-79	11 (2.7)	415
80-89	23 (5.1)	447
90-105	23 (5.1)	451
>105	44 (9.2)	479
Missing	9 (9.4)	96
		Chi-sq p <0.001

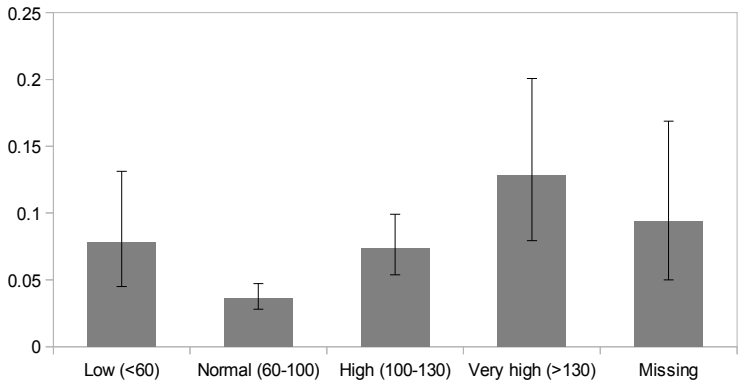


	Dead	Total
Pulse value quintiles		

<60	12 (7.8)	154
60-99	54 (3.7)	1478
100-139	44 (7.9)	558
140-179	9 (13.8)	65
>179	0	12
Missing	9 (9.4)	96
		Chi-sq p <0.001

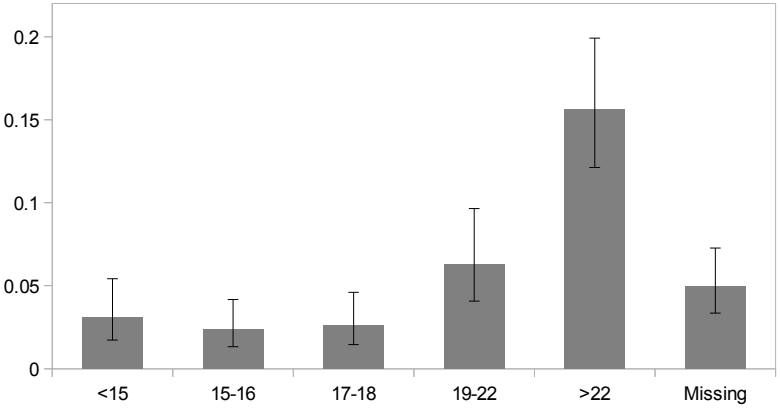


	Dead	Total
Pulse *a priori* groups		
Low (<60)	12 (7.8)	154
Normal (60-100)	54 (3.7)	1478
High (100-130)	38 (7.3)	518
Very high (>130)	15 (12.8)	117
Missing	9 (9.4)	96
		Chi-sq p <0.001

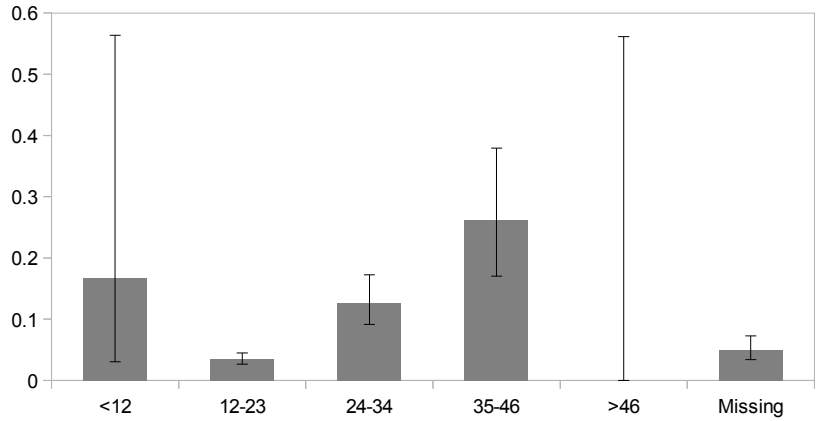


	Dead	Total
Respiratory rate population quintiles		
<15	11 (3.1)	357

15-16	11 (2.4)	467
17-18	11 (2.6)	421
19-22	19 (6.3)	301
>22	52 (15.6)	333
Missing	24 (5.0)	484
		Chi-sq p <0.001

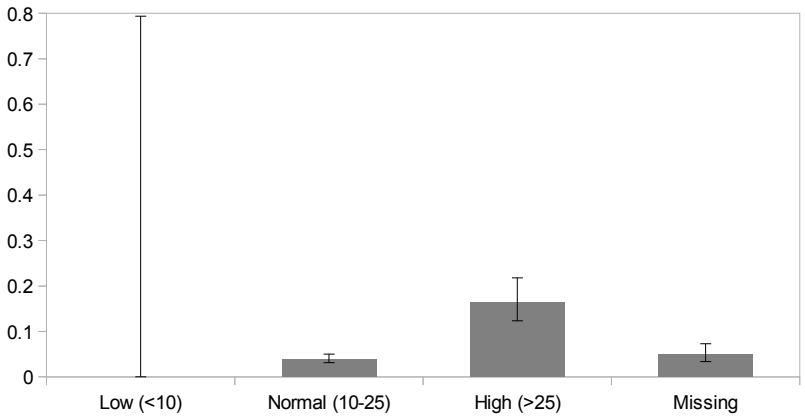


	Dead	Total
Respiratory rate value quintiles		
<12	1 (16.7)	6
12-23	53 (3.4)	1544
24-34	33 (12.6)	261
35-46	17 (26.2)	65
>46	0	3
Missing	24 (5.0)	484
		Chi-sq p <0.001

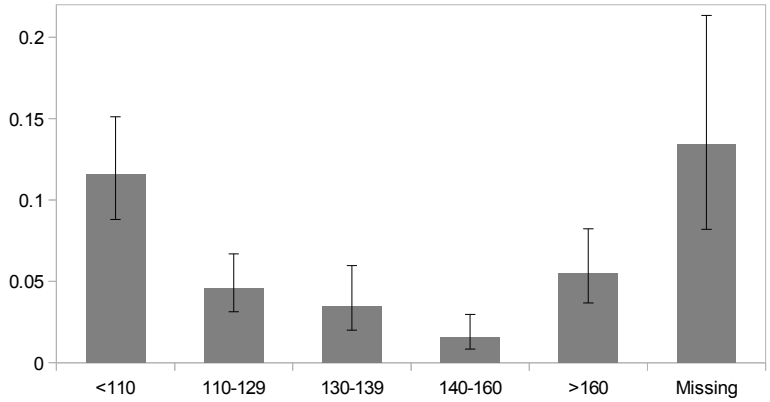


	Dead	Total
Respiratory rate *a priori* groups		

Low (<10)	0	1
Normal (10-25)	65 (4.0)	1642
High (>25)	39 (16.5)	236
Missing	24 (5.0)	484
Total	128	2363
		Chi-sq p <0.001

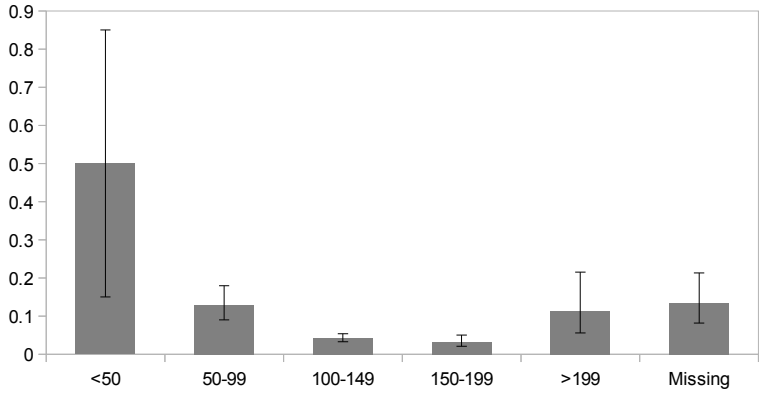


	Dead	Total
Systolic BP population quintiles		
<110	46 (11.6)	397
110-129	25 (4.6)	545
130-139	12 (3.5)	346
140-160	9 (1.6)	573
>160	22 (5.5)	398
Missing	14 (13.5)	104
		Chi-sq p <0.001

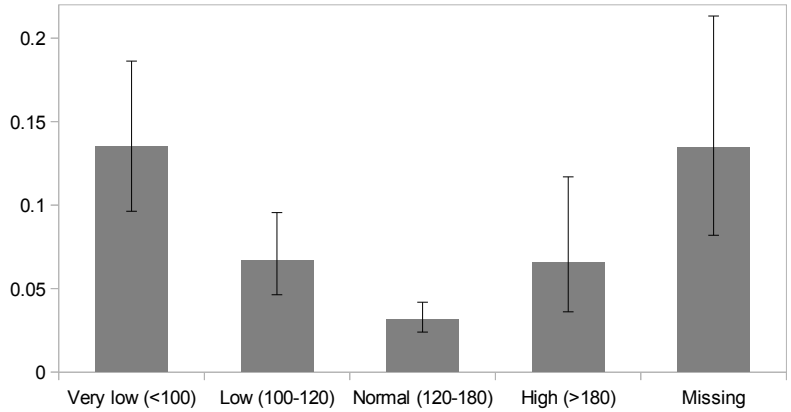


	Dead	Total
Systolic BP value quintiles		

<50	2 (50)	4
50-99	28 (12.8)	218
100-149	58 (4.2)	1390
150-199	19 (3.2)	585
>199	7 (11.3)	62
Missing	14 (13.5)	104
		Chi-sq p <0.001

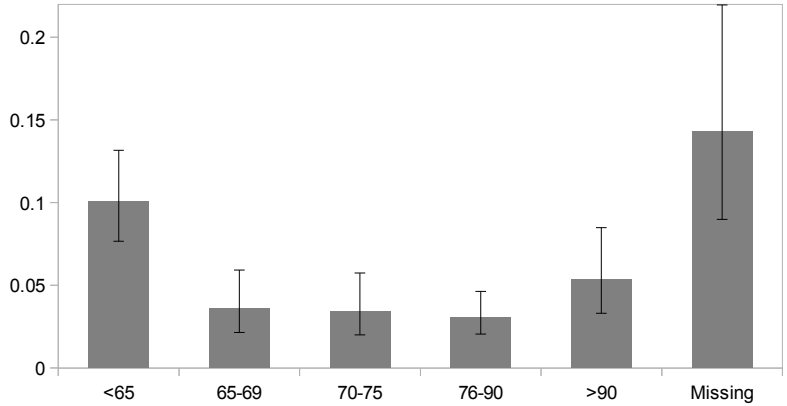


	Dead	Total
Systolic BP *a priori* groups		
Very low (<100)	30 (13.5)	222
Low (100-120)	27 (6.7)	404
Normal (120-180)	47 (3.2)	1481
High (>180)	10 (6.6)	152
Missing	14 (13.5)	104
Total	128	2363
		Chi-sq p <0.001

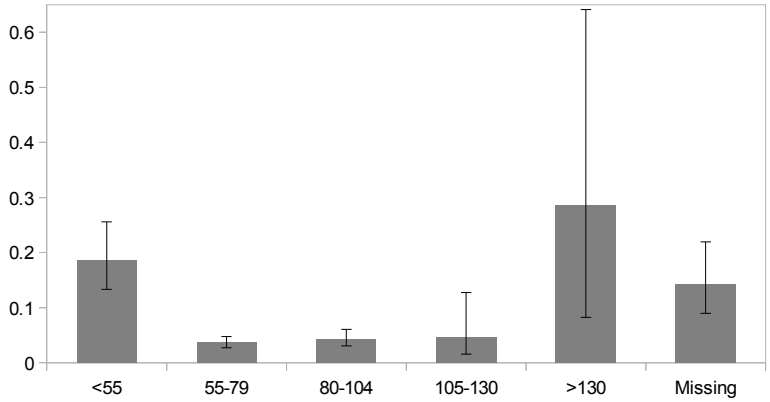


	Dead	Total
Diastolic BP population quintiles		

<65	47 (10.1)	466
65-69	14 (3.6)	391
70-75	13 (3.4)	381
76-90	22 (3.1)	713
>90	16 (5.3)	300
Missing	16 (14.3)	112
		Chi-sq p <0.001

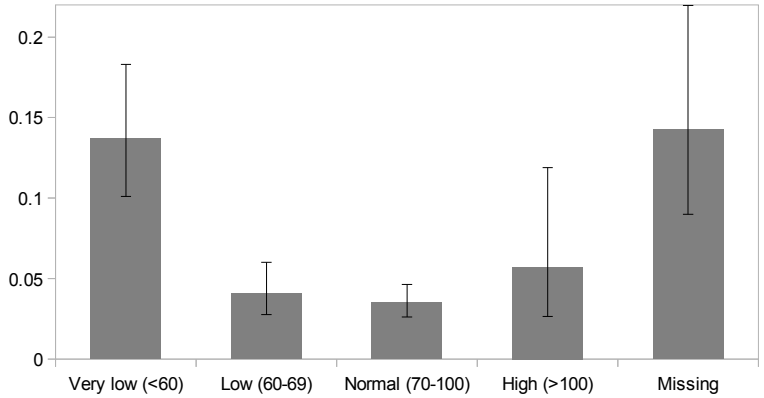


	Dead	Total
Diastolic BP value quintiles		
<55	29 (18.7)	155
55-79	48 (3.6)	1327
80-104	30 (4.3)	697
105-130	3 (4.6)	65
>130	2 (28.6)	7
Missing	16 (14.3)	112
		Chi-sq p <0.001

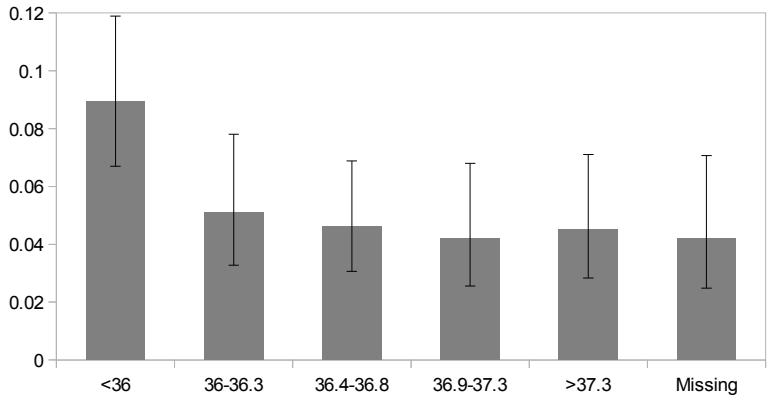


	Dead	Total
Diastolic BP *a priori* groups		
Very low (<60)	37 (13.8)	270

Low (60-69)	24 (4.1)	587
Normal (70-100)	45 (3.5)	1289
High (>100)	6 (5.7)	105
Missing	16 (14.3)	112
		Chi-sq p <0.001

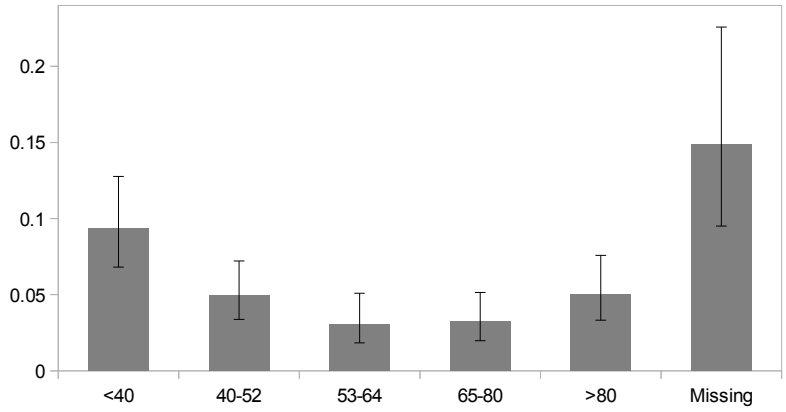


	Dead	Total
Temperature population quintiles		
<36	42 (9.0)	469
36-36.3	19 (5.1)	373
36.4-36.8	22 (4.6)	477
36.9-37.3	15 (4.2)	358
>37.3	17 (4.5)	377
Missing	13 (4.2)	309
		Chi-sq p
		0.012

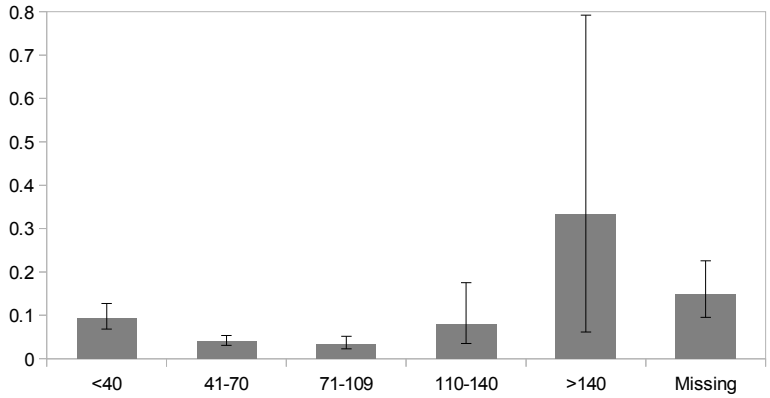


	Dead	Total
Pulse pressure population quintiles		
<40	35 (9.4)	373

40-52	25 (5.0)	505
53-64	14 (3.1)	456
65-80	16 (3.2)	498
>80	21 (5.0)	417
Missing	17 (14.9)	114
		Chi-sq p <0.001

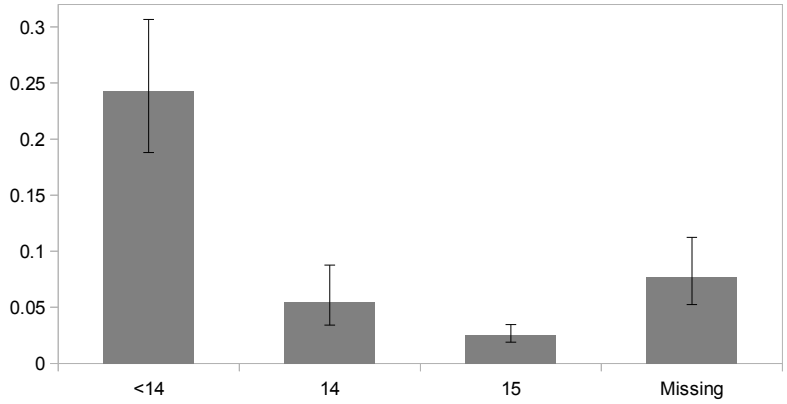


	Dead	Total
Pulse pressure value quintiles		
<40	35 (9.4)	373
41-70	48 (4.1)	1174
71-109	22 (3.5)	637
110-140	5 (8.1)	62
>140	1 (33.3)	3
Missing	17 (14.9)	114
		Chi-sq p <0.001

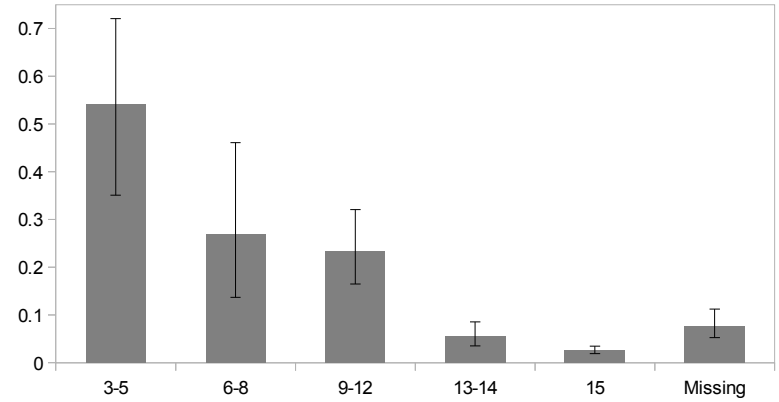


	Dead	Total
GCS population tertiles		
<14	48 (24.2)	198

14	16 (5.5)	290
15	40 (2.6)	1564
Missing	24 (7.7)	311
		Chi-sq p <0.001

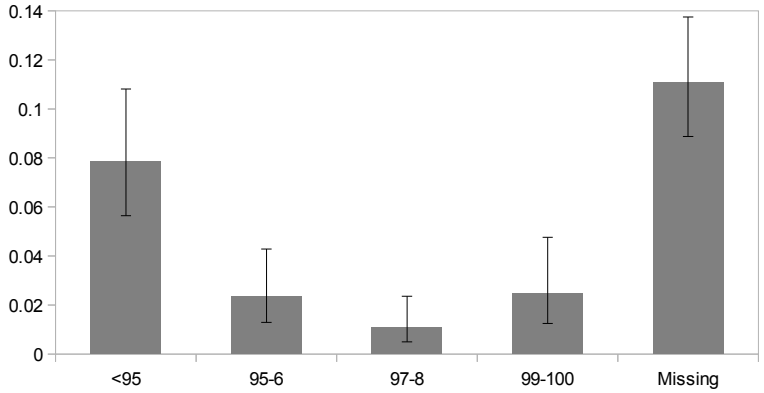


	Dead	Total
GCS *a priori* groups		
3-5	13 (54.2)	24
6-8	7 (26.9)	26
9-12	26 (23.4)	111
13-14	18 (5.5)	327
15	40 (2.6)	1564
Missing	24 (7.7)	311
		Chi-sq p <0.001

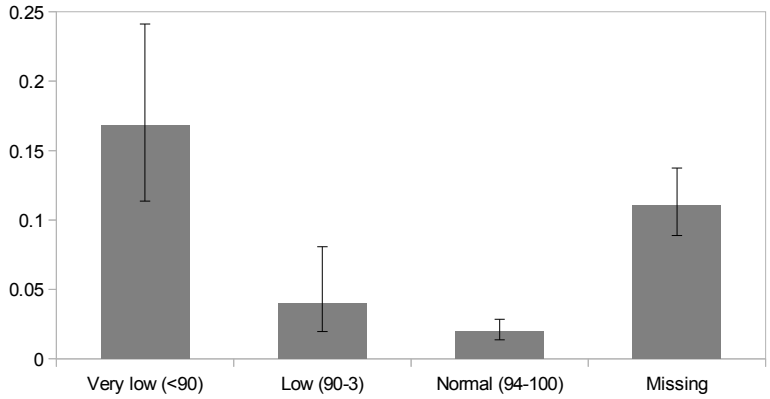


	Dead	Total
Oxygen saturations breathing air population quartiles		
<95	33 (7.8)	421

95-6	10 (2.4)	424
97-8	6 (1.1)	551
99-100	8 (2.5)	326
Missing	71 (11.1)	641
		Chi-sq p <0.001

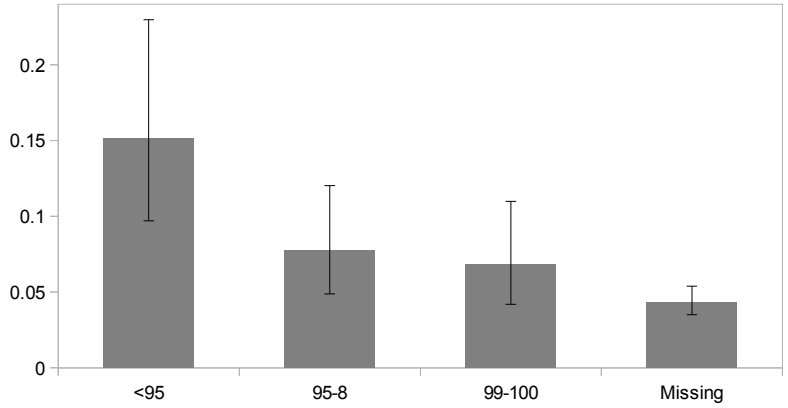


	Dead	Total
Oxygen saturations breathing air *a priori* groups		
Very low (<90)	22 (16.8)	131
Low (90-3)	7 (4.0)	174
Normal (94-100)	28 (2.0)	1417
Missing	71 (11.1)	641
		Chi-sq p <0.001

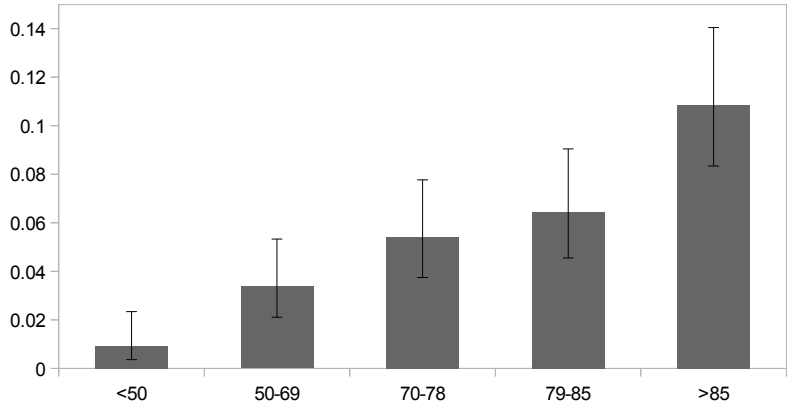


	Dead	Total
Oxygen saturations breathing supplemental oxygen population tertiles		

<95	17 (15.2)	112
95-8	17 (7.7)	220
99-100	15 (6.9)	219
Missing	79 (4.4)	1812
		Chi-sq p <0.001

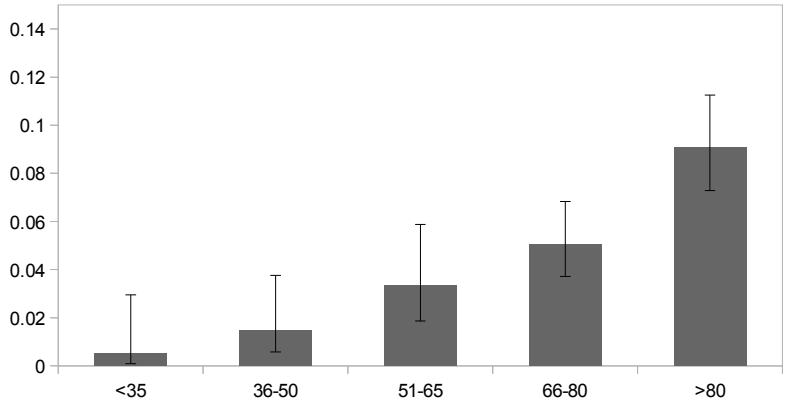


	Dead	Total
Age population quintiles		
<50	4 (1)	435
50-69	17 (3.4)	504
70-78	27 (5.4)	498
79-85	30 (6.4)	466
>85	50 (10.9)	460
		Chi-sq p <0.001



	Dead	Total
Age value quintiles		

<35	1 (0.5)	188
36-50	4 (1.5)	269
51-65	11 (3.3)	329
66-80	39 (5.1)	772
>80	73 (9.1)	805
		Chi-sq p <0.001



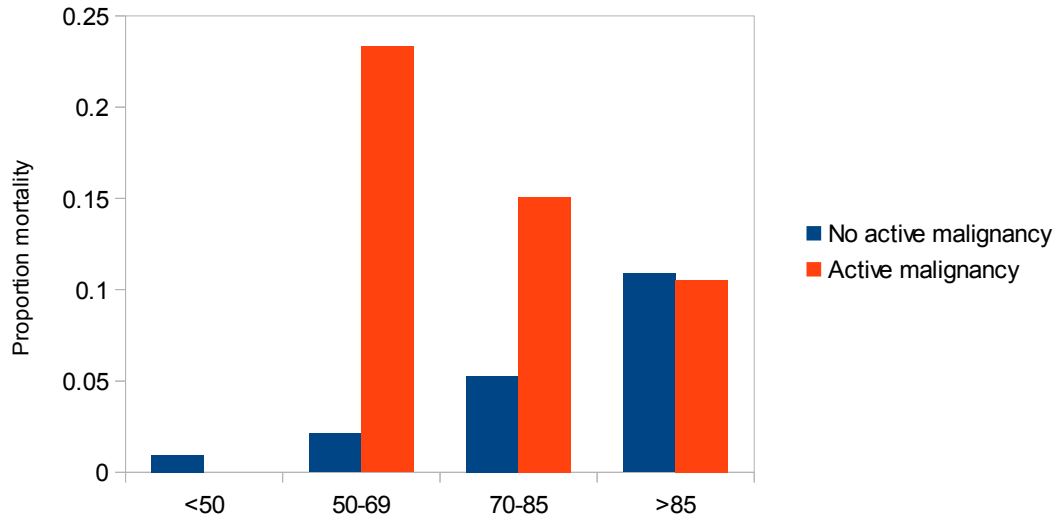
Appendix 7:
Interactions between significant variables in predicting death in 7 days

Interaction term	Wald	Degrees of freedom	p
Active malignancy by pulse	0.268	2	0.875
Active malignancy by respiratory rate	0.832	2	0.660
Active malignancy by SBP	5.479	3	0.140
Active malignancy by DBP	2.338	2	0.311
Active malignancy by SaO2	0.710	2	0.701
Active malignancy by respiratory disease	0.400	1	0.527
Age by pulse	3.200	6	0.783
Age by respiratory rate	8.305	6	0.217
Age by SBP	6.399	9	0.699
Age by DBP	5.860	6	0.439
Age by GCS	5.507	9	0.788
Age by SaO2	6.288	6	0.392
Age by pulse pressure	6.066	9	0.733
Age by respiratory disease	2.547	4	0.636
Pulse by respiratory rate	1.050	4	0.902
Pulse by SBP	9.920	6	0.128
Pulse by DBP	4.567	4	0.335
Pulse by GCS	6.665	6	0.353
Pulse by SaO2	6.598	4	0.159
Pulse by pulse pressure	1.815	6	0.936
Pulse by respiratory disease	0.753	2	0.686
Respiratory rate by SBP	8.490	6	0.204
Respiratory rate by DBP	4.960	4	0.291
Respiratory rate by GCS	3.066	6	0.801
Respiratory rate by SaO2	4.998	4	0.288
Respiratory rate by pulse pressure	8.751	6	0.188
Respiratory rate by respiratory disease	2.454	2	0.293
SBP by DBP	1.508	6	0.959
SBP by GCS	6.292	9	0.710
SBP by SaO2	4.613	6	0.594
SBP by pulse pressure	6.154	8	0.630
SBP by respiratory disease	4.917	3	0.178
DBP by GCS	2.019	6	0.918
DBP by SaO2	2.978	4	0.561
DBP by pulse pressure	9.131	6	0.166
DBP by respiratory disease	1.128	2	0.569
GCS by SaO2	5.845	6	0.441
GCS by pulse pressure	7.627	9	0.572
GCS by respiratory disease	4.408	3	0.221
SaO2 by pulse pressure	3.246	6	0.777
SaO2 by respiratory disease	0.363	2	0.834

Graphs below present proportion of deaths by interaction groups.

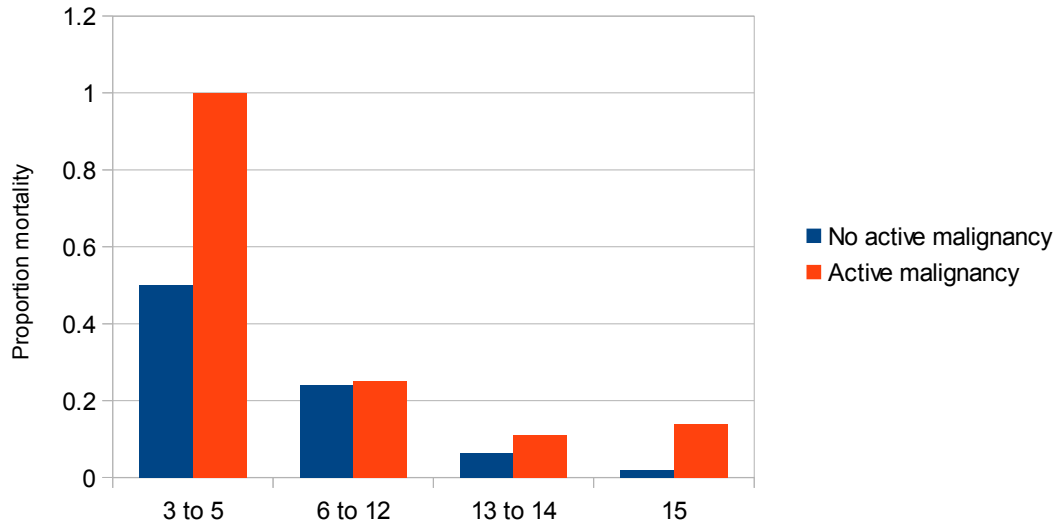
Active malignancy by age

	B	Sig.	Exp(B)
Active malignancy present	-5.808	<0.001	0.003
Age		.166	
Age 50-69	*5.603*	*<0.001*	*271.246*
Age 70-85	*-0.686*	*0.036*	*0.504*
Age >85	*-0.810*	*0.006*	*0.445*
Active malignancy by age		<0.001	
Active malignancy present by age 50-69	*-14.575*	*<0.001*	*0*
Active malignancy present by age 70-85	*-1.315*	*0.044*	*0.269*
Active malignancy present by age >85	*0.029*	*0.961*	*1.029*



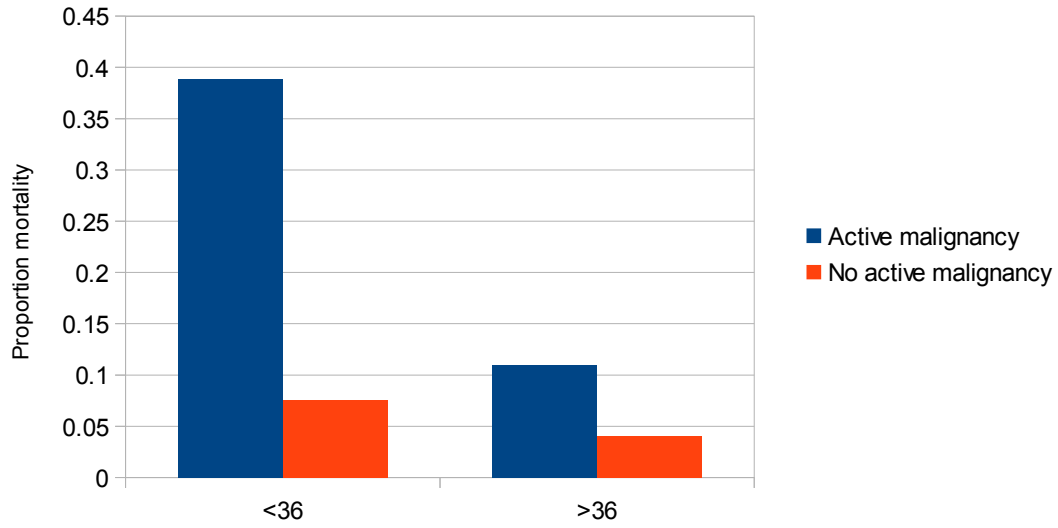
Active malignancy by GCS

	B	Sig.	Exp(B)
Active malignancy present	-3.874	<0.001	0.021
GCS		<0.001	
GCS 3-5	*9.230*	*<0.001*	*10201*
GCS 6-12	*1.731*	*<0.001*	*5.644*
GCS 13-14	*0.461*	*0.215*	*1.585*
Active malignancy by GCS		<0.001	
Active malignancy present by GCS 3-5	*-10.688*	*<0.001*	*0*
Active malignancy present by GCS 6-12	*2.009*	*0.030*	*7.459*
Active malignancy present by GCS 13-14	*1.431*	*0.054*	*4.185*



Active malignancy by temperature

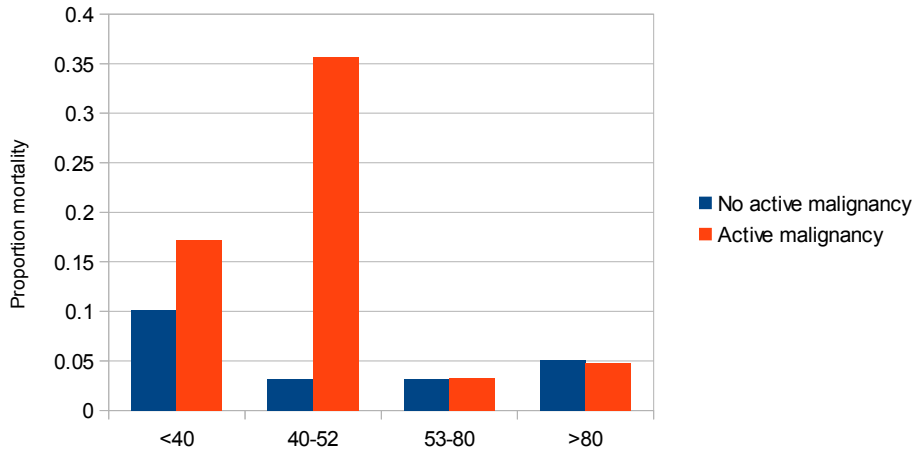
	B	Sig.	Exp(B)
Active malignancy present	5.271	<0.001	194.706
Temperature <36	4.224	<0.001	68.328
Active malignancy present and temperature <36	6.917	<0.001	1009.38



Active malignancy by pulse pressure

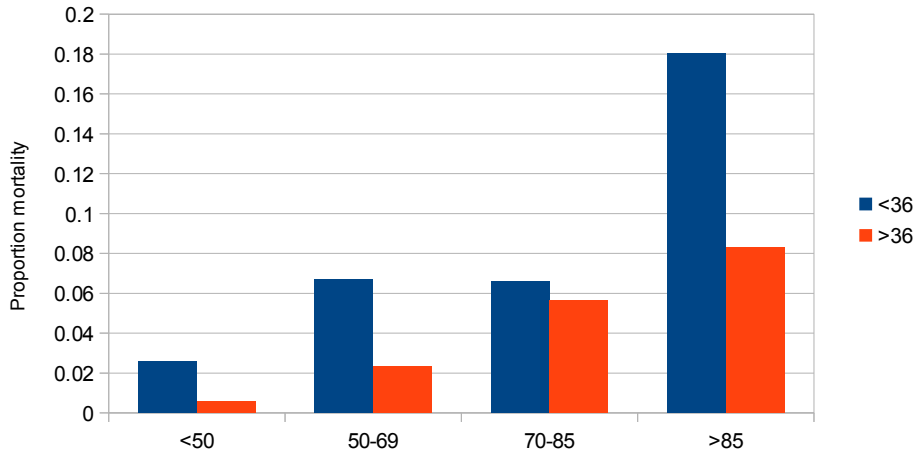
	B	Sig.	Exp(B)
Active malignancy present	-4.795	<0.001	0.008
Pulse pressure		<0.001	
Pulse pressure <40	*0.995*	*0.001*	*2.705*
Pulse pressure 40-53	*2.752*	*<0.001*	*15.672*
Pulse pressure >80	*1.216*	*0.004*	*3.375*
Active malignancy by pulse pressure		<0.001	
Active malignancy present by pulse pressure <40	*0.236*	*0.703*	*1.267*

Active malignancy present by pulse pressure 40-52	*-4.854*	*<0.001*	*0.008*
Active malignancy present by pulse pressure >80	*-1.069*	*0.202*	*0.343*



Temperature by age

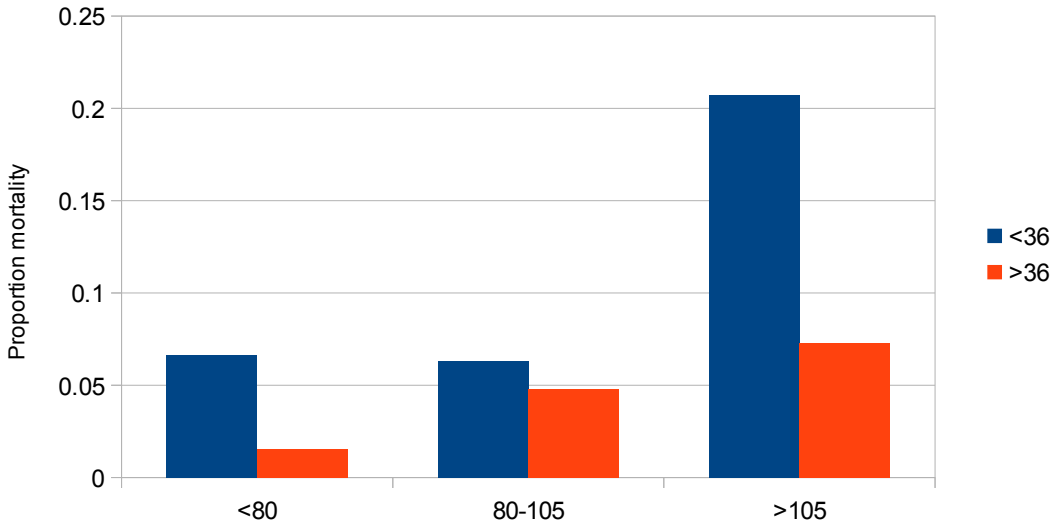
	B	Sig.	Exp(B)
Temperature <36	2.569	<0.001	13.047
Age (ref <50)		<0.001	
50-69	*-1.469*	*<0.001*	*0.230*
70-85	*-2.335*	*<0.001*	*0.097*
>85	*-0.990*	*<0.001*	*0.371*
Temperature by age		<0.001	
Age 50-69 by temperature <36	*-2.978*	*<0.001*	*0.051*
Age 70-85 by temperature <36	*-4.004*	*<0.001*	*0.018*
Age >85 by temperature <36	*-2.929*	*<0.001*	*0.053*



Pulse by temperature

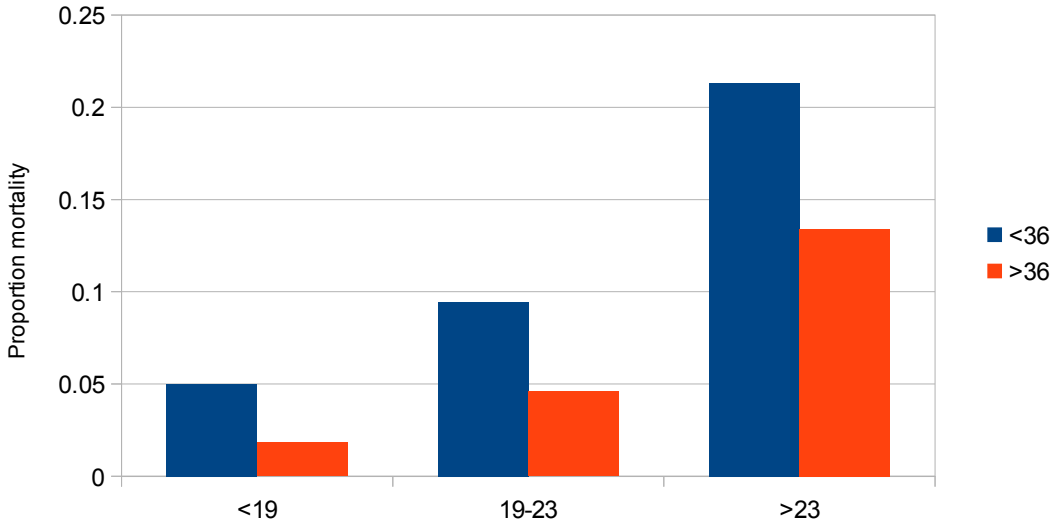
	B	Sig.	Exp(B)
Pulse		<0.001	
Pulse 80-105	*0.334*	*0.009*	*1.396*

Pulse >105	*3.985*	*<0.001*	*53.808*
Temperature <36	3.410	<0.001	30.255
Pulse by temperature		<0.001	
Pulse 80-105 by temperature <36	*0.421*	*0.1*	*1.523*
Pulse >105 by temperature <36	*6.806*	*<0.001*	*903.281*



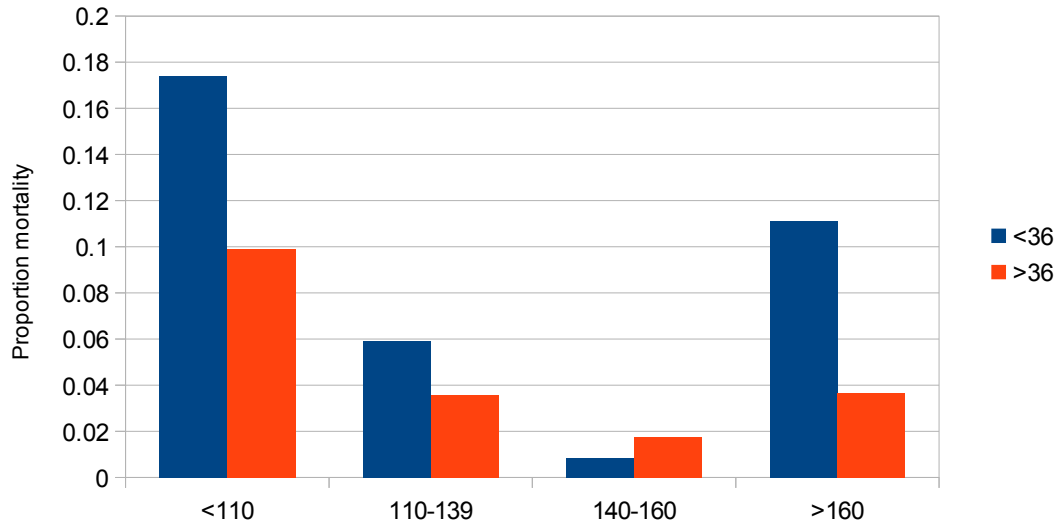
Respiratory rate by temperature

	B	Sig.	Exp(B)
Temperature <36	3.174	<0.001	23.891
Respiratory rate (ref <19)		<0.001	
19-23	*0.746*	*<0.001*	*2.108*
>23	*4.657*	*<0.001*	*105.311*
Temperature by respiratory rate		<0.001	
Respiratory rate 19-23 by temperature <36	*0.248*	*0.342*	*1.281*
Respiratory rate >23 by temperature <36	*5.586*	*<0.001*	*266.543*



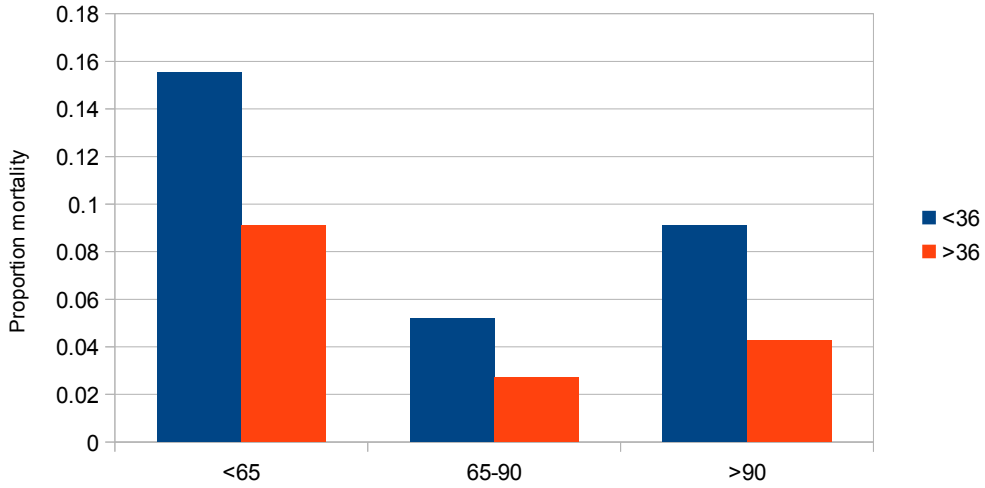
SBP by temperature

	B	Sig.	Exp(B)
Temperature <36	2.211	<0.001	9.122
SBP (ref 140-60)		<0.001	
<110	*-4.393*	*<0.001*	*0.012*
110-139	*-1.882*	*<0.001*	*0.152*
>160	*-3.031*	*<0.001*	*0.048*
Temperature by SBP		<0.001	
SBP <110 by temperature <36	*-2.950*	*0.007*	*0.052*
SBP 110-139 by temperature <36	*-1.562*	*<0.001*	*0.210*
SBP >160 by temperature <36	*-1.684*	*<0.001*	*0.186*



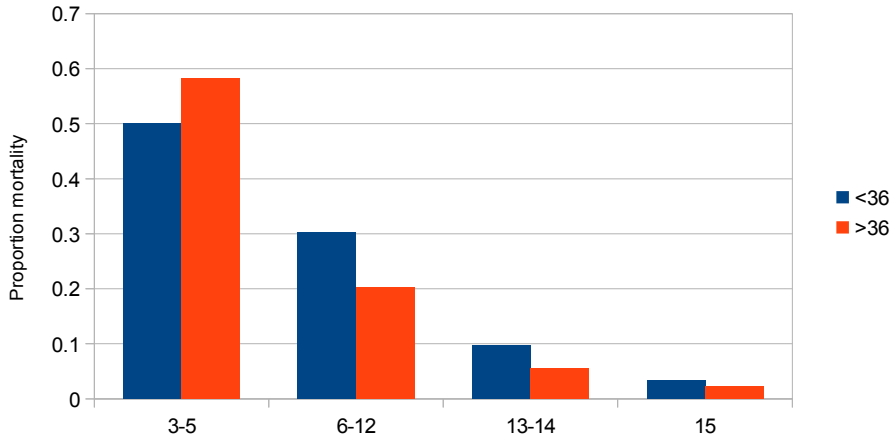
DBP by Temperature

	B	Sig.	Exp(B)
Temperature <36	2.990	<0.001	19.877
DBP (ref 65-90)		<0.001	
<65	*1.306*	*<0.001*	*3.693*
>90	*4.229*	*<0.001*	*68.652*
DBP by temperature		<0.001	
DBP <65 by temperature <36	*0.090*	*0.727*	*1.095*
DBP >90 by temperature <36	*5.281*	*<0.001*	*196.587*



Temperature by GCS

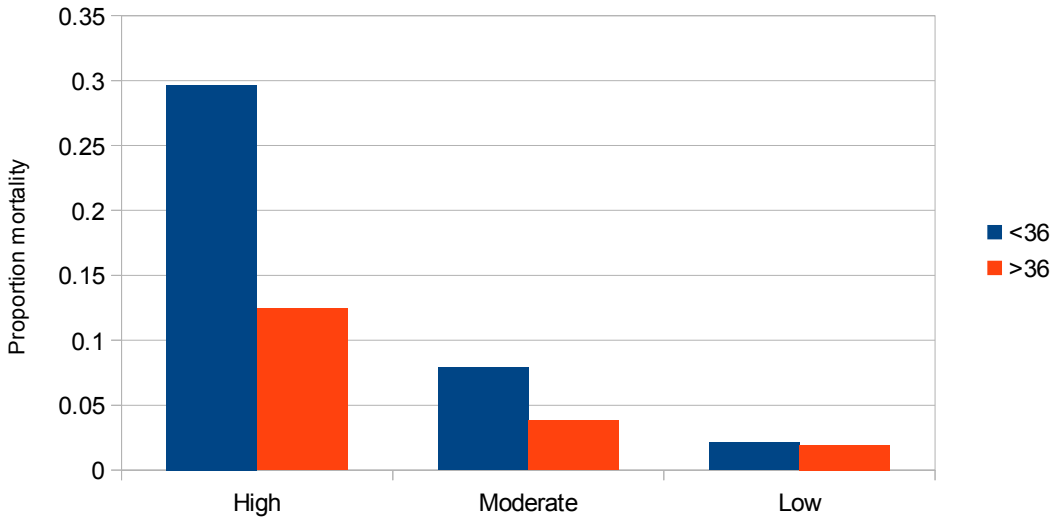
	B	Sig.	Exp(B)
Temperature <36	-1.459	0.010	0.232
GCS (ref 15)		<0.001	
3-5	*9.569*	*<0.001*	*14315*
6-12	*2.554*	*<0.001*	*12.858*
13-14	*1.008*	*<0.001*	*2.739*
GCS by temperature		0.002	
3-5 by temperature <36	*-8.358*	*<0.001*	*0*
6-12 by temperature <36	*0.035*	*0.942*	*1.036*
13-14 by temperature <36	*0.166*	*0.698*	*1.181*



SaO2 by temperature

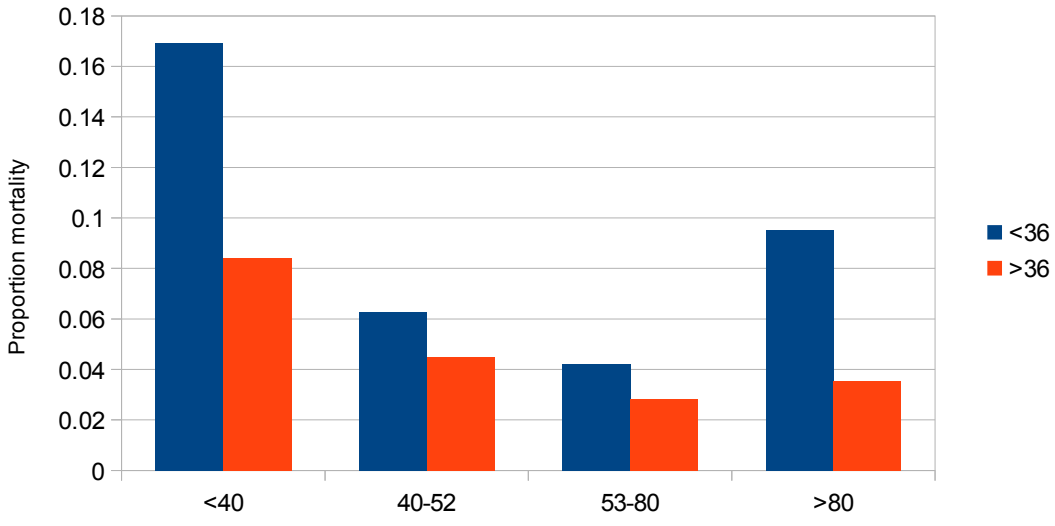
	B	Sig.	Exp(B)
Temperature <36	3.186	<0.001	24.2
SaO2 (ref low risk)		<0.001	
High risk	*4.488*	*<0.001*	*88.928*
Moderate risk	*0.532*	*<0.001*	*1.703*
Temperature by SaO2		<0.001	

High risk by temperature <36	*5.876*	*<0.001*	*356.496*
Moderate risk by temperature <36	*0.281*	*0.276*	*1.325*



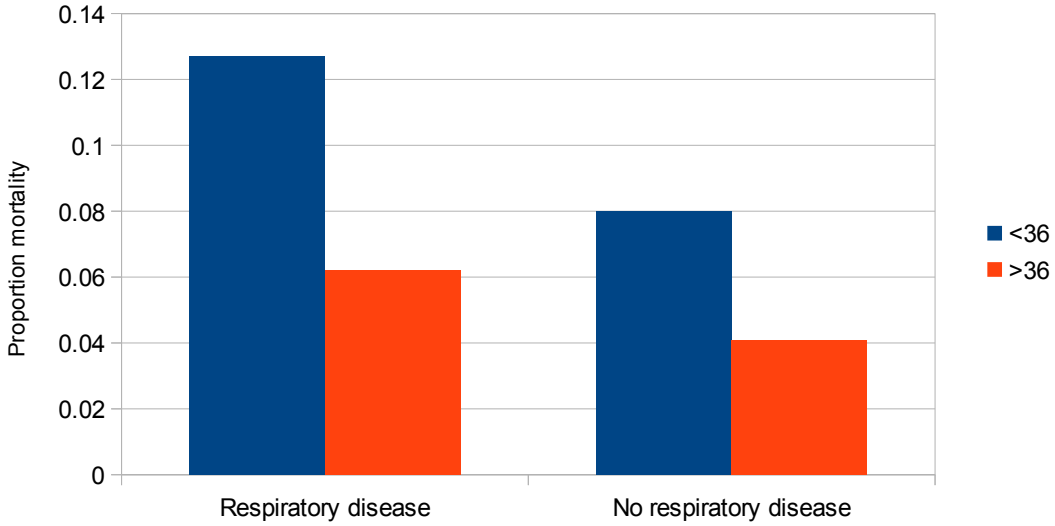
Temperature by pulse pressure

	B	Sig.	Exp(B)
Temperature <36	2.191	<0.001	8.948
Pulse pressure		<0.001	
Pulse pressure <40	*1.396*	*<0.001*	*4.037*
Pulse pressure 40-52	*1.009*	*<0.001*	*2.743*
Pulse pressure >80	*1.696*	*<0.001*	*5.453*
Pulse pressure by respiratory disease		<0.001	
Pulse pressure <40 by temperature <36	*0.702*	*0.020*	*2.017*
Pulse pressure 40-52 by temperature <36	*0.432*	*0.131*	*1.541*
Pulse pressure >80 by temperature <36	*1.440*	*<0.001*	*4.222*



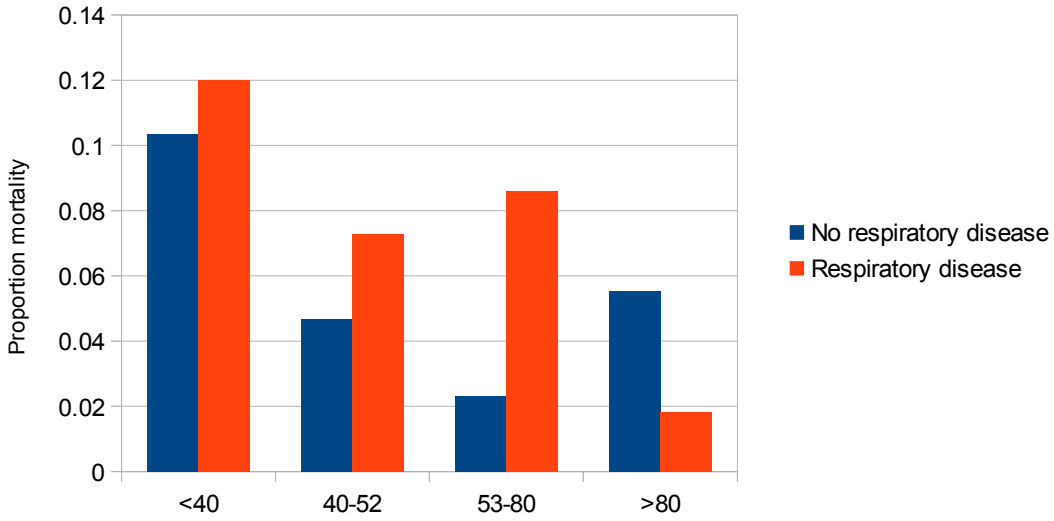
Temperature by respiratory disease

	B	Sig.	Exp(B)
Respiratory disease present	3.972	<0.001	53.107
Temperature <36	3.326	<0.001	27.824
Respiratory disease present and temperature <36	4.599	<0.001	99.382



Pulse pressure by respiratory disease

	B	Sig.	Exp(B)
Respiratory disease present	3.608	<0.001	36.888
Pulse pressure		<0.001	
Pulse pressure <40	*2.862*	*<0.001*	*17.491*
Pulse pressure 40-52	*1.419*	*<0.001*	*4.132*
Pulse pressure >80	*1.341*	*<0.001*	*3.824*
Pulse pressure by respiratory disease		<0.001	
Pulse pressure <40 by respiratory disease present	*3.447*	*<0.001*	*31.391*
Pulse pressure 40-52 by respiratory disease present	*1.408*	*0.002*	*4.088*
Pulse pressure >80 by respiratory disease present	*0.773*	*0.069*	*2.166*


0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
Proportion mortality
<40
40-52
53-80
>80
No respiratory disease
Respiratory disease

Appendix 8: Patients with outcomes of interest in the derivation cohort

Patients sustaining inevitable death (n=15)

Reference	Age/gender	Description
271	91F	Presented unresponsive, died day 0
767	86M	Presented collapse, died day 6
1143	89F	Presented unwell, died day 7
728	95F	Presented with fall, died day 5
248	67M	Presented difficulty breathing, died day 0
242	69M	Presented heart failure, died day 1
752	81F	Presented difficulty breathing, died day 1
945	75M	Presented hypothermic, died day 6
409	73F	Lung carcinoma, presented short of breath, died day 1
375	83F	Presented respiratory failure, died day 7
1016	66M	Presented collapse, died day 2
430	95F	Presented short of breath, died day 0
508	94F	Presented unwell, died day 0
568	79F	Presented with possible cardiac arrest prehospital, died day 1
730	86M	Presented unwell, died day 0

Patients sustaining preventable death (n=5)

Reference	Age/gender	Description
765	74M	Presented unwell, CPAP, died day 1
490	84M	Presented abdominal pain, died day 2
475	75M	Presented short of breath, acute MI, died day 4
398	86M	Presented unwell, pneumonia, died day 7
379	73F	Presented short of breath, died day 2 from dissecting aneurysm not visible on CTPA day 1

Patients in whom death was prevented (n=79)

Reference	Age/gender	Description
232	57M	Presented with potential abdominal problem, cardiac catheter day 0
233	84M	Presented difficulty breathing, CPAP day 0
277	56M	Presented difficulty breathing, appendicectomy day 1
303	20M	Presented with potential pneumonia, chest drain day 0
305	58F	Presented abdominal pain, appendicectomy day 1
316	36M	Presented chest pain, thrombolysis day 0
327	73M	Presented short of breath, CPAP day 0
346	56M	Presented difficulty breathing, CPAP day 1
373	74F	Presented difficulty breathing, CPAP day 0
493	74F	Presented short of breath, BIPAP day 0
449	38F	Presented unwell, ICU
534	46M	Presented fitting, ICU
535	65M	Presented chest pain, cardiac catheter
543	84M	Presented vomiting blood, OGD with adrenaline injection
553	66F	Presented exacerbation COPD, BiPAP day 0
591	41F	Presented unwell, ICU day 2

623	36M	Presented epigastric pain, appendicectomy day 1
628	70F	Presented unwell, ICU day 0
486	57M	Presented overdose, NAC
506	21F	Presented overdose, NAC
338	94F	Presented overdose, NAC
693	76F	Presented chest pain, cardiac cath day 3
629	80M	Presented chest pain, IV antibiotics sepsis
1070	69F	Presented abdominal pain, transfusion for upper GI bleed
560	94F	Presented collapse, cardioversion day 0
621	72M	Presented with potential chest infection, IV antibiotics given for sepsis of unknown source
638	42F	Presented chest/abdominal pain, CPAP day 1
749	76M	Presented diabetic, IV antibiotics for sepsis
779	69F	Presented fall, IV antibiotics and HDU day 4 for LRTI
257	89M	Presented pyrexic, IV antibiotics urosepsis (died after day 7)
370	77M	Presented fall, IV antibiotics urosepsis day 2
473	77M	Presented short of breath, CPAP day 1
928	56M	Presented short of breath, BiPAP day 0
695	68F	Presented short of breath, BiPAP day 0
589	23F	Presented pneumonia, IV antibiotics and fluids
548	48F	Presented collapse, HDU IV antibiotics/antivirals for potential encephalitis
234	61M	Presented collapse, IV antibiotics urosepsis
608	49F	Presented asthma, IV MgSO4
351	85F	Presented fast AF, IV antibiotics for sepsis of unknown source
551	39M	Presented collapse, IV antibiotics HDU for DKA
530	73F	Presented fall, 2L IVI for postural hypotension
935	63F	Presented chest pain, cardiac catheter day 5
753	26F	Presented unresponsive, naloxone
1066	68M	Presented chest pain, thrombolysis day 0
737	89M	Presented fall, IV antibiotics pneumonia
1109	39M	Presented overdose, NAC
722	40F	Presented asthma, IV MgSO4/aminophylline
794	74M	Presented collapse, IV antibiotics sepsis of unknown source
1136	39F	Presented overdose, NAC
538	79F	Presented unwell, IV antibiotics sepsis day 2
396	75M	Presented abdo pain, IV antibiotics sepsis of unknown source
549	73M	Presented short of breath, cardiac catheter day 6
771	18M	Presented hyperglycaemia, IVI/insulin DKA
633	76M	Presented high temp, IV antibiotics urosepsis
1097	76F	Presented difficulty breathing, BiPAP day 0
786	86M	Presented CVA, intubated post-fit
1181	92F	Presented unwell, IV antibiotics sepsis day 2
476	83F	Presented unwell, IV antibiotics sepsis secondary to pneumonia
776	37M	Presented fall, ICU day 0
696	23M	Presented hyperglycaemic, HDU/IVI/insulin DKA
1152	42F	Presented OD, flumazenil
667	69F	Presented chest/abdo pain, IV antibiotics biliary sepsis

953	33F	Presented collapse, neurosurgery for SAH day 2
1146	79M	Presented social problems, IVI rhabdomyolysis
1129	90M	Presented with potential MI, pacemaker day 4
700	30M	Presented overdose, NAC
931	55M	Presented chest pain, cardiac catheter day 1
763	55M	Presented vomiting blood, OGD/adrenaline
710	75M	Presented collapse, ICU day 0
660	42F	Presented overdose of insulin, IV dextrose
679	40M	Presented overdose, ICU day 0
1159	50F	Presented overdose, NAC
922	31M	Presented intoxicated, naloxone
1048	78F	Presented with potential GI bleed, 4unit transfusion
949	82F	Presented near faint, adenosine cardioversion
1094	72F	Presented acutely short of breath, IV GTN/frusemide
947	71F	Presented palpitations, metoprolol cardioversion
442	93F	Presented short of breath, IV frusemide/nitrate/MgSO4
417	82M	Presented abdominal pain, IV pamidronate for hypercalcaemia

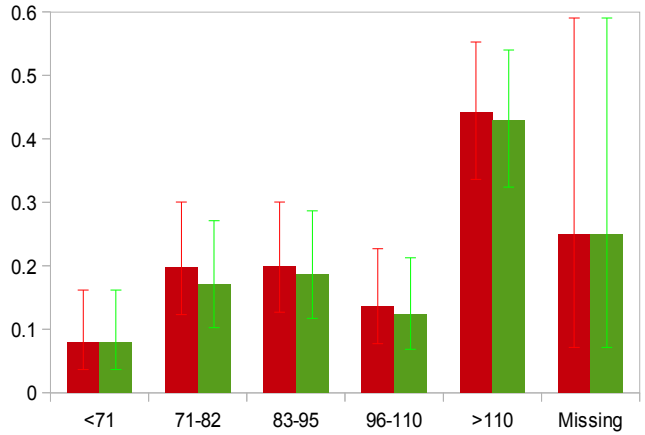
Appendix 9: Univariate analysis of continuous variables in predicting potentially prevented and potentially preventable death at 7 days

Tables present numbers (percentages) with each outcome.
Each graph presents proportions of potentially prevented (green) and potentially prevented plus potentially preventable (red) deaths, with 95% confidence intervals.

Pulse rate
By quintiles of population

	Potentially preventable + potentially prevented	Potentially prevented	Total
<71	6 (8)	6 (8)	76

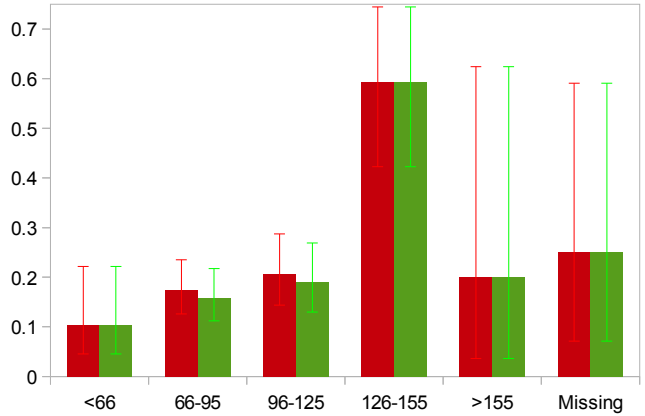
71-82	15 (20)	13 (17)	76
83-95	16 (20)	15 (19)	80
96-110	11 (14)	10 (12)	81
>110	34 (44)	33 (43)	77
Missing	2 (25)	2 (25)	8
Total	84 (21)	79 (20)	398
Chi-sq p	<0.001	<0.001	



By quintiles of value

	Potentially preventable + potentially prevented	Potentially prevented	Total
<66	5 (10)	5 (10)	48
66-95	32 (17)	29 (16)	184

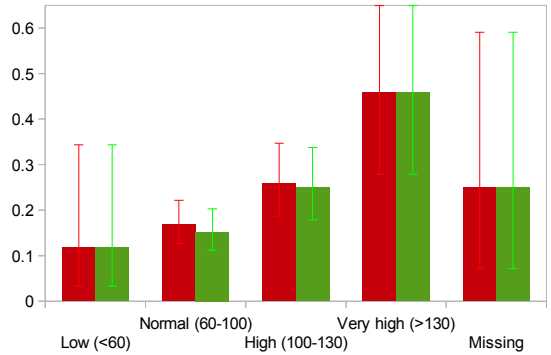
96-125	25 (21)	23 (19)	121
126-155	19 (59)	19 (59)	32
>155	1 (20)	1 (20)	5
Missing	2 (25)	2 (25)	8
Total	84 (21)	79 (20)	398
Chi-sq p	<0.001	<0.001	



By groups a priori from DAVROS study

	Potentially preventable+ potentially prevented	Potentially prevented	Total
Low (<60)	2 (12)	2 (12)	17
Normal (60-100)	40 (17)	36 (15)	237

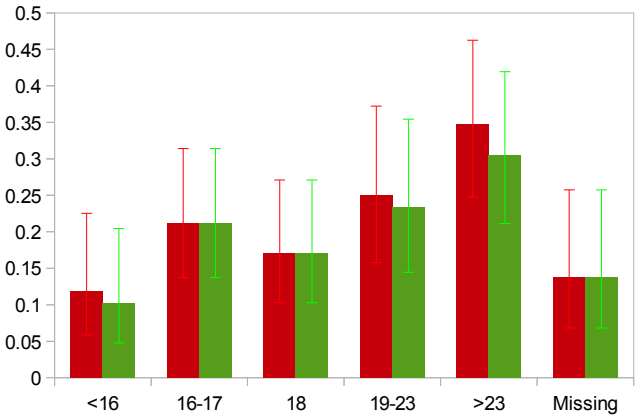
High (100-130)	29 (26)	28 (25)	112
Very high (>130)	11 (46)	11 (46)	24
Missing	2 (25)	2 (25)	8
Total	84 (21)	79 (20)	398
Chi-sq p	0.008	0.003	



Respiratory rate
By quintiles of population

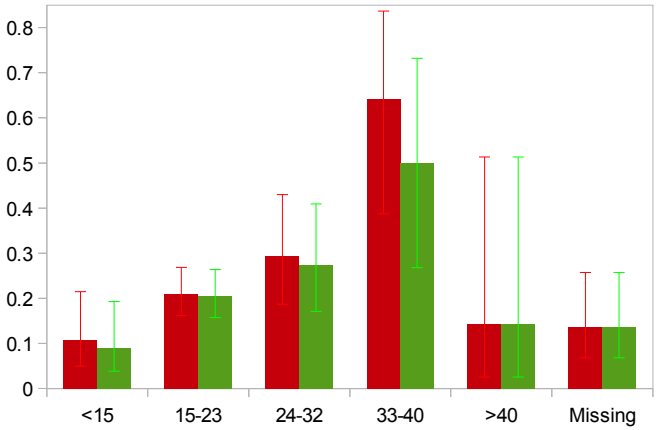
	Potentially preventable + potentially prevented	Potentially prevented	Total
<16	7 (12)	6 (10)	59
16-17	17 (21)	17 (21)	80
18	13 (17)	13 (17)	76
19-23	15 (25)	14 (23)	60
>23	25 (35)	22 (31)	72
Missing	7 (14)	7 (14)	51
Total	84 (21)	79 (20)	398

Chi-sq p	0.016	0.056



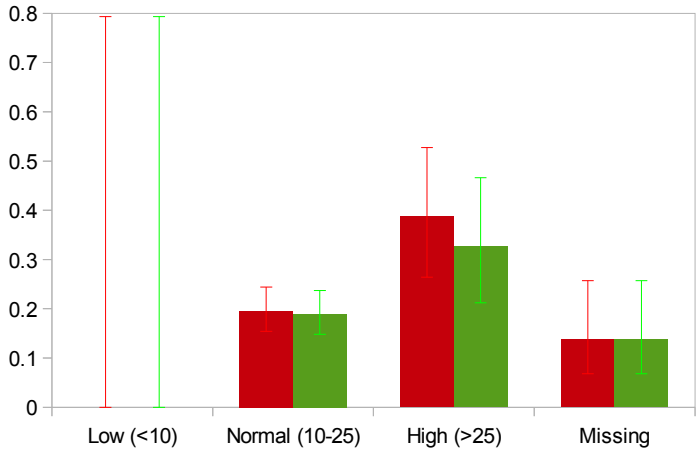
By quintiles of value

	Potentially preventable + potentially prevented	Potentially prevented	Total
<15	6 (11)	5 (9)	56
15-23	46 (21)	45 (21)	219
24-32	15 (29)	14 (27)	51
33-40	9 (64)	7 (50)	14
>40	1 (14)	1 (14)	7
Missing	7 (14)	7 (14)	51
Total	84 (21)	79 (20)	398
Chi-sq p	<0.001	0.009	



By groups a priori from DAVROS study

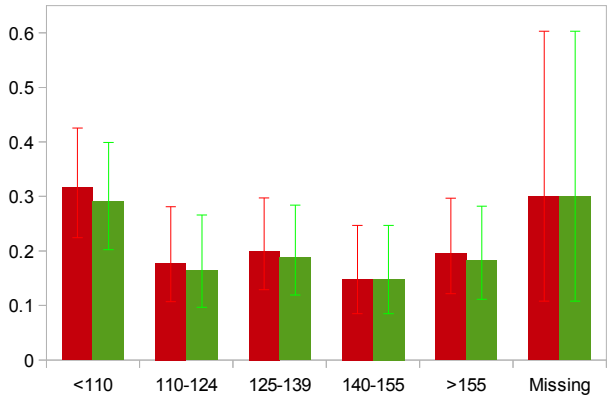
	Potentially preventable + potentially prevented	Potentially prevented	Total
Low (<10)	0	0	1
Normal (10-25)	58 (20)	56 (19)	297
High (>25)	19 (39)	16 (33)	49
Missing	7 (14)	7 (14)	51
Total	84 (21)	79 (20)	398
Chi-sq p	0.009	0.083	



Systolic blood pressure
By quintiles of population

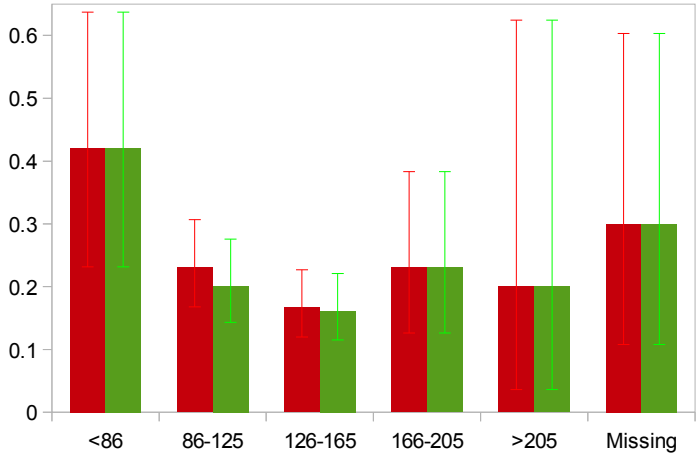
	Potentially preventable + potentially prevented	Potentially prevented	Total
<110	25 (32)	23 (29)	79
110-124	13 (18)	12 (16)	73
125-139	17 (20)	16 (19)	85
140-155	11 (15)	11 (15)	74
>155	15 (19)	14 (18)	77
Missing	3 (30)	3 (30)	10
Total	84 (21)	79 (20)	398

Chi-sq p	0.149	0.237



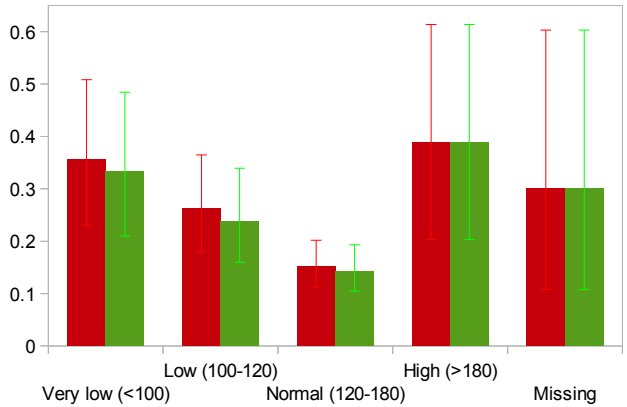
By quintiles of value

	Potentially preventable + potentially prevented	Potentially prevented	Total
<86	8 (42)	8 (42)	19
86-125	32 (23)	28 (20)	139
126-165	31 (17)	30 (16)	186
166-205	9 (23)	9 (23)	39
>205	1 (20)	1 (20)	5
Missing	3 (30)	3 (30)	10
Total	84 (21)	79 (20)	398
Chi-sq p	0.150	0.133	



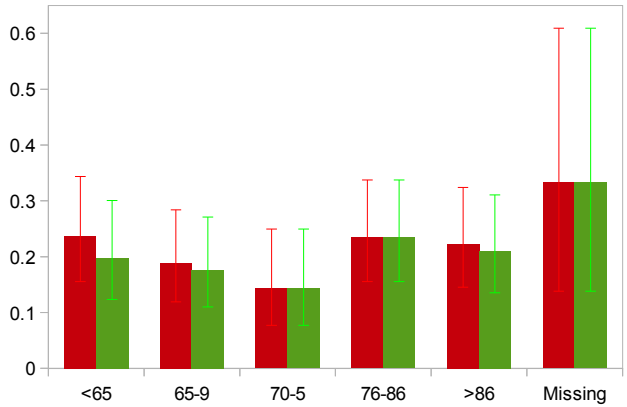
By groups a priori from DAVROS study

	Potentially preventable + potentially prevented	Potentially prevented	Total
Very low (<100)	15 (36)	14 (33)	42
Low (100-120)	22 (26)	20 (24)	84
Normal (120-180)	37 (15)	35 (14)	244
High (>180)	7 (39)	7 (39)	18
Missing	3 (30)	3 (30)	10
Total	84 (21)	79 (20)	398
Chi-sq p	0.003	0.005	



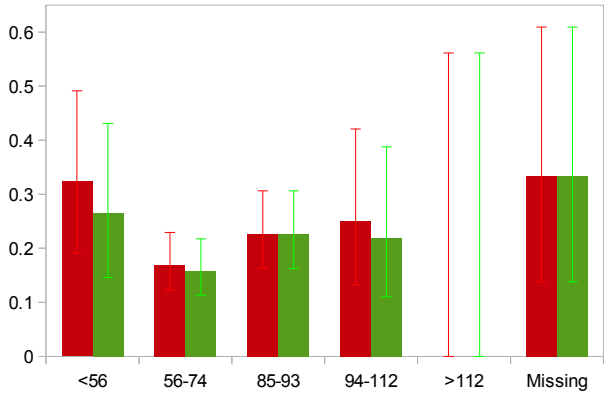
Diastolic blood pressure
By quintiles of population

	Potentially preventable + potentially prevented	Potentially prevented	Total
<65	18 (24)	15 (20)	76
65-9	16 (19)	15 (18)	85
70-75	9 (14)	9 (14)	63
76-86	19 (23)	19 (23)	81
>86	18 (22)	17 (21)	81
Missing	4 (33)	4 (33)	12
Total	84 (21)	79 (20)	398
Chi-sq p	0.588	0.611	



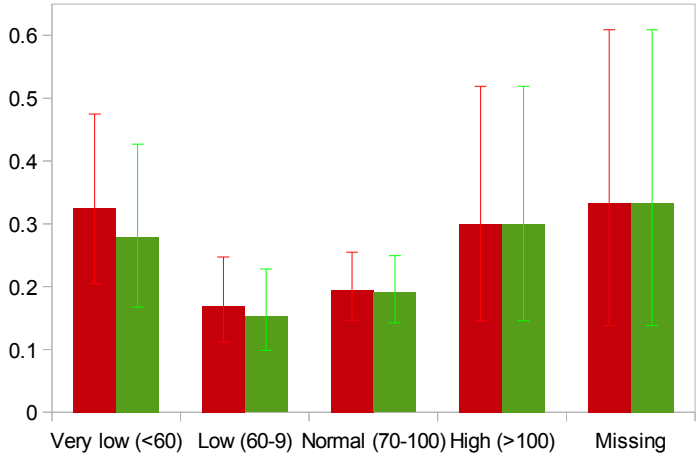
By quintiles of value

	Potentially preventable + potentially prevented	Potentially prevented	Total
<56	11 (32)	9 (26)	34
56-74	32 (17)	30 (16)	189
85-93	29 (23)	29 (23)	128
94-112	8 (25)	7 (22)	32
>112	0	0	3
Missing	4 (33)	4 (33)	12
Total	84 (21)	79 (20)	398
Chi-sq p	0.227	0.342	



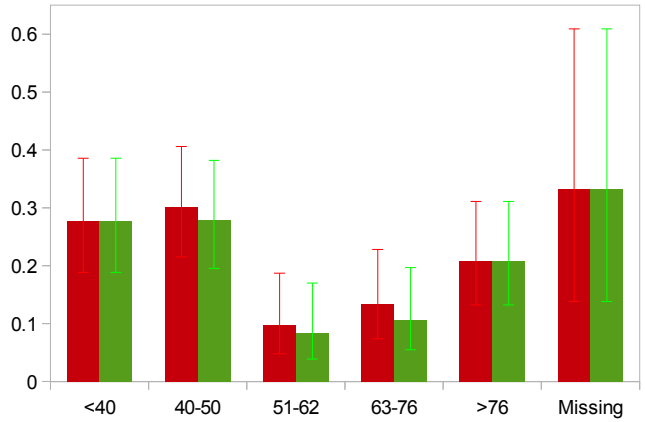
By groups a priori from DAVROS study

	Potentially preventable + potentially prevented	Potentially prevented	Total
Very low (<60)	14 (33)	12 (28)	43
Low (60-69)	20 (17)	18 (15)	118
Normal (70-100)	40 (20)	39 (19)	205
High (>100)	6 (30)	6 (30)	20
Missing	4 (33)	4 (33)	12
Total	84 (21)	79 (20)	398
Chi-sq p	0.138	0.194	



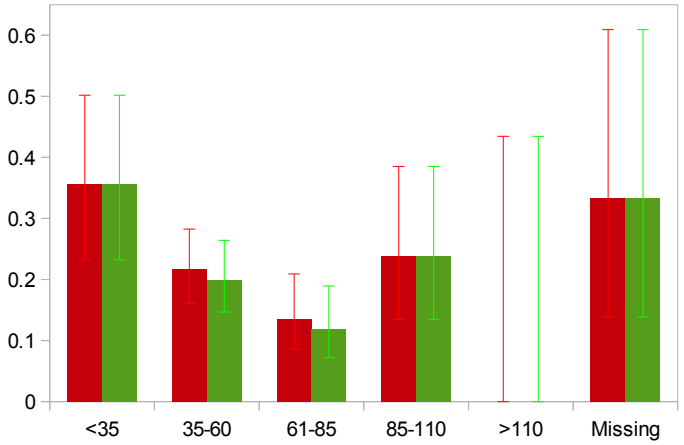
Pulse pressure
By quintiles of population

	Potentially preventable + potentially prevented	Potentially prevented	Total
<40	21 (28)	21 (28)	76
40-50	26 (30)	24 (28)	86
51-62	7 (10)	6 (8)	72
63-76	10 (13)	8 (11)	75
>76	16 (21)	16 (21)	77
Missing	4 (33)	4 (33)	12
Total	84 (21)	79 (20)	398
Chi-sq p	0.005	0.003	



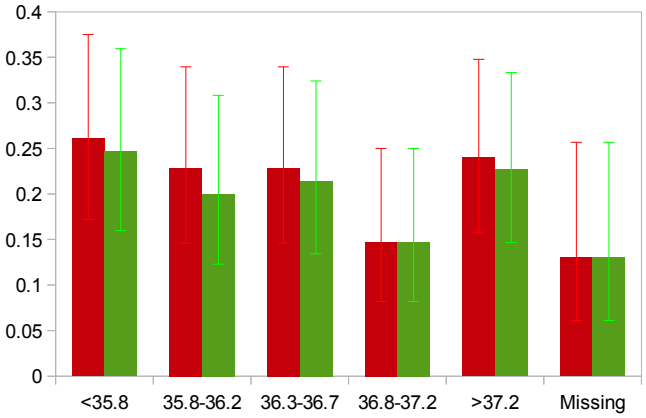
By quintiles of value

	Potentially preventable + potentially prevented	Potentially prevented	Total
<35	16 (36)	16 (36)	45
35-60	38 (22)	35 (20)	176
61-85	16 (14)	14 (12)	118
85-110	10 (24)	10 (24)	42
>110	0	0	5
Missing	4 (33)	4 (33)	12
Total	84 (21)	79 (20)	398
Chi-sq p	0.031	0.012	



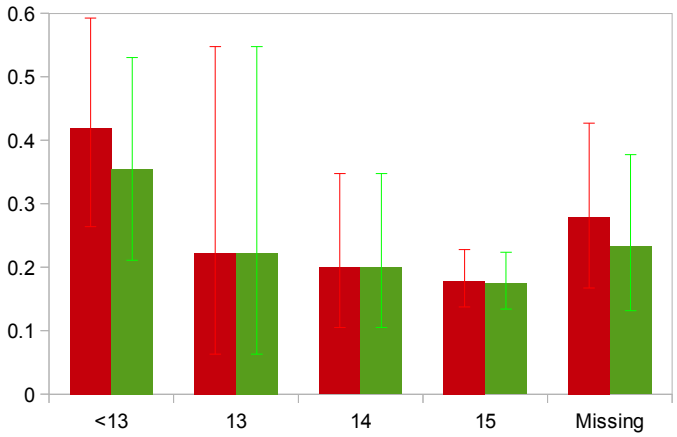
Temperature
By quintiles of population

	Potentially preventable + potentially prevented	Potentially prevented	Total
<35.8	18 (26)	17 (25)	69
35.8-36.2	16 (23)	14 (20)	70
36.3-36.7	16 (23)	15 (21)	70
36.8-37.2	10 (15)	10 (15)	68
>37.2	18 (24)	17 (23)	75
Missing	6 (13)	6 (13)	46
Total	84 (21)	79 (20)	398
Chi-sq p	0.400	0.557	

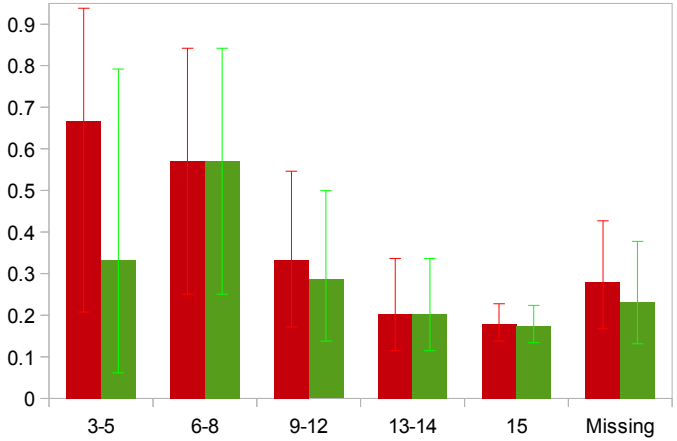


Glasgow Coma Scale
By quartiles of population

	Potentially preventable + potentially prevented	Potentially prevented	Total
<13	13 (42)	11 (35)	31
13	2 (22)	2 (22)	9
14	8 (20)	8 (20)	40
15	49 (18)	48 (17)	275
Missing	12 (28)	10 (23)	43
Total	84 (21)	79 (20)	398
Chi-sq p	0.026	0.192	

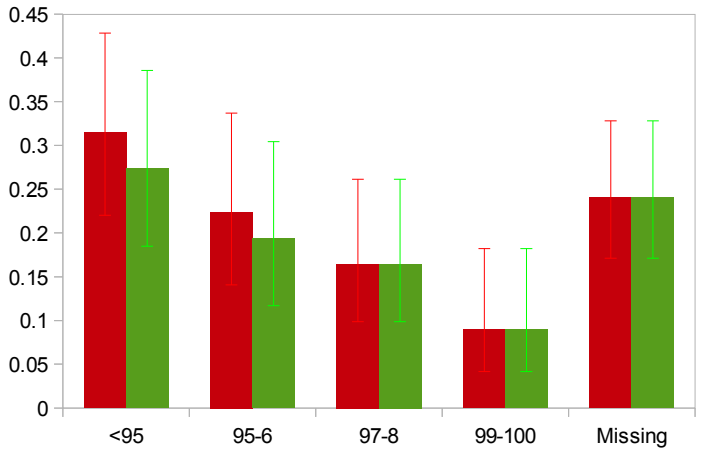

0.6
0.5
0.4
0.3
0.2
0.1
0
<13
13
14
15
Missing

	Potentially preventable + potentially prevented	Potentially prevented	Total
3-5	2 (66)	1 (33)	3
6-8	4 (57)	4 (57)	7
9-12	7 (33)	6 (29)	21
13-14	10 (20)	10 (20)	49
15	49 (18)	48 (17)	275
Missing	12 (28)	10 (23)	43
Total	84 (21)	79 (20)	398
Chi-sq p	0.015	0.118	



Oxygen saturations breathing air
By quartiles of population

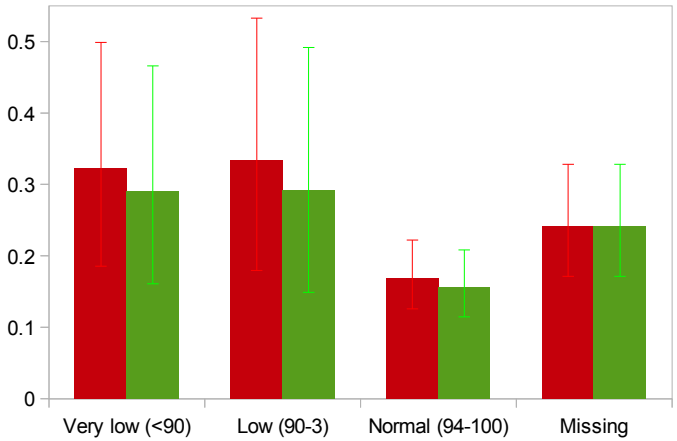
	Potentially preventable + potentially prevented	Potentially prevented	Total
<95	23 (32)	20 (27)	73
95-6	15 (22)	13 (19)	67
97-8	13 (16)	13 (16)	79
99-100	6 (9)	6 (9)	67
Missing	27 (24)	27 (24)	112
Total	84 (21)	79 (20)	398
Chi-sq p	0.015	0.05	



By groups a priori from DAVROS study

	Potentially preventable + potentially	Potentially prevented	Total

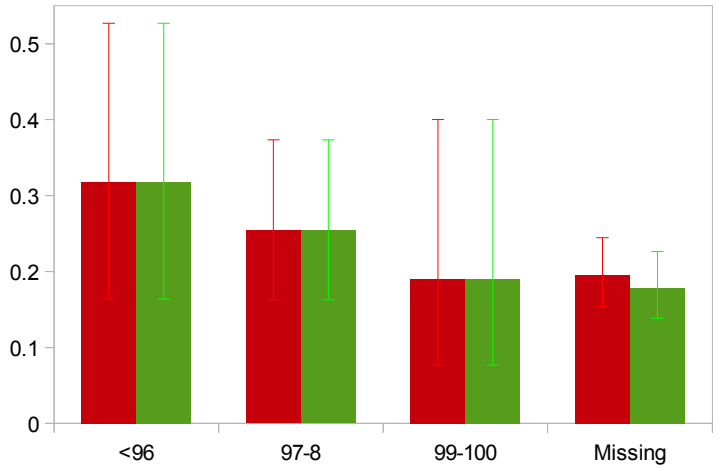
	prevented		
Very low (<90)	10 (32)	9 (29)	31
Low (90-3)	8 (33)	7 (29)	24
Normal (94-100)	39 (17)	36 (16)	231
Missing	27 (24)	27 (24)	112
Total	84 (21)	79 (20)	398
Chi-sq p	0.056	0.076	



Oxygen saturations breathing supplemental oxygen
By tertiles of population

	Potentially preventable + potentially prevented	Potentially prevented	Total
<96	7 (32)	7 (32)	22

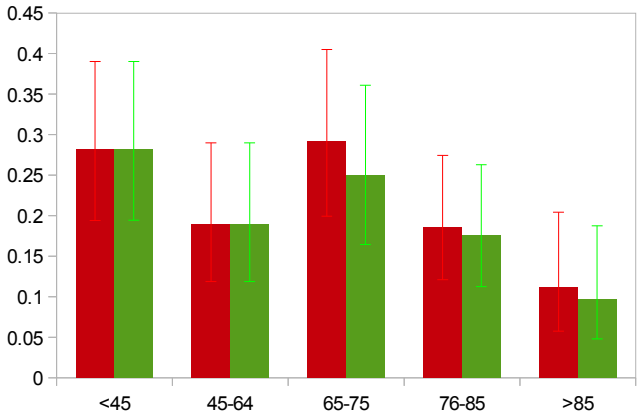
96-99	16 (25)	16 (25)	63
100	4 (19)	4 (19)	21
Missing	57 (20)	52 (18)	292
Total	84 (21)	79 (20)	398
Chi-sq p	0.439	0.264	



Age
By quintiles of population

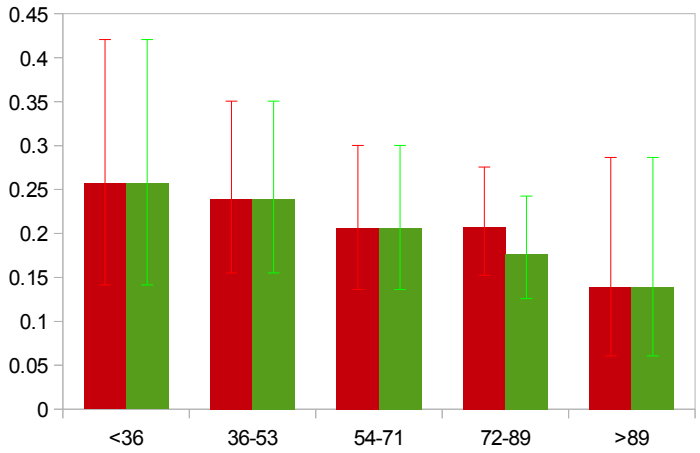
	Potentially preventable + potentially prevented	Potentially prevented	Total
<45	22 (28)	22 (28)	78
45-64	15 (19)	15 (19)	79
65-75	21 (29)	18 (25)	72
76-85	18 (19)	17 (18)	97
>85	8 (11)	7 (10)	72

Total	84 (21)	79 (20)	398
Chi-sq p	0.039	0.047	



By quintiles of value

	Potentially preventable + potentially prevented	Potentially prevented	Total
<36	9 (26)	9 (26)	35
36-53	17 (24)	17 (24)	71
54-71	19 (21)	19 (21)	92
72-89	34 (21)	29 (18)	164
>89	5 (14)	5 (14)	36
Total	84 (21)	79 (20)	398
Chi-sq p	0.747	0.587	


0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
<36
36-53
54-71
72-89
>89

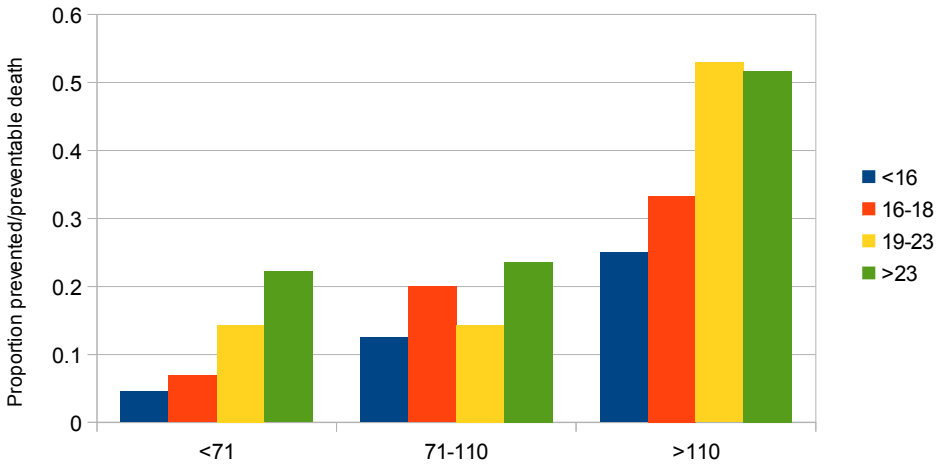
Appendix 10:
Interactions between significant variables in predicting potentially preventable and potentially prevented death

Interaction term	Wald	Degrees of freedom	p
Pulse by respiratory rate	19.290	6	0.004
Pulse by SBP	19.200	6	0.004
Pulse by pulse pressure	9.008	4	0.061
Pulse by GCS	2.449	4	0.654
Pulse by SaO2	20.088	4	<0.001
Respiratory rate by SBP	12.699	9	0.177
Respiratory rate by pulse pressure	16.431	6	0.012
Respiratory rate by GCS	5.453	6	0.487
Respiratory rate by SaO2	14.071	6	0.029
SBP by pulse pressure	32.734	4	<0.001
SBP by GCS	2.424	6	0.877
SBP by SaO2	20.303	6	0.002
Pulse pressure by GCS	5.261	4	0.262
Pulse pressure by SaO2	17.221	4	0.002
GCS by SaO2	7.420	4	0.115

Graphs below present proportion of outcomes of interest by interaction groups.

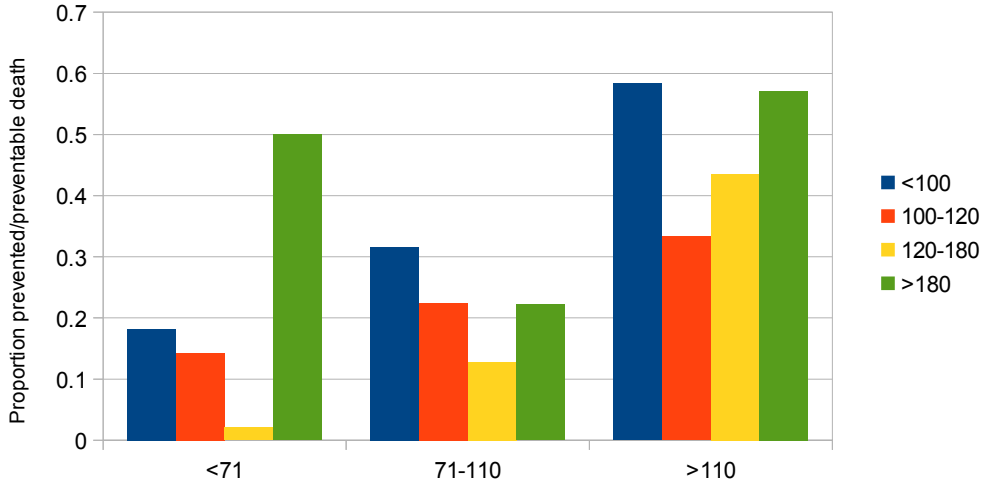
Pulse by respiratory rate

	B	Sig.	Exp(B)
Pulse (ref <71)		<0.001	
71-110	-0.467	0.197	0.627
>110	*2.403*	*<0.001*	*11.060*
Respiratory rate (ref <16)		0.001	
16-18	*-1.957*	*0.002*	*0.141*
19-23	-1.017	0.181	0.362
>23	*-1.005*	*0.147*	*0.366*
Pulse by respiratory rate		0.004	
71-110 and 16-18	*0.122*	*0.891*	*1.130*
71-110 and 19-23	*-1.731*	*0.119*	*0.177*
71-110 and >23	*-1.189*	*0.251*	*0.304*
>110 and 16-18	*-5.801*	*0.003*	*0.003*
>110 and 19-23	*-6.683*	*0.001*	*0.001*
>110 and >23	*-7.001*	*<0.001*	*0.001*



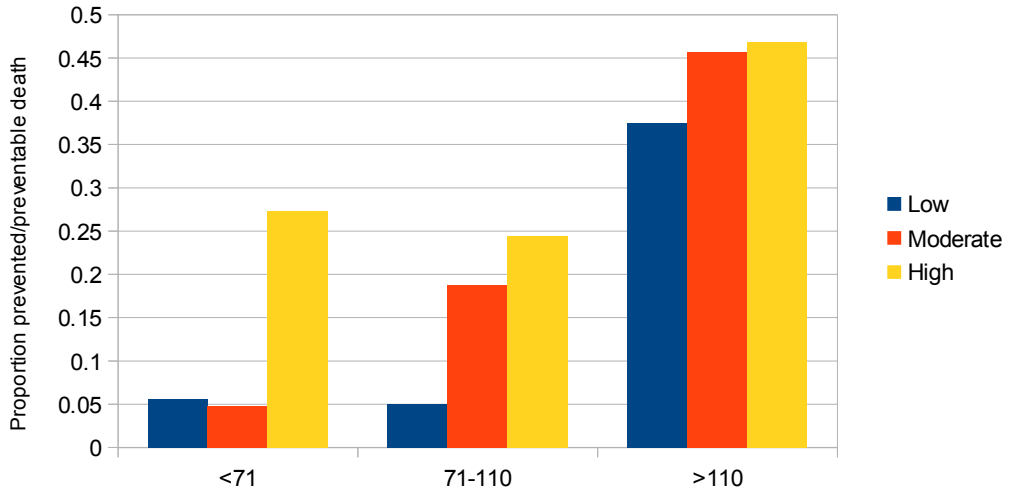
Pulse by SBP

	B	Sig.	Exp(B)
Pulse (ref <71)		<0.001	
71-110	-1.870	0.001	0.154
>110	*-0.559*	*0.381*	*0.572*
SBP (ref 121-180)		<0.001	
<100	*1.442*	*0.001*	*4.230*
100-120	*0.752*	*0.060*	*2.122*
>180	*4.677*	*<0.001*	*107.470*
Pulse by SBP		0.004	
71-100 and <100	*-0.830*	*0.461*	*0.436*
71-100 and 100-120	*-1.117*	*0.286*	*0.327*
71-100 and >180	*-10.620*	*<0.001*	*0*
>110 and <100	*-1.307*	*0.279*	*0.270*
>110 and 100-120	*-2.134*	*0.060*	*0.118*
>110 and >180	*-10.720*	*<0.001*	*0*



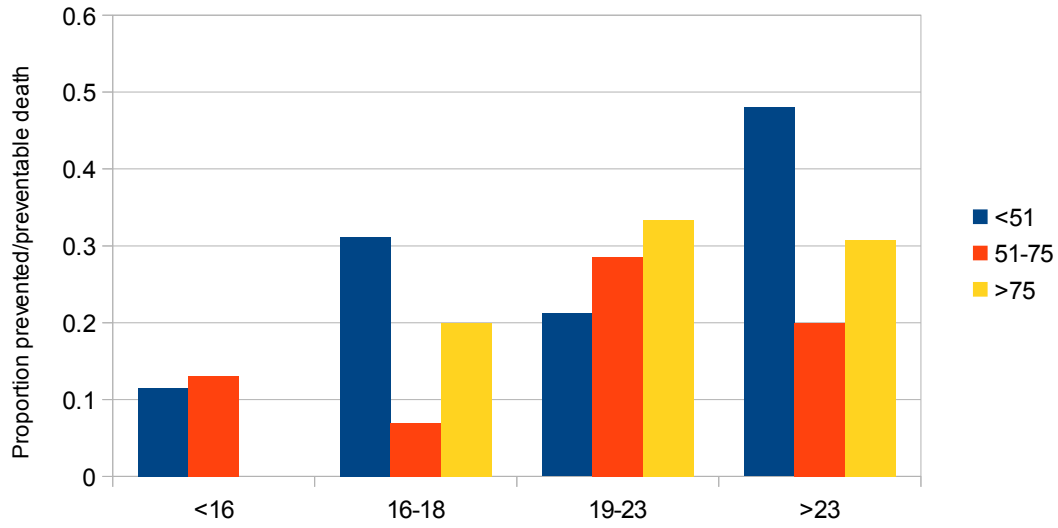
Pulse by SaO2

	B	Sig.	Exp(B)
Pulse (ref <71)		<0.001	
71-110	-0.678	0.057	0.507
>110	*2.121*	*<0.001*	*8.343*
SaO2 (ref low risk)		<0.001	
High risk	*-0.324*	*0.554*	*0.723*
Moderate risk	*-1.413*	*0.002*	*0.243*
Pulse by SaO2		<0.001	
71-100 and high risk	*-0.716*	*0.464*	*0.488*
71-100 and moderate risk	*1.222*	*0.128*	*3.393*
>110 and high risk	*-5.193*	*<0.001*	*0.006*
>110 and moderate risk	*-2.745*	*0.056*	*0.064*



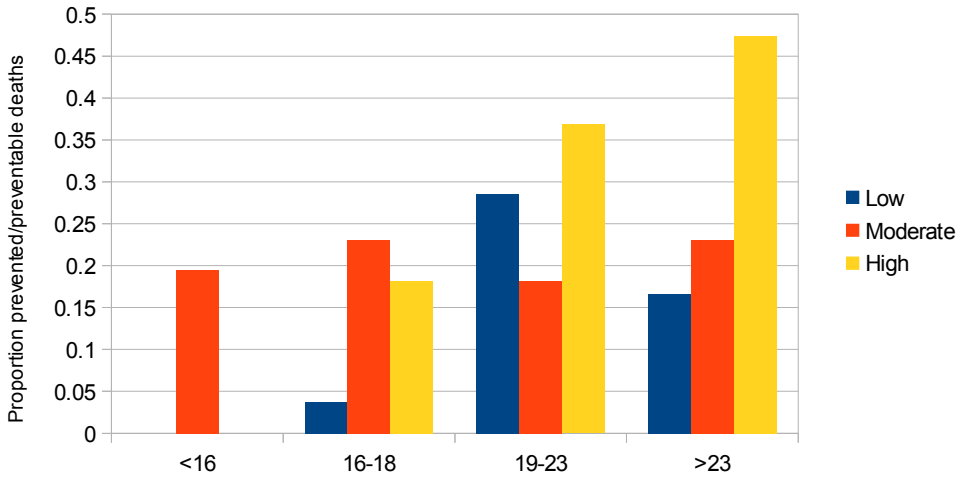
Respiratory rate by pulse pressure

	B	Sig.	Exp(B)
Respiratory rate (ref <16)		<0.001	
16-18	*-0.632*	*0.065*	*0.532*
19-23	*0.805*	*0.073*	*2.237*
>23	*0.719*	*0.085*	*2.053*
Pulse pressure (ref 51-75)		*0.013*	
<51	*0.272*	*0.345*	*1.313*
>75	*1.052*	*0.004*	*2.864*
Respiratory rate by pulse pressure		0.012	
16-18 and <51	*1.412*	*0.057*	*4.106*
16-18 and >75	*0.010*	*0.992*	*1.010*
19-23 and <51	*-1.085*	*0.227*	*0.338*
19-23 and >75	*-0.414*	*0.737*	*0.661*
>23 and <51	*1.521*	*0.067*	*4.577*
>23 and >75	*0.483*	*0.660*	*1.621*



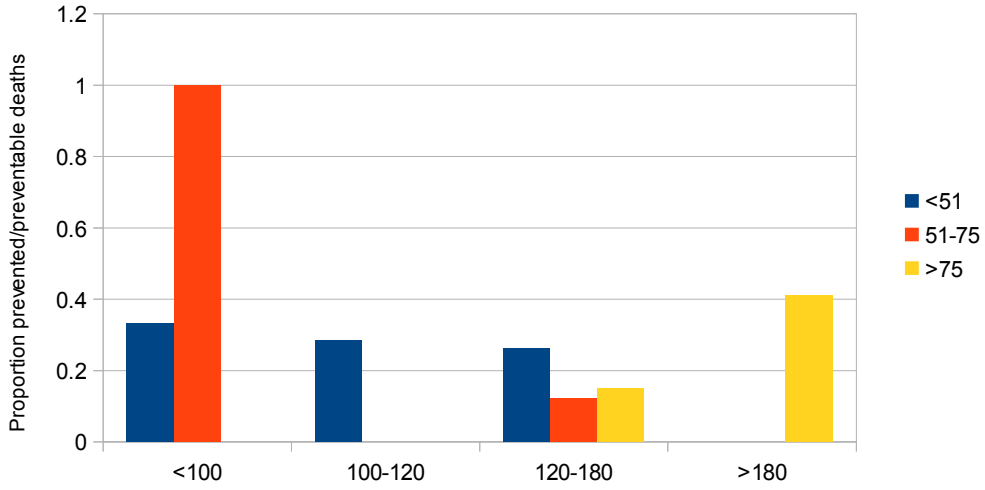
Respiratory rate by SaO2

	B	Sig.	Exp(B)
Respiratory rate (ref <16)		<0.001	
16-18	*-0.534*	*0.176*	*0.586*
19-23	*1.074*	*0.058*	*2.927*
>23	*1.127*	*0.045*	*3.085*
SaO2 (ref low risk)		<0.001	
High risk	*-0.496*	*0.285*	*0.609*
Moderate risk	*-1.349*	*<0.001*	*0.260*
Respiratory rate by SaO2		0.029	
16-18 and high risk	*0.036*	*0.975*	*1.037*
16-18 and moderate risk	*0.506*	*0.560*	*1.658*
19-23 and high risk	*-2.813*	*0.065*	*0.060*
19-23 and moderate risk	*-3.028*	*0.034*	*0.048*
>23 and high risk	*-2.685*	*0.068*	*0.068*
>23 and moderate risk	*-2.487*	*0.085*	*0.083*



SBP by pulse pressure

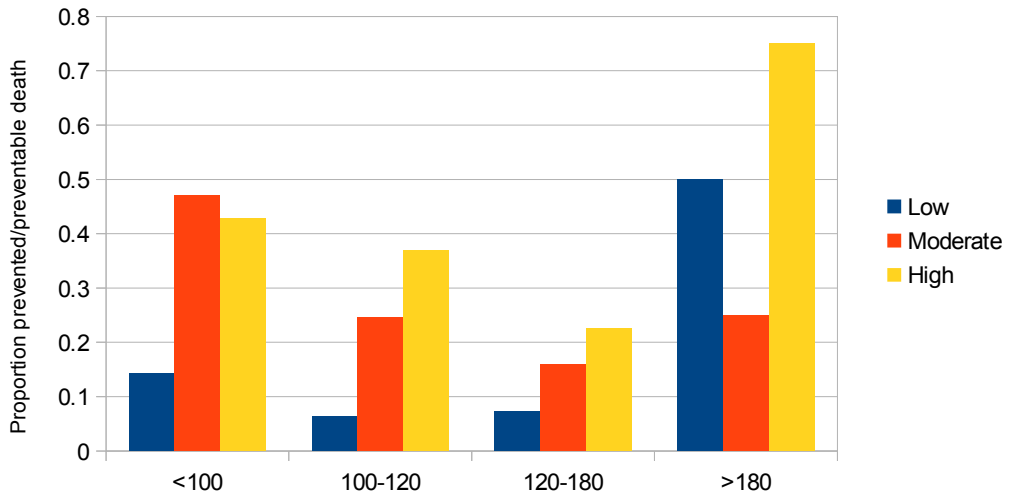
	B	Sig.	Exp(B)
SBP (ref 121-180)		<0.001	
<100	*15.556*	*1*	*>1000*
100-120	*-12.779*	*0.999*	*>1000*
>180	*3.544*	*1*	*34.59*
Pulse pressure (ref 51-75)		1	
<51	*11.992*	*1*	*>1000*
>75	*5.39*	*1*	*219.174*
Pulse pressure by SBP		<0.001	
<100 and <51	*-22.827*	*1*	*>1000*
100-120 and <51	*19.356*	*0.999*	*>1000*
>180 and <51	*47.714*	*1*	*>1000*
>180 and >75	*20.608*	*1*	*>1000*



SBP by SaO2

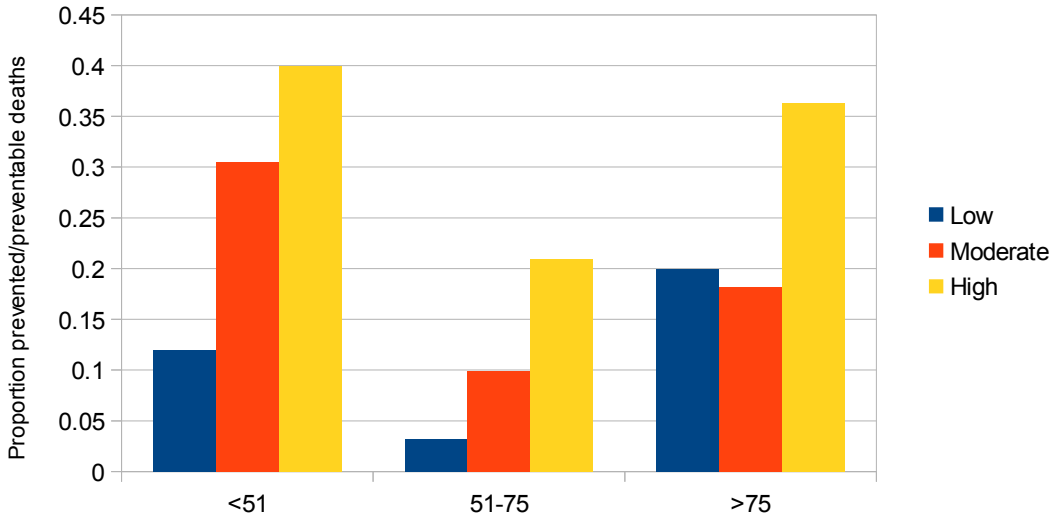
	B	Sig.	Exp(B)

SBP (ref 121-180)		<0.001	
<100	*1.384*	*0.001*	*3.990*
100-120	*0.556*	*0.147*	*1.744*
>180	*4.694*	*<0.001*	*109.311*
SaO2 (ref low risk)		0.001	
High risk	*1.091*	*0.661*	*2.976*
Moderate risk	*-0.667*	*0.604*	*0.513*
SBP by SaO2		0.002	
<100 and high risk	*-0.179*	*0.878*	*0.836*
100-120 and high risk	*0.531*	*0.624*	*1.7*
>180 and high risk	*-0.406*	*0.967*	*0.666*
<100 and moderate risk	*0.433*	*0.701*	*1.541*
100-120 and moderate risk	*0.320*	*0.749*	*1.377*
>180 and moderate risk	*-5.867*	*0.252*	*0.003*



Pulse pressure by SaO2

	B	Sig.	Exp(B)
Pulse pressure (ref 51-75)		<0.001	
<51	*0.835*	*0.003*	*2.306*
>75	*2.340*	*<0.001*	*10.384*
SaO2 (ref low risk)		*<0.001*	
High risk	*0.894*	*0.121*	*2.444*
Moderate risk	*-0.963*	*0.006*	*0.382*
Pulse pressure by SaO2		0.002	
<51 and high risk	*0.266*	*0.685*	*1.305*
<51 and moderate risk	*0.281*	*0.685*	*1.325*
>75 and high risk	*0.701*	*0.661*	*2.016*
>75 and moderate risk	*-2.257*	*0.026*	*0.105*


0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Proportion prevented/preventable deaths
<51
51-75
>75
Low
Moderate
High

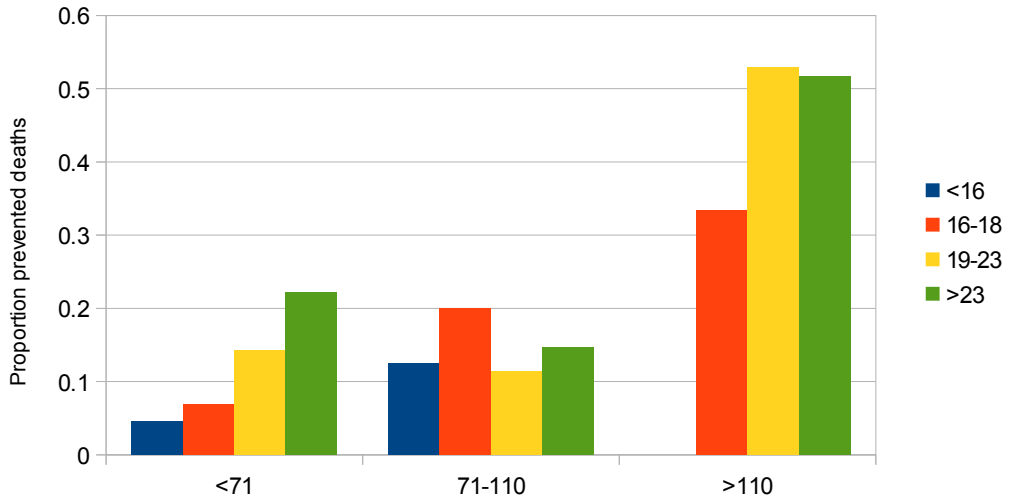
Appendix 11
Interactions between significant variables in predicting potentially prevented death

Interaction term	Wald	Degrees of freedom	p
Pulse by respiratory rate	16.646	6	0.011
Pulse by SBP	21.388	6	0.002
Pulse by pulse pressure	8.874	4	0.064
Pulse by SaO2	3.595	4	0.464
Respiratory rate by SBP	13.286	9	0.150
Respiratory rate by pulse pressure	14.386	6	0.026
Respiratory rate by SaO2	16.403	6	0.012
SBP by pulse pressure	36.601	4	<0.001
SBP by SaO2	22.282	6	0.001
Pulse pressure by SaO2	19.012	4	0.001

Graphs below present proportion of prevented deaths by interaction groups.

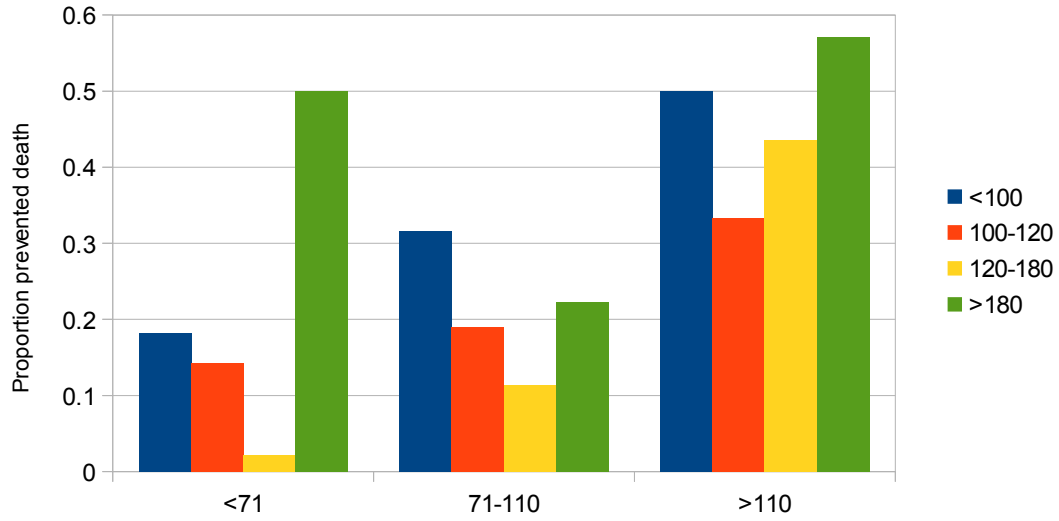
Pulse by respiratory rate

	B	Sig.	Exp(B)
Pulse (ref <71)		<0.001	
71-110	*-0.861*	*0.018*	*0.423*
>110	*1.689*	*0.006*	*5.413*
Respiratory rate (ref <16)		0.003	
16-18	*-1.295*	*0.03*	*0.274*
19-23	*-0.220*	*0.769*	*0.803*
>23	*-0.397*	*0.556*	*0.672*
Pulse by respiratory rate		0.011	
71-110 and 16-18	*0.061*	*0.941*	*1.063*
71-110 and 19-23	*-2.220*	*0.048*	*0.109*
71-110 and >23	*-1.818*	*0.079*	*0.162*
>110 and 16-18	*-3.600*	*0.049*	*0.027*
>110 and 19-23	*-4.746*	*0.010*	*0.009*
>110 and >23	*-5.062*	*0.004*	*0.006*



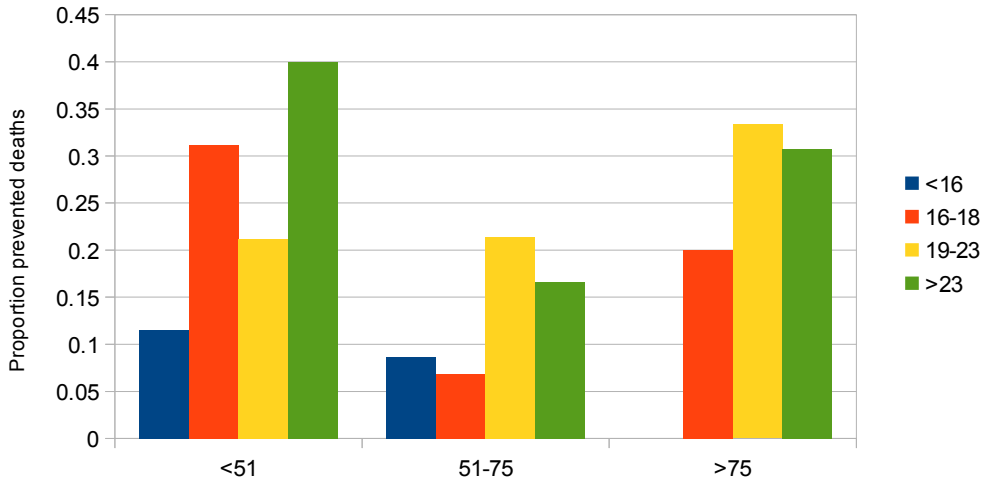
Pulse by SBP

	B	Sig.	Exp(B)
Pulse (ref <71)		<0.001	
71-110	*-2.119*	*<0.001*	*0.120*
>110	*-0.817*	*0.197*	*0.442*
SBP (ref 120-180)		*<0.001*	
<100	*1.364*	*0.002*	*3.911*
100-120	*0.726*	*0.070*	*2.068*
>180	*4.943*	*<0.001*	*140.137*
Pulse by SBP		0.002	
71-100 and <100	*-0.709*	*0.529*	*0.492*
71-100 and 100-120	*-1.195*	*0.258*	*0.303*
71-100 and >180	*-11.174*	*<0.001*	*0*
>110 and <100	*-1.664*	*0.163*	*0.189*
>110 and 120-180	*-2.134*	*0.06*	*0.118*
>110 and >180	*-11.396*	*<0.001*	*0*



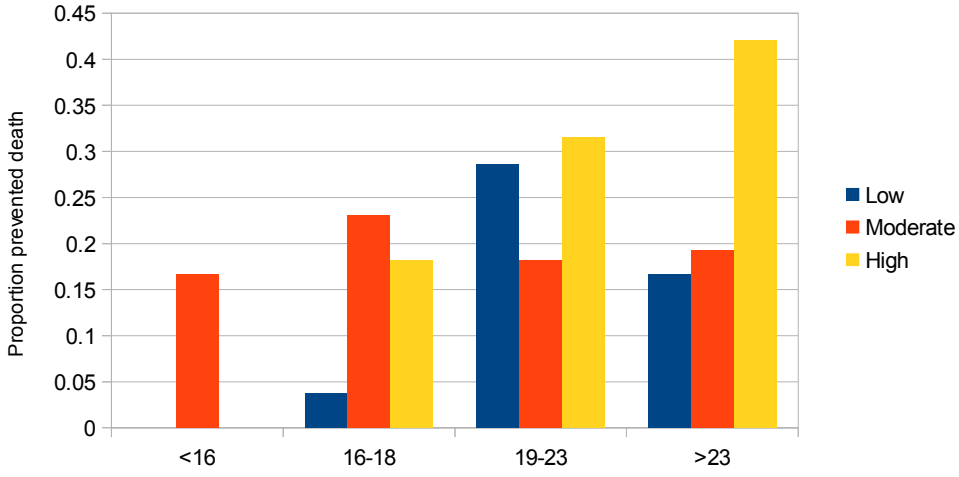
Respiratory rate by pulse pressure

	B	Sig.	Exp(B)
Respiratory rate (ref <16)		<0.001	
16-18	*-0.631*	*0.067*	*0.532*
19-23	*0.780*	*0.089*	*2.181*
>23	*0.601*	*0.155*	*1.824*
Pulse pressure (ref 51-75)		*0.002*	
<51	*0.324*	*0.253*	*1.383*
>75	*1.286*	*<0.00 1*	*3.620*
Respiratory rate by pulse pressure		0.026	
16-18 and <51	*1.211*	*0.102*	*3.357*
16-18 and >75	*-0.283*	*0.766*	*0.754*
19-23 and <51	*-1.013*	*0.255*	*0.363*
19-23 and >75	*-0.30*	*0.813*	*0.741*
>23 and <51	*1.130*	*0.170*	*3.095*
>23 and >75	*0.429*	*0.704*	*1.535*



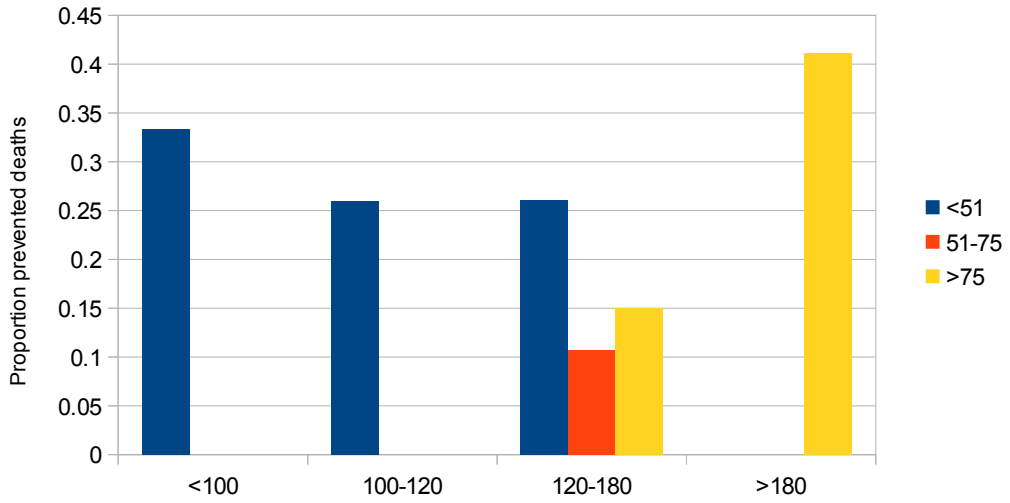
Respiratory rate by SaO2

	B	Sig.	Exp(B)
Respiratory rate (ref <16)		<0.001	
16-18	*-0.511*	*0.197*	*0.6*
19-23	*1.100*	*0.063*	*3.003*
>23	*1.096*	*0.063*	*2.993*
SaO2 (ref low risk)		<0.001	
High risk	*-0.710*	*0.134*	*0.492*
Moderate risk	*-1.544*	*<0.00 1*	*0.213*
Respiratory rate by SaO2		0.012	
16-18 and high risk	*0.012*	*0.992*	*1.012*
16-18 and moderate risk	*0.641*	*0.460*	*1.898*
19-23 and high risk	*-3.259*	*0.041*	*0.038*
19-23 and moderate risk	*-3.091*	*0.041*	*0.045*
>23 and high risk	*-3.146*	*0.041*	*0.043*
>23 and moderate risk	*-2.714*	*0.075*	*0.066*



SBP by pulse pressure

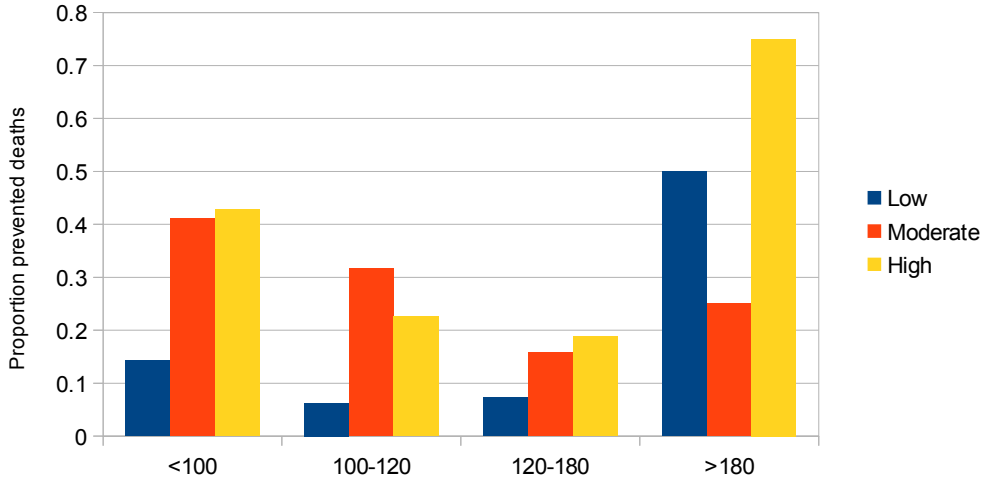
	B	Sig.	Exp(B)
SBP (ref 121-180)		<0.001	
<100	*-12.604*	*1*	*0*
100-120	*-12.722*	*0.999*	*0*
>180	*31.858*	*0.999*	*>1000*
Pulse pressure (ref 51-75)		1	
<51	*43.796*	*0.999*	*>1000*
>75	*5.503*	*1*	*245.417*
Pulse pressure by SBP		<0.001	
<100 and <51	*19.428*	*1*	*>1000*
100-120 and <51	*19.074*	*0.999*	*>1000*
>180 and <51	*132.355*	*0.999*	*>1000*
>180 and >75	*20.458*	*1*	*>1000*



SBP by SaO2

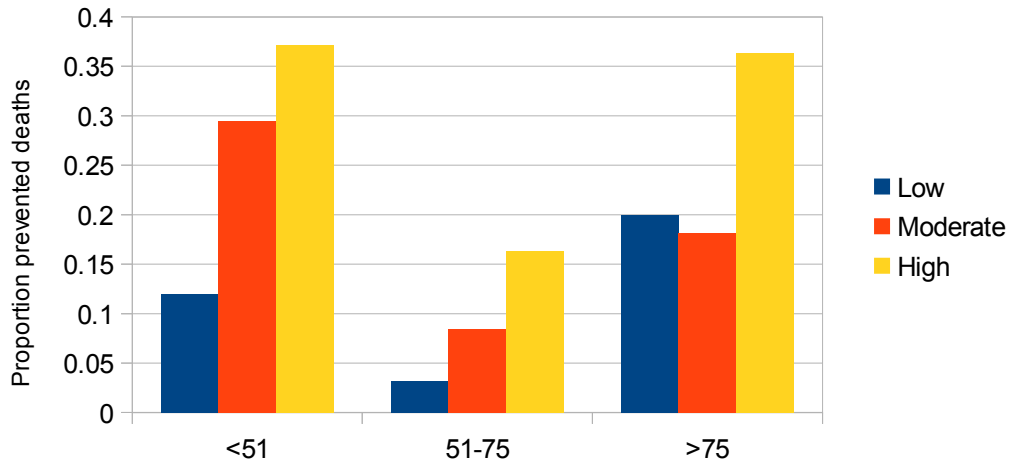
	B	Sig.	Exp(B)
SBP (ref 121-180)		<0.001	
<100	*1.379*	*0.001*	*3.971*
100-120	*0.522*	*0.177*	*1.686*
>180	*5.017*	*<0.00 1*	*150.961*
SaO2 (ref low risk)		<0.001	
High risk	*0.980*	*0.743*	*2.663*
Moderate risk	*-0.847*	*0.580*	*0.429*
SBP by SaO2		0.001	
<100 and high risk	*0.035*	*0.976*	*1.036*
100-120 and high risk	*0.525*	*0.631*	*1.690*
>180 and high risk	*-0.190*	*0.987*	*0.827*
<100 and moderate risk	*0.195*	*0.862*	*1.216*
100-120 and moderate risk	*0.219*	*0.827*	*1.245*

>180 and moderate risk	*-6.240*	*0.306*	*0.002*



Pulse pressure by SaO2

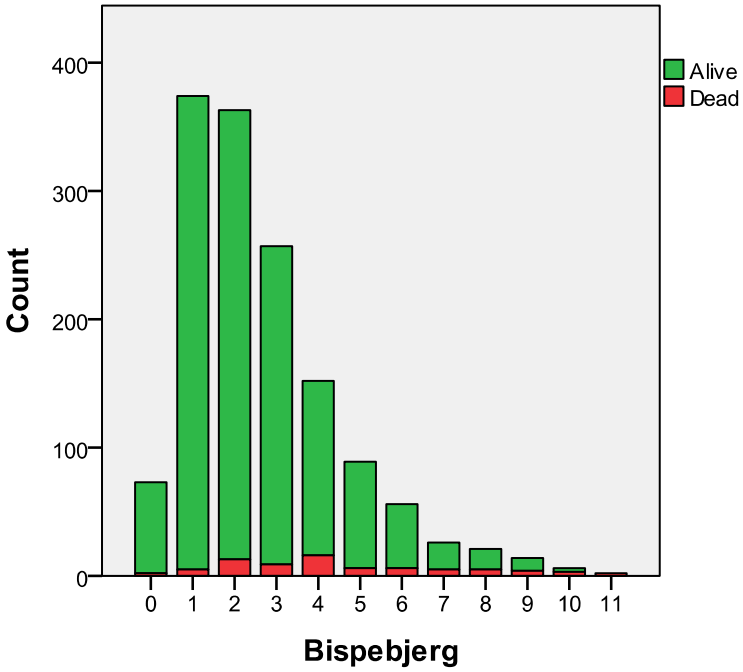
	B	Sig.	Exp(B)
Pulse pressure (ref 51-75)		<0.001	
<51	*0.880*	*0.002*	*2.412*
>75	*2.552*	*<0.001*	*12.834*
SaO2 (ref low risk)		<0.001	
High risk	*0.838*	*0.168*	*2.313*
Moderate risk	*-1.046*	*0.004*	*0.351*
Pulse pressure by SaO2		0.001	
<51 and high risk	*0.360*	*0.633*	*1.433*
<51 and moderate risk	*0.325*	*0.641*	*1.384*
>75 and high risk	*1.093*	*0.520*	*2.982*
>75 and moderate risk	*-2.239*	*0.031*	*0.107*



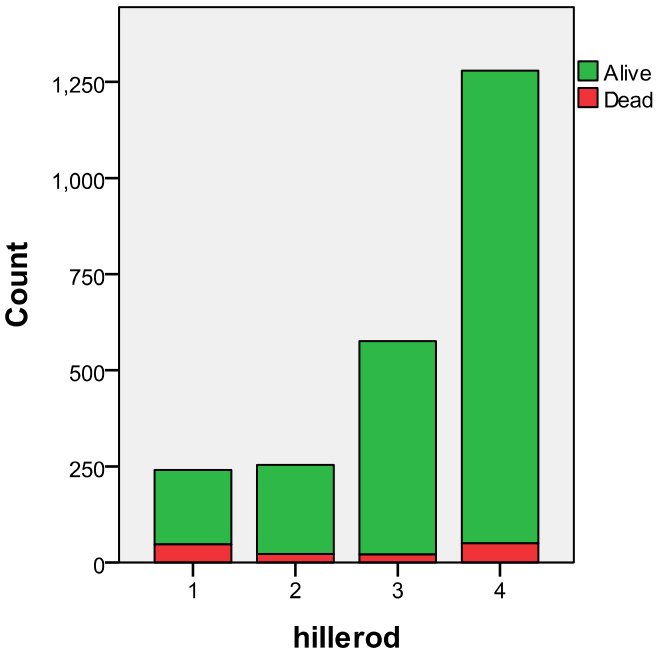
Appendix 12

Frequency distributions of other scores predicting death at 7 days in the validation cohort

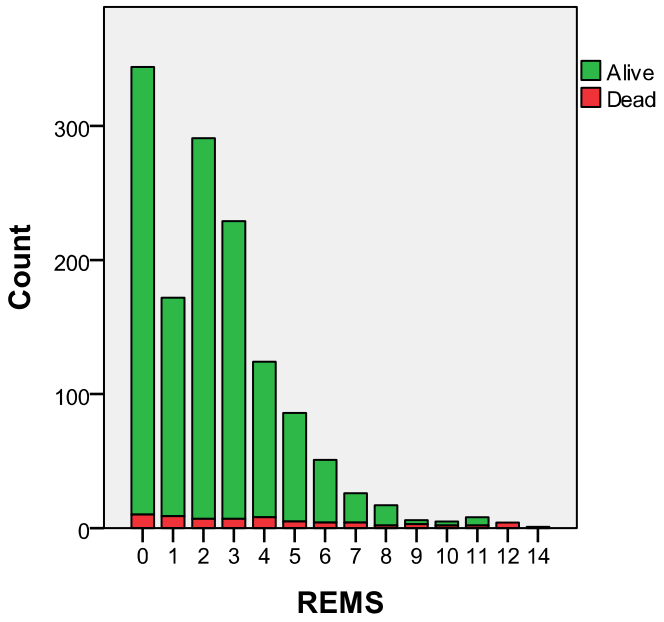
Bispebjerg score



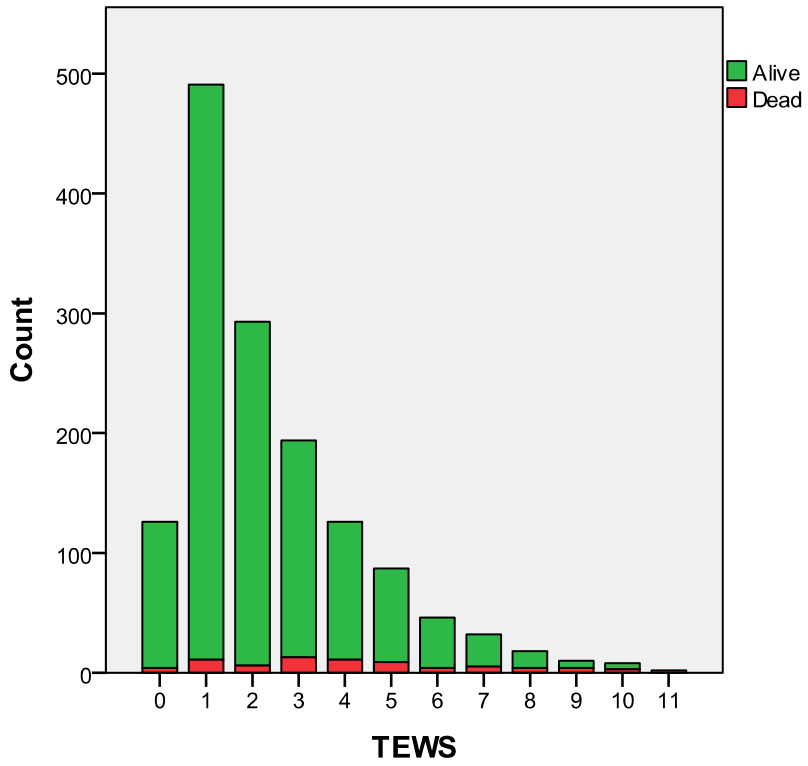
Hillerod score



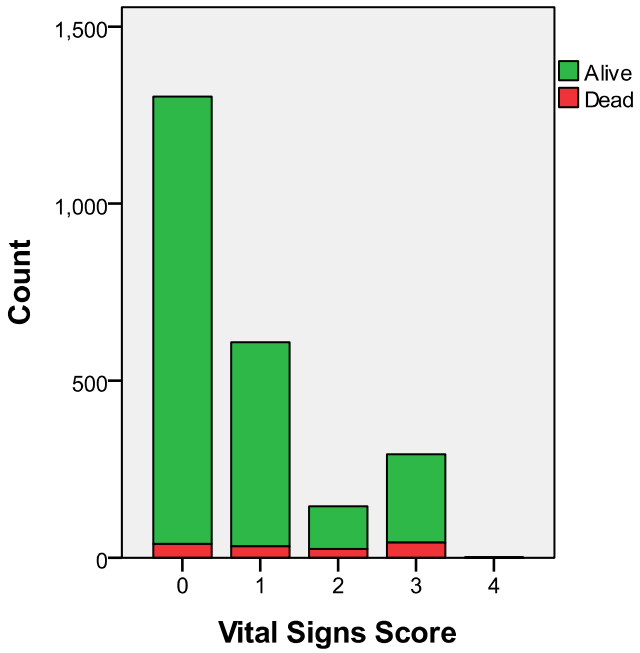
Rapid Emergency Medicine Score (REMS)



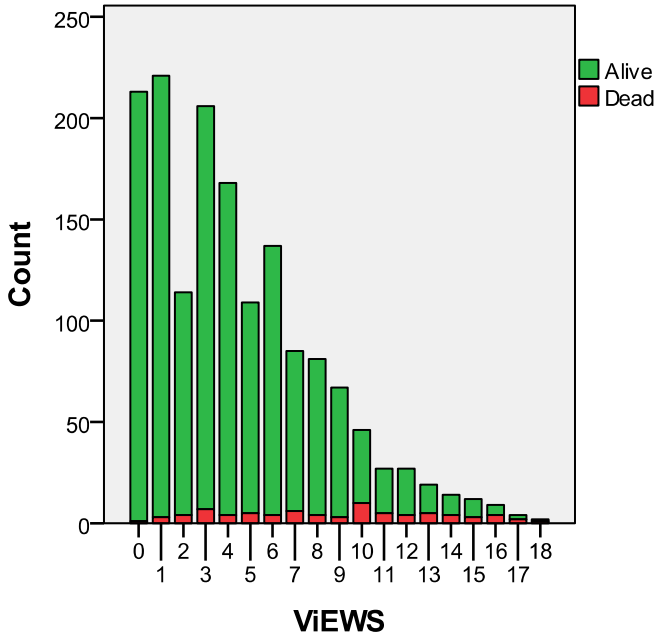
Triage Early Warning Score (TEWS)



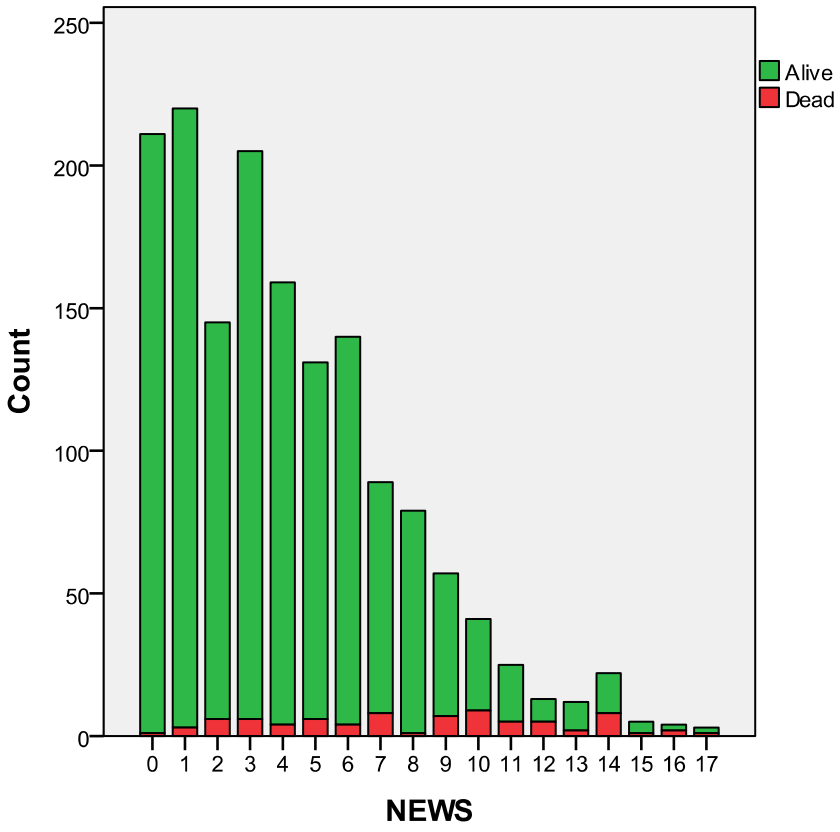
Vital Signs Score



ViEWS



NEWS


250
200
150
100
50
0
Count
Alive
Dead
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
NEWS

Appendix 13: Patients with outcomes of interest in the validation cohort

Patients sustaining inevitable death (n=2)

Ref number	Age/gender	Description
5613	65M	Presented difficulty in breathing, died day 5
5590	94F	Presented chest infection, died day 3

Patients sustaining potentially preventable death (n=1)

Ref number	Age/gender	Description
5177	92F	Presented unwell, treated for UTI, discharged day 2, died day 4

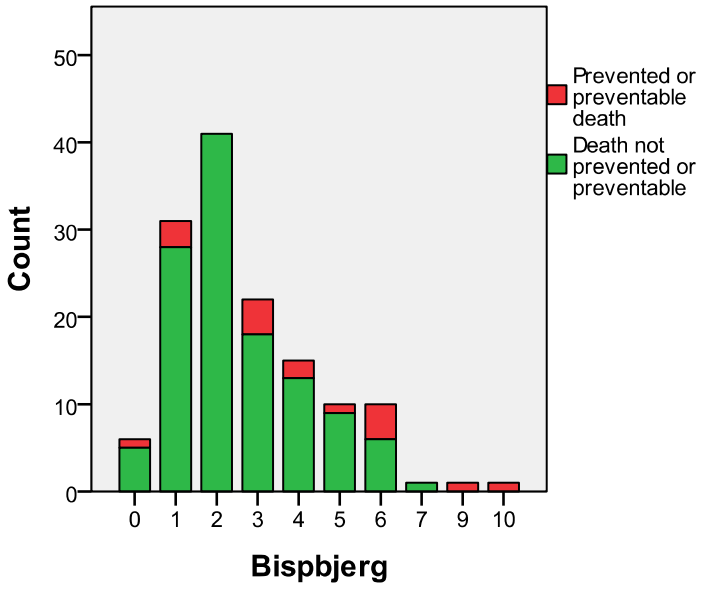
Patients where death was prevented (n=35)

Ref number	Age/gender	Description
5231	39F	Presented abdominal pain, appendicectomy day 3
5416	40M	Presented difficulty breathing, SAH clipped day 2
5331	77F	Presented difficulty breathing, IV antibiotics for sepsis
5512	78F	Presented diabetic problem, IV antibiotics for urosepsis
5586	43F	Presented with overdose, ICU
5320	86M	Presented with faint, CPR day 1
5391	35M	Presented collapse ?hypoglycaemia, IV dextrose for insulin OD
5517	42M	Presented unwell, IV insulin for DKA
5217	91M	Presented fall, IV fluids for rhabdomyolysis
5225	77F	Presented difficulty breathing, NIV for exacerbation asthma
5504	84M	Presented angina/chest pain, PCI day 7
5187	55M	Presented severe breathing problems, BiPAP day 0
5214	78F	Presented difficulty breathing, ICU IPPV day 0 for pneumonia
5213	67M	Presented difficulty breathing, ICU vasopressors day 0 for sepsis
5142	66M	Presented epigastric pain, PCI day 1
5552	67F	Presented short of breath, IV antibiotics for urosepsis day 2
5427	70M	Presented unwell, transfusion and IV fluids for GI bleed
5630	82M	Presented chest infection, IV antibiotics for sepsis from LRTI
5506	84M	Presented head injury, IV antibiotics for sepsis
5390	86M	Presented unresponsive, IV antibiotics/anticonvulsants
5387	48F	Presented ?fitting, ICU for status epilepticus
5406	80F	Presented chest pain, PCI
5337	72M	Presented hypothermia, ICU
5608	63F	Presented asthma, IPPV for COPD day 0
5438	65F	Presented pain in side, IV antibiotics pneumonia day 0
5488	93F	Presented chest pain, IV antibiotics pneumonia day 1
5554	55F	Presented abdominal pain/vomiting, IV antibiotics pneumonia
5307	87F	Presented fall, IV antibiotics for urosepsis
5475	40M	Presented difficulty breathing, HDU for CPAP for pneumonia day 0
5429	47M	Presented overdose, endotracheal tube
5182	31M	Presented abdominal pain, appendicectomy day 1
5526	35M	Presented collapse, ETT and ICU day 0 for ?encephalitis
5181	25F	Presented ?Addisonian/infection, IV steroids for Addisonian crisis
5334	56M	Presented chest pain, chemical cardioversion of AF
5408	73M	Presented ?obstruction, flatus tube for decompression sigmoid volvulus

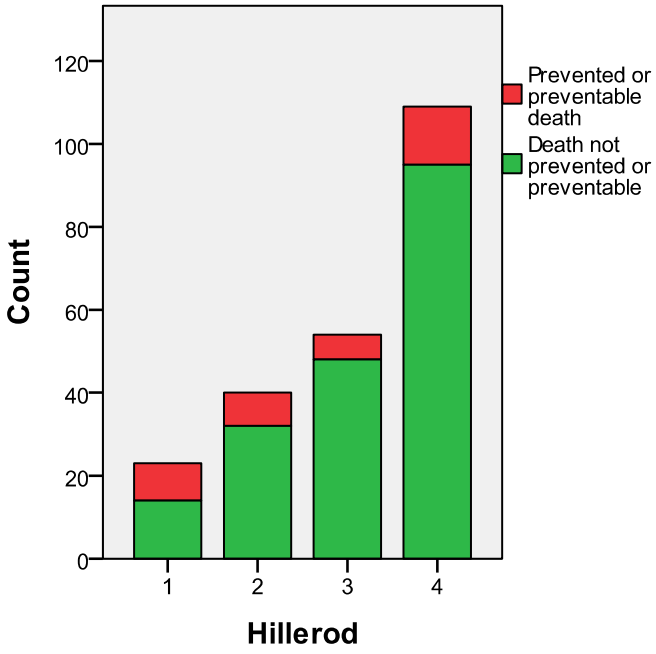
Appendix 14

Frequency distributions of other scores predicting potentially preventable or potentially prevented death in 7 days

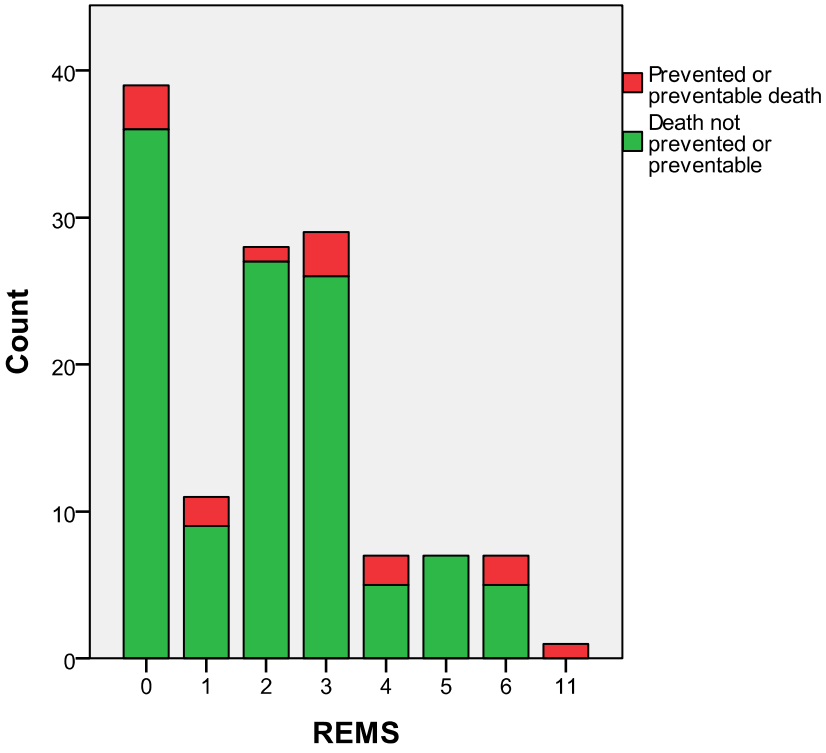
Bispbjerg



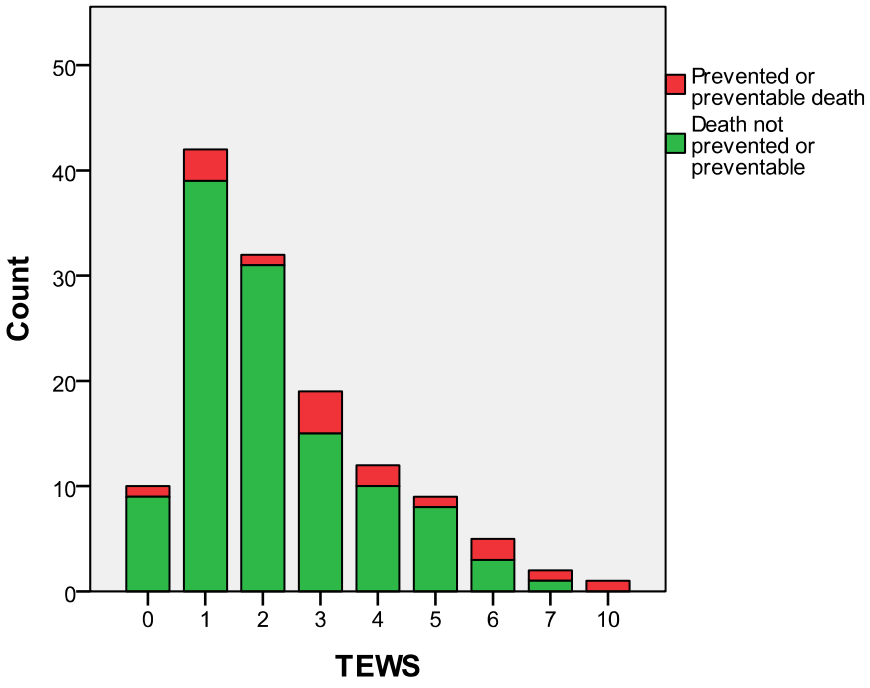
Hillerod



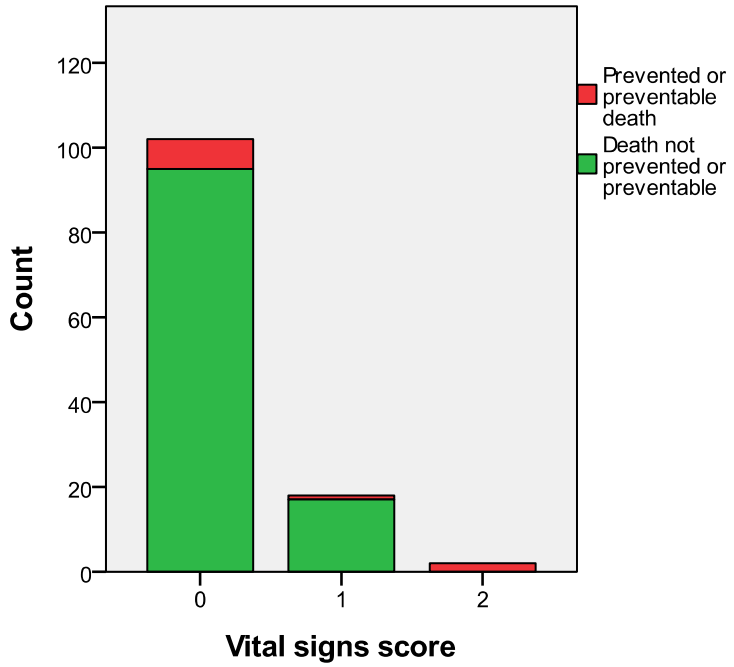
REMS


Count
40
30
20
10
0
0
1
2
3
4
5
6
11
REMS
Prevented or preventable death
Death not prevented or preventable

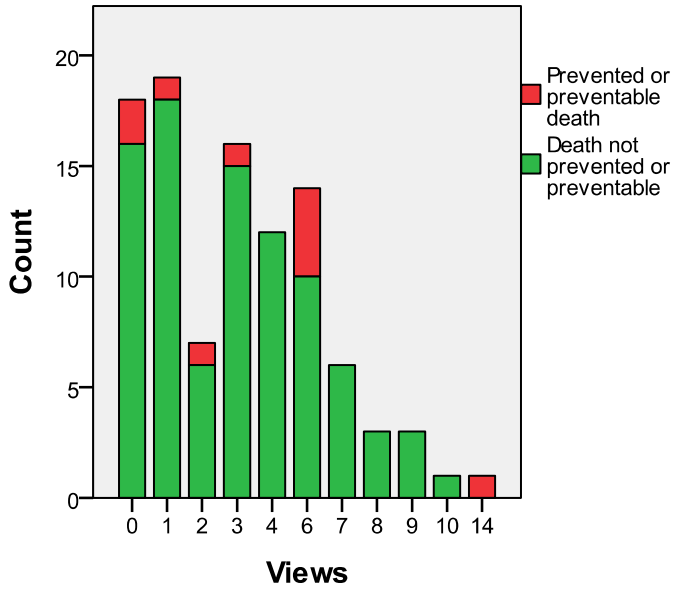
TEWS


Count
50
40
30
20
10
0
0
1
2
3
4
5
6
7
10
TEWS
Prevented or preventable death
Death not prevented or preventable

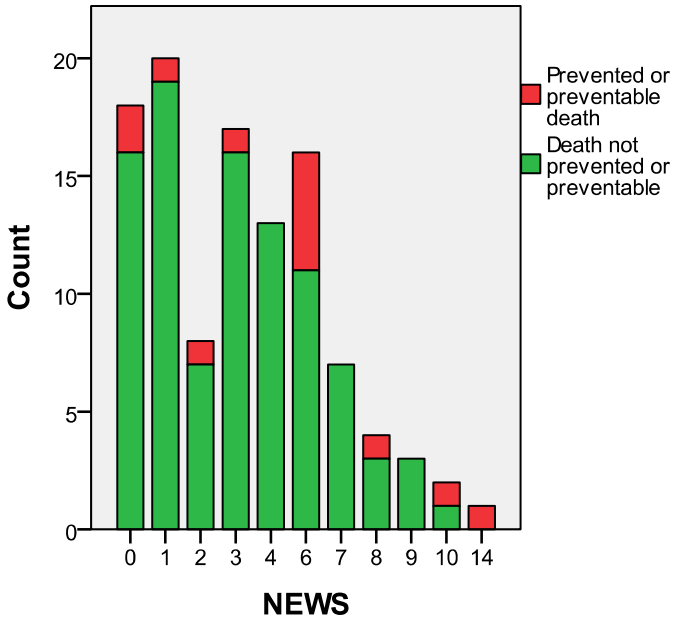
Vital Signs score



ViEWS



NEWS


Count
20
15
10
5
0
0 1 2 3 4 6 7 8 9 10 14
NEWS
Prevented or preventable death
Death not prevented or preventable

Appendix 15: Nominal group briefing document

Dear participant

Thank you for agreeing to help with my PhD studies. I have developed a bedside score to try to predict which Emergency Department non-trauma patients will require life-saving intervention during their first week in hospital. This phase of the project aims to develop a consensus of senior emergency physicians about the appropriate systematic responses to patients at different levels of risk.

Background

Why assess severity?
The general principle of healthcare provision in the UK reflects the "greatest happiness for the greatest number"(6). Accurate assessment of illness severity has implications both for the individual patient and the health care system. Patients admitted to critical care areas via non-critical care areas (such as standard wards) have higher rates of mortality than those admitted directly from the ED (9-11), with significant numbers of patients (23/122 (9) and 144/343 (11)) admitted to critical care more than 24 hours after their ED attendance. NCEPOD has documented widespread failings in the identification of sick patients and their escalation to appropriately senior staff (14), and the National Early Warning Score has been launched by the Royal College of Physicians in response to the problem (15).

What is severity?
Severity of illness can be defined in a number of ways – risk of death in either the short or long term, magnitude of symptoms such as pain or nausea, effect on functional status and deviation from either clinical or laboratory "norms". I argue that patient acuity is not necessarily concurrent with illness severity. I suggest that the point of an emergency care system is to provide prompt care to those patients likely to benefit in a time-sensitive manner from interventions. These are not inherently the same patients who are at highest risk of death – some of these patients will progress to death irrespective of interventions. Nor are they the patients with the highest overall benefit from healthcare interventions – the young man with a testicular teratoma has massive potential for benefiting from treatment, but this will not be affected by whether he is seen within one hour or six hours in the Emergency Department. A measure of patient acuity for the emergency department should therefore reflect patients whose outcome will be improved with prompt care and/or those whose outcome will worsen without this care.

Which patients should we identify?
Unfortunately within emergency care little evidence exists as to the time-sensitive nature of many interventions (16). Although, as has been argued in the context of critical care outreach, it is intuitively appealing to define a deteriorating patient and respond rapidly (17), meta-analysis of tools to do this in an in-patient setting has failed to identify a benefit in terms of patient outcome (18-19). The introduction of a medical emergency team in one hospital was associated with a decrease in mortality amongst surgical patients but a sustained increase in mortality in medical patients, highlighting the non-congruence of risk of death with the potential to benefit from early medical intervention (20). Recent commentary on clinical

decision rules has drawn attention to the disconnect between the identification of patients at risk of a particular outcome and the potential of those patients to benefit from available interventions (21).

What currently exists?

The Royal College of Physicians recently launched a National Early Warning Score developed by a multidisciplinary consensus panel. This was in response to "the multiplicity of early warning systems used in different hospitals in the UK …causing a lack of consistency in detecting deterioration of patients' conditions and calling for urgent medical help". The NEWS is advocated, for the purposes of standardisation "during the initial prehospital and/or hospital assessment of a patient and throughout the patient's hospital stay", although the development group did not include emergency physicians or nurses as stakeholders and the group was unable to identify any relevant literature relating to Emergency Department patients (15). It uses a number of easily measurable physiological parameters to score a patient's risk of deterioration and suggests clinical responses to each risk category as shown:

News score	**Clinical risk**	**Frequency of monitoring**	**Clinical response**
0	Low	Minimum 12 hrly	Continue routine monitoring
Aggregate 1-4		Minimum 4-6 hrly	Inform RN who must assess patient and decide if increased frequency of monitoring of escalation of clinical care required
Individual parameter 3 Aggregate 5-6	Medium	Minimum 1 hrly	RN to inform medical team urgently Urgent assessment by clinician with core competencies in assessing acutely ill patients Clinical care in environment with monitoring facilities
Aggregate 7 or more	High	Continuous monitoring	RN to inform medical team immediately at SpR level Emergency assessment by clinician with critical care competencies including advanced airway skills Consider transfer to level 2 or 3 care

The proposed score

This has been developed from a set of patients presenting at the Northern General Hospital, Sheffield, to identify patients at high risk of needing a life-saving intervention. The score is:

Variable	Points
Pulse	
71-110	2
>110	10
Systolic BP	
<100	4
100-120	1
>180	3
GCS	
3-8	13
9-12	3

Why are we asking you to do this?

The score has been validated to demonstrate its ability to predict preventable mortality (ie to identify those patients likely to benefit in a time-sensitive way from emergency care). We would like to use your expertise to explore how the score might be operationalised in a working ED, and what the resource implications of this might be.

Points for discussion

Below are 10 brief vignettes of information that would be available at the point of triage, with the proposed score calculated. All of them relate to a 70 year old man who has presented with lightheadedness and "weakness". Please consider:

- Where in the ED should this patient be accommodated (resus, majors, waiting room etc). This may include streaming to primary care if you think appropriate.
- How often should this patient have observations repeated?
- Which staff should be informed about the patient (nurse in charge, doctor in charge, any nurse, any doctor, no specific need to inform anyone)?
- How soon should the patient be assessed by a doctor or nurse practitioner?

Case 1

P70, BP 130/72, GCS 15. Score 0.

Location in ED	
Frequency of routine observations	
Staff to be informed	
Time to assessment	

Case 2

P61, BP 120/65, GCS 15. Score 1.

Location in ED	
Frequency of routine observations	
Staff to be informed	
Time to assessment	

Case 3

P84, BP 111/64, GCS 15. Score 3.

Location in ED	
Frequency of routine observations	
Staff to be informed	
Time to assessment	

Case 4

P63, BP 93/69, GCS 14. Score 4.

Location in ED	
Frequency of routine observations	
Staff to be informed	
Time to assessment	

Case 5

P102, BP 99/56, GCS 14. Score 6.

Location in ED	
Frequency of routine observations	
Staff to be informed	
Time to assessment	

Case 6

P35, BP 70/40, GCS 12. Score 7.

Location in ED	
Frequency of routine observations	
Staff to be informed	
Time to assessment	

Case 7

P134, BP 149/80, GCS 15. Score 10.

Location in ED	
Frequency of routine observations	
Staff to be informed	
Time to assessment	

Case 8

P123, BP 199/94, GCS 15. Score 13.

Location in ED	
Frequency of routine observations	
Staff to be informed	
Time to assessment	

Case 9

P129, BP 82/57, GCS 12. Score 17.

Location in ED	
Frequency of routine observations	
Staff to be informed	
Time to assessment	

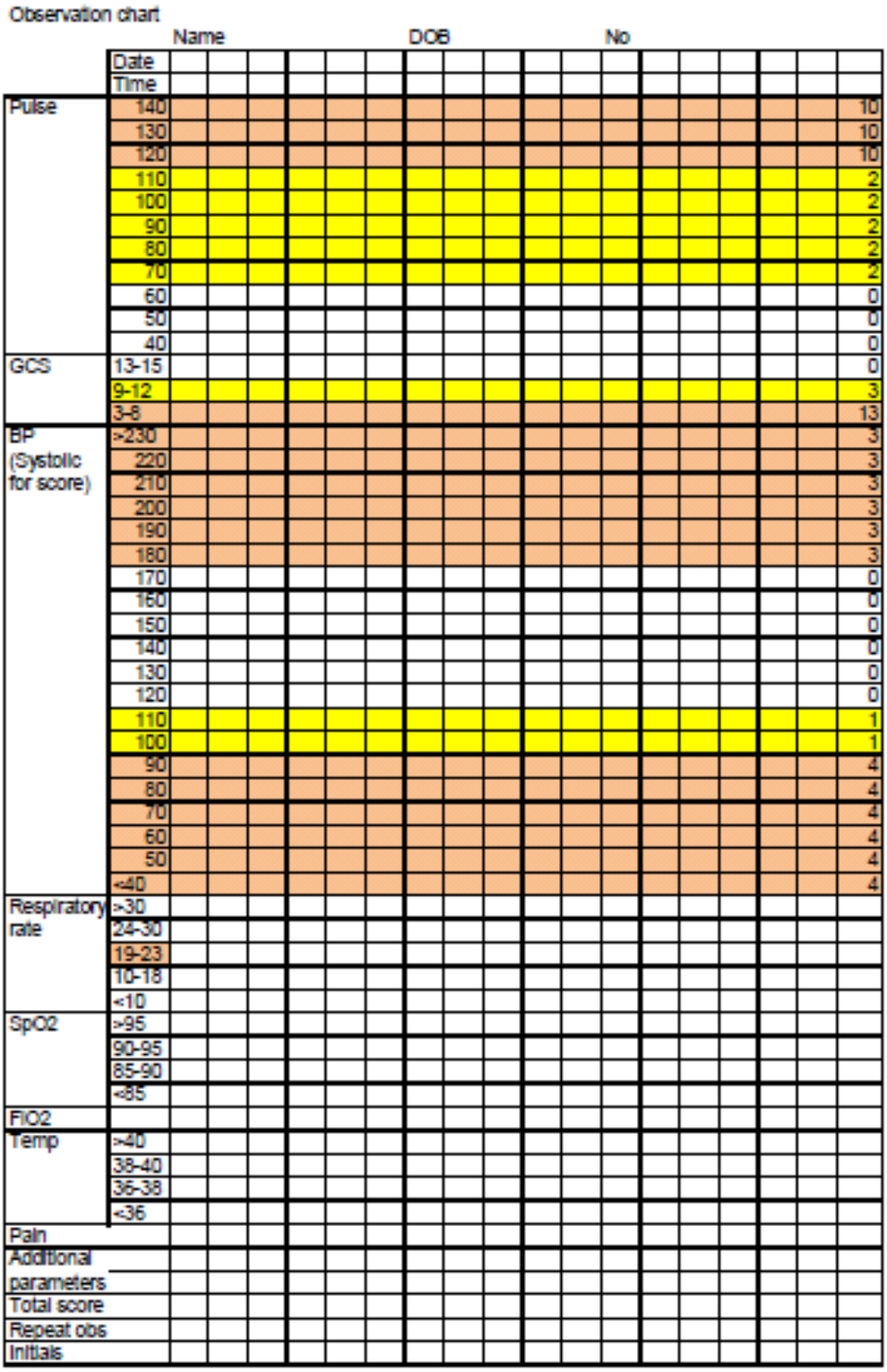
Case 10

P114, BP 133/92, GCS 8. Score 23.

Location in ED	
Frequency of routine observations	
Staff to be informed	
Time to assessment	

Appendix 16: Potential presentations of the score for clinical practice

1. Paper-based observation chart



2. Web-based triage form

NHS

Lancashire Teaching Hospitals Trust

Royal Preston Hospital

30/08/2014

11:22

Smith, John **1/01/1980** **111999**

Presenting complaint: Unwell adult

Triage notes: Collapsed playing football. 3m LOC.

Pulse: 130 SBP: 125 DBP: 60 GCS: 12

RR: 22 SaO2: 97 FiO2: RA Temperature: 38.1

Pain score: 0

Analgesia:

Risk score 13

Resus bed x3800

Inform nurse and doctor in charge x3310

References

1. HSCIC. Hospital Episode Statistics: Accident and Emergency Attendances in England – 2012-13: Health and Social Care Information Centre2014.
2. Blunt I, Bardsley M, Dixon J. Trends in emergency admissions in England 2004–2009: is greater efficiency breeding inefficiency? London: Nuffield Trust2010.
3. Hospital Episode Statistics. Adult Critical Care in England – April 2010 to March 2011: Experimental Statistics. London: Health and Social Care Information Centre2012.
4. Larrey D. Memoirs of military surgery, and campaigns of the French armies: Vol 2. Baltimore: Joseph Cushing; 1814.
5. Anonymous. Dealing with disaster. Lancet. 1952;ii:1117-8.
6. Mill JS. Utilitarianism. London1863.
7. Frykberg ER, Hutton PM, Balzer RH. Disaster in Beirut: an application of mass casualty principles. Military Medicine. 1987;152(11):563-6.
8. Lennquist S. Management of Major Accidents and Disasters: An Important Responsibility for the Trauma Surgeons. Journal of Trauma. 2007;62:1321-9.
9. Parkhe M, Myles PS, Leach DS, Maclean AV. Outcome of emergency department patients with delayed admission to an intensive care unit. Emergency Medicine Australasia. 2002;14:50-7.
10. McQuillan P, Pilkington S, Allan A, Taylor B, Short A, Morgan G, et al. Confidential inquiry into quality of care before admission to intensive care. BMJ. 1998;316:1853-8.
11. Renaud B, Brun-Buisson C, Santin A, Coma E, Noyez C, Fine MJ, et al. Outcomes of Early, Late, and No Admission to the Intensive Care Unit for Patients Hospitalized with Community-acquired Pneumonia. Academic Emergency Medicine. 2012;19:294-303.
12. Frykberg ER, Tepas JJ. Terrorist Bombings: Lessons Learned From Belfast to Beirut. Annals of Surgery. 1988;208(5):569-76.
13. Brody H. From an Ethics of Rationing to an Ethics of Waste Avoidance. New England Journal of Medicine. 2012;366:1949-51.
14. Findlay GP, Shotton H, Kelly K, Mason M. Time to Intervene? A review of patients who underwent cardiopulmonary resuscitation as a result of an in-hospital cardiorespiratory arrest. London: NCEPOD2012.
15. Royal College of Physicians. National Early Warning Score (NEWS): Standardising the assessment of acute illness severity in the NHS. London: Royal College of Physicians2012.
16. Moll HA. Challenges in the validation of triage systems at emergency departments. Journal of Clinical Epidemiology. 2010;63:384-8.
17. Hillman KM. Rapid response systems: you won't know there is a problem until you measure it. Critical Care. 2011;15:1001.
18. Gao H, McDonnell A, Harrison DA, Moore T, Adam S, Daly K, et al. Systematic review and evaluation of physiological track and trigger warning systems for identifying at-risk patients on the ward. Intensive Care Med. 2007;33:667-9.
19. Chan PS, Jain R, Nallmothu BK, Berg RA, Sasson C. Rapid Response Teams: A Systematic Review and Meta-analysis. Archives of Internal Medicine. 2010;170:18-26.

20. Jones D, Opdam H, Egi M, Goldsmith D, Bates S, Gutteridge G, et al. Long-term effect of a Medical Emergency Team on mortality in a teaching hospital. Resuscitation. 2007;74:235-41.
21. Grady D, Berkowitz SA. Why Is a Good Clinical Prediction Rule So Hard to Find? Archives of Internal Medicine. 2011;171:1701-2.
22. Goldhill D, White S, Sumner A. Physiological values and procedures in the 24 h before ICU admission from the ward. Anaesthesia. 1999;54:529-34.
23. Fine MJ, Auble TE, Yealy DM, Hanusa BH, Weissfeld LA, Singer DE, et al. A prediction rule to identify low-risk patients with community acquired pneumonia. New England Journal of Medicine. 1997;336:243-50.
24. Wahlgren N, Ahmed N, Eriksson N, Aichner F, Bluhmki E, Davalos A, et al. Multivariable Analysis of Outcome Predictors and Adjustment of Main Outcome Results to Baseline Data Profile in Randomized Controlled Trials Safe Implementation of Thrombolysis in Stroke-MOnitoring STudy (SITS-MOST). Stroke. 2008;39:3316-22.
25. Feldman MJ, Verbeek PR, Lyons DG, Chad SJ, Craig AM, Schwartz B. Comparison of the Medical Priority Dispatch System to an Out-of-hospital Patient Acuity Score. Academic Emergency Medicine. 2006;13:954-60.
26. Hinchey P, Myers B, Zalkin J, Lewis R, Garner D. Low acuity EMS dispatch criteria can reliably identify patients without high-acuity illness or injury. Prehospital Emergency Care. 2007;11:42-8.
27. Wilson S, Cooke M, Morrell R, Bridge P, Allan T. A systematic review of the evidence supporting the use of priority dispatch of emergency ambulances. Prehospital Emergency Care. 2002;6:42-9.
28. Manchester Triage Group. Emergency Triage. 2nd ed. London: BMJ books; 2005.
29. Wuerz RC. Emergency Severity Index triage category is associated with six-month survival. Academic Emergency Medicine. 2001;8(1):61-4.
30. Murray M, Bullard M, Grafstein E. Revisions to the Canadian Emergency Department Triage and Acuity Scale Implementation Guidelines. Canadian Journal of Emergency Medicine. 2004;6(6):421-7.
31. Beveridge R, Ducharme J, Janes L, Beaulieu S, Walter S. Reliability of the Canadian Emergency Department Triage and Acuity Scale: Interrater Agreement. Annals of Emergency Medicine. 1999;34(2):155-9.
32. Atzema CL, Austin PC, Tu JV, Schull MJ. ED triage of patients with acute myocardial infarction: predictors of low acuity triage. American Journal of Emergency Medicine. 2010;28:694-702.
33. Cooke MW, Jinks S. Does the Manchester triage system detect the critically ill? Journal of Accident and Emergency Medicine. 1999;16:179-81.
34. Farrohknia N, Castrén M, Ehrenberg A, Lind L, Oredsson S, Jonsson H, et al. Emergency Department Triage Scales and Their Components: A Systematic Review of the Scientific Evidence. Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine. 2011;19:42.
35. Storm-Versloot MN, Vermeulen H, van Lammeren N, Luitse JS, Goslings JC. Influence of the Manchester Triage System on waiting time, treatment time, length of stay and patient satisfaction; a before and after study. Emergency Medicine Journal. 2012;doi:10.1136/emermed-2012-201099
36. George S, Read S, Westlake L, Williams B, Fraser-Moodie A, Pritty P. Evaluation of nurse triage in a British accident and emergency department. British Medical Journal. 1992;304:876-8.

37. Terris J, Leman P, O'Connor N, Wood R. Making an IMPACT on emergency department flow: improving patient processing assisted by consultant at triage. Emergency Medicine Journal. 2004;21:537-41.
38. Subash F, Dunn F, McNicholl B, Marlow J. Team triage improves emergency department efficiency. Emergency Medicine Journal. 2004;21:542-4.
39. Choi Y, Wong T, Lau C. Triage rapid initial assessment by doctor (TRIAD) improves waiting time and processing time of the emergency department. Emergency Medicine Journal. 2006;23:262-5.
40. NICE Short Clinical Guidelines Technical Team. Acutely ill patients in hospital: recognition of and response to acute illness in adults in hospital. London: National Institute for Health and Clinical Excellence2007.
41. Iles S, Hodges A, Darley J, Frampton C, Epton M, Beckert L, et al. Clinical experience and pre-test probability scores in the diagnosis of pulmonary embolism. Quarterly Journal of Medicine. 2003;96:211-5.
42. de Vries EN, Prins HA, Crolla RMPH, den Outer AJ, van Andel G, van Helden SH, et al. Effect of a Comprehensive Surgical Safety System on Patient Outcomes. New England Journal of Medicine. 2010;363:1928-37.
43. Considine J, Thomas S, Potter R. Predictors of critical care admission in emergency department patients triaged as low to moderate urgency. Journal of Advanced Nursing. 2009;65:818-27.
44. Aujesky D, Obrosky DS, Stone RA, Auble TE, Perrier A, Cornuz J, et al. Derivation and Validation of a Prognostic Model for Pulmonary Embolism. American Journal of Respiratory and Critical Care Medicine. 2005;172:1041-6.
45. Christian MD, Hawryluck L, Wax RS, Cook T, Lazar NM, Herridge MS, et al. Development of a triage protocol for critical care during an influenza pandemic. Canadian Medical Association Journal. 2006;175(11):1377-81.
46. Ardagh M. Criteria for prioritising access to healthcare resources in New Zealand during an influenza pandemic or at other times of overwhelming demand. New Zealand Medical Journal. 2006;119(1232).
47. Shaughnessy A, Slawson D, Bennett J. Becoming an information master: a guidebook to the medical information jungle. Journal of Family Practice. 1994;39:489-99.
48. Ebell M. AHRQ White Paper: Use of Clinical Decision Rules for Point-of-Care Decision Support. Medical Decision Making. 2010;30:712-21.
49. Jaspers MWM, Smeulers M, Vermeulen H, Peute LW. Effects of clinical decision-support systems on practitioner performance and patient outcomes: a synthesis of high-quality systematic review findings. Journal of the American Medical Informatics Association. 2011;18:327-34.
50. Sim I, Gorman P, Greenes RA, Haynes RB, Kaplan B, Lehmann H, et al. Clinical Decision Support Systems for the Practice of Evidence-based Medicine. Journal of the American Medical Informatics Association. 2001;8:527-34.
51. Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. BMJ. 2005;330:765.
52. Stiell IG, Wells GA. Methodologic Standards for the Development of Clinical Decision Rules in Emergency Medicine. Annals of Emergency Medicine. 1999;33:437-47.

53. McGinn TG, Guyatt GH, Wyer PC, Naylor CD, Stiell IG, Richardson WS. Users' Guides to the Medical Literature: XXII: How to Use Articles About Clinical Decision Rules. JAMA. 2000;284:79-84.
54. Poses RM, Bekes C, Copare FJ, Scott WE. The answer to "What are my chances, Doctor?" depends on whom is asked: prognostic disagreement and inaccuracy for critically ill patients. Critical Care Medicine. 1989;17(8):827-33.
55. Marks R, Simons R, Blizzard R, Browne D. Predicting outcome in intensive therapy units - a comparison of Apache II with subjective assessments. Intens Care Med. 1991;17:159-63.
56. States JD. The Abbreviated and the Comprehensive Research Injury Scales. STAPP Car Crash Journal. 1969;13:282-94.
57. Kuwabara K, Matsuda S, Fushimi K, Ishikawa KB, Horiguchi H, Fujimori K. Probability of survival, early critical care process, and resource use in trauma patients. American Journal of Emergency Medicine. 2010;28:673-81.
58. Baxt WG, Upenieks V. The Lack of Full Correlation Between the Injury Severity Score and the Resource Needs of Injured Patients. Annals of Emergency Medicine. 1990;19:1396-400.
59. Wallis L, Carley S, Hodgetts C. A procedure based alternative to the injury severity score for major incident triage of children: results of a Delphi consensus process. Emergency Medicine Journal. 2006;23:291-5.
60. Smith GCS, Pell JP. Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials. BMJ. 2003;327:1459-61.
61. Donald C, Duncan R, Thakore S. Predictors of the need for rapid sequence intubation in the poisoned patient with reduced Glasgow coma score. Emergency Medicine Journal. 2009;26:510-2.
62. Hillman K, Alexandrou E, Flabouris M, Brown D, Flabouris A, Parr M, et al. Clinical outcome indicators in acute hospital medicine. Clinical Intensive Care. 2000;11:89-94.
63. Johnston SC, Rothwell PM, Nguyen-Huynh MN, Giles MF, Elkins JS, Bernstein AL, et al. Validation and refinement of scores to predict very early stroke risk after transient ischaemic attack. Lancet. 2007;369:283-92.
64. Knaus WA, Draper EA, Wagner D, Zimmerman J. APACHE II: a severity of disease classification system. Critical Care Medicine. 1985;13(10):818-29.
65. Lim W, van der Eerden M, Laing R, Boersma W, Karalus N, Town G, et al. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. Thorax. 2003;58:377-82.
66. Antman E, Cohen M, Bernink P. The TIMI Risk Score for Unstable Angina/Non–ST Elevation MI. A Method for Prognostication and Therapeutic Decision Making. JAMA. 2000;284(835-42).
67. Williams B, Wright R, Murphy J, Brilakis E, Reeder G, Jaffe A. A new simplified immediate prognostic risk score for patients with acute myocardial infarction. Emergency Medicine Journal. 2006;23:186-92.
68. Hasdai D, Califf RM, Thompson TD, Hochman JS, Ohman EM, Pfisterer M, et al. Predictors of Cardiogenic Shock After Thrombolytic Therapy for Acute Myocardial Infarction. Journal of the American College of Cardiology. 2000;35:136-43.

69. Renaud B, Labarère J, Coma E, Santin A, Hayon J, Gurgui M, et al. Risk stratification of early admission to the intensive care unit of patients with no major criteria of severe community-acquired pneumonia: development of an international prediction rule. Critical Care. 2009;13(R54).
70. Blatchford O, Murray WR, Blatchford M. A risk score to predict need for treatment for upper gastrointestinal haemorrhage. Lancet. 2000;356:1318-21.
71. Jimenez D, Aujesky D, Moores L, Gomez V, Lobo JL, Uresandi F, et al. Simplification of the Pulmonary Embolism Severity Index for Prognostication in Patients With Acute Symptomatic Pulmonary Embolism. Archives of Internal Medicine. 2010;170:1383-9.
72. Nicholl J. Case-mix adjustment in non-randomised observational evaluations: the constant risk fallacy. Journal of Epidemiology and Community Health. 2007;61:1010-3.
73. Karnofsky D, Burchenal J. The Clinical Evaluation of Chemotherapeutic Agents in Cancer. In: MacLeod C, editor. Evaluation of Chemotherapeutic Agents: Columbia University Press; 1949. p. 196.
74. Shapiro NI, Wolfe RE, Moore RB, Smith E, Burdick E, Bates DW. Mortality in Emergency Department Sepsis (MEDS) score: A prospectively derived and validated clinical prediction rule. Crit Care Med. 2003;31:670-5.
75. Challen K, Bright J, Bentley A, Walter D. Physiological-social score (PMEWS) vs. CURB-65 to triage pandemic influenza: a comparative validation study using community-acquired pneumonia as a proxy. BMC Health Services Research. 2007;7:33.
76. Goodacre S, Wilson R, Shephard N, Nicholl J. Derivation and validation of a risk adjustment model for predicting seven day mortality in emergency medical admissions: mixed prospective and retrospective cohort study. BMJ. 2012;344:e2904.
77. Fleming S, Thompson M, Stevens R, Heneghan C, Plüddemann A, Maconochie I, et al. Normal ranges of heart rate and respiratory rate in children from birth to 18 years of age: a systematic review of observational studies. Lancet. 2011;377:1011-8.
78. Tarassenko L, Clifton DA, Pinsky MR, Hravnak MT, Woods JR, Watkinson PJ. Centile-based early warning scores derived from statistical distributions of vital signs. Resuscitation. 2011;82:1013-8.
79. Moons KG, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? BMJ. 2009;338:b375.
80. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A Simulation Study of the Number of Events per Variable in Logistic Regression Analysis. Journal of Clinical Epidemiology. 1996;49:1373-9.
81. Vittinghoff E, McCulloch CE. Relaxing the Rule of Ten Events per Variable in Logistic and Cox Regression. American Journal of Epidemiology. 2007;165:710-8.
82. Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. BMJ. 2009;338:b605.
83. Quinn JV, Stiell IG, McDermott DA, Sellers KL, Kohn MA, Wells GA. Derivation of the San Francisco Syncope Rule to Predict Patients With Short-Term Serious Outcomes. Annals of Emergency Medicine. 2004;43:224-32.

84. Yarnold PR, Soltysik RC, Bennett CL. Predicting in-hospital mortality of patients with AIDS-related Pneumocystis carinii pneumonia: an example of hierarchically optimal classification tree analysis. Statistics in medicine. 1997;16:1451-63.
85. Mark DB, Shaw L, Harrell FE, Hlatky MA, Lee KL, Bengtson JR, et al. Prognostic Value of a Treadmill Exercise Score in Outpatients with Suspected Coronary Artery Disease. New England Journal of Medicine. 1991;325:849-53.
86. Seymour CW, Kahn JM, Cooke CR, Watkins TR, Heckbert SR, Rea TD. Prediction of Critical Illness During Out-of-Hospital Emergency Care. JAMA. 2010;304:747-54.
87. Christensen D, Jensen NM, Maaløe R, Rudolph SS, Belhage B, Perrild H. Nurse-administered early warning score system can be used for emergency department triage. Danish Medical Bulletin. 2011;58:A4221.
88. Barfod C, Lauritzen MMP, Danker JK, Sölétormos G, Forberg JL, Berlac PA, et al. Abnormal vital signs are strong predictors for intensive care unit admission and in-hospital mortality in adults triaged in the emergency department - a prospective cohort study. Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine. 2012;20:28.
89. Kellett J, Deane B, Gleeson M. Derivation and validation of a score based on Hypotension, Oxygen saturation, low Temperature, ECG changes and Loss of independence (HOTEL) that predicts early mortality between 15 min and 24 h after admission to an acute medical unit. Resuscitation. 2008;78:52-8.
90. Carmichael HA, Robertson E, Austin J, McCruden D, Messow CM, Belcher PR. A new approach to scoring systems to improve identification of acute medical admissions that will require critical care. Scottish Medical Journal. 2011;56:195-202.
91. Olsson T, Terent A, Lind L. Rapid Emergency Medicine score: a new prognostic tool for in-hospital mortality in nonsurgical emergency department patients. Journal of Internal Medicine. 2004;255:579-87.
92. Kellett J, Deane B. The Simple Clinical Score predicts mortality for 30 days after admission to an acute medical unit. Quarterly Journal of Medicine. 2006;99:771-81.
93. Rosedale K, Smith ZA, Davies H, Wood D. The effectiveness of the South African Triage Score (SATS) in a rural emergency department. South African Medical Journal. 2011;101:537-40.
94. Bruijns SR, Wallis LA, Burch VC. A prospective evaluation of the Cape triage score in the emergency department of an urban public hospital in South Africa. Emergency Medicine Journal. 2008;25:398-402.
95. Sun Y, Heng BH, Tay SY, Seow E. Predicting Hospital Admissions at Emergency Department Triage Using Routine Administrative Data. Academic Emergency Medicine. 2011;18:844-50.
96. Merz TM, Etter R, Mende L, Barthelmes D, Wiegand J, Martinolli L, et al. Risk assessment in the first fifteen minutes – a prospective cohort study of a simple physiological scoring system in the emergency department. Critical Care. 2011;15:R25.
97. Prytherch DR, Smith GB, Schmidt PE, Featherstone PI. ViEWS—Towards a national early warning score for detecting adult inpatient deterioration. Resuscitation. 2010;81:932-7.
98. Mikulich O, Callaly E, Bennett K, O'Riordan D, Silke B. The increased mortality associated with a weekend emergency admission is due to

increased illness severity and altered case-mix. Acute Medicine. 2011;10:182-7.
99. Lee JY, Oh SH, Peck EH, Lee JM, Park KN, Kim SH, et al. The validity of the Canadian Triage and Acuity Scale in predicting resource utilization and the need for immediate life-saving interventions in elderly emergency department patients. Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine. 2011;19:68.
100. Prah Ruger J, Lewis LM, Richter CJ. Identifying high-risk patients for triage and resource allocation in the ED. American Journal of Emergency Medicine. 2007;25:794-8.
101. Armagan E, Yilmaz Y, Olmez O, Simsek G, Gul C. Predictive value of the modified Early Warning Score in a Turkish emergency department. European Journal of Emergency Medicine. 2008;15:338-40.
102. Groarke J, Gallagher J, Stack J, Aftab A, Dwyer C, McGovern R, et al. Use of an admission early warning score to predict patient morbidity and mortality and treatment success. Emergency Medicine Journal. 2008;25:803-6.
103. Burch V, Tarr G, Morroni C. Modified early warning score predicts the need for hospital admission and inhospital mortality. Emergency Medicine Journal. 2008;25:674-8.
104. Heitz CR, Gaillard JP, Blumstein H, Case D, Messick C, Miller CD. Performance of the Maximum Modified Early Warning Score to Predict the Need for Higher Care Utilization Among Admitted Emergency Department Patients. Journal of Hospital Medicine. 2010;5:E45-52.
105. Perera Y, Ranasinghe P, Adikari A, Welivit W, Perera W, Wijesundara W, et al. The value of the Modified Early Warning Score and biochemical parameters as predictors of patient outcome in acute medical admissions: a prospective study. Acute Medicine. 2011;10:126-32.
106. Li JYZ, Yong TY, Hakendorf P, Roberts S, O'Brien L, Sharma Y, et al. Simple clinical score is associated with mortality and length of stay of acute general medical admissions to an Australian hospital. Internal Medicine Journal. 2012;42:160-5.
107. Kellett J, Emmanuel A, Deane B. Who will be sicker in the morning? Changes in the Simple Clinical Score the day after admission and the subsequent outcomes of acutely ill unselected medical patients. European Journal of Internal Medicine. 2011;22:375-81.
108. Subbe C, Jishi F, Hibbs R. The Simple Clinical Score: a tool for benchmarking of emergency admissions in acute internal medicine. Clinical Medicine. 2010;10:352-7.
109. Subbe CP, Gauntlett W, Kellett JG. Collaborative Audit of Risk Evaluation in Medical Emergency Treatment (CARE-MET I) — An international pilot. European Journal of Internal Medicine. 2010;21:222-5.
110. Cochrane Collaboration. Considerations and recommendations for figures in Cochrane reviews: graphs of statistical data2008.
111. L'Abbe K, Detsky AS, O'Rourke K. Meta-analysis in clinical research. Annals of Internal Medicine. 1987;107:224-33.
112. Box GE, Draper N. Empirical model building and response surfaces. New York: John Wiley & Sons; 1987.
113. Sklar DP, Crandall CS, Loeliger E, Edmunds K, Paul I, Helitzer DL. Unanticipated Death After Discharge Home From the Emergency Department. Annals of Emergency Medicine. 2007;49(6):735-45.

114. Gabayan GZ, Derose SF, Asch SM, Yiu S, Lancaster EM, Poon KT, et al. Patterns and Predictors of Short-Term Death After Emergency Department Discharge. Annals of Emergency Medicine. 2011;58:551-8.
115. Cabello J, Burls A, Emparanza J, Bayliss S, Quinn T. Oxygen therapy for acute myocardial infarction. Cochrane Database of Systematic Reviews. 2010;2010.
116. Villanueva C, Colomo A, Bosch A, Concepción M, Hernandez-Gea V, Aracil C, et al. Transfusion Strategies for Acute Upper Gastrointestinal Bleeding. New England Journal of Medicine. 2013;368:11-21.
117. Haukoos JS, Witt MD, Lewis RJ. Derivation and reliability of an instrument to estimate medical benefit of emergency treatment. American Journal of Emergency Medicine. 2010;28:404-11.
118. Edbrooke DL, Minelli C, Mills GH, Iapichino G, Pezzi A, Corbella D, et al. Implications of ICU triage decisions on patient mortality: a cost-effectiveness analysis. Critical Care. 2011;15:R56.
119. Cardoso LT, Grion CM, Matsuo T, Anami EH, Kauss IA, Seko L, et al. Impact of delayed admission to intensive care units on mortality of critically ill patients: a cohort study. Critical Care. 2011;15:R28.
120. Platts-Mills TF, Travers D, Biese K, McCall B, Kizer S, LaMantia M, et al. Accuracy of the Emergency Severity Index Triage Instrument for Identifying Elder Emergency Department Patients Receiving an Immediate Life-saving Intervention. Academic Emergency Medicine. 2010;17:238-43.
121. Gray A, Goodacre S, Nicholl J, Masson M, Sampson F, Elliott M, et al. The Development of a Simple Risk Score to Predict Early Outcome in Severe Acute Acidotic Cardiogenic Pulmonary Edema: The 3CPO Score. Circulation Heart Failure. 2010;3:111-7.
122. Vergouwe Y, Steyerberg EW, Eijkemans MJ, Habbema JDF. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. Journal of Clinical Epidemiology. 2005;58:475-83.
123. Hanley JA, McNeil BJ. A Method of Comparing the Areas under Receiver Operating Characteristic Curves Derived from the Same Cases. Radiology. 1983;148:839-43.
124. Obuchowski NA. Sample size calculations in studies of test accuracy. Statistical methods in medical research. 1998;7:371-92.
125. Bouamra O, Wrotchford A, Hollis S, Vail A, Woodford M, Lecky F. Outcome prediction in trauma. Injury. 2006;37:1092-7.
126. O'Driscoll B, Howard L, Davison A. BTS guideline for emergency oxygen use in adult patients. Thorax. 2008;63(suppl VI):vi1-68.
127. Goodacre S, Turner J, Nicholl J. Prediction of mortality among emergency medical admissions. Emergency Medicine Journal. 2006;23:372-5.
128. Man SY, Chan K, Wong F, Wong K, Yim C, Mak P, et al. Evaluation of the performance of a modified acute physiology and chronic health evaluation (APACHE II) scoring system for critically ill patients in emergency departments in Hong Kong. Academic Emergency Medicine. 2007;14(5):S85-6.
129. Cattermole GN, Mak SP, Liow CE, Ho MF, Hung KYG, Keung KM, et al. Derivation of a prognostic score for identifying critically ill patients in an emergency department resuscitation room. Resuscitation. 2009;80:1000-5.
130. Rycroft-Malone J. Formal consensus: the development of a national clinical guideline. Quality in Health Care. 2001;10:238-44.

131. Jones J, Hunter D. Consensus methods for medical and health services research. British Medical Journal. 1995;311:376-80.
132. Berthelot S, Lang ES, Quan H, Stelfox HT. Identifying Emergency-Sensitive Conditions for the Calculation of an Emergency Care Inhospital Standardized Mortality Ratio. Annals of Emergency Medicine. 2014;63:418-24.
133. Churpek MM, Yuen TC, Edelson DP. Predicting clinical deterioration in the hospital: The impact of outcome selection. Resuscitation. 2013;84:564-8.
134. Gabayan GZ, Asch SM, Hsia RY, Zingmond D, Liang L-J, Han W, et al. Factors Associated With Short-Term Bounce-Back Admissions After Emergency Department Discharge. Annals of Emergency Medicine. 2013;62:136-44.
135. Jones D, Mitchell I, Hillman K, Story D. Defining clinical deterioration. Resuscitation. 2013;84:1029-34.
136. Kline JA, Stubblefield WB. Clinician Gestalt Estimate of Pretest Probability for Acute Coronary Syndrome and Pulmonary Embolism in Patients With Chest Pain and Dyspnea. Annals of Emergency Medicine. 2014;63:275-80.
137. Green SM, Schriger DL, Yealy DM. Methodologic Standards for Interpreting Clinical Decision Rules in Emergency Medicine: 2014 Update. Annals of Emergency Medicine. 2014;doi 10.1016/j.annemergmed.2014.01.016.
138. Goodacre S, Cohen J, Bradburn M, Stevens J, Gray A, Benger J, et al. The 3Mg trial: a randomised controlled trial of intravenous or nebulised magnesium sulphate versus placebo in adults with acute severe asthma. Health Technology Assessment. 2014;18(22).
139. Cookson R, Claxton K, editors. The Humble Economist: Tony Culyer on Health, Health Care and Social Decision Making. York: University of York; 2012.
140. Girling AJ, Hofer TP, Wu J, Chilton PJ, Nicholl JP, Mohammed MA, et al. Case-mix adjusted hospital mortality is a poor proxy for preventable mortality: a modelling study. BMJ Quality and Safety. 2012;21:1052-6.
141. Hogan H, Healey F, Neale G, Thomson R, Vincent C, Black N. Preventable deaths due to problems in care in English acute hospitals: a retrospective case record review study. BMJ Quality and Safety. 2012;21:737-45.
142. Kellett J, Woodworth S, Wang F, Huang W. Changes and their prognostic implications in the abbreviated VitalpacTM early warning score (ViEWS) after admission to hospital of 18,853 acutely ill medical patients. Resuscitation. 2013;84:13-20.
143. Croskerry P. From Mindless to Mindful Practice — Cognitive Bias and Clinical Decision Making. New England Journal of Medicine. 2013;368:2445-8.
144. Gaddis GM, Greenwald P, Huckson S. Toward Improved Implementation of Evidence-based Clinical Algorithms: Clinical Practice Guidelines, Clinical Decision Rules, and Clinical Pathways. Academic Emergency Medicine. 2007;14:1015-22.
145. Sussman S. Handbook of Program Development for Health Behavior Research and Practice. London: Sage; 2001.
146. Taylor WJ, Brown M, Aati O, Weatherall M, Dalbeth N. Do Patient Preferences for Core Outcome Domains for Chronic Gout Studies Support

the Validity of Composite Response Criteria? Arthritis Care and Research. 2013;65:1259-64.
147. Ryan M, Gerard K. Using discrete choice experiments to value health care programmes: current practice and future research reflections. Applied Health Economics and Health Policy. 2003;2:55-64.
148. Gerard K, Lattimer V, Turnbull J, Smith H, George S, Brailsford S, et al. Reviewing emergency care systems 2: measuring patient preferences using a discrete choice experiment. Emergency Medicine Journal. 2004;21:692-7.
149. O'Cathain A, Knowles E, Maheswaran R, Turner J, Hirst E, Goodacre S, et al. Hospital characteristics affecting potentially avoidable emergency admissions: National ecological study. Health Services Management Research. 2014;DOI: 10.1177/0951484814525357.
150. Armitage G, Newell R, Wright J. Reporting drug errors in a British acute hospital trust. Clinical Governance. 2007;12:102-14.
151. Lazarides MK, Arvanitis DP, Drista H, Staramos DN, Dayantas JN. POSSUM and APACHE II Scores Do Not Predict the Outcome of Ruptured Infrarenal Aortic Aneurysms. Annals of Vascular Surgery. 1997;11:155-8.
152. Tambyraja A, Murie J, Chalmers R. Predictors of Outcome After Abdominal Aortic Aneurysm Rupture: Edinburgh Ruptured Aneurysm Score. World Journal of Surgery. 2007;31:2243-7.
153. Tambyraja AL, Lee AJ, Murie JA, Chalmers RT. Prognostic scoring in ruptured abdominal aortic aneurysm: A prospective evaluation. Journal of Vascular Surgery. 2008;47:282-6.
154. Tambyraja A, Fraser S, Murie J, Chalmers R. Validity of the Glasgow Aneurysm Score and the Hardman Index in predicting outcome after ruptured abdominal aortic aneurysm repair. British Journal of Surgery. 2005;92:573-3.
155. Leo E, Biancari F, Nesi F, Pogany G, Bartolucci R, Pasquale FD, et al. Risk-scoring methods in predicting the immediate outcome after emergency open repair of ruptured abdominal aortic aneurysm. American Journal of Surgery. 2006;192:19-23.
156. Prance S, Wilson Y, Cosgrove C, Walker A, Wilkins D, Ashley S. Ruptured Abdominal Aortic Aneurysms: Selecting Patients for Surgery. European Journal of Vascular and Endovascular Surgery. 1999;17:129-32.
157. Neary W, Crow P, Foy C, Prytherch D, Heather B, Earnshaw J. Comparison of POSSUM scoring and the Hardman Index in selection of patients for repair of ruptured abdominal aortic aneurysm. British Journal of Surgery. 2003;90:421-5.
158. Calderwood R, Halka T, Haji-Michael P, Welch M. Ruptured abdominal aortic aneurysm: is it possible to predict outcome? International Angiology. 2004;23:47-53.
159. Sharif MA, Lee B, Makar RR, Loan W, Soong CV. Role of the Hardman Index in Predicting Mortality for Open and Endovascular Repair of Ruptured Abdominal Aortic Aneurysm. Journal of Endovascular Therapy. 2007;14:528-35.
160. Karkos CD, Karamanos D, Papazoglou KO, Kantas AS, Theochari EG, Kamparoudis AG, et al. Usefulness of the Hardman index in predicting outcome after endovascular repair of ruptured abdominal aortic aneurysms. Journal of Vascular Surgery. 2008;48:788-94.
161. Moreau R, Soupison T, Vauquelin P, Derrida S, Beaucour H, Sicot C. Comparison of two simplified severity scores (SAPS and APACHE II) for

patients with acute myocardial infarction. Critical Care Medicine. 1989;17(5):409-13.
162. Alemi F, Rice J, Hankins R. Predicting In-Hospital Survival of Myocardial Infarction: A Comparative Study of Various Severity Measures. Medical Care. 1990;28:762-75.
163. Macdonald SP, Nagree Y, Fatovich DM, Flavell HL, Loutsky F. Comparison of two clinical scoring systems for emergency department risk stratification of suspected acute coronary syndrome. Emergency Medicine Australasia. 2011;23:717-25.
164. Bazzino O, Díaz R, Tajer C, Paviotti C, Mele E, Trivi M, et al. Clinical predictors of in-hospital prognosis in unstable angina: ECLA 3. American Heart Journal. 1999;137:322-31.
165. Chang W-C, Kaul P, Fu Y, Westerhout CM, Granger CB, Mahaffey KW, et al. Forecasting mortality: dynamic assessment of risk in ST-segment elevation acute myocardial infarction. European Heart Journal. 2006;27:419-26.
166. Rawlings C, Oglesby K, Turner J, Sen A. Comparison of two clinical scoring systems in risk stratification of non-ST elevation acute coronary syndrome patients in predicting 30-day outcomes. Emergency Medicine Journal. 2012;29:40-2.
167. Gale C, Manda S, Weston C, Birkhead J, Batin P, Hall A. Evaluation of risk scores for risk stratification of acute coronary syndromes in the Myocardial Infarction National Audit Project (MINAP) database. Heart. 2009;95:221-7.
168. Brieger D, Fox K, FitzGerald G, Eagle KA, Budaj A, Avezum Á, et al. Predicting freedom from clinical events in non-ST elevation acute coronary syndromes: the Global Registry of Acute Coronary Events. Heart. 2009;95:888-94.
169. Soderholm M, Deligani MM, Choudhary M, Bjork J, Ekelund U. Ability of risk scores to predict a low complication risk in patients admitted for suspected acute coronary syndrome. Emergency Medicine Journal. 2012;29:644-9.
170. Chin CT, Chen AY, Wang TY, Alexander KP, Mathews R, Rumsfeld JS, et al. Risk adjustment for in-hospital mortality of contemporary patients with acute myocardial infarction: The Acute Coronary Treatment and Intervention Outcomes Network (ACTION) Registry®–Get With The Guidelines (GWTG)™ acute myocardial infarction mortality model and risk score. American Heart Journal. 2011;161:113-22.
171. Goldman L, Cook EF, Johnson PA, Brand DA, Rouan GW, Lee TH. Prediction of the Need for Intensive Care in Patients Who Come to Emergency Departments with Acute Chest Pain. New England Journal of Medicine. 1996;334:1498-504.
172. Durairaj L, Reilly B, Das K, Smith C, Acob C, Husain S, et al. Emergency Department Admissions to Inpatient Cardiac Telemetry Beds: A Prospective Cohort Study of Risk Stratification and Outcomes. American Journal of Medicine. 2001;110:7-11.
173. Limkakeng A, Gilbler WB, Pollack C, Hoekstra JW, Sites F, Shofer FS, et al. Combination of Goldman risk and initial cardiac troponin I for emergency department chest pain patient risk stratification. Academic Emergency Medicine. 2001;8(7):696-702.
174. Hollander JE, Sites FD, Pollack CV, Shofer FS. Lack of Utility of Telemetry Monitoring for Identification of Cardiac Death and Life-

Threatening Ventricular Dysrhythmias in Low-Risk Patients With Chest Pain. Annals of Emergency Medicine. 2004;43:71-6.
175. Manini AF, Dannemann N, Brown DF, Butler J, Bamberg F, Nagurney JT, et al. Limitations of risk score models in patients with acute chest pain. American Journal of Emergency Medicine. 2009;27:43-8.
176. de Araujo Goncalves P, Ferreira J, Aguiar C, Seabra-Gomes R. TIMI, PURSUIT, and GRACE risk scores: sustained prognostic value and interaction with revascularization in NSTE-ACS. European Heart Journal. 2005;26:865-72.
177. Yan AT, Jong P, Yan RT, Tan M, Fitchett D, Chow C-M, et al. Clinical trial–derived risk model may not generalize to real-world patients with acute coronary syndrome. American Heart Journal. 2004;148:1020-7.
178. Rahimi K, Watzlawek S, Thiele H, Secknus M-A, Hayerizadeh B-F, Niebauer J, et al. Incidence, time course, and predictors of early malignant ventricular arrhythmias after non-ST-segment elevation myocardial infarction in patients with early invasive treatment. European Heart Journal. 2006;27:1706-11.
179. Lyon R, Morris AC, Caesar D, Gray S, Gray A. Chest pain presenting to the Emergency Department—to stratify risk with GRACE or TIMI? Resuscitation. 2007;74:90-3.
180. Sinclair H, Paterson M, Walker S, Beckett G, Fox K. Predicting Outcome in Patients with Acute Coronary Syndrome: Evaluation of B-Type Natriuretic Peptide and the Global Registry of Acute Coronary Events (GRACE) Risk Score. Scottish Medical Journal. 2007;52:8-13.
181. Yan AT, Yan RT, Tan M, Casanova A, Labinaz M, Sridhar K, et al. Risk scores for risk stratification in acute coronary syndromes: useful but simpler is not necessarily better. European Heart Journal. 2007;28:1072-8.
182. Lev EI, Kornowski R, Vaknin-Assa H, Porter A, Teplitsky I, Ben-Dor I, et al. Comparison of the Predictive Value of Four Different Risk Scores for Outcomes of Patients With ST-Elevation Acute Myocardial Infarction Undergoing Primary Percutaneous Coronary Intervention. American Journal of Cardiology. 2008;102:6-11.
183. Elbarouni B, Goodman SG, Yan RT, Welsh RC, Kornder JM, DeYoung JP, et al. Validation of the Global Registry of Acute Coronary Event (GRACE) risk score for in-hospital mortality in patients with acute coronary syndrome in Canada. American Heart Journal. 2009;158:392-9.
184. Goodacre S, Bradburn M, Fitzgerald P, Cross E, Collinson P, Gray A, et al. The RATPAC (Randomised Assessment of Treatment using Panel Assay of Cardiac markers) trial: a randomised controlled trial of point-of-care cardiac markers in the emergency department. Health Technology Assessment. 2011;15(23):1-110.
185. Goodacre SW, Bradburn M, Mohamed A, Gray A. Evaluation of Global Registry of Acute Cardiac Events and Thrombolysis in Myocardial Infarction scores in patients with suspected acute coronary syndrome. American Journal of Emergency Medicine. 2012;30:37-44.
186. Gale C, Manda S, Batin P, Weston C, Birkhead J, Hall A. Predictors of in-hospital mortality for patients admitted with ST-elevation myocardial infarction: a real-world study using the Myocardial Infarction National Audit Project (MINAP) database. Heart. 2008;94:1407-12.
187. Normand S-LT, Glickman ME, Sharma R, McNeil BJ. Using Admission Characteristics to Predict Short-term Mortality From Myocardial Infarction

in Elderly Patients: Results From the Cooperative Cardiovascular Project. JAMA. 1996;275:1322-8.
188. Hess EP, Brison RJ, Perry JJ, Calder LA, Thiruganasambandamoorthy V, Agarwal D, et al. Development of a Clinical Prediction Rule for 30-Day Cardiac Events in Emergency Department Patients With Chest Pain and Possible Acute Coronary Syndrome. Annals of Emergency Medicine. 2012;59:115-25.
189. Boersma E, Peiper KS, Steyerberg E, Wilcox RG, Chang W-C, Lee KL, et al. Predictors of outcome in patients with acute coronary syndromes without persistent ST-segment elevation: result from an international trial of 9461 patients. Circulation. 2000;101(22):2557-67.
190. Brilakis ES, Wright RS, Kopecky SL, Mavrogiorgos NC, Reeder GS, Rihal CS, et al. Association of the PURSUIT risk score with predischarge ejection fraction, angiographic severity of coronary artery disease, and mortality in a nonselected, community-based population with non-ST–elevation acute myocardial infarction. American Heart Journal. 2003;146:811-8.
191. Selker HP, Griffith JL, D'Agostino RB. A Time-Insensitive Predictive Instrument for Acute Myocardial Infarction Mortality: A Multicenter Study. Medical Care. 1991;29:1196-211.
192. Rathore SS, Weinfurt KP, Gross CP, Krumholz HM. Validity of a Simple ST-Elevation Acute Myocardial Infarction Risk Index: Are Randomized Trial Prognostic Estimates Generalizable to Elderly Patients? Circulation. 2003;107:811-6.
193. Das R, Dorsch M, Lawrance R, Kilcullen N, Sapsford R, Robinson M, et al. External validation, extension and recalibration of Braunwald's simple risk index in a community-based cohort of patients with both STEMI and NSTEMI. International Journal of Cardiology. 2006;107:327-32.
194. Wiviott SD, Morrow DA, Frederick PD, Giugliano RP, Gibson CM, McCabe CH, et al. Performance of the Thrombolysis In Myocardial Infarction Risk Index in the National Registry of Myocardial Infarction-3 and -4: A Simple Index That Predicts Mortality in ST-Segment Elevation Myocardial Infarction. Journal of the American College of Cardiology. 2004;44:783-9.
195. Ilkhanoff L, O'Donnell CJ, Camargo CA, O'Halloran TD, Giugliano RP, Lloyd-Jones DM. Usefulness of the TIMI Risk Index in Predicting Short- and Long-Term Mortality in Patients With Acute Coronary Syndromes. American Journal of Cardiology. 2005;96:773-7.
196. Bradshaw PJ, Ko DT, Newman AM, Donovan LR, Tu JV. Validation of the Thrombolysis In Myocardial Infarction (TIMI) risk index for predicting early mortality in a population-based cohort of STEMI and non-STEMI patients. Canadian Journal of Cardiology. 2007;23:51-6.
197. García-Almagro FJ, Gimeno JR, Villegas M, Hurtado J, Teruel F, Cerdán MC, et al. Prognostic value of the Thrombolysis in Myocardial Infarction risk score in a unselected population with chest pain. Construction of a new predictive model. American Journal of Emergency Medicine. 2008;26:439-45.
198. Samaha FF, Kimmel SE, Kizer JR, Goyal A, Wade M, Boden WE. Usefulness of the TIMI Risk Score in Predicting Both Short- and Long-Term Outcomes in the Veterans Affairs Non–QWave Myocardial Infarction Strategies In-Hospital (VANQWISH) Trial. American Journal of Cardiology. 2002;90:922-6.

199. Foussas SG, Zairis MN, Lyras AG, Patsourakos NG, Tsirimpis VG, Katsaros K, et al. Early Prognostic Usefulness of C-Reactive Protein Added to the Thrombolysis In Myocardial Infarction Risk Score in Acute Coronary Syndromes. American Journal of Cardiology. 2005;96:533-7.
200. Conway Morris A, Caesar D, Gray S, Gray A. TIMI risk score accurately risk stratifies patients with undifferentiated chest pain presenting to an emergency department. Heart. 2006;92:1333-4.
201. Pollack CV, Sites FD, Shofer FS, Sease KL, Hollander JE. Application of the TIMI Risk Score for Unstable Angina and Non-ST Elevation Acute Coronary Syndrome to an Unselected Emergency Department Chest Pain Population. Academic Emergency Medicine. 2006;13:13-8.
202. Soiza R, Leslie S, Williamson P. Risk stratification in acute coronary syndromes—does the TIMI risk score work in unselected cases? Quarterly Journal of Medicine. 2006;99(2):81-7.
203. Chase M, Brown AM, Robey JL, Zogby KE, Shofer FS, Chmielewski L, et al. Application of the TIMI risk score in ED patients with cocaine-associated chest pain. American Journal of Emergency Medicine. 2007;25:1015-8.
204. Jaffery Z, Hudson MP, Jacobsen G, Nowak R, McCord J. Modified Thrombolysis in Myocardial Infarction (TIMI) risk score to risk stratify patients in the emergency department with possible acute coronary syndrome. Journal of Thrombosis and Thrombolysis. 2007;24:137-44.
205. Karounos M, Chang AM, Robey JL, Sease KL, Shofer FS, Follansbee C, et al. TIMI risk score: does it work equally well in both males and females? Emergency Medicine Journal. 2007;24:471-4.
206. Body R, Carley S, McDowell G, Ferguson J, Mackway-Jones K. Can a modified thrombolysis in myocardial infarction risk score outperform the original for risk stratifying emergency department patients with chest pain? Emergency Medicine Journal. 2009;26:95-9.
207. Campbell CF, Chang AM, Sease KL, Christopher Follansbee, McCusker CM, Shofer FS, et al. Combining Thrombolysis in Myocardial Infarction risk score and clear-cut alternative diagnosis for chest pain risk stratification. American Journal of Emergency Medicine. 2009;27:37-42.
208. Morrow DA, Antman EM, Charlesworth A, Cairns R, Murphy SA, de Lemos JA, et al. TIMI Risk Score for ST-Elevation Myocardial Infarction: A Convenient, Bedside, Clinical Score for Risk Assessment at Presentation. Circulation. 2000;102:2031-7.
209. Januzzi JL, Newby LK, Murphy SA, Pieper K, Antman EM, Morrow DA, et al. Predicting a late positive serum troponin in initially troponin-negative patients with non–ST-elevation acute coronary syndrome: Clinical predictors and validated risk score results from the TIMI IIIB and GUSTO IIA studies. American Heart Journal. 2006;151:360-6.
210. Tabak YP, Sun X, Johannes RS, Gupta V, Shorr AF. Mortality and Need for Mechanical Ventilation in Acute Exacerbations of Chronic Obstructive Pulmonary Disease: Development and Validation of a Simple Risk Score. Archives of Internal Medicine. 2009;169:1595-602.
211. Shorr AF, Sun X, Johannes RS, Yaitanes A, Tabak YP. Validation of a Novel Risk Score for Severity of Illness in Acute Exacerbations of COPD. Chest. 2011;140:1177-83.
212. Tsai C-L, Clark S, Camargo CA. Risk stratification for hospitalization in acute asthma: the CHOP classification tree. American Journal of Emergency Medicine. 2010;28:803-8.

213. Chang CL, Sullivan GD, Karalus NC, Mills GD, McLachlan JD, Hancox RJ. Predicting early mortality in acute exacerbation of chronic obstructive pulmonary disease using the CURB65 score. Respirology. 2011;16:146-51.
214. Steer J, Norman E, Afolabi O, Gibson G, Bourke S. Dyspnoea severity and pneumonia as predictors of in-hospital mortality and early readmission in acute exacerbations of COPD. Thorax. 2012;67:117-21.
215. Kelly A-M, Kerr D, Powell C. Is severity assessment after one hour of treatment better for predicting the need for admission in acute asthma? Respiratory Medicine. 2004;98:777-81.
216. Rodrigo G, Rodrigo C. A New Index for Early Prediction of Hospitalization in Patients With Acute Asthma. American Journal of Emergency Medicine. 1997;15:8-13.
217. Gralnek IM, Dulai GS. Incremental value of upper endoscopy for triage of patients with acute non-variceal upper-GI hemorrhage. Gastrointestinal endoscopy. 2004;60:9-14.
218. Chen I-C, Hung M-S, Chiu T-F, Chen J-C, Hsiao C-T. Risk scoring systems to predict need for clinical intervention for patients with nonvariceal upper gastrointestinal tract bleeding. American Journal of Emergency Medicine. 2007;25:774-9.
219. Masaoka T, Suzuki H, Hori S, Aikawa N, Hibi T. Blatchford scoring system is a useful scoring system for detecting patients with upper gastrointestinal bleeding who do not need endoscopic intervention. Journal of Gastroenterology and Hepatology. 2007;22:1404-8.
220. Stanley A, Ashley D, Dalton H, Mowat C, Gaya D, Thompson E, et al. Outpatient management of patients with low-risk upper-gastrointestinal haemorrhage: multicentre validation and prospective evaluation. Lancet. 2009;373:42-7.
221. Chandra S, Hess EP, Agarwal D, Nestler DM, Montori VM, Song LMWK, et al. External validation of the Glasgow-Blatchford Bleeding Score and the Rockall Score in the US setting. American Journal of Emergency Medicine. 2012;30:673-9.
222. Farooq FT, Lee MH, Das A, Dixit R, Wong RCK. Clinical triage decision vs risk scores in predicting the need for endotherapy in upper gastrointestinal bleeding. American Journal of Emergency Medicine. 2012;30:129-34.
223. Romagnuolo J, Barkun AN, Enns R, Armstrong D, Gregor J. Simple Clinical Predictors May Obviate Urgent Endoscopy in Selected Patients With Nonvariceal Upper Gastrointestinal Tract Bleeding. Archives of Internal Medicine. 2007;167:265-70.
224. Kollef MH, O'Brien JD, Zuckerman GR, Shannon W. BLEED: A classification tool to predict outcomes in patients with acute upper and lower gastrointestinal hemorrhage. Critical Care Medicine. 1997;25:1125-32.
225. Bordley DR, Mushlin AI, Dolan JG, Richardson WS, Barry M, Polio J, et al. Early Clinical Signs Identify Low-Risk Patients With Acute Upper Gastrointestinal Hemorrhage. JAMA. 1985;253:3282-5.
226. Church NI, Dallal HJ, Masson J, Mowat NAG, Johnston DA, Radin E, et al. Validity of the Rockall scoring system after endoscopic therapy for bleeding peptic ulcer: a prospective cohort study. Gastrointestinal endoscopy. 2006;63:606-12.
227. Sarwar S, Dilshad A, Khan AA, Alam A, Butt AK, Tariq S, et al. Predictive value of Rockall score for rebleeding and mortality in patients

with variceal bleeding. Journal of the College of Physicians and Surgeons of Pakistan. 2007;17:253-6.
228. Sanders D, Carter M, Goodchap R, Cross S, Gleeson D, Lobo A. Prospective Validation of the Rockall Risk Scoring System for Upper GI Hemorrhage in Subgroups of Patients With Varices and Peptic Ulcers. American Journal of Gastroenterology. 2002;97(3):630-5.
229. Das A, Ben-Menachem T, Farooq FT, Cooper GS, Chak A, Sivak MV, et al. Artificial Neural Network as a Predictive Instrument in Patients With Acute Nonvariceal Upper Gastrointestinal Hemorrhage. Gastroenterology. 2008;134:65-74.
230. Strate LL, Saltzman JR, Ookubo R, Mutinga ML, Syngal S. Validation of a Clinical Prediction Rule for Severe Acute Lower Intestinal Bleeding. American Journal of Gastroenterology. 2005;100:1821-7.
231. Auble TE, Hsieh M, McCausland JB, Yealy DM. Comparison of Four Clinical Prediction Rules for Estimating Risk in Heart Failure. Annals of Emergency Medicine. 2007;50:127-35.
232. Lee DS, Austin PC, Rouleau JL, Liu PP, Naimark D, Tu JV. Predicting mortality among patients hospitalized for heart failure. JAMA. 2003;290:2581-7.
233. Peterson PN, Rumsfeld JS, Liang L, Albert NM, Hernandez AF, Peterson ED, et al. A Validated Risk Score for In-Hospital Mortality in Patients With Heart Failure From the American Heart Association Get With the Guidelines Program. Circulation Cardiovascular Quality and Outcomes. 2010;3:25-32.
234. Le Conte P, Coutant V, N'Guyen JM, Baron D, Touze MD, Potel G. Prognostic Factors in Acute Cardiogenic Pulmonary Edema. American Journal of Emergency Medicine. 1999;17:329-32.
235. Fiutowski M, Waszyrowski T, Krzeminska-Pakula M, Kasprzak JD. Pulmonary edema prognostic score predicts in-hospital mortality risk in patients with acute cardiogenic pulmonary edema. Heart Lung. 2008;37:46-53.
236. Muller MP, McGeer AJ, Hassan K, Marshall J, Christian M. Evaluation of Pneumonia Severity and Acute Physiology Scores to Predict ICU Admission and Mortality in Patients Hospitalized for Influenza. PLos One. 2010;5:e9563.
237. Mulrennan S, Tempone SS, Ling ITW, Williams SH, Gan G-C, Murray RJ, et al. Pandemic Influenza (H1N1) 2009 Pneumonia: CURB-65 Score for Predicting Severity and Nasopharyngeal Sampling for Diagnosis Are Unreliable. PLos One. 2010;5:e12849.
238. Adeniji KA, Cusack R. The Simple Triage Scoring System (STSS) successfully predicts mortality and critical care resource utilization in H1N1 pandemic flu: a retrospective analysis. Critical Care. 2011;15:R39.
239. Meek K, Toosie K, Stabile BE, Elbassir M, Murrell Z, Lewis RJ, et al. Simplified Admission Criterion for Predicting Severe Complications of Gallstone Pancreatitis. Archives of Surgery. 2000;135:1048-54.
240. Gan I, May G, Raboud J, Tilley J, Enns R. Pancreatitis in HIV Infection: Predictors of Severity. American Journal of Gastroenterology. 2003;98(6):1278-83.
241. Halonen KI, Leppäniemi AK, Lundin JE, Puolakkainen PA, Kemppainen EA, Haapianen RK. Predicting Fatal Outcome in the Early Phase of Severe Acute Pancreatitis by Using Novel Prognostic Methods. Pancreatology. 2003;3:309-15.

242. Gürleyik G, Emir S, Kiliçoglu G, Arman A, Saglam A. Computed Tomography Severity Index, APACHE II Score, and Serum CRP Concentration for Predicting the Severity of Acute Pancreatitis. Journal of the Pancreas. 2005;6:562-7.
243. Taylor SL, Morgan DL, Denson KD, Lane MM, Pennington LR. A comparison of the Ranson, Glasgow, and APACHE II scoring systems to a multiple organ system score in predicting patient outcome in pancreatitis. American Journal of Surgery. 2005;189:219-22.
244. Papachristou GI, Papachristou DJ, Avula H, Slivka A, Whitcomb DC. Obesity Increases the Severity of Acute Pancreatitis: Performance of APACHE-O Score and Correlation with the Inflammatory Response. Pancreatology. 2006;6:279-85.
245. Yeung YP, Lam BYK, Yip AWC. APACHE system is better than Ranson system in the prediction of severity of acute pancreatitis. Hepatobiliary and Pancreatic Disease International. 2006;5:294-9.
246. Ueda T, Takeyama Y, Yasuda T, Matsumura N, Sawa H, Nakajima T, et al. Simple scoring system for the prediction of the prognosis of severe acute pancreatitis. Surgery. 2007;141:51-8.
247. Garcea G, Gouda M, Hebbes C, Ong SL, Neal CP, Dennison AR, et al. Predictors of Severity and Survival in Acute Pancreatitis: Validation of the Efficacy of Early Warning Scores. Pancreas. 2008;37:e54-61.
248. Singh VK, Wu BU, Bollen TL, Repas K, Maurer R, Johannes RS, et al. A Prospective Evaluation of the Bedside Index for Severity in Acute Pancreatitis Score in Assessing Mortality and Intermediate Markers of Severity in Acute Pancreatitis. American Journal of Gastroenterology. 2009;104:966-71.
249. de Beaux A, Palmer K, Carter D. Factors influencing morbidity and mortality in acute pancreatitis; an analysis of 279 cases. Gut. 1995;37:121-6.
250. Meek K, Virgilio Cd, Murrell Z, Stabile BE, Elbassir M, Renslo R, et al. Correlation between Admission Laboratory Values, Early Abdominal Computed Tomography, and Severe Complications of Gallstone Pancreatitis. American Journal of Surgery. 2000;180:556-60.
251. van den Biezenbos A, Kruyt P, Bosscha K, van Leeuwen M, Feldberg M, van der Schouw Y, et al. Added value of CT criteria compared to the clinical SAP score in patients with acute pancreatitis. Abdominal imaging. 1998;23:622-6.
252. Shindo Y, Sato S, Maruyama E, Ohashi T, Ogawa M, Imaizumi K, et al. Comparison of severity scoring systems A-DROP and CURB-65 for community-acquired pneumonia. Respirology. 2008;13:731-5.
253. Jeong KY, Kim K, Kim TY, Lee CC, Jo SO, Rhee JE, et al. Prognostic value of N-terminal pro-brain natriuretic peptide in hospitalised patients with community-acquired pneumonia. Emergency Medicine Journal. 2011;28:122-7.
254. Angus DC, Marrie TJ, Obrosky DS, Clermont G, Dremsizov TT, Coley C, et al. Severe Community-acquired Pneumonia Use of Intensive Care Services and Evaluation of American and British Thoracic Society Diagnostic Criteria. American Journal of Respiratory and Critical Care Medicine. 2002;166:717-23.
255. Buising KL, Thursky KA, Black JF, Macgregor L, Street A, Kennedy M, et al. Reconsidering what is meant by severe pneumonia: a prospective

comparison of severity scores for community acquired pneumonia. Thorax. 2006;61:419-24.
256. Kontou P, Kuti JL, Nicolau DP. Validation of the Infectious Diseases Society of America/American Thoracic Society criteria to predict severe community-acquired pneumonia caused by Streptococcus pneumoniae. American Journal of Emergency Medicine. 2009;27:968-74.
257. Valencia M, Badia JR, Cavalcanti M, Ferrer M, Agusti C, Angrill J, et al. Pneumonia Severity Index Class V Patients With Community-Acquired Pneumonia: Characteristics, Outcomes, and Value of Severity Scores. Chest. 2007;132:515-22.
258. Feldman C, Alanee S, Yu V, Richards G, Ortqvist A, Rello J, et al. Severity of illness scoring systems in patients with bacteraemic pneumococcal pneumonia: implications for the intensive care unit care. Clinical Microbiology and Infection. 2009;15:850-7.
259. Phua J, See K, Chan Y, Widjaja L, Aung N, Ngerng W, et al. Validation and clinical implications of the IDSA/ATS minor criteria for severe community-acquired pneumonia. Thorax. 2009;64:598-603.
260. Man SY, Graham CA, Chan SSW, Mak PSK, Yu AHY, Cheung CSK, et al. Disease severity prediction for nursing home-acquired pneumonia in the emergency department. Emergency Medicine Journal. 2011;28:1046-50.
261. Loh L-C, Khoo S-K, Quah S-Y, Visvalingam V, Radhakrishnan A, Vijayasingham P, et al. Adult community-acquired pneumonia in Malaysia: Prediction of mortality from severity assessment on admission. Respirology. 2004;9:379-86.
262. Buising KL, Thursky KA, Black JF, MacGregor L, Street AC, Kennedy MP, et al. Identifying severe community-acquired pneumonia in the emergency department: A simple clinical prediction tool. Emergency Medicine Australasia. 2007;19:418-26.
263. Bauer T, Ewig S, Marre R, Suttorp N, Welte T. CRB-65 predicts death from community-acquired pneumonia. Journal of Internal Medicine. 2006;260(1):93-101.
264. Capelastegui A, Espana P, Quintana J, Areitio I, Gorordo I, Egurrola M, et al. Validation of a predictive rule for the management of community-acquired pneumonia. European Respiratory Journal. 2006;27:151-7.
265. Barlow G, Nathwani D, Davey P. The CURB65 pneumonia severity score outperforms generic sepsis and early warning scores in predicting mortality in community-acquired pneumonia. Thorax. 2007;62:253-9.
266. Man SY, Lee N, Ip M, Antonio GE, Chau SS, Mak P, et al. Prospective comparison of three predictive rules for assessing severity of community-acquired pneumonia in Hong Kong. Thorax. 2007;62:348-53.
267. Schaaf B, Kruse J, Rupp J, Reinert R, Droemann D, Zabel P, et al. Sepsis severity predicts outcome in community-acquired pneumococcal pneumonia. European Respiratory Journal. 2007;30:517-24.
268. Chalmers J, Singanayagam A, Hill A. Systolic blood pressure is superior to other haemodynamic predictors of outcome in community acquired pneumonia. Thorax. 2008;63:698-702.
269. Kruger S, Ewig S, Marre R, Papassotiriou J, Richter K, von Baum H, et al. Procalcitonin predicts patients at low risk of death from community-acquired pneumonia across all CRB-65 classes. European Respiratory Journal. 2008;31:349-55.
270. Schuetz P, Koller M, Christ-Crain M, Steyerberg E, Stolz D, Muller C, et al. Predicting mortality with pneumonia severity scores: importance of

model recalibration to local settings. Epidemiology and Infection. 2008;136:1628-37.
271. Zuberi F, Khan J. Prospective comparison of prediction rules of mortality risk for CAP in a developing country. International Journal of Tuberculosis and Lung Disease. 2008;12:447-52.
272. Chalmers J, Singanayagam A, Murray M, Scally C, Fawzi A, Hill A. Risk factors for complicated parapneumonic effusion and empyema on presentation to hospital with community-acquired pneumonia. Thorax. 2009;64:592-7.
273. Menendez R, Martinez R, Reyes S, Mensa J, Filella X, Marcos M, et al. Biomarkers improve mortality prediction by prognostic scales in community-acquired pneumonia. Thorax. 2009;64:587-91.
274. El-Solh AA, Alhajhusain A, Jaoude PA, Drinka P. Validity of Severity Scores in Hospitalized Patients With Nursing Home-Acquired Pneumonia. Chest. 2010;138:1371-6.
275. Ewig S, de Roux A, Bauer T, Garcia E, Mensa J, Niederman M, et al. Validation of predictive rules and indices of severity for community acquired pneumonia. Thorax. 2004;59:421-7.
276. Aujesky D, Auble TE, Yealy DM, Stone RA, Obrosky DS, Meehan TP, et al. Prospective comparison of three validated prediction rules for prognosis in community-acquired pneumonia. American Journal of Medicine. 2005;118:384-92.
277. España PP, Capelastegui A, Gorordo I, Esteban C, Oribe M, Ortega M, et al. Development and Validation of a Clinical Prediction Rule for Severe Community-acquired Pneumonia. American Journal of Respiratory and Critical Care Medicine. 2006;174:1249-56.
278. Ananda-Rajah MR, Charles PG, Melvani S, Burrell LL, Johnson PD, Grayson ML. Comparing the pneumonia severity index with CURB-65 in patients admitted with community acquired pneumonia. Scandinavian Journal of Infectious Diseases. 2008;40:293-300.
279. Charles PG, Wolfe R, Whitby M, Fine MJ, Fuller AJ, Stirling R, et al. SMART-COP: A Tool for Predicting the Need for Intensive Respiratory or Vasopressor Support in Community-Acquired Pneumonia. Clinical Infectious Diseases. 2008;47:375-84.
280. Schuetz P, Stolz D, Mueller B, Morgenthaler NG, Struck J, Mueller C, et al. Endothelin-1 precursor peptides correlate with severity of disease and outcome in patients with community acquired pneumonia. BMC Infectious Diseases. 2008;8(22).
281. Chalmers JD, Singanayagam A, Scally C, Hill AT. Admission D-dimer Can Identify Low-Risk Patients With Community-Acquired Pneumonia. Annals of Emergency Medicine. 2009;53:633-8.
282. Huang DT, Angus DC, Kellum JA, Pugh NA, Weissfeld LA, Struck J, et al. Midregional Proadrenomedullin as a Prognostic Tool in Community-Acquired Pneumonia. Chest. 2009;136:823-31.
283. Parsonage M, Nathwani D, Davey P, Barlow G. Evaluation of the performance of CURB-65 with increasing age. Clinical Microbiology and Infection. 2009;15:858-64.
284. Yandiola PPE, Capelastegui A, Quintana J, Diez R, Gorordo I, Bilbao A, et al. Prospective Comparison of Severity Scores for Predicting Clinically Relevant Outcomes for Patients Hospitalized With Community-Acquired Pneumonia. Chest. 2009;135:1572-9.

285. Chen J-H, Chang S-S, Liu JJ, Chan R-C, Wu J-Y, Wang W-C, et al. Comparison of clinical characteristics and performance of pneumonia severity score and CURB-65 among younger adults, elderly and very old subjects. Thorax. 2010;65:971-7.
286. Schuetz P, Wolbers M, Christ-Crain M, Thomann R, Falconnier C, Widmer I, et al. Prohormones for prediction of adverse medical outcome in community-acquired pneumonia and lower respiratory tract infections. Critical Care. 2010;14:R106.
287. Albrich WC, Dusemund F, Rüegger K, Christ-Crain M, Zimmerli W, Bregenzer T, et al. Enhancement of CURB65 score with proadrenomedullin (CURB65-A) for outcome prediction in lower respiratory tract infections: Derivation of a clinical algorithm. BMC Infectious Diseases. 2011;11:112.
288. Jones BE, Jones J, Bewick T, Lim WS, Aronsky D, Brown SM, et al. CURB-65 Pneumonia Severity Assessment Adapted for Electronic Decision Support. Chest. 2011;140:156-63.
289. Labarère J, Schuetz P, Renaud B, Claessens Y-E, Albrich W, Mueller B. Validation of a Clinical Prediction Model for Early Admission to the Intensive Care Unit of Patients With Pneumonia. Academic Emergency Medicine. 2012;19:994-1003.
290. Park JH, Wee JH, Choi SP, Oh SH. The value of procalcitonin level in community-acquired pneumonia in the ED. American Journal of Emergency Medicine. 2012;30:1248-54.
291. Carusone SBC, Walter SD, Brazil K, Loeb MB. Pneumonia and Lower Respiratory Infections in Nursing Home Residents: Predictors of Hospitalization and Mortality. Journal of the American Geriatric Society. 2007;55:414-9.
292. Flanders WD, Tucker G, Krishnadasan A, Martin D, Honig E, McClellan WM. Validation of the pneumonia severity index: importance of study-specific recalibration. Journal of General and Internal Medicine. 1999;14:333-40.
293. Feagan BG, Marrie TJ, Lau CY, Wheeler SL, Wong CJ, Vandervoort MK. Treatment and outcomes of community-acquired pneumonia at Canadian hospitals. Canadian Medical Association Journal. 2000;162:1415-20.
294. Dedier J, Singer DE, Chang Y, Moore M, Atlas SJ. Processes of Care, Illness Severity, and Outcomes in the Management of Community-Acquired Pneumonia at Academic Hospitals. Archives of Internal Medicine. 2001;161:2099-104.
295. Mody L, Sun R, Bradley S. Community-Acquired Pneumonia in Older Veterans: Does the Pneumonia Prognosis Index Help? Journal of the American Geriatric Society. 2002;50:434-8.
296. Querol-Ribelles JM, Tenias JM, Grau E, Querol-Borras JM, Climent JL, Gomez E, et al. Plasma d-Dimer Levels Correlate With Outcomes in Patients With Community-Acquired Pneumonia. Chest. 2004;126:1087-92.
297. Migliorati P, Boccoli E, Bracci L, Sestini P, Melani A. A survey on hospitalised community-acquired pneumonia in Italy. Monaldi Archives of Chest Disease. 2006;65:82-8.
298. Sanders KM, Marras TK, Chan CK. Pneumonia severity index in the immunocompromised. Canadian Respiratory Journal. 2006;13:89-93.
299. Etzion O, Novack V, Avnon L, Porath A, Dagan E, Riesenberg K, et al. Characteristics of low-risk patients hospitalized with community-acquired pneumonia. European Journal of Internal Medicine. 2007;18:209-14.

300. Renaud B, Coma E, Hayon J, Gurgui M, Longo C, Blancher M, et al. Investigation of the ability of the Pneumonia Severity Index to accurately predict clinically relevant outcomes: a European study. Clinical Microbiology and Infection. 2007;13:923-31.
301. Chen C-Z, Fan P-S, Lin C-C, Lee C-H, Hsiue T-R. Repeated Pneumonia Severity Index Measurement After Admission Increases its Predictive Value for Mortality in Severe Community-acquired Pneumonia. Journal of the Formosan Medical Association. 2008;108:219-23.
302. Garau J, Baquero F, Perez-Trallero E, Perez J-L, Martin-Sanchez A, Garcia-Rey C, et al. Factors impacting on length of stay and mortality of community-acquired pneumonia. Clinical Microbiology and Infection. 2008;14:322-9.
303. Restrepo MI, Mortensen EM, Velez JA, Frei C, Anzueto A. A Comparative Study of Community-Acquired Pneumonia Patients Admitted to the Ward and the ICU. Chest. 2008;133:610-7.
304. Pilotto A, Addante F, Ferrucci L, Leandro G, D'Onofrio G, Corritore M, et al. The Multidimensional Prognostic Index Predicts Short- and Long-Term Mortality in Hospitalized Geriatric Patients With Pneumonia. Journal of Gerontology A Biological Sciences Medical Sciences. 2009;64A:880-7.
305. Aujesky D, Obrosky DS, Stone RA, Auble TE, Perrier A, Cornuz J, et al. A Prediction Rule to Identify Low-Risk Patients With Pulmonary Embolism. Archives of Internal Medicine. 2006;166:169-75.
306. Palmieri V, Gallotta G, Rendina D, De Bonis S, Russo V, Postiglione A, et al. Troponin I and right ventricular dysfunction for risk assessment in patients with nonmassive pulmonary embolism in the Emergency Department in combination with clinically based risk score. Internal and Emergency Medicine. 2008;3:131-8.
307. Choi W-H, Kwon SU, Jwa YJ, Kim JA, Choi Y-H, Chang JH, et al. The Pulmonary Embolism Severity Index in Predicting the Prognosis of Patients With Pulmonary Embolism. Korean Journal of Internal Medicine. 2009;24:123-7.
308. Nordenholz K, Ryan J, Atwood B, Heard K. Pulmonary embolism risk stratification: pulse oximetry and pulmonary embolism severity index. Journal of Emergency Medicine. 2011;40:95-102.
309. Agterof MJ, Schutgens REG, Moumli N, Eijkemans MJC, van der Griend R, Tromp EAM, et al. A prognostic model for short term adverse events in normotensive patients with pulmonary embolism. American Journal of Hematology. 2011;86:646-9.
310. Nguyen HB, Banta JE, Cho TW, Van Ginkel C, Burroughs K, Wittlake WA, et al. Mortality predictions using current physiological scoring systems in patients meeting criteria for early goal-directed therapy and the severe sepsis resuscitation bundle. Shock. 2008;30:23-8.
311. Chou T-NK, Lee Y-T, Lai Y-Y, Chao W-N, Yang C, Chen C-C, et al. Prognostic factors for primary septicemia and wound infection caused by Vibrio vulnificus. American Journal of Emergency Medicine. 2010;28:424-31.
312. Howell MD, Donnino M, Talmor D, Clardy P, Long N, Shapiro N. Performance of severity of illness scoring systems in Emergency Department patients with infection. Academic Emergency Medicine. 2007;14:709-14.
313. Crowe CA, Kulstad EB, Mistry CD, Kulstad CE. Comparison of severity of illness scoring systems in the prediction of hospital mortality in severe

sepsis and septic shock. Journal of Emergencies, Trauma and Shock. 2010;3:342-7.
314. Jones AE, Saak K, Kline JA. Performance of the Mortality in emergency department sepsis score for predicting hospital mortality among patients with severe sepsis and septic shock. American Journal of Emergency Medicine. 2008;26:689-92.
315. Lee C-C, Chen S-Y, Tsai C-L, Wu S-C, Chiang W-C, Wang J-L, et al. Prognostic value of mortality in emergency department sepsis score, procalcitonin, and C-reactive protein in patients with sepsis at the emergency department. Shock. 2008;29(3):322-7.
316. Sankoff JD, Goyal M, Gaieski DF, Deitch K, Davis CB, Sabel AL, et al. Validation of the Mortality in Emergency Department Sepsis (MEDS) score in patients with the systemic inflammatory response syndrome (SIRS). Critical Care Medicine. 2008;36:421-6.
317. Vorwerk C, Loryman B, Coats T, Stephenson J, Gray L, Reddy G, et al. Prediction of mortality in adult emergency department patients with sepsis. Emergency Medicine Journal. 2009;26:254-8.
318. Ghanem-Zoubi NO, Vardi M, Laor A, Weber G, Bitterman H. Assessment of disease-severity scoring systems for patients with sepsis in general internal medicine departments. Critical Care. 2011;15:R95.
319. Hermans M, Leffers P, Jansen L, Keulemans Y, Stassen P. The Value of the Mortality in Emergency Department Sepsis (MEDS) score, CRP, and Lactate in Predicting 28-day Mortality of Sepsis in a Dutch Emergency Department. Emergency Medicine Journal. 2012;29:295-300.
320. Kofoed K, Eugen-Olsen J, Petersen J, Larsen K, Andersen O. Predicting mortality in patients with systemic inflammatory response syndrome: an evaluation of two prognostic models, two soluble receptors, and a macrophage migration inhibitory factor. European Journal of Clinical Microbiology and Infectious Disease. 2008;27:375-83.
321. Chen S-C, Huang C-C, Tsai S-J, Yen C-H, Lin D-B, Wang P-H, et al. Severity of disease as main predictor for mortality in patients with pyogenic liver abscess. American Journal of Surgery. 2009;198:164-72.
322. Thiel SW, Rosini JM, Shannon W, Doherty JA, Micek ST, Kollef MH. Early Prediction of Septic Shock in Hospitalized Patients. Journal of Hospital Medicine. 2010;5:19-25.
323. Kulkarni SV, Naik AS, Subramanian N. APACHE-II scoring system in perforative peritonitis. American Journal of Surgery. 2007;194:549-52.
324. Ertan T, Yoldas O, Kýlýc YA, Kýlýc M, Göcmen E, Koc M, et al. External validation of prognostic models among cancer patients undergoing emergency colorectal surgery. American Journal of Surgery. 2008;195:439-41.
325. Mishra A, Sharma D, Raina V. A simplified prognostic scoring system for peptic ulcer perforation in developing countries. Indian Journal of Gastroenterology. 2003;22:49-53.
326. Mäkelä JT, Kiviniemi H, Laitinen S. Prognostic Factors of Perforated Sigmoid Diverticulitis in the Elderly. Digestive Surgery. 2005;22:100-6.
327. Notash AY, Salimi J, Rahimian H, Feshakari MSH, Abbasi A. Evaluation of Mannheim peritonitis index and mutiple organ failure score in patients with peritonitis. Indian Journal of Gastroenterology. 2005;24:197-200.

328. Biondo S, Ramos E, Fraccalvieri D, Kreisler E, Rague JM, Jaurrieta E. Comparative study of left colonic Peritonitis Severity Score and Mannheim Peritonitis Index. British Journal of Surgery. 2006;93:616-22.
329. Poon JT, Chan B, Law WL. Evaluation of P-POSSUM in Surgery for Obstructing Colorectal Cancer and Correlation of the Predicted Mortality With Different Surgical Options. Diseases of the Colon and Rectum. 2005;48:493-8.
330. Del Rosso A, Ungar A, Maggi R, Giada F, Petix N, De Santo T, et al. Clinical predictors of cardiac syncope at initial evaluation in patients referred urgently to a general hospital: the EGSYS score. Heart. 2008;94:1620-6.
331. Hing R, Harris R. Relative utility of serum troponin and the OESIL score in syncope. Emergency Medicine Australasia. 2005;17:31-8.
332. Dipaola F, Costantino G, Perego F, Borella M, Galli A, Cantoni G, et al. San Francisco Syncope Rule, Osservatorio Epidemiologico sulla Sincope nel Lazio risk score, and clinical judgment in the assessment of short-term outcome of syncope. American Journal of Emergency Medicine. 2010;28:432-9.
333. Numeroso F, Mossini G, Spaggiari E, Cervellin G. Syncope in the emergency department of a large northern Italian hospital: incidence, efficacy of a short-stay observation ward and validation of the OESIL risk score. Emergency Medicine Journal. 2010;27:653-8.
334. Quinn J, McDermott D, Stiell I, Kohn M, Wells G. Prospective Validation of the San Francisco Syncope Rule to Predict Patients With Serious Outcomes. Annals of Emergency Medicine. 2006;47:448-54.
335. Cosgriff TM, Kelly A-M, Kerr D. External validation of the San Francisco Syncope Rule in the Australian context. Canadian Journal of Emergency Medicine. 2007;9:157-61.
336. Sun BC, Mangione CM, Merchant G, Weiss T, Shlamovitz GZ, Zargaraff G, et al. External Validation of the San Francisco Syncope Rule. Annals of Emergency Medicine. 2007;49:420-7.
337. Birnbaum A, Esses D, Bijur P, Wollowitz A, Gallagher EJ. Failure to Validate the San Francisco Syncope Rule in an Independent Emergency Department Population. Annals of Emergency Medicine. 2008;52(2):151-9.
338. Bray JE, Coughlan K, Bladin C. Can the ABCD Score be dichotomised to identify high-risk patients with transient ischaemic attack in the emergency department? Emergency Medicine Journal. 2007;24:92-5.
339. Tsivgoulis G, Spengos K, Manta P, Karandreas N, Zambelis T, Zakopoulos N, et al. Validation of the ABCD Score in Identifying Individuals at High Early Risk of Stroke After a Transient Ischemic Attack. Stroke. 2007;37:2892-7.
340. Sciolla R, Melis F. Rapid Identification of High-Risk Transient Ischemic Attacks Prospective Validation of the ABCD Score. Stroke. 2008;39:297-302.
341. Ay H, Arsava EM, Johnston SC, Vangel M, Schwamm LH, Furie KL, et al. Clinical- and Imaging-Based Prediction of Stroke Risk After Transient Ischemic Attack: The CIP Model. Stroke. 2009;40:181-6.
342. Giles MF, Albers GW, Amarenco P, Arsava EM, Asimos AW, Ay H, et al. Early stroke risk and ABCD2 score performance in tissue- vs time-defined TIA. Neurology. 2011;77:1222-8.
343. Stead LG, Suravaram S, Bellolio MF, Enduri S, Rabinstein A, Gilmore RM, et al. An Assessment of the Incremental Value of the ABCD2 Score in

the Emergency Department Evaluation of Transient Ischemic Attack. Annals of Emergency Medicine. 2011;57:46-51.
344. Man SY, Chan KM, Wong FY, Wong KY, Yim CL, Mak PS, et al. Evaluation of the performance of a modified Acute Physiology and Chronic Health Evaluation (APACHE II) scoring system for critically ill patients in emergency departments in Hong Kong. Resuscitation. 2007;74:259-65.
345. Silke B, Kellett J, Rooney T, Bennett K, O'Riordan D. An improved medical admissions risk system using multivariable fractional polynomial logistic regression modelling. Quarterly Journal of Medicine. 2010;103:23-32.
346. Cosentini R, Folli C, Cazzaniga M, Aliberti S, Piffer F, Grazioli L, et al. Usefulness of simplified acute physiology score II in predicting mortality in patients admitted to an emergency medicine ward. Internal and Emergency Medicine. 2009;4:241-7.
347. Paterson R, MacLeod D, Thetford D, Beattie A, Graham C, Lam S, et al. Prediction of in-hospital mortality and length of stay using an early warning scoring system: clinical audit. Clinical Medicine. 2006;6(3):281-4.
348. Emmanuel A, Ismail A, Kellett J. Assessing the need for hospital admission by the cape triage discriminator presentations and the simple clinical score. Emergency Medicine Journal. 2010;27:852-55.
349. Duckitt R, Buxton-Thomas R, Walker J, Cheek E, Bewick V, Venn R, et al. Worthing physiological scoring system: derivation and validation of a physiological early-warning system for medical admissions. An observational, population-based single-centre study. British Journal of Anaesthesia. 2007;98:769-74.
350. Barrett TW, Martin AR, Storrow AB, Jenkins CA, Harrell FE, Russ S, et al. A Clinical Prediction Model to Estimate Risk for 30-Day Adverse Events in Emergency Department Patients With Symptomatic Atrial Fibrillation. Annals of Emergency Medicine. 2011;57:1-12.
351. Eizadi-Mood N, Saghaei M, Jabalameli M. Predicting outcomes in organophosphate poisoning based on APACHE II and modified APACHE II scores. Human and Experimental Toxicology. 2007;26:573-8.
352. Mood NE, Sabzghabaee AM, Khalili-Dehkordi Z. Applicability of different scoring systems in outcome prediction of patients with mixed drug poisoning-induced coma. Indian Journal of Anaesthesia. 2011;55:599-604.
353. Aronin SI, Peduzzi P, Quagliarello VJ. Community-Acquired Bacterial Meningitis: Risk Stratification for Adverse Clinical Outcome and Effect of Antibiotic Timing. Annals of Internal Medicine. 1998;129:862-9.
354. Elbaz G, Etzion O, Delgado J, Porath A, Talmor D, Novack V. Hypothermia in a desert climate: severity score and mortality prediction. American Journal of Emergency Medicine. 2008;26:683-88.
355. Corcoran AT, Smaldone MC, Gibbons EP, Walsh TJ, Davies BJ. Validation of the Fournier's Gangrene Severity Index in a Large Contemporary Series. Journal of Urology. 2008;180:944-8.
356. Davies J, Eddleston M, Buckley N. Predicting outcome in acute organophosphorus poisoning with a poison severity score or the Glasgow coma scale. Quarterly Journal of Medicine. 2008;101:371-9.
357. Smith EE, Shobha N, Dai D, Olson DM, Reeves MJ, Saver JL, et al. Risk Score for In-Hospital Ischemic Stroke Mortality Derived and Validated Within the Get With The Guidelines–Stroke Program. Circulation. 2010;122:1496-504.

358. Fonarow GC, Saver JL, Smith EE, Broderick JP, Kleindorfer DO, Sacco RL, et al. Relationship of National Institutes of Health Stroke Scale to 30-Day Mortality in Medicare Beneficiaries With Acute Ischemic Stroke. Journal of the American Heart Association. 2012;1:42-50.
359. Dutta P, Bhansali A, Masoodi SR, Bhadada S, Sharma N, Rajput R. Predictors of outcome in myxoedema coma: a study from a tertiary care centre. Critical Care. 2008;12(R1).