

Failure to Rescue

Title Page

Is “failure to rescue” derived from administrative data in England a nurse sensitive patient safety indicator for surgical care? Observational study

Peter Griffiths

Simon Jones²

Alex Bottle³

1 Faculty of Health Sciences, University of Southampton peter.griffiths@soton.ac.uk

Room E4015, Building 67
Highfield Campus
SO17 1BJ
+44(0)2380597877

2. Dr Foster Intelligence, London

3. Dr Foster Unit, Imperial College, London

This is a pre-print version of a paper accepted for publication in the International Journal of Nursing Studies.

Failure to Rescue

Is “failure to rescue” derived from administrative data in England a nurse sensitive patient safety indicator for surgical care? Observational study

Abstract

Background: ‘*Failure to rescue*’ - death after a treatable complication - is used as a nursing sensitive quality indicator in the USA. It is associated with the size of the nursing workforce relative to patient load, for example patient to nurse ratio, although assessments of nurse sensitivity have not previously considered other staff groups. This study aims to assess the potential to derive failure to rescue and a proxy measure, based on long length of stay, from English hospital administrative data. By exploring change in coding practice over time and measuring associations between failure to rescue and factors including staffing, we assess whether these are useful nurse sensitive indicators.

Design: Cross sectional observational study of routinely collected administrative data.

Participants: Discharge data from 66,100,672 surgical admissions to 146 general acute hospital trusts in England (1997 to 2009).

Results: Median percentage of surgical admissions with at least one secondary diagnosis recorded increased from 26% in 1997/8 to 40% in 2008/9. Regression analyses showed that mortality based failure to rescue was significantly associated ($p < 0.05$) with several hospital characteristics previously associated with quality, including staffing levels. Lower rates of failure to rescue were associated with a greater number of nurses per bed and doctors per bed in a bivariate analysis. Higher total clinically qualified staffing (doctors + nurses) per bed and a higher number of doctors relative to the number of nurses were both associated with lower mortality based failure to rescue in the fully adjusted analysis ($p < 0.05$); however, the extended stay based measure showed the opposite relationship.

Conclusion: Failure to rescue can be derived from English administrative data and may be a valid quality indicator. This is the first study to assess the association between failure to rescue and medical staffing. The suggestion that it is particularly sensitive to nursing is not clearly supported, nor is the suggestion that the number of patients with an extended hospital stay is a good proxy.

Failure to Rescue

[Hospital Mortality, Quality Indicators, Health Care*](#), [Surgical Procedures, Operative/adverse effects, Nursing Service, Hospital, Treatment Failure, nursing care quality, hospital staffing, nursing workforce.](#)

346 Words (abstract) 3370 (main text)

What is known

- The rate of failure to rescue (death of a hospital patient after a treatable complication), derived from routine administrative data, is recognized and used as a nurse sensitive patient safety indicator by the United States
- Sensitivity to nursing is supported by its association with nurse staffing levels, but medical staffing has not previously been studied
- The validity of the indicator outside the US has been questioned

What this paper adds

- Failure to rescue can be derived from English hospital data
- Failure to rescue is associated with the number of doctors and nurses per bed
- The higher proportion of doctors in the workforce, the lower the rate of failure to rescue

Introduction

'Failure to rescue' refers to the death of a hospital patient after a treatable complication (Silber et al., 1992). The rate of failure to rescue, derived from routine administrative data, is recognized and used as patient safety indicator by the United States (US) Agency for Healthcare Research and Quality (PSI 4, now renamed "Death among Surgical Inpatients with Serious Treatable Complications" (Agency for Healthcare Research and Quality, 2007). It holds the promise of being more sensitive to the quality of care in a hospital than either conventional mortality or complication rates (Silber et al., 2007). Failure to rescue has been identified as being particularly sensitive to the quality of nursing (Clarke and Aiken, 2003) and endorsed as a nurse sensitive quality measure (National Quality Forum, 2004) but it has not been widely used or reported outside North America. In this paper we assess the feasibility of deriving failure to rescue indicators for surgical patients from English hospital administrative data, which have previously been assessed as unsuitable for the purpose, primarily because secondary diagnoses are not sufficiently well recorded (McKee et al., 1999). We also assess the relationship between failure to rescue and a number of markers of hospital quality including staging by both nurses and physicians.

Mortality rates are widely used to indicate the quality of care in hospitals, but variation in mortality is largely due to factors unrelated to hospital care (Mant, 2001). Rates must be adjusted to reflect differences in the underlying risk of the population that is treated if valid comparisons are to be made between hospitals (Iezzoni, 1997). However, different risk adjustment models

Failure To Rescue

give different estimates of individual risk of death and identify different hospitals as performing outside normal limits (Iezzoni, 1997). Failure to rescue is proposed as an alternative, or complementary, indicator. It is hypothesized that the ability of a hospital to successfully treat (rescue) a patient who suffers a complication is strongly related to the quality of care provided, whereas the occurrence of the complication is more closely related to the patient's underlying risk (Silber et al., 1995). Because failure to rescue indicators consider only patients who have developed a serious but treatable complication, they offer a partial solution to the problems of risk adjustment, because the population is more homogenous and the underlying risk of death is less variable, since all patients included in the denominator are severely ill (McDonald et al., 2007, Silber et al., 2007, Silber et al., 1995). There is empirical evidence that failure to rescue rates are more closely associated with hospital characteristics including nurse staffing levels and less influenced by patient characteristics than either complications or mortality (Silber et al., 2007, Silber et al., 1995, Silber et al., 1995).

The potential significance of this measure is reflected in recent reports and research into responses to deteriorating patients in acute care that emphasise the numerous potential points of failure prior to initiating appropriate intervention including:

- not *taking* observations
- not *recording* observations
- not *recognising* early signs of deterioration
- not *communicating* observations (Clarke, 2004, Luettel et al., 2007)

Failure To Rescue

Because of the role of nurses in early identification of deterioration, failure to rescue has been widely advocated as a nursing sensitive outcome indicator in hospitals (Clarke and Aiken, 2003, Griffiths et al., 2008, Naylor, 2007), because observation is compromised when staffing is not adequate. An association between low levels of nurse staffing and high levels of failure to rescue is supported by meta-analysis of observational studies. The increased odds of failure to rescue is estimated as 16% per additional patient per nurse (Kane et al., 2007). Other characteristics which have been associated with hospital quality, such as nursing skill mix, hospital size and teaching status (Aiken et al., 2002, Hartz et al., 1989, Jarman et al., 1999) have also been associated with lower rates of failure to rescue (Aiken et al., 2002, Silber et al., 2007) but numbers of doctors, an important part of the hospital staff in many countries outside the USA, have not been widely studied, even though doctors too have a role in surveillance and the medical response to deterioration is also likely to be an important determinant of outcome and higher medical staffing levels have also been associated with mortality (Bond et al., 1999, Jarman et al., 1999) .

Because the indicators need to identify a group of patients who experience particular complications, the validity of such indicators can be compromised if coding of secondary diagnoses in the administrative data set is poor. In the absence of codes to indicate diagnoses that are present on admission, the indicators also rely on complex exclusion rules in order to eliminate pre-existing comorbidity. Because of the difficulty doing this for medical cases

Failure To Rescue

(Horwitz et al., 2007, Moriarty et al., 2010), use of the indicators has generally been recommended for surgical cases only.

The under recording of secondary diagnoses in administrative databases is a known issue. Previously, McKee et al reported that English hospital data from 1996/7 and 1997/8 were unsuitable for deriving failure to rescue measures, primarily because of low rates of coding (McKee et al., 1999). Doubts about both the accuracy and completeness of coding have continued to be raised in the UK and other countries (Casez et al., 2010, Leibson et al., 2008, Williams and Mann, 2002). Although a recent systematic review suggested improvements and overall acceptable accuracy for coding in the UK, studies revealed substantial variation between hospitals and have mainly focused on the primary diagnosis / procedure (Burns et al., 2011).

While classic studies of failure to rescue have looked at mortality in a sub group of patients presumed to have treatable suffered complications (e.g.Aiken et al., 2002, Silber et al., 2007) an alternative has been proposed. The alternative approach, is predicated on the recognition that death is not the only possible result of a “failure” to rescue. If failure to rescue results in serious deterioration that in turn leads to extended hospital stay, then stays that fall well outside the norm can be used as a proxy indicator of failure to rescue. This has previously been used on UK data (Rafferty et al., 2007).

Thus this study is an exploratory study that aims to assess the potential for deriving failure to rescue indicators and a proxy measure, based on exceptionally long length of stay, from English hospital administrative data by exploring change in coding practice over time and measuring associations between failure to rescue and factors that suggest how the indicator will perform as a quality measure. These factors include the association between failure to rescue and depth of coding (number of complications recorded) and staffing by doctors, nurses and support workers.

METHODS

Our assessment is based on the approaches undertaken by McKee et al and Silber et al (2007).

Specifically we consider:

- Whether coding of secondary diagnoses has increased since the previous assessment (McKee et al., 1999) – indicating improved potential for deriving mortality based failure to rescue indicators
- Whether failure to rescue rates are associated with coding practices (rates of secondary diagnostic coding and rates of complications coded) in order to determine potential for bias due to variation in depth of coding between hospitals
- Whether failure to rescue rates correlate from year to year, indicating that it is a *relatively* stable property, as would be expected if rates are reliable indicators of a property of the hospital
- Whether failure to rescue is associated with the Hospital Standardised Mortality Rate (HSMR)(Jarman et al., 1999) – to assess whether there are plausible associations with other recognised mortality indicators and whether failure to rescue indicators give additional information
- Whether FTR rates are (relatively) invariant with age (McKee et al., 1999) in order to identify the need for additional risk adjustment.

Failure To Rescue

- Whether rates are associated with hospital factors that are generally supported as linked to quality of care (e.g. nurse staffing levels, teaching status of hospital) – to assess the validity of the claim that the indicator reflects some dimension of underlying ‘quality’ and the potential sensitivity to nursing.

We use the AHRQ definition, now renamed “Death among Surgical Inpatients with Serious Treatable Complications”. We refer to this as FTR-A. This indicator counts deaths among a subset of surgical patients experiencing certain complications likely to occur after admission, including renal failure, venous thromboembolism, health care acquired infection, haemorrhage and gastrointestinal ulceration (Agency for Healthcare Research and Quality, 2007). Exclusion rules aim to reduce the chances of including diagnoses that were present on admission. The AHRQ specification defines eligible surgical admissions and uses ICD-9 diagnostic codes to identify complications. As there is large variation in day surgery rates across the country, which are not necessarily related to case mix, we included day cases in our data to ensure comparable populations. The National Health Service (NHS) uses ICD10 codes and its own specific data dictionary. We mapped the AHRQ specification onto the NHS data dictionary (<http://www.datadictionary.nhs.uk/>) and the ICD-9 codes for complications to ICD-10 codes by inspection of the ICD-10 codebook (V2 including updates to end of 2009). We supplemented our ICD-9 to ICD-10 mapping by an inspection of codes generated by an automated translation from ICD-9 to ICD-10. We did not rely on the automated translation exclusively because of known issues with such mapping (Schulz et al., 1998). See supplemental material for detail of the mappings and final specification.

Failure To Rescue

We defined a 'long stay' based indicator (FTR-L) based on all patients with abnormally long hospital stays, an approach previously taken in England (Rafferty et al., 2007). Here, failure to rescue cases were any patients with in-patient hospital stays longer than the 75th percentile for that patient's Healthcare Resource Group (HRG). HRGs are standard groupings of clinically similar treatments or sets of patients that use common levels of healthcare resource, similar to Diagnosis Related Groups (DRG). HRGs were assigned using the NHS HRG classifications (v3.5).

We used bivariate correlations and multiple Poisson regression models to determine associations between FTR rates and quality related organisational factors (validation variables). Because failure to rescue has been identified as an indicator of the adequacy of the nursing workforce based on both logical argument and empirical data (Clarke and Aiken, 2003, Kane et al., 2007) we hypothesised that a valid failure to rescue indicator would be related to nurse staffing levels. However, hospital safety is not influenced solely by nurse staffing, and studies which have been used to establish failure to rescue as nurse sensitive have not considered the impact of staffing by other professional groups. An analysis of associations between organisational factors and mortality in UK hospitals found that mortality was related to the number of doctors but not the number of registered nurses (Jarman et al., 1999), which opens up the possibility that a valid measure of failure to rescue might be more strongly influenced by medical staffing (or other clinical staff) than nursing. Therefore we aimed to assess the relationship between FTR rates and all clinical staffing. For staffing we assessed relationships between FTR rates and the numbers of

Failure To Rescue

registered nurses doctors (physicians) and support staff per bed. In the UK, NHS hospital doctors of all grades are employees of the hospital, in contrast to some other countries such as the USA where senior doctors are generally independent practitioners. The number of doctors in the workforce therefore includes both training grade qualified medical practitioners (junior doctors / residents) and fully qualified specialists (consultants / attending physicians).

Other factors were selected for assessment because they had been identified or suggested by previous studies as being associated with safety related quality of care (Aiken et al., 2002, Bond et al., 1999, Griffiths et al., 2009, Hartz et al., 1989, Hayes et al., 2006, Jarman et al., 1999, McKee et al., 1999, Silber et al., 2007) and relevant data could be obtained from readily available administrative data sources. We assessed relationships between failure to rescue teaching status of the hospital (teaching status is associated with lower rates of failure to rescue), nursing staff stability (a measure of staff turnover, associated with a less stable workforce and potentially poorer safety) and bed occupancy (high bed occupancy has been associated with lower levels of safety).

We identified a number of factors outside the immediate control of the hospital that have been identified as significant and used in previous assessments of the link between organisational characteristics and mortality in England (Jarman et al., 1999). These include the number of community based general medical practitioners, proportion of deaths occurring in hospital, number of hospital discharges and location of the hospital in London. These factors are not necessarily related to the quality of hospital care but need to be controlled for in order to ensure

Failure To Rescue

that the independent effects of quality related factors can be properly estimated. For example, the proportion of deaths outside hospital is likely to be influenced by the provision of community palliative care services. London is generally included as a control factor in models examining health services in England because it is not typical of the rest of the country in terms of the population and the characteristics of hospitals and other health services. We also controlled for percentage of patients with a failure to rescue complication, since this number will be influenced by the extent to which these specific secondary diagnoses are coded in a trust and therefore provides a control for any bias related to different coding practices [see supplemental material for full list of variables]. We refer to these as control variables.

The dependant variable in the model was the count of the number of FTR events (numerator) with the number of people at risk as the denominator (a poisson rate). The number of people at risk was included as an offset in the model. The independent variables were standardised as z scores prior to being used in the model so that model coefficients give an accurate and comparable indication of the relative influence of each variable on the dependent variable (FTR), although this does render the coefficient themselves harder to interpret. We initially entered all the independent validation variables in a regression model so that we could assess the contribution of each variable, given the value of the others. This was important in order to ensure that we could gain a clear picture of which validation variables FTR was related to. For example nurse staffing in teaching hospitals tend to be related. Each is hypothesised to have an independent association with failure to rescue but this can only be properly assessed by

considering both variables simultaneously. Backwards stepwise regression, based on minimising the Bayesian Information Criterion, was used to remove variables that contributed little to the model (Cameron and Trivedi, 1998).

Regression model diagnostics suggested multicollinearity because the staffing level variables were highly correlated with each other. For example, hospitals with more nurses per bed also tended to have more doctors per bed. While multicollinearity does not alter overall model estimates, it does lead to unstable estimation of individual regression coefficients for the correlated variables, meaning that it is not possible to identify their independent effects accurately (Farrar and Glauber, 1967). Therefore, in order to assess the impact of hospital staffing and the relative influence of the two staff groups previously associated with variations in overall mortality or failure to rescue our regression models considered total numbers of professionally qualified clinical staff (primarily doctors and nurses) per bed and the relative numbers (ratios) of each staff group to nurses in order to give an indication of the relationship with both staffing levels and skill mix. While not ideal, this approach is warranted based on the potential common mechanism of action in reducing failure to rescue from both staff groups (surveillance), the degree of overlap in their work, potential substitution of nurses for doctors (Goryakin et al., 2011) and previous research identifying skill mix (albeit in the nursing team) as a factor in hospital mortality. All analysis was undertaken using R 2.10.1 software.

Data sources

Failure To Rescue

To calculate failure to rescue rates, we used hospital discharge data from the National Health Service (NHS) Commissioning Data Sets (CDS) data from April 1997 to March 2009 to identify all admissions for surgical procedures to general acute National Health Service (NHS) hospitals in England (146 hospital trusts in 2008/9 – a trust may comprise several hospital sites). The CDS provides a record of admission method diagnoses, procedures, discharge dispositions and patient demographic details for all episodes of NHS care. We obtained the HSMR for these hospital trusts from health care information provider Dr Foster Intelligence who calculate rates for a group of diagnoses which account for 80% of all hospital deaths using a regression methodology to adjust for case mix (Bottle and Aylin, 2008). We obtained HRG length of stay data from the same source. The Dr Foster Intelligence calculations for average length of stay do not include day cases, hence less than 25% of our patients experienced stays above the 25th percentile. Workforce data and other hospital and locality characteristics for the regression analysis were obtained from Dr Foster Intelligence and the NHS Information Centre (see supplemental material for full list of data and sources).

Results

Between 1997/8 and 2008/9, there were 66,100,672 eligible surgical admissions (day case and inpatient) of whom 442,462 (0.7%) died and 4,993,863 (7.6%) experienced a long hospital stay, above the 25th percentile for their HRG. The median percentage of surgical admissions with at least one secondary diagnosis recorded increased from 26% in year 1997/8 to 40% in 2008/9. Overall 2,496,356 patients (3.8%) had an eligible complication for FTR-A of whom 226,237 died (9.1%). Overall failure to rescue rates were thus 9.1% (FTR-A) and 7.6% (FTR-L) (table

Failure To Rescue

2). FTR rates for seem to be stable from year to **year by the end of the period studied**. The inter-year correlations between rates for 2007/8 and 2008/9 for the 146 hospital trusts were high for both FTR indicators (FTR-A $r=0.92$, FTR-L $r=0.94$).

FTR-A shows no correlation with FTR-L ($r<0.1$) and neither are correlated with HSMR. Neither of the FTR indicators are significantly correlated with coding depth (average number of all secondary diagnoses coded per hospital trust) but FTR-A has a small negative correlation with the average number of FTR complications coded ($r= -0.17$) and FTR-L has a medium positive correlation the complication rate ($r=0.31$). See **Table 3**. The rate of FTR-A increased sharply with age. There was a 12-fold increase in FTR-A rates from the youngest (18-39) to oldest (75+) age groups. FTR-L did not show a clear age related trend (**Table 4**).

Mortality based FTR rates (FTR-A) were associated with a number of hospital characteristics that have been previously linked to quality (**see table 5**). There were significant ($p<0.05$) bivariate associations, in the expected direction, between FTR-A and most of the validation variables, including both doctors and nurses per bed. In the multiple regression models, hospitals with more professionally qualified clinical staff per bed were associated with lower rates of failure to rescue. Higher bed occupancy was associated with higher rates of failure to rescue. However, some results ran counter to expectation. For example, a more stable nursing workforce was associated with *higher* rates of failure to rescue. While bivariate associations showed that more nurses were associated with lower rates of FTR-A, a higher nurse to doctor ratio was

Failure To Rescue

associated with higher rates of failure to rescue in the multiple regression models based on total number of professionally qualified clinical staff per bed. FTR-L was also associated with quality related characteristics in some respects in regression models. For example, a more stable nursing workforce was associated with lower failure to rescue and higher bed occupancy was associated with higher levels of failure to rescue. However, again there were some unexpected findings for example hospitals with more professionally qualified clinical staff per bed were associated with *higher* rates of failure to rescue.

DISCUSSION

Our results point to improved coding practice in English hospital data and a relatively stable failure to rescue rate derived from them. We have observed several associations between failure to rescue and presumed markers of quality, including clinical staffing levels which have been previously associated with hospital mortality and failure to rescue. This suggests that the FTR-A indicator we derived from English data may well be a valid measure of quality. However, the claim that failure to rescue is specifically and particularly sensitive to nurse staffing levels is not supported, although we were unable to model the effects of nurse staffing and medical staffing simultaneously.

Failure To Rescue

Rates of secondary coding have increased over time since McKee concluded that it was impossible to derive a valid mortality based failure to rescue indicator from 1996-1997 English data (McKee et al., 1999). This may be as a result of changing hospital funding arrangements whereby some complex cases (ie those with comorbidities) attract a higher rate of reimbursement to hospitals. patients with In contrast to McKee *et al.*, (McKee et al., 1999)we found strong year on year correlations for both our failure to rescue rates in recent years. This suggests a degree of stability consistent with the rate reflecting an underlying characteristic of hospital performance. FTR rates were not significantly correlated with a hospital's coding depth (average number of secondary diagnoses coded), which suggests an absence of systematic bias caused by coding practices. However, rates are correlated with the number of FTR complications that are coded. This may reflect residual bias or differences in case mix, which needs to be considered in any comparative analysis.

We found several associations between the failure to rescue indicators and presumed markers of quality that have been demonstrated in previous work (Aiken et al., 2002, Bond et al., 1999, Griffiths et al., 2009, Hartz et al., 1989, Hayes et al., 2006, Jarman et al., 1999, McKee et al., 1999, Silber et al., 2007), providing some evidence of validity. Our bivariate correlations showed an association between the number of nurses per bed and lower rates of failure to rescue.

Previous US work showed that higher nurse staffing levels were associated with lower levels of mortality based failure to rescue and this has been the basis of the indicator being identified as nurse sensitive (Clarke and Aiken, 2003, National Quality Forum, 2004). However, these US

Failure To Rescue

studies did not consider physician staffing, although other studies have identified associations between higher numbers of doctors and lower risk adjusted mortality in both the US and UK (Bond et al., 1999, Jarman et al., 1999). Because of the staffing model of US hospitals, any relationship between medical staffing and **outcomes in that country is likely to be confounded by a 'teaching hospital'** effect, since it is these hospitals that have relatively large resident medical staff. Our study is the first to look at the relationship between failure to rescue and a range of hospital staff groups, while controlling for this teaching hospital effect. We saw bivariate associations with FTR-A for both doctors per bed and nurses per bed and, in our regression models, higher levels of professionally qualified clinical staff (doctors plus nurses) were associated with lower levels of FTR-A, but a higher nurse to doctor ratio was associated with higher rates of failure to rescue (that is, fewer doctors relative to the number of nurses is associated with higher failure to rescue). This suggests that the numbers of doctors in the workforce might be the more significant factor and that a skill mix rich in doctors, in tandem with a larger overall clinical workforce, is key. Certainly these findings cast doubt on failure to rescue being a specifically nurse sensitive indicator and also points to a significant limitation in generalising findings about the relationship between the size of the nursing workforce and outcomes in the US to settings with different medical staffing models.

However, not all the relationships observed are clearly or plausibly indicative of variations in quality. Although we adjusted for some organisational and system characteristics outside the control of the hospital (for example percentage of cases with comorbidity and level of GP

Failure To Rescue

coverage) it seems likely that there are other confounding variables that were not accounted for in our models. This emphasises the need for further risk adjustment for case mix, including age and comorbidity. FTR-A rates increase sharply with age. A relationship between FTR and age is perhaps unsurprising. Although the logic of the indicator suggests that development of the complication is more strongly related to underlying risk than is its successful treatment, some additional risk associated with age might be expected. The AHRQ risk adjustment model for deaths among surgical patients with treatable complications indicates that risk increases with age (Agency for Healthcare Research and Quality, 2011) although the most closely comparable data available suggest that risk increases approximately two fold, from 7% for the 18-39 age group to 15% for the 65-75 (Agency for Healthcare Research and Quality, 2007), whereas the risk increases six fold across the same age groups in our sample. While the absolute rates should not be directly compared, due to differences in the population included in the indicators, the reason for this difference in age gradient warrants some further consideration. Inclusion of day cases in our sample may be a partial explanation. A number of low risk young patients who may not appear in the US indicator denominator might be included in calculating our rates. It seems likely that this ‘deflation’ of the indicator would affect older people less as they are less likely to have day surgery. **However, given the profile of day surgery rates the numbers are likely to be low** and it is possible that the difference is at least in part a result in genuine differences in outcomes which may be related to care quality.

Failure To Rescue

Although FTR-L performed well in some respects, the absence of any association with the mortality based FTR indicator FTR-A suggests that is not acting as a proxy measure of failure to rescue, as originally conceived (Rafferty et al., 2007). Furthermore the rate of extended stay, measured by FTR-L, is potentially linked to a number of factors beyond the control of the hospital, such as provision of community services. Other approaches to identifying extended stays as ‘proxy’ for failure to rescue have been taken. Rafferty et al used median times 1.25 as the cut-off for extended stay, which meant that more than a third of all patients were included as ‘cases’ in the indicator. (Rafferty et al., 2007). Our indicator included less than 8% of patients. In contrast, US based work has recognised prolonged lengths of stay as a composite of inefficiency and complications. The cut-off point for prolongation was therefore based upon the point after which the probability of staying longer is increased, rather than decreased as would be expected as the patient nears readiness for discharge (Silber et al., 2009). Importantly, this approach counts early deaths as ‘prolonged’, since otherwise these deaths increase the appearance of efficiency and quality as measured by the indicator. Typically 40% or more of patients are classified as experiencing ‘prolonged’ stays according to this approach (Silber et al., 2009). The **potential** significance of the prolonged stay indicator needs to be reconsidered and a composite measure of complications and inefficiency may have value, but we conclude that it is not a valid indicator of failure to rescue. **While we could have formally assessed the properties of these alternative specifications they would be affected by similar extraneous factors and it seems incorrect as a matter of principle to apply the construct of failure to rescue to an indicator that incorporates where more than 30% of all patients have an adverse outcome that is, in fact, a**

potentially short delay to discharge. Furthermore, given the initial premise of FTR indicators – that the occurrence of complications is primarily driven by patient level factors – such an indicator would require proper risk adjustment before being used as a quality measure.

The absence of correlation between FTR-A and HSMR also suggests that they are measuring different things. As both are based on mortality rates this does seem surprising at first. However, HSMR comprises a ‘basket’ of patients with conditions which account for the majority of deaths in hospital. These are predominantly medical cases, whereas FTR-A includes a very different group of patients, many of whom may belong to low mortality diagnostic groups. Thus the two indicators are potentially complementary.

A criticism that has been levelled against mortality based failure to rescue indicators is that they do not properly distinguish comorbidities (which exist before admission) from complications that develop afterwards (Horwitz et al., 2007, Moriarty et al., 2010). This issue is most significant when the administrative database does not include a flag to indicate which secondary diagnoses were present on admission. Our study was limited, as the data do not include a present on admission flag and we were unable to assess the accuracy of identification of complications, which requires case note audit. Previous research has suggested that the problem is less severe for surgical patients than for medical patients (Horwitz et al., 2007, Moriarty et al., 2010) and that hospital rankings (that is, relative assessments of quality) may be relatively unaffected by the inclusion of some pre-existing conditions (Mattke et al., 2004). Certainly, one claimed

Failure To Rescue

advantage of the failure to rescue indicator over mortality rates - that it includes a relatively homogenous group of patients all of whom are at relatively high risk - is not lost by the inclusion of comorbidities. However, the distinctive significance of the measure, as an indicator of the hospital's timely and appropriate response to emergent complications, is lessened. Further validation work involving audit of case notes is warranted in order to determine the proportion of cases flagged by the indicator that are true 'failures to rescue' from a complication that arose on hospital.

Conclusions

We conclude that there is potential to derive mortality based failure to rescue indicators for surgical patients from routine administrative data in England. Such indicators may offer some advantages over standardised mortality measures, such as HSMR, for surgical patients. Our FTR-A indicator, based on the AHRQ definition, is a potentially valid quality indicator which can complement HSMR, but like overall mortality it needs to be properly risk adjusted to facilitate benchmarking and comparison between hospitals. Our findings reiterate the importance of adequate staffing in assuring patient safety; however, failure to rescue does not appear to be a specifically nurse sensitive indicator. A key element of future validation of the indicator is to determine its responsiveness, relative to HSMR and other risk adjusted mortality measures, to the successful implementation of quality improvement initiatives and interventions such as physiological track and trigger tools or early warning systems (Gao et al., 2007) that aim to reduce failure to rescue. Future work examining the relationships between hospital staffing and

Failure To Rescue

patient safety should pay more attention to staffing of hospitals by groups other than nurses, including hospital doctors.

Funding

This work was supported by the Policy Research Programme in the Department of Health in England, which commissioned this as an independent study [016/0058]. The views expressed are not necessarily those of the Department. The study was also supported in kind by the provision of data and time by the Dr Foster Unit of Imperial College and Dr Foster Intelligence. The Dr Foster Unit at Imperial is affiliated with the Imperial Centre for Patient Safety and Service Quality at Imperial College Healthcare NHS Trust, which is funded by the National Institute of Health Research. The Department of Primary Care and Social Medicine is grateful for support from the National Institute for Health Research Biomedical Research Centre Funding Scheme.

Competing Interests

None

Author Contributions

PG conceived and designed the study jointly with SJ and AB. AB and SJ extracted data. AB & SJ mapped the AHRQ indicators to English coding and ICD 10 with advice from PG on clinical codes. SJ undertook statistical analysis and AB, PG and SJ interpreted the results. PG and SJ drafted the paper and AB, PG and SJ commented on drafts and approved the final version. SJ is guarantor for the extraction of data and analysis, PG for other aspects of the paper including the design, interpretation of results and decision to publish.

References

- 1 Agency for Healthcare Research and Quality, 2011. Patient Safety Indicators (PSI) Risk
2 Adjustment Coefficients for the PSI Version 4.3. Agency for Healthcare Research and
3 Quality,.
- 4 Agency for Healthcare Research and Quality, 2007. Patient Safety Indicators (PSI) Version 3.1
5 Comparative Data Agency for Healthcare Research and Quality,.
- 6 Aiken, L.H., Clarke, S.P., Sloane, D.M., Sochalski, J., Silber, J.H., 2002. Hospital Nurse Staffing
7 and Patient Mortality, Nurse Burnout, and Job Dissatisfaction. *Journal of the American*
8 *Medical Association* 288 (16), 1987-1993.
- 9 Bond, C., Raehl, C.L., Pitterle, M.E., Franke, T., 1999. Health care professional staffing, hospital
10 characteristics, and hospital mortality rates. *Pharmacotherapy* 19 (2), 130-138.
- 11 Bottle, A., Aylin, P., 2008. Intelligent information: a national system for monitoring clinical
12 performance. *Health Services Research* 43 (1 Pt 1), 10-31.
- 13 Burns, E.M., Rigby, E., Mamidanna, R., Bottle, A., Aylin, P., Ziprin, P., Faiz, O.D., 2011.
14 Systematic review of discharge coding accuracy. *Journal of Public Health*.
- 15 Cameron, A., Trivedi, P., 1998. *Regression analysis of count data*. Cambridge University Press.
- 16 Casez, P., Labarère, J., Sevestre, M.-A., Haddouche, M., Courtois, X., Mercier, S.,
17 Lewandowski, E., Fauconnier, J., François, P., Bosson, J.-L., 2010. ICD-10 hospital

- 1 discharge diagnosis codes were sensitive for identifying pulmonary embolism but not
2 deep vein thrombosis. *Journal of Clinical Epidemiology* 63 (7), 790-797.
- 3 Clarke, S.P., 2004. Failure to rescue: lessons from missed opportunities in care. *Nursing Inquiry*
4 11 (2), 67-71.
- 5 Clarke, S.P., Aiken, L.H., 2003. Failure to Rescue: Needless deaths are prime examples of the
6 need for more nurses at the bedside. [Article]. *American Journal of Nursing* 103 (1), 42-
7 47.
- 8 Farrar, D.E., Glauber, R.R., 1967. Multicollinearity in Regression Analysis: The Problem
9 Revisited. *The Review of Economics and Statistics* 49 (1), 92-107.
- 10 Gao, H., McDonnell, A., Harrison, D.A., Moore, T., Adam, S., Daly, K., Esmonde, L., Goldhill,
11 D.R., Parry, G.J., Rashidian, A., 2007. Systematic review and evaluation of physiological
12 track and trigger warning systems for identifying at-risk patients on the ward. *Intensive*
13 *Care Medicine* 33 (4), 667-679.
- 14 Goryakin, Y., Griffiths, P., Maben, J., 2011. Economic evaluation of nurse staffing and nurse
15 substitution in health care: A scoping review. *International Journal of Nursing Studies* 48
16 (4), 501-512.
- 17 Griffiths, P., Jones, S., Maben, J., Murrells, T., 2008. State of the Art Metrics for Nursing: a
18 rapid appraisal. King's College London, London.
- 19 Griffiths, P., Renz, A., Hughes, J., Rafferty, A.M., 2009. Impact of organisation and
20 management factors on infection control in hospitals: a scoping review. *Journal of*
21 *Hospital Infection* 73, 1-14.

- 1 Hartz, A.J., Krakauer, H., Kuhn, E.M., Young, M., Jacobsen, S.J., Gay, G., Muenz, L., Katzoff,
2 M., Bailey, R.C., Rimm, A.A., 1989. Hospital characteristics and mortality rates. *New*
3 *England Journal of Medicine* 321 (25), 1720-1725.
- 4 Hayes, L.J., O'Brien-Pallas, L., Duffield, C., Shamian, J., Buchan, J., Hughes, F., Spence
5 Laschinger, H.K., North, N., Stone, P.W., 2006. Nurse turnover: A literature review.
6 *International Journal of Nursing Studies* 43 (2), 237-263.
- 7 Horwitz, L.I., Cuny, J.F., Cerese, J., Krumholz, H.M., 2007. Failure to Rescue: Validation of an
8 Algorithm Using Administrative Data. *Medical Care* 45 (4), 283-287.
- 9 Iezzoni, L.I., 1997. The risks of risk adjustment. *JAMA* 278 (19), 1600-1607.
- 10 Jarman, B., Gault, S., Alves, B., Hider, A., Dolan, S., Cook, A., Hurwitz, B., Iezzoni, L.I., 1999.
11 Explaining differences in English hospital death rates using routinely collected data. *BMJ*
12 318 (7197), 1515-1520.
- 13 Kane, R., Shamliyan, T., Mueller, C., Duval, S., Wilt, T., 2007. The Association of Registered
14 Nurse Staffing Levels and Patient Outcomes: Systematic Review and Meta-Analysis.
15 *Medical Care* 45 (12), 1195-1204.
- 16 Leibson, C.L., Needleman, J., Buerhaus, P., Heit, J.A., Melton III, L.J., Naessens, J.M., Bailey,
17 K.R., Petterson, T.M., Ransom, J.E., Harris, M.R., 2008. Identifying in-hospital venous
18 thromboembolism (VTE): a comparison of claims-based approaches with the Rochester
19 Epidemiology Project VTE cohort. *Medical Care* 46 (2), 127.
- 20 Luettel, D., Beaumont, K., Healey, F., 2007. Recognising and responding appropriately to early
21 signs of deterioration in hospitalised patients. National Patient Safety Agency.

Failure To Rescue

- 1 Mant, J., 2001. Process versus outcome indicators in the assessment of quality of health care.
2 International Journal for Quality in Health Care 13 (6), 475-480.
- 3 Mattke, S., Needleman, J., Buerhaus, P., Stewart, M., Zelevinsky, K., 2004. Evaluating the role
4 of patient sample definitions for quality indicators sensitive to nurse staffing patterns.
5 Medical Care 42 (2), II.
- 6 McDonald, K.M., Davies, S.M., Geppert, J.J.D., Romano, P.S., 2007. Why Rescue the
7 Administrative Data Version of the "Failure to Rescue" Quality Indicator. Medical Care
8 45 (4), 277-279.
- 9 McKee, M., Coles, J., James, P., 1999. 'Failure to rescue' as a measure of quality of hospital care:
10 the limitations of secondary diagnosis coding in English hospital data. Journal of Public
11 Health 21 (4), 453-458.
- 12 Moriarty, J.P., Finnie, D.M., Johnson, M.G., Huddleston, J.M., Naessens, J.M., 2010. Do pre-
13 existing complications affect the failure to rescue quality measures? Quality and Safety in
14 Health Care 19 (1), 65-68.
- 15 National Quality Forum, 2004. National Voluntary Consensus Standards for Nursing-Sensitive
16 Care: An Initial Performance Measure Set. National Quality Forum, Washington DC.
- 17 Naylor, M.D., 2007. Advancing the Science in the Measurement of Health Care Quality
18 Influenced by Nurses. Medical Care Research and Review 64 (2_suppl), 144S-169.
- 19 Rafferty, A.M., Clarke, S.P., Coles, J., Ball, J., James, P., McKee, M., Aiken, L.H., 2007.
20 Outcomes of variation in hospital nurse staffing in English hospitals: cross-sectional

- 1 analysis of survey data and discharge records. *International Journal of Nursing Studies* 44
2 (2), 175-182.
- 3 Schulz, S., Zaiss, A., Brunner, R., Spinner, D., Klar, R., 1998. Conversion problems concerning
4 automated mapping from ICD-10 to ICD-9. *Methods of Information in Medicine* 37 (3),
5 254-259.
- 6 Silber, J.H., Romano, P.S., Rosen, A.K., Wang, Y., Even-Shoshan, O., Volpp, K.G., 2007.
7 Failure-to-Rescue: Comparing Definitions to Measure Quality of Care. *Medical Care* 45
8 (10), 918-925.
- 9 Silber, J.H., Rosenbaum, P.R., Rosen, A.K., Romano, P.S., Itani, K.M., Cen, L., Mi, L., Halenar,
10 M.J., Even-Shoshan, O., Volpp, K.G., 2009. Prolonged hospital stay and the resident duty
11 hour rules of 2003. *Medical Care* 47 (12), 1191-1200.
- 12 Silber, J.H., Rosenbaum, P.R., Ross, R.N., 1995. Comparing the Contributions of Groups of
13 Predictors: Which Outcomes Vary With Hospital Rather Than Patient Characteristics.
14 *Journal of the American Statistical Association* 90 (429), 7-18.
- 15 Silber, J.H., Rosenbaum, P.R., Schwartz, J.S., Ross, R.N., Williams, S.V., 1995. Evaluation of
16 the complication rate as a measure of quality of care in coronary artery bypass graft
17 surgery. *JAMA: the journal of the American Medical Association* 274 (4), 317.
- 18 Silber, J.H., Williams, S.V., Krakauer, H., Schwartz, J.S., 1992. Hospital and patient
19 characteristics associated with death after surgery: a study of adverse occurrence and
20 failure to rescue. *Medical Care* 30 (7), 615-629.

Failure To Rescue

- 1 Williams, J., Mann, R., 2002. Hospital episode statistics: time for clinicians to get involved?
- 2 Clinical Medicine, Journal of the Royal College of Physicians 2 (1), 34-37.
- 3