Surgeon-specific mortality data disguise wider failings in delivery of safe surgical services

Stephen Westaby*, Ravi De Silva, Mario Petrou, Simon Bond and David Taggart

Abstract

OBJECTIVES: Feedback of clinical outcome data to clinicians can promote and enhance patient safety. Surgeon-specific mortality data (SSMD) have been released to the public for a number of specialties. This implies that one individual is culpable for all deaths. Debate continues about SSMD because of risk-averse behaviour. In the USA, improved outcome measures derived from phase of care mortality analysis (POCMA) and the failure to rescue (FTR) are replacing SSMD, but they have not been tested in Europe.

METHODS: Using POCMA and FTR analysis, we studied hospital deaths in 1558 cardiac surgical patients between 2009 and 2013. Comorbidity and urgency status were used to calculate modified logistic EuroSCORE (MLE). The circumstances of death were critically reviewed by a panel of four experienced surgeons. Death certificate information and autopsy were taken into account. Deaths were then classified: Class 1 surgeon dependent, Class 2 FTR or Class 3 where multiple factors conspired to cause death.

RESULTS: There were 51 deaths providing 3.3% mortality, as predicted by MLE. In the 86% who underwent autopsy, no surgical error was identified. Most deaths in each group were related to high-risk status, age, frailty, comorbidity and urgency. FTR was the predominant factor occurring in 45%. Though difficult operations were implicated in 37%, no deaths occurred in the operating theatre. Some FTR deaths occurred in low-risk patients. Scrutiny of FTR deaths provided important information that could be used for quality improvement.

CONCLUSIONS: The study showed that most deaths cannot be prevented by the operating surgeon. They occurred through issues of patient comorbidity, lack of process or infrastructure. This casts doubt on SSMD publication alone as a tool for quality improvement. In contrast, POCMA and FTR highlight problems of process, and are more likely to promote advances in surgical care.

Keywords: Failure to rescue • Phase of care • Public reporting • Surgeon-specific mortality risk assessment

INTRODUCTION

It has long been accepted that feedback of outcome data to clinicians is vital to promote and enhance patient safety [1]. The most widely recognized example of this is the publication of named surgeon-specific mortality data (SSMD) [2]. SSMD are also used to illustrate the openness and transparency of the National Health Service (NHS). Yet debate continues about the real value of SSMD because of its potential adverse consequences [3, 4]. In the public eye SSMD implies that only one individual has the scope to determine patient outcome. This approach may inadvertently undermine the importance of team work while best evidence suggests that collective team experience and multidisciplinary input have considerably more influence on outcome [5, 6]. As a result, SSMD may also divert attention away from key aspects of process and management that impact directly on mortality rates [7, 8]. Finally, it is widely acknowledged in the surgical community that an inevitable consequence of SSMD is ‘risk-averse behaviour’ whereby the sickest patients with potentially the most to gain are less likely to be offered an operation [3, 4]. It is this group of patients who largely determine postoperative mortality and consume considerable resources in the intensive care unit [9].

Recently two important new concepts have been introduced in North America to improve understanding as to why patients die after cardiac surgery. First, ‘phase of care mortality analysis’ (POCMA) is a quality improvement tool based on the concept that a patient’s clinical course is a series of events, therapeutic decisions, interventions and treatment responses that are inter-dependent. During this pathway, deaths are triggered by a seminal event (root cause), which is the catalyst for a cascade of processes that culminate in death [10]. To identify the seminal event, the preoperative, intraoperative, intensive care, postoperative ward and discharge phases of care are scrutinized carefully [10]. Secondly, ‘failure to rescue’ (FTR) recognizes that postoperative...
complications principally relate to existing comorbidity and occur at similar rates in all tertiary care centres [11]. It is their expeditious and effective treatment that differentiates mortality rates for different institutions [11–13]. As such FTR is now accepted as an important quality metric for cardiac, vascular, general and trauma surgery programmes in US hospitals.

To our knowledge, POCMA and FTR have not been systematically employed in the analysis of deaths after cardiac surgery in Europe. We therefore reviewed the deaths of cardiac surgical patients in a large tertiary UK teaching hospital and asked the questions ‘why do patients die after cardiac surgery’ and ‘which outcome measures can provide quality improvement in the patient’s best interests?’. We specifically tried to identify deaths that were due to intraoperative events and those deemed FTR and therefore potentially avoidable. We then questioned whether attributing deaths solely to an individual surgeon can be justified by clinical evidence?

PATIENTS AND METHODS

We prospectively collected demographic, clinical and outcome data for 1558 consecutive patients operated by three cardiac surgeons over a 4-year period (2009–13). The operations performed are given in Table 1.

All data were submitted continuously to the Society of Cardiothoracic Surgeons UK Database for analysis and SSMD then published in a public portal [14]. Where urgency status permitted, all patients were subject to multidisciplinary team review by cardiologists, surgeons and anaesthetists before acceptance for surgery. After surgery all patients were managed primarily by consultant anaesthetists in the cardiac intensive care unit (ICU). After return to single rooms on the ward, the surgeon resumed responsibility for day-to-day care. As for the Society database postoperative mortality was defined as all deaths that occurred during the hospital episode in which the operation was performed [14]. Salvage and emergency cases or trauma with catastrophic preoperative status requiring immediate operation are not included in the study. After surgery all patients were subject to multidisciplinary team review by cardiologists, surgeons and anaesthetists before acceptance for surgery. After return to single rooms on the ward, the surgeon resumed responsibility for day-to-day care. As for the Society database postoperative mortality was defined as all deaths that occurred during the hospital episode in which the operation was performed [14]. Salvage and emergency cases or trauma with catastrophic preoperative state requiring immediate operation are not included in the study.

We determined 30-day mortality whether or not the patient died in our centre. We cross-checked all patients names with national mortality records to check that we had not missed any patient who had died outside the centre. Hospital or Coroner’s autopsy were routinely requested and occurred in 86% of cases. Whenever possible, the surgeon or a member of the team attended the autopsy with the specific objective of identifying or ruling out a potential surgical cause for poor outcome. Death certificate information was obtained for every case. After autopsy reports became available, deaths were reviewed in detail at the monthly Morbidity and Mortality meeting attended by the same staff as the preoperative review.

POCMA was used to determine the root cause of death [10]. Before the operation, comorbidity and urgency status were used to calculate modified logistic EuroSCORE (MLE). This was the risk assessment tool used by our society at the time of the study. The operation notes and perfusion records were scrutinized to determine whether the procedure had been straightforward and uneventful or complicated by predictable or unpredictable events. Myocardial ischaemic time, duration of cardiopulmonary bypass and overall operation time together with the need for reoperation for bleeding were recorded. The intraoperative quality assurance measures, transoesophageal echocardiography (for all valve operations) and bypass graft flow imaging (in a large number of coronary artery bypass grafts (CABG)) were taken into account.

We reviewed the postoperative care to ascertain factors that contributed to death. Complications examined included reoperation for bleeding, cardiac tamponade, pneumothorax or haemothorax, dysrhythmia or cardiac arrest, prolonged ventilation, ICU readmission, pneumonia, renal failure + dialysis, sternal wound infection, systemic sepsis, coma, delirium or stroke, gastrointestinal bleeding or ischaemia, anticoagulation events, multiorgan failure or accidents including air embolism and tracheostomy tube displacement. Time and site of death were recorded. All findings were confirmed with the hospital data managers and cross referenced with the Society of Cardiothoracic Surgeons of Great Britain and Northern Ireland and National Institute for Clinical Outcomes Research databases.

The circumstances of death were compiled then critically reviewed by the three operating surgeons. They were then reviewed by an experienced surgeon (M.P.), whose patients were not included in the study, and finally classified by the panel of four with a collective experience of 60 years at consultant level. From the findings, we determined whether the trigger for death was surgeon dependent (intraoperative) or FTR from a de novo postoperative complication [11]. Because death is often related to

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Sex</th>
<th>Number of patients</th>
<th>Median patient age</th>
<th>Deaths</th>
<th>Percentage deaths</th>
<th>Average predicted MLE (%)</th>
<th>Median predicted MLE (%)</th>
<th>Median postop hospital stay</th>
</tr>
</thead>
<tbody>
<tr>
<td>CABG</td>
<td>Male</td>
<td>733</td>
<td>66</td>
<td>16</td>
<td>2.2</td>
<td>2.1</td>
<td>1.3</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>138</td>
<td>70</td>
<td>6</td>
<td>4.3</td>
<td>3.1</td>
<td>2.3</td>
<td>7</td>
</tr>
<tr>
<td>AVR</td>
<td>Male</td>
<td>190</td>
<td>69</td>
<td>2</td>
<td>1.1</td>
<td>2.7</td>
<td>1.7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>113</td>
<td>70</td>
<td>2</td>
<td>1.8</td>
<td>2.1</td>
<td>2.6</td>
<td>7</td>
</tr>
<tr>
<td>AVR + CABG</td>
<td>Male</td>
<td>121</td>
<td>75</td>
<td>4</td>
<td>3.3</td>
<td>5.5</td>
<td>3.7</td>
<td>9</td>
</tr>
<tr>
<td>MV surgery</td>
<td>Male</td>
<td>37</td>
<td>78</td>
<td>3</td>
<td>8.1</td>
<td>6.4</td>
<td>5.8</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>10</td>
<td>57</td>
<td>0</td>
<td>0</td>
<td>1.4</td>
<td>0.9</td>
<td>8</td>
</tr>
<tr>
<td>MV + CABG</td>
<td>Male</td>
<td>14</td>
<td>68</td>
<td>0</td>
<td>0</td>
<td>4.7</td>
<td>3.0</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>7</td>
<td>77</td>
<td>0</td>
<td>0</td>
<td>6.8</td>
<td>4.7</td>
<td>11</td>
</tr>
<tr>
<td>Other</td>
<td>Male</td>
<td>116</td>
<td>65</td>
<td>7</td>
<td>6.0</td>
<td>7.2</td>
<td>3.1</td>
<td>8</td>
</tr>
<tr>
<td>surgery</td>
<td>Female</td>
<td>73</td>
<td>67</td>
<td>10</td>
<td>13.7</td>
<td>5.5</td>
<td>4.2</td>
<td>9</td>
</tr>
<tr>
<td>Total group</td>
<td>1558</td>
<td>68</td>
<td>51</td>
<td>3.3</td>
<td>3.3</td>
<td>3.3</td>
<td>1.8</td>
<td>6</td>
</tr>
</tbody>
</table>

pre-existing patient-dependent factors, we included a third ‘multifactorial’ category. The definition of these categories is as follows:

Class 1 surgeon dependent: mortality directly related to the surgical procedure, where inappropriate patient selection, a de novo intraoperative event, a technical error or postoperative bleeding initiated the pathway to death.

Class 2 failure to rescue: mortality following an uneventful operation where a new potentially treatable and retrievable postoperative complication not present on hospital admission caused death.

Class 3 multifactorial: mortality where multiple factors including frailty, comorbidity, critical preoperative state or surgical complexity, contributed to death despite an uneventful and effective surgical procedure.

When unpredictably complicated surgery resulted in prolonged ischaemic and cardiopulmonary bypass time, long operating theatre time or reoperation for bleeding, the death was deemed ‘surgeon related’ as were cases where autopsy suggested a surgical cause. We attributed deaths from intraoperative stroke or multiorgan failure secondary to low cardiac output state originating in the operating theatre as surgeon related. Even when an important comorbidity, missed preoperatively, caused death we categorized mortality as surgeon related because FTR relates to a de novo postoperative complication.

Statistical methods

Due to the nature of the data and sample sizes that are small in some sub-groups, data have been presented as median values with interquartile ranges (IQRs). The nature of the study did not set out to explore a hypothesis that was framed by a statistical test or analysis.

RESULTS

From a total of 1558 patients there were 51 postoperative deaths (3.3%) (Table 1). None of the patients died in the operating theatre. The overall mortality matches the predicted average MLE for the entire patient group (3.3%). The median age for all patients was 68 years but 79 years for the 51 who died. The median MLE was appreciably higher for those who died (4.6%) compared with the entire patient cohort (1.8%). The MLE for the whole patient group and by class of death is shown in Fig. 1. Females had a higher actual mortality rate in all surgical categories with the greatest disparity for more complex procedures such as mitral valve surgery with CABGs. Twenty-two of the 51 deaths (43%) occurred in patients undergoing the commonest operation, isolated CABG and Class I CABG mortality was principally related to urgent status. The number of deaths was closely aligned with that predicted by MLE (22.0 vs 19.2). Half were attributed to FTR as were half of the deaths after aortic valve replacement.

With reference to the sub-groups, 19 deaths (37.3%) were deemed Class 1, 23 (45.1%) Class 2 and 9 (17.6%) Class 3 (Table 2). Thus, FTR accounted for the largest group. Class I deaths occurred in only 1.2% of the whole patient cohort. Of note, none of the autopsies revealed a surgical technical error as precipitating factor for low cardiac output state and death. Therefore, despite risk status and surgical complexity adverse intraoperative events were rare.

Class 3, the multifactorial group had the oldest and sickest patients with median age 80 and MLE 7.0. In contrast, Class 1 patients were younger (median age: 71) with less comorbidity (median MLE: 2.9). The Class 2 FTR patients fell between the other categories with median age 75 and median MLE 4.7. This is consistent with FTR patients having significant comorbidity, which predisposes to postoperative complications. The median time to death for each category was Class I: 5 days (IQR: 3–14), Class 2: 9 days (IQR: 6–14 days) and Class 3: 4 days (IQR: 3–12 days).

As expected the median postoperative length of stay increased with surgical complexity. Whereas the median postoperative stay was 6 days for the entire group, those undergoing combined aortic valve replacement with CABG and mitral valve surgery with CABG had median stays of 9 and 11 days, respectively.

DISCUSSION

Although increasingly used as a quality metric in North America, to our knowledge this is the first time POCMA and FTR have been used in an analysis of SSMD in patients undergoing major surgery in Europe. The most striking observation in the current study is that we estimate FTR to account for 45% of deaths after cardiac surgery while mortality related to intraoperative events accounted for 37%. We also identified a third category of mortality whereby death could not convincingly be classified solely within the first two categories, but seemed to be due to a combination of factors.
Our assertion is that POCMA cannot always identify an individual root cause as suggested [10]. Equally, some deaths cannot be prevented by improved treatment in the operating theatre, ICU or ward environment. Assuming our findings are typical of most cardiac surgery units in the UK (with an overall 3.4% mortality in 30 000 UK cardiac surgery patients per year) this would potentially amount to 459 avoidable deaths per year [14]. By simply attributing these deaths to a named surgeon, the system disguises numerous deficiencies in NHS infrastructure that could and should be addressed to improve outcomes [9].

Publication of outcome data is a prerequisite for transparency and accountability in any healthcare system [1–4]. However for either to be truly valid and meaningful requires an explanation of why the patient died rather than arbitrary allocation of responsibility, whatever the circumstances, to one team member. For this study, we were stringent in allocating blame to the surgeon. Even so, detailed clinical review, POCMA and autopsy findings showed that surgical judgement or technical performance was rarely implicated in poor outcome (~1% of all cases). Indeed in 86% of patients who underwent autopsy, none revealed any surgical technical error that obviously or directly explained the patient’s death. Equally risk profiling does not protect the operator as this becomes less accurate when age, critical preoperative state and urgency of surgery combine. In contrast, many treatable FTR events happened either suddenly or out of working hours or were not recognized early enough for the surgeon to intervene. Some occurred in low-risk patients. Because an individual surgeon’s number of operations is relatively small (<200/year), one or two FTR events can double mortality and trigger defensive practice. This encourages the strategy of sending prolonged intensive care patients back to the referring hospital in order to conceal surgeon- and centre-specific mortality from the database.

In all three categories, deaths occurred predominantly in elderly high-risk patients who required urgent operation. A major fallibility of risk scoring systems in the elderly surgical population is that biological frailty, defined by Fried as a syndrome of decreased physical reserve and resistance to stress, is not included [15]. Furthermore, frailty is distinct from comorbidity. As many as 27% of patients defined as frail have no associated comorbidity but frailty itself is significantly correlated with 30-day mortality after cardiac surgery [16]. Considering that other significant comorbidities that are prevalent in the elderly (neurological impairment, liver disease, porcelain aorta or previous chest radiation) are not taken into account, emphasizes that risk assessment measures may be misleading in the population that sustains most deaths. Such nuances are missed in the case mix-adjusted surgeon mortality rates published in the UK [17]. Our findings are consistent with those from the National Confidential Enquiry into Patient Outcome and Death (NCEPOD) ‘Review of care received by elderly patients undergoing surgery’ (2010) [9]. This showed that only 38% of the deceased patients had received good treatment and highlighted delays between admission and surgery. Deaths were predominantly attributed to issues of process and infrastructure rather than to surgical errors.

The term FTR was adopted in relation to death from a new postoperative complication not present at the time of hospital admission [11]. Ghafieri sub-categorized FTR into two components: first, time taken to recognize a life threatening complication and secondly, the ability to manage that problem effectively [12]. All tertiary care centres have similar complication rates closely related to patient comorbidity. In a recent analysis of 45 904 risk-stratified cardiac surgical patients from 33 US hospitals, Reddy confirmed that low mortality centres had superior ability to rescue from postoperative events [11]. Whereas the complication rates were similar (ranging from 19 to ~23%) FTR rates varied from 7 to 14% between low and high mortality centres. Differences were greatest in the treatment of cardiac arrest, renal failure, sepsis, prolonged ventilation and pneumonia. Similar differences in quality of care are now widely recognized in NHS hospitals as reflected in the Francis Inquiry and Keogh Mortality Review (2013) [18, 19]. Team consistency, ICU staffing levels, nurse education, job satisfaction and burn-out have all been identified as factors that underpin FTR [20]. Aitken et al. [20] showed that each additional patient to nurse was associated with 7% increase in the odds of FTR. West demonstrated a clear relationship between the effectiveness of human resources management and hospital mortality rates [21]. When the factors that underpin FTR are recognized, quality improvement initiatives follow. Through these developments US hospital patients are said to be far less likely to die from FTR events than their UK counterparts [22]. Our analysis supports this.

High-risk cardiac surgical patients need round the clock observation and expert care by appropriately trained staff [23]. Yet several factors conspire against this in the NHS. Overnight and at weekends, patients are managed predominantly by anaesthetic or surgical trainees with variable operative or intensive care experience. This is because of ‘streamlining of training’ in the UK and limitations imposed by the European Working Time Directive. This not only impacts adversely on team consistency and collective experience but means that trainees are less likely to recognize or deal effectively with unexpected but remediable postoperative events. As a result, some patients reach an irretrievable state before skilled intervention occurs. Inevitably, a number of deaths in our study were attributable to out-of-hours events, which involved inexperienced or temporary staff. Along these lines, Aylin et al. [24] reported a significant trend towards higher mortality for coronary bypass and complex surgery at the end of the working week or weekend. There are important differences between US and UK surgical practice. The USA emphasizes the importance of early intervention, while as reported by NCEPOD, delay is problematic in the NHS and likely to increase risk and the prospect of waiting-list deaths [9]. Rescue circulatory support equipment is uniformly available in the US but funded only for transplant centres in the UK.

Cardiac SSMD were released in the UK (2006) following the Bristol inquiry. After this, UK mortality rates were said to fall below US and European levels despite the broadly acknowledged weaknesses in NHS hospital manpower and infrastructure. It is impossible to define whether this occurred through genuine improvements in care or as a result of risk-averse behaviour [3]. Nor is there clear evidence that the public understands or engages with risk-adjusted data. Nonetheless, publication of SSMD has been imposed on other surgical specialties and interventional cardiology without local systems to confirm, refute or address outlying performance. This has caused concern even among the most staunch protagonists of outcome reporting. In the wake of the Francis Inquiry into excess deaths in a UK hospital, transparency and quality improvement have emerged as key NHS policies [18]. Unfortunately, we believe that continued publication of SSMD has the potential to undermine these initiatives. Two decades ago after release of SSMD by default in the USA, Donald Berwick, then President of the Institute for Healthcare Improvement, argued that negative outcomes were due to failures of process and systems, not individuals. He cautioned against the name and shame culture stating that ‘mortality is a property inherent in the system and only by changing systems can better results be expected’. Subsequently (2007) the Society of Thoracic
Surgeons in the USA established a Quality Management Task Force that specifically advised reporting of outcome data at programme level not SSMD. They advocated composite hospital quality scores based on analysis of structure, process and outcome. Berwick was subsequently commissioned by the British Prime Minister to examine NHS patient safety after the Francis Inquiry. In his report he reiterated, when things go wrong it is usually due to failings of the system and there is no logic in penalizing individuals [25].

In summary, our findings cast doubt on the validity of SSMD through which surgeons, alone, are implicated in deaths for which they are not responsible. It is arguable that SSMD also disguise failings in service delivery by diverting focus away from actual causes of death and deficiencies in infrastructure. Risk-averse behaviour can reduce mortality but it is not in the interest of the patient or profession. We emphasize that provision of skilled staff, consistency of team work and access to circulatory support equipment are not in the remit of an individual NHS surgeon. Our experience is likely to be similar to that of other acute specialties in the UK. While we support the public reporting of outcome data, we propose that POCMA and FTR rates should also be introduced to improve safety and transparency. Reduction in FTR rates should be a clear objective for any centre and as for the USA, public comparisons should be at programme level. Although this may prove expensive, it has the potential to drive genuine quality improvement.

LIMITATIONS OF THE STUDY

Risk assessment and mortality analysis are inexact sciences. They are both subject to bias, subjective interpretation and gaming. In this study, classification of mortality was as objective as possible and supported by independent autopsy findings (86% of deaths) and intraoperative quality control measures. These included transesophageal echocardiography for all valve patients, or bypass graft flow imaging in a large number of CABG. Causes of death were verified by an experienced surgeon whose own deaths were not included in the study. We did not include an anaesthetist in the analytic process: the patients who died. We acknowledge that this could be regarded as a weakness in the analysis though the combined experience of the surgeons exceeded 60 years and 20,000 operations. Finally, it is sometimes difficult to separate deaths from multiple factors (high-risk status, urgency, frailty), which conspire with frequent complications (e.g. dysrhythmia, renal impairment and infection) to cause death, as surgeon related or FTR. For this reason we have introduced the multifactorial category to mortality analysis. For many of these patients, death was likely with even minor complications but risk-benefit consideration still supported surgical intervention.

ACKNOWLEDGEMENTS

We thank our data manager Colin Evans for his considerable efforts in cross checking and verifying our recorded data. We also thank Sue Francis for help in the preparation of the manuscript.

Conflict of interest: none declared.

REFERENCES


