# Title page

## Working title

Mortality among deteriorating ward patients referred to critical care: a prospective observational cohort study in 48 NHS hospitals

## Authors

* Steve Harris
* Mervyn Singer
* Colin Sanderson
* David Harrison
* Kathy Rowan

# Introduction

There are more than 160 acute hospitals in England that care for more than 11 million overnight hospital admissions per annum. Each patient spends a mean of 5 days on a hospital ward.[Centre, 2012, #17089] These inpatients undergo a process of continual triage, and those who deteriorate are referred to critical care. This interface between the ward and critical care has been a priority area for the English National Health Service (NHS), but available data derive from qualitative work such as the NHS National Patient Safety Agency report[Luettel et al., 2007, #734], relatively small retrospective studies[McQuillan et al., 1998, #605], or voluntary reporting systems[Cullinane et al., 2005, #15159].

International reports indicate that episodes of deterioration are common, and have poor outcome. Inpatient referrals to critical care outreach teams (CCOT), or their equivalent, appear to run between 25–50 per 1,000 hospital admissions.[Bell et al., 2006, #17133; Buist et al., 2007, #17325; Jones et al., 2009, #17139] For such patients, hospital mortality rates are reported as 26% (Australia), and 30-day mortality as 28% (Israel).[Buist et al., 2007, #17325; Simchen et al., 2004, #705] This greatly exceeds the 8·9% inpatient 30-day mortality recently reported in Scottish NHS hospitals.[Clark et al., 2014, #86524]

Critical care provision in the NHS is constrained despite a one-third increase in funding for beds in 2000. In 2010, the United Kingdom (UK) was ranked 24 out of 28 European countries in terms of provision of critical care beds per capita population.[Rhodes et al., 2012, #15692] Similar results are found when the comparison is with North American health care systems.[Wunsch et al., 2008, #761] This ought to imply that access to critical care for the deteriorating ward patient is correspondingly constrained, and that delay to admission to critical care is a problem.

We set out to measure the incidence, the disposition, and the outcome of the deteriorating ward patient referred to critical care, and the factors affecting the decision making around, and the timing of, admission.

# Methods

## Study design and participants

The (SPOT)light study was a prospective observational cohort study of the deteriorating ward patient referred to critical care. The physiological status of the patient at the time of the first bedside assessment by critical care was prospectively recorded along with the recommendation made at the end of the assessment. By linking the records generated at the time of the bedside assessment, to the Intensive Care National Audit & Research Centre’s Case Mix Programme Database (ICNARC CMPD), the fact and timing of admission to critical care was identified. Similarly, by linking to the NHS Information Service then vital status up to 1 year was recorded.

Patients were eligible if they were inpatients on general hospital wards who had been referred to, and assessed by, critical care. The assessment had to be performed at the bedside by a member of the critical care team. This was defined broadly to include members of the critical care outreach service, or members of the critical care medical or nursing staff. Only the first assessment for a given episode of illness was eligible. Cardiac arrests, planned admissions, and visits that were retrievals of patients where a decision to admit had already been made were not eligible.

Patient demographics, and the date, time and location of the visit were recorded, along with the level of care at the time of the visit.[2000, #1009] Available patient physiology (vital signs, arterial blood gas and laboratory measurements) at the time of, or immediately preceding, the visit was abstracted along with organ support, antibiotic therapy, and a subjective assessment of the likelihood of sepsis, and its source. The assessor was then asked to report the level of care recommended, and the outcome of the decision to admit to critical care. Treatment limitation orders were recorded for those not admitted.

## Procedures

The study was registered on the National Institute of Health Research (NIHR) research portfolio, and hospitals were eligible if they participated in the CMP. Research teams at each hospital attended a Dataset Familiarisation Course, and a data collection manual containing definitions of items to be collected was provided. The Clinical Trials Unit at ICNARC provided support for research queries during the study.

Hospitals were asked to report all consecutive ward referrals to the critical care team. Where possible, data collection was to be contemporaneous, and hospitals were requested to identify and submit missed referrals. We used the proportion of emergency ward admissions in the ICNARC CMP that were successfully linked to the (SPOT)light database to quality control the study. Data quality was judged on a monthly basis, and all data from individual months falling below a minimum standard of 80% data linkage were excluded. Reporting was via a secure online web portal which performed real-time field and record level validation. Further on-line validation reports were completed by all hospitals before the database was locked in September 2012. Fact and date of death were then requested from the NHS Information Service. CCOT provision was reported by participating hospitals, and contemporaneous CMP data and Hospital Episode Statistics (HES) were used to define critical care provision, occupancy, and hospital characteristics.

## Statistical analysis

The primary outcome was 90-day mortality. Sample size was calculated to evaluate mortality increases from delay to admission using estimates from 2007 ICNARC CMP data. The target sample size was 12,075–20,125 patients referred to critical care which allowed for delays to occur in 10–40% of admissions and mortality effect sizes of 5–10%.

Physiology measurements at the time of the ward assessment were abstracted. From these, the ICNARC physiology score, the NHS National Early Warning Score and the Sequential Organ Failure Assessment (SOFA) score were calculated with missing values given zero weights as recommended.[Physicians, 2012, #9726; Harrison et al., 2007, #1640; Vincent et al., 1996, #719] The NEWS score additionally defines three risk classes (Low, Medium, and High) designed to trigger an escalating clinical response.

The patient’s existing dependency at assessment was defined using the UK Critical Care Minimum Dataset (CCMDS) levels of care: Levels 0 and 1 are most commonly provided on normal wards while Levels 2 and 3 are within high dependency (HDU) and intensive care units (ICU) respectively.{2013f} The level of care recommended was recorded using the same scale, along with the decision to admit. Prompt admission was defined as one occuring within four hours of ward assessment, in line with recently published UK guidelines.{2013f}

The indicator of critical care unit occupancy was the difference between the maximum number of beds reported to ICNARC, and the number of actively treated patients occupying those beds at the time the ward patient was assessed. Bed pressure was defined as being high (zero or fewer beds), medium (one or two beds), or low (three or more beds).

Incidence models were stratified by NEWS risk class. The unit of analysis was a study day so that daily fluctuations in lagged critical care occupancy could be examined. Estimation was via generalised estimating equations (GEE) with hospitals as clusters, and day-by-day correlations modelled using a first order autoregressive structure. Decision to admit was modelled multi-level logisitc regression with patients nested within hospitals. Cox proportional hazards were used to model survival with a shared frailty. The proportional hazards assumption was checked by inspecting plots of smoothed exponentiated standardised Schoënfeld residuals, and re-entering terms using time-varying co-efficients where necessary. We reported random effects using the Median Odds Ratio (MOR), and the Median Hazard Ratio (MHR) using the bootstrap to generate 95% confidence intervals. These statistics represent the median difference when comparing outcomes from any two randomly selected hospitals.[Bengtsson and Dribe, 2010, #60939]

Categorical data were reported as counts and percentages, and continuous data as mean (SD) or median (IQR) values. Effect measures are reported with their 95% confidence intervals.

## Role of the funding source

The study was centrally funded by the Wellcome Trust, sponsored by ICNARC, and supported at NHS hospitals through the National Institute of Health Research service support costs. The funders of the study had no role in the study design; gathering, analysis, and interpretation of the data; writing of the report; and decision to submit the report for publication. The corresponding author had full access to all the data (including statistical reports and tables); takes responsibility for the integrity of the data and the accuracy of the data analysis; and takes final responsibility for the decision to submit for publication.

# Results

48 hospitals reported 20893 visits over 435 months. 2694 visits (12.9%) did not meet the inclusion criteria including 1860 (8.9%) repeat visits, and 586 (2.8%) patients recently discharged from critical care. Data linkage was incomplete (< 80%) for 66 (15%) study-months excluding a further 2440 (11.7%) visits. Therefore 15759 patients were recruited to the study, of which 15158 (96.1%) completed follow-up without error and were available for analysis (Figure 1a). Final data linkage (ward visits to critical care admissions) was 93% complete.

## Participating hospitals

The participating hospitals comprised 10 teaching and 38 general hospitals that each collected data between September 2010 and December 2011 for a median of 8 months (IQR 5–9 months). Each hospital contributed a median of 252 patients (IQR 162–380) equivalent to 6 patients referred to critical care (IQR 3–9) per 1,000 overnight admissions. Critical Care Outreach Teams (CCOT) operated 24 hours/day and 7 days/week in 14 (29%) hospitals, less than 24 hours/day in 19 (40%) hospitals, and less than 7 days/week in 13 (27%) hospitals. Two hospitals had no CCOT.

There was a median of 12 (IQR 9–18) critical care beds per hospital (mixed Level 2 [typically intensive monitoring or single organ support], and Level 3 [ventilated or multiple organ support]), most often in a single physical location (45 hospitals).

## Incidence of critical care referrals

We estimated the mean baseline (non-teaching hospital with 50,000 admissions per year and 24/7 CCOT provision) incidence of referrals to critical care at 46 (95%CI 50–54) patients per month of which 17 (95%CI 17–18) patients met the NEWS high risk criteria at assessment. As critical care outreach provision decreased then the number of patients assessed fell (supplementary Table 1). Winter was busier (IRR 1.22, 95%CI 1.14–1.31), and weekends quieter (IRR 0.87, 95%CI 0.82-0.92) than non-winter weekday periods. When we included a measure of case finding in the models (cases assessed per 1000 overnight hospital admissions), referral incidence increased initially but may have begun to plateau for the hospitals with referral rates in the highest quartile (supplementary Figure 1).

## Patient characteristics

Table 1 shows the baseline data for all ward patients assessed. Sepsis was reported in (9296, 61%) patients. Of these, the respiratory system was considered to be the source in half (4772, 51%). Organ failure, defined as a SOFA score greater than or equal to two, was present in 5164 (34%) of patients. 1427 patients (9%) were in respiratory failure, 2931 (19%) were in renal failure, and 4636 (31%) were shocked. Organ support at the time of assessment was, however, uncommon (870 patients [6%]).

Overall, 2708 (18%) patients died during the 7-days following ward assessment. Mortality was heavily front-loaded with 1539 (57%) of these deaths occuring in the first 48 hours (supplementary Figure 2). There was a clear correlation between physiological severity and early (7-day) mortality using either ward based (NEWS) or critical care scoring systems (SOFA, ICNARC) (supplementary Figure 3). For example, the 7-day mortality was 9% (328 deaths), 15% (15% deaths), and 26% (1734 deaths)for NEWS low, medium and high risk classes respectively.

## Patient pathways following ward assessment

We extracted two clinical decisions at the point of the bedside assessment. We asked whether the bedside assessor would recommend the patient for critical care (5321 patients [35%]) or ward care (9837 patients [65%]) (Figure 1b). Additionally, we recorded the decision to admit (Figure 1c), and classified patients as either initially refused (9471 patients [62%]) without treatment limits, initially refused with treatment limits (2141 patients [14%]), or immediately accepted to critical care (3546 patients [23%]). Amongst those recommended for critical care, the proportions initially refused without treatment limits, initially refused with limits, or immediately accepted were 30% (1601 patients), 6% (345 patients), and 63% (3375 patients) respectively.

We then tracked patient pathways over the following week.

### Patients with treatment limits refused critical care

These 2141 patients had a 7-day mortality of 41% (881 deaths). The initial decision to refuse admission was reversed in just 76 patients (4%) of whom 26 (34%) died within the week.

Compared to those refused without a treatment limitation order, this cohort was older (77 vs 66 years, 95%CI for difference 11–12 years), and more acutely unwell (3.6 vs 2.8 SOFA points, 95%CI for difference 0.7–0.9). The final 90-day mortality was substantial 65% (1402 deaths), nonetheless 506 patients (24%) survived a year despite the decision.

We did not consider these patients further.

### Patients without treatment limits refused critical care

Patients without treatment limits initially refused critical care had a 7-day mortality of 12% (1102 deaths). Most deaths (799 deaths 73%) occured on the ward without critical care but 303 deaths (27%) followed delayed admission. Within the week, the initial decision to refuse was reversed in 1745 patients (18%), so a total of 2544 (27%) patients died or were admitted to critical care.

7-day mortality was higher for those recommended but refused (17.6% versus 10.4% [risk difference 7.1% 95%CI 5.1–9.1%]), and for those where the initial decision to refuse was reversed (17.4% versus 10.3% [risk difference 7.0% 95%CI 5.1–9.0%]).

### Patients immediately accepted to critical care

The 3546 (23%) patients accepted to critical care at the initial bedside assessment had a 7-day mortality of 20% (725 deaths). Just 42 (6%) of those deaths occurred before admission was arranged, but a further 254 patients (9%) were never admitted but survived regardless.

Those offered admission were marginally younger (64.1 versus 65.6 years, 95%CI 0.8–2.2 years), and more acutely unwell (4.1 versus 2.8 SOFA points, 95%CI 1.2–1.3).

The median delay between assessment and admission for these patients was 2 hours (IQR 1–4) compared to 12 hours (IQR 5–29) for those refused without treatment limits (median additional delay 9 hours, IQR 9–10) (Figure 2). This meant that 2277 (74%) admissions were achieved within 4 hours for those initially accepted versus 256 (16%) for those initially refused (risk difference 58%, 95%CI 56–60%).

## Critical care occupancy

There were 1198 (8%) patients assessed when the unit was full, 3757 (25%) assessed when there were either one or two beds available, and 10197 (67%) assessed when there were more than two beds available (Table 2). Critical care occupancy fluctuated with time of the day, day of the week, and season of the year (supplementary Figure 4).

As occupancy at the time of assessment increased, patients were less likely to be immediately, or ever, accepted to critical care; less likely to be admitted promptly; and, more likely to die on the ward without critical admission (Table 2, Cochran-Armitage test for trend p<0.0001 for all measures).

## Determinants of a decision to admit

We built a multi-level (patients nested within hospitals) logistic regression model to examine factors associated with a decision to admit for patients without treatment limits (Table 3a). As with the univariate comparisons above, older patients were less likely to be admitted (OR 0.60 for patients over 80 years, 95%CI 0.53–0.69), and more acutely unwell patients were more likely to be admitted (OR 1.07 per ICNARC physiology point (e.g. an increase in heart rate from 105 to 115), 95%CI 1.06–1.07). Similarly, patients already receiving organ support (OR 1.83 95%CI 1.55–2.16), or clinically judged to be peri-arrest (OR 6.32 95%CI 5.18–7.70) were also more likely to be admitted.

Patients referred out-of-hours (7pm-7am), during the weekend, or during the winter were more likely to be offered critical care (95%CI for odds ratios from 1.04 to 1.33), but those assessed when the critical care unit was at, or near capacity were less likely to be admitted (OR 0.70, 95%CI 0.57–0.86, and OR 0.87 95%CI 0.77–0.98 respectively). We estimated that in this sample had there been no limitations on capacity then an additional 122 patients (95%CI 53-186) would have been admitted.

The model also demonstrated significant hospital level variation with a Median Odds Ratio (MOR) of 2.11 (95%CI 1.81-2.42) which differed little to that estimated excluding patient predictors (MOR 2.18 [95%CI 1.82–2.60]). The MOR summarises the differences when comparing decision making for patients in any two randomly selected hospitals, and consistency when excluding patient level predictors suggests that it is a true hospital level difference.

## Determinants of prompt admission

The modelling was repeated but now with the delivery of admission to critical care within 4 hours (a prompt admission) as the outcome, and the decision to admit as an additional predictor (Table 3b). We excluded an additional 358 (2.4%) patients where surgery between assessment and admission inevitably delayed that admission.

The associations between prompt admission and patient level predictors were broadly similar with younger and sicker patients being admitted more promptly (Table TTT). However, patients assessed during the winter, while being more likely to be offered critical care, were now less likely to be admitted promptly (OR 0.76, 95%CI 0.64–0.90). The strongest predictor of prompt admission was a decision to admit at the initial bedside assessment (OR 69 95%CI 59–81). Even though the decision to admit was included in the model, occupancy still had a marked effect (OR 0.27 [95%CI 0.19–0.37] for the 1170 (8%) patients assessed when critical care was full). Hospital level variation persisted: MOR 1.89 (95%CI 1.63–2.21).

## Determinants of 90-day mortality

Amongst patients without treatment limits, there were 372 deaths by the end of the first day, 1742 by the end of the first week, 3130 by the end of the first month, and 3946 by 90 days. The risk of death at these time points was 3%, 13%, 24%, and 30% respectively (Kaplan-Meier failure function).

A series of models were fitted with 90-day survival as the dependent variable for patients without treatment limits. The final best model (Table 4) incorporated a time-varying effect for acute physiological severity such that its effect was attenuated after the first week (supplementary Figure 5).

Other patient level risk factors were consistent with the existing literature on outcomes in similar patients: older patients, and those with sepsis (other than genito-urinary) had worse outcomes.[Harrison, 2007] Patients assessed during the winter months, over the weekend, and out-of-hours did not have a significantly worse adjusted survival than baseline.

Critical care occupancy did not affect adjusted mortality in the multi-level model (e.g. HR 1.03, 95%CI 0.90–1.17 for patients assessed when the critical care was full). A single level model, constructed in case occupancy was mediated rather than confounded by the effect of hospital, similarly could not exclude a null effect (HR 1.07, 95%CI 1.00–1.15).

The full multi-level model demonstrated significant hospital level variation in survival (MHR 1.28, 95%CI 1.22–1.34) which was little altered by adjustment for patient level risk factors (MHR 1.29, 95%CI 1.22–1.35).

# Discussion

This is the largest prospective study of deteriorating ward patients to date.{>>see research in context<<} We describe the events in the week following bedside assessment by critical care of more than 15000 ward patients in 48 acute NHS hospitals.

Nearly half (45%) of the patients were defined as being at high risk as per the latest guidelines[@Anonymous:2012un], and one-third (33%) were assessed when the critical care unit was under-strain (two or fewer beds available). One in twelve (8%) were assessed when the unit was completely full.

The decision making that occured at the bedside assessment was affected by occupancy, but also varied significantly between hospitals after adjustment for measured severity of illness. There was evidence of rationing of critical care in that we observed a bias against admitting the elderly, and saw that decisions to admit were less likely when capacity was limited.

The consequences of critical care strain, and inter-hospital variation in decision making were two-fold. Firstly, affected patients were less likely to be cared for during their acute illness in critical care, and more likely to die on the ward without a trial of critical care. This was true even for the cohort of patients for whom the bedside assessor thought critical care was justified. Secondly, patients who did receive critical care had their admission delayed - both directly (prompt admission was less likely), and indirectly (by altering the admission pathway to a ‘watch-and-wait’ strategy with a late rather than a direct admission).

Overall, three-quarters of patients accepted to critical care were admitted within the 4-hour recommendation. This fell to around one-half for the parent cohort of those recommended. The single most important determinant of a prompt admission was the bedside decision, but critical care occupancy was also important both via its effect on decision making, and in delaying the delivery of that decision.

Among patients without treatment limits, we observed an overall 90-day mortality of 30%. This is comparable to the mortality observed in unselected critical care admissions[@Harrison:2014ei; @Harrison:2007jt] — even though only a third of our cohort received critical care. It is also double that of recent inpatient cohorts that did not exclude patients with treatment limitations.[@Clark:2014hu]

The mortality burden was heavily front-loaded with around half of deaths in the first week, and half of these in the first 48 hours. This strongly suggests that the opportunity for intervention in these patients is very limited. Importantly, the 7-day mortality burden was elevated even for patients defined as low risk by NEWS class, and for actively treated patients recommended for ward rather than critical care. Otherwise said, while the bedside assessment could effectively identify patients at higher risk, it did not define a group where the immediate risk of death, or critical care admission, was reassuringly low.

Again, hospital level variation, already seen to be part of the decision making process, was also a significant determinant of 90 day survival.

* themes to pick-up on
  + inter-hospital variation
  + occupancy

Aspects of the study stand independent of these limitations. Regardless, of the effect of prompt admission to critical care,

We have identified a cohort of hospital patients at very high risk. This risk is heavily front-loaded, and the window for intervention is short. The bedside assessment is an effective but imperfect triage tool, as the mortality in those initially refused admission is high. Given that we already excluded patients with treatment limitations, it is of concern that nearly half of these early deaths occur without a trial of critical care.

A substantial proportion of patients recommended for critical care are not offered a bed, and this proportion increases when capacity is limited. Expanding critical care bed numbers would first and foremost benefit this group. This is an opportunity to create a virtuous circle. Earlier admission may lead to shorter stays thereby improving flow through critical care as well as outcomes. Identifying those patients who should be admitted promptly is already the top priority for both clinicians and patients.{Reay2014} What we have contributed we hope, is firm evidence in support of this.