European Heart Journal – Quality of Care and Clinical Outcomes (2017) **3**, 11–19
doi:10.1093/ehjqcco/qcw027

REVIEW

Measuring the quality of perioperative care in cardiac surgery

Tim G. Coulson[1]*, Daniel V. Mullany[2], Christopher M. Reid[1], Michael Bailey[1], and David Pilcher[3,4]

[1]School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia; [2]Critical Care Research Group, The Prince Charles Hospital and University of Queensland, Brisbane, Australia; [3]Department of Intensive Care, The Alfred Hospital, 55 Commercial Rd, Melbourne, Victoria 3004, Australia; and [4]ANZICS Centre for Outcome and Resource Evaluation, Ievers Terrace, Carlton, Melbourne, Victoria, Australia

Received 15 March 2016; revised 3 May 2016; online publish-ahead-of-print 19 May 2016

Quality of care is of increasing importance in health and surgical care. In order to maintain and improve quality, we must be able to measure it and identify variation. In this narrative review, we aim to identify measures used in the assessment of quality of care in cardiac surgery and to evaluate their utility. The electronic databases Pubmed/MEDLINE, EMBASE, Cochrane Central Register of Controlled Trials (CENTRAL), Cochrane Database of Systematic Reviews, and CINAHL were searched for original published studies using the terms 'cardiac surgery' and 'quality or outcome or process or structure' as either keywords in the title or text or MeSH terms. Secondary searches and identification of references from original articles were carried out. We found a total of 54 original articles evaluating measurements of quality. While structure, process, and outcome indicators remain the mainstay of quality measurement, new and innovative methods of risk assessment have improved reliability and discrimination. Continuous assessment provides a promising method of both maintaining and improving quality of care. Future studies should focus on long-term and patient-centred outcomes, such as quality-of-life measures.

Keywords Quality of health care • Surgery • Cardiac • Safety and quality

The concept of quality of care

As individuals we all wish to receive the best health care possible. Similarly, healthcare providers have a strong interest in delivering this. To understand what this means, and therefore to define quality, we have to consider what key factors we want from our care, both as individuals and as a population. The quality of care has classically been defined as 'the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge';[1] however, the concept of quality of care is more complex than this. In their publication 'Crossing the Quality Chasm', the Institute of Medicine (IOM) state health systems should be safe, effective, patient-centred, timely, equitable, and efficient.[2] In reality, this requires two factors: access to care and effective care given.[3]

The delivery of high-quality care in health systems is of increasing importance.[4] Rising costs of medical care, reductions in available resources, and health scandals have focused media and public attention on healthcare quality. Many countries have statutory organizations to monitor quality such as the Care Quality Commission in the UK or the Australian Commission on Safety and Quality in Health Care. In addition to this, many medical specialties have national quality monitoring organizations [for example the National Institute for Cardiovascular Outcomes Research (NICOR) in the UK] to determine and monitor standards of care. Reduction in errors, provision of best practice, and standardization have the potential to make a huge impact on overall patient outcomes and may even be more effective than the (sometimes-marginal) benefit achieved by developing new treatments. However, before we can improve quality, we need an effective way to measure and identify variation between providers.

Measuring quality of care

Quality of care is measured using a systems approach divided into structural, process, and outcomes measures.[5] Structural measures broadly assess resourcing such as hospital facilities and staff characteristics. Structural measures may be useful where there are few or no process or outcome measures and potentially used to set a minimum standard. Process measures describe what is done and how it is done. These focus on using evidence-based treatment, known to have a positive effect on patient outcome. More recently, learning curves and team function in complex interventions have been studied using human factors and systems engineering methods.

* Corresponding author. Tel: +61 3 9076 2000, Fax: +61 3 9076 3780, Email: timcoulson@doctors.org.uk

Outcome measures are perhaps the most intuitive and relevant to our practice. They may be usefully divided into four groups: functional status, medical outcome, patient satisfaction, and cost.[6]

The structure, process, and outcome measures of quality are all necessarily interrelated. Structural measures are valid only when they affect the ability to deliver a process. For process measures to be credible, they should be related to improved outcome, and for outcome measures to be useful, they should be improved by alteration of the process.[7] Proving a causal link between process and outcome has been problematic.

Measurement of quality in cardiac surgery

Cardiac surgery was an early specialty to measure quality of care on a large scale. Data collection for calculating risk-adjusted mortality rates in the USA began in 1987.[8] Since then, the database has expanded significantly. Current quality measures collected by the Society of Thoracic Surgeons (STS) in the USA include a composite score for each provider generated from a variety of process and outcome measures. Individual measures used to rate quality include outcomes (risk-adjusted mortality, length of stay, and major morbidity including stroke, new renal failure, re-operation, deep sternal wound infection, and prolonged intubation), processes [use of internal mammary artery (IMA) and pre- and postoperative medication], and one structural measure (participation in the database).[9] In the UK, NICOR reports annually, providing data on process measures (IMA use and mitral repair) and mortality outcomes.[10] A more comprehensive intermittent report comprising similar measures as for the STS is also produced (with the exception of a composite quality score).[11] The Australia and New Zealand Cardiothoracic Society Cardiac Surgery Database (ANZSCTS-CSD) reports mortality and morbidity outcomes, although not all states participate in reports and fewer morbidity outcomes are reported.[12] Large cardiac surgery databases are available in most countries some with linkage to social security via a unique identifier.

In addition to the above, many providers internally monitor their own performance.[13] Although there is potential for bias, where review is independent and multidisciplinary it may allow for a more contemporaneous recognition of problems, potentially addressing concerns before they cause further harm.

Aims

In this article, we carry out a narrative review of the literature to identify measures of quality in cardiac surgery and evaluate the evidence behind their utility for identifying changes in performance with respect to time or other providers. We aim to provide an overview for the non-expert of the measures available for monitoring quality of care, as well as providing an evidence-based assessment of their relative merits.

Methods

Data sources and study selection

The electronic databases MEDLINE, EMBASE, PubMed, Cochrane Central Register of Controlled Trials (CENTRAL), Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effectiveness, CINAHL, and Web of Science were searched for original published studies. The most recent search was on 1 November 2015. To maximize sensitivity of the search strategy and identify all studies, the authors combined the terms: 'cardiac surgery' and 'quality or outcome or process or structure' as keywords in title or text. Initially, titles and/or abstracts were screened. The reference lists of retrieved articles were reviewed to identify additional, potentially relevant studies.

Selection criteria

Observational studies examining the use of any quality of care measures in patients who underwent cardiac surgery were included. Studies published in the year 1990 or later were included. Studies that involved thoracic organ transplantation, forms of mechanical circulatory support, such as ventricular assist devices or extracorporeal membrane oxygenation, robotic surgery, and those involving percutaneous valve surgery were excluded. With duplicate studies, those with accumulating sample sizes or increased lengths of follow-up, only the most recent complete reports were included. Studies were limited to human subjects and published in the English language. Abstracts, case reports, conference presentations, editorials, and expert opinions were excluded. Articles focusing solely on quality improvement, rather than quality measurement, were beyond the scope of this review and were also excluded.

Articles were then grouped based on category of assessed measure and method of assessment.

Discussion

Literature search

The initial searches identified 2175 sources. After initial selection of articles, a total of 429 articles were identified. Detailed analysis reduced this to a total of 51 articles. An additional three articles were identified using secondary searches and references (Supplementary material online, *Figure S1* and *Table S1*). Quality measures assessed are shown in *Table 1*.

Outcome measures

Mortality

Mortality is the most widely collected metric. It is a solid, irrefutable, logical, and important outcome. It is usually but not always adjusted for patient-level risk factors to allow for difference in provider case mix. However, despite its widespread use, mortality is not necessarily a good measure of quality, and while a provider with increased mortality rates may reflect a reduction in quality, the lack of a detectable mortality difference does not exclude a difference in quality.

Risk adjustment is usually carried out using preoperative patient variables not under the control of the provider. Estimated risk of death is calculated using a validated or locally derived risk score calculated using multivariate analysis of large databases (e.g. Euroscore, AusScore, STS, RACHS score).[14–16] Outlier status may vary according to risk-adjustment model used.[17] Using administrative databases (usually set up to determine costs of care and funding models) is also possible, but these may be prone to missing data.[18] While these scores provide a useful prediction for populations, higher-risk patients may have exaggerated risk scores, resulting in overestimation of risk of death in this group.[19] This may result in a lower than appropriate risk-adjusted mortality rate. In addition, gaming

Table 1 Quality measures

Structural	Volume Staff/skill set
Process	Medical treatment: Perioperative beta-blockers Antibiotic prophylaxis Anti-platelet therapy Lipid-lowering therapy Discharge beta-blockers Postoperative glucose Surgical treatment: IMA use Technical performance Other: TOE use Epiaortic ultrasound use
Outcome	Mortality Morbidity: New renal failure Deep sternal wound infection Prolonged perioperative ventilation Stroke Reoperation Reintubation Transfusion Length of stay Readmission rate Other: Failure to rescue Quality of life MV repair Cost of care Inotrope/mechanical support
Other	Composite: STS composite score ARC

IMA, internal mammary artery; TOE, transoesophageal echocardiography; STS, Society of Thoracic Surgeons; ARC, acute risk change; MV, mitral valve.

by 'upcoding' variables may result in lower than appropriate risk-adjusted mortality.[20] Where surgery and populations are less standard, such as in adult congenital heart surgery, risk adjustment becomes more complex and controversial.[21–23] Similarly, more rare but important comorbidities (such as severe liver disease) may not be included.

A quality measure would ideally be able to detect a variation in provider quality rapidly in order to identify and rectify the cause. The statistical ability to detect a difference in binary outcome depends on the frequency of the outcome interest (in this case death) and the total number of tests (the number of surgical procedures). Dimmick *et al.* simulated the number of cases required to detect a mortality in coronary artery bypass grafting surgery of either double or 1.5 times the benchmark mortality.[24] They found a minimum caseload of 219 cases needed to detect double the mortality, or 744 to detect 1.5 times the mortality. This was based on a mortality rate of 3.5%. With mortality rates for some procedures now at 1–2%,[12] these numbers would be significantly increased. With a large case volume required to detect increased mortality, the time to accrue these cases becomes longer, delaying the detection of cases markedly. This may be particularly relevant to less commonly performed procedures.[25] Cases of worsened quality of care may be short-lived, and the staff and practices involved may have moved on before detection and investigation.

Mortality is commonly defined as death while in hospital during the immediate admission or within 30 days in or out of hospital.[26] By these definitions, a patient who is transferred to another hospital, survives 30 days and subsequently dies as an inpatient, would not be necessarily counted as an operative death. Similarly, patients readmitted for a complication of surgery who subsequently die would not necessarily be recorded as operative mortality. This may have an effect on provider quality ratings.[27,28] For this reason, and others, longer-term mortality outcomes are being included in registry databases,[11] and linkage to national mortality data has occurred.

Typically, the provider who is assessed is the hospital or individual cardiac surgeon. More recently, the role of the anaesthetist in cardiac surgery mortality outcomes has been investigated in the USA and UK in large datasets. While one major study found a significant difference between anaesthetist-adjusted mortality rates,[29] the other found that the difference was negligible.[30] This may reflect differences between UK and US anaesthesia practice or may be a result of different statistical methodology.

The rate of preventable death compared with risk-adjusted mortality has been investigated, and no correlation was found.[31] This suggests risk-adjusted mortality alone may not be a useful marker for performance. Finally, survival at all costs should not necessarily be a goal to aspire to. That is to say, we should consider the possibility and impact of altered quality of life postoperatively as a result of morbidity and disability.

Morbidity

The occurrence of complications after surgery is an important aspect of quality of care and results in increased cost, patient suffering, and considerable inconvenience. Complications are associated with reduced long-term survival[32] after surgery. In addition, they are likely to contribute towards poorer quality of life, even if survival is not reduced.[33,34] Morbidity may be a more sensitive measure for identifying outlier units due to a higher incidence and may help in the early identification of poorly performing providers.[35] Similarly to mortality rates, the adjustment for risk is important in providing a fair assessment of quality.[36] Morbidity measures used to assess quality should have a proven association with long-term disability and/or mortality.

Re-exploration for bleeding has been identified as a risk factor for mortality and other postoperative complications.[37,38] Wolfe *et al.* have analysed re-exploration using the cumulative sum (CUSUM) analysis. In this way, they were able to identify an excess event rate, investigate the event, and rectify the causative factor.[39] Re-exploration rate is routinely monitored by many cardiac databases.

Blood product has been used both as a quality metric and as a target for quality improvement by some centres.[40,41] A targeted reduction in blood product use was associated with a reduction in other complications. Risk adjustment and standardization of transfusion thresholds are likely to be important. A recent large randomized trial showed that a restrictive transfusion threshold was associated with an increased risk of death (although no cause–effect relationship could be demonstrated).[42]

Prolonged mechanical ventilation,[33,43] deep sternal wound infection,[34,44] postoperative renal dysfunction,[45] and stroke[46] are all associated with short- and long-term mortality, as well as poorer quality-of-life outcomes. Some centres have also reported continuous monitoring of these complications.[47] While there is correlation between morbidity and mortality (risk adjusted), morbidity appears to provide additional information over and above mortality alone.[48]

Length of stay

Length of stay has been suggested as a quality measure. At first examination, this appears reasonable. Long hospital stays are usually undesirable for the provider and patient and are associated with complications[49] and costs. Liu *et al*. reported a prolonged length of stay (>80th centile) to be a useful measure as a surrogate for mortality in paediatric cardiac surgery (where procedure numbers were low);[50] however, other investigators have noted poorer correlation with mortality.[49] Length of stay may be affected by many clinical and non-clinical factors, resulting in marked variation between hospitals. If length of stay is to be used, it should be used in the context of other important outcome measures.

Readmission rate

Readmission to hospital within 30 days of discharge may be a marker of quality in surgery and appears to be associated with other quality measures such as mortality rate, hospital volume, and surgical quality score.[51] However, this is controversial (partly as a result of the affect this metric has on funding), and some have suggested that variability in readmission rate in cardiac surgery is attributable to statistical noise.[52] In addition, it may be affected by systematic differences between hospitals in recording of data for readmissions to other hospitals.[53]

Other outcomes

Mitral valve repair versus replacement for isolated P2 prolapse is used by some organizations as a marker of quality.[10] Outcomes that may be useful as quality markers and have been evaluated include reintubation,[54] unplanned readmission to ICU,[55] and cost of care.[56] In addition to these, other outcomes related to quality of care and preventable death may be measured, for example, new intra-aortic balloon pump use, postoperative inotropic support, or postoperative cardiac arrest.[57]

The overall aim of medical quality improvement programmes is to get the best possible outcome for the patient. While mortality and morbidity are clearly important, we need to consider more patient-centred outcomes.[58] Patients are likely to be equally interested in quality of life as an outcome.[59] Effective methods of measuring quality of life are available[60] and are likely to become more important in quality measurement as outcomes research evolves, although significantly increased resources are required to collect these data.

Failure to rescue was reported as a measure of quality more than 20 years ago.[61] Initial findings suggested that it was a useful measure that related well to hospital characteristics rather than patient characteristics. More recently, it has been studied in cardiac surgery and has been suggested to be particularly useful in providing a mechanism for increased mortality in hospitals, and therefore a timeframe for improvement.[62–65] Other studies have also used a timeframe focus for quality measurement, including: ‘phase of care’ mortality analysis,[66] separating events into four different time periods around surgery and acute risk change (ARC), focusing on the immediate perioperative period (discussed later).[67]

Process measures

Medical and surgical treatments that are associated with improved outcomes should be chosen as process measures.[68] This may be difficult when there are multiple potential measures to choose from. In addition, evidence may be different for different subsets of surgical procedures within the same database. In the absence of direct evidence that a particular set of process measures are ideal for measuring quality of care, expert opinion is used. Ideally, the measures should be broadly applicable and have specific, easily documented and unequivocal contraindications to reduce the impact of gaming.[69] The STS records a variety of measures including antibiotic administration, perioperative beta-blockade, discharge beta-blockade, anti-platelet medications at discharge, lipid-lowering therapy at discharge, and IMA use. The latter five are then used in the STS composite score.[70] Process measures have the advantage that they do not require risk adjustment.

Further to traditionally used process measures, some additional measures have been suggested. Technical performance scores have been assessed for paediatric congenital surgery. They have been found to be associated with morbidity and mortality but have not yet been assessed as quality measures.[71] The use of postoperative glucose control as a process quality measure has been rejected due to lack of association with outcome.[72] Other processes associated with preventable death may also be used, for example antibiotic administration, transoesophageal echocardiography, and epiaortic ultrasound use.[57] Further processes such as optimal intraoperative management, teamwork, and effective handover may influence outcome and are potential new process measures.

Structural measures

Case volume is the most commonly assessed structural measure. The premise is that higher-volume centres will have greater experience and will therefore perform better, or that referral occurs to better providers in market-based healthcare. However, the volume–quality relationship is a controversial one. While some studies have confirmed a relationship between volume and quality, many have not.[73–84] Nursing ratios may be a useful structural measure and are associated with other quality measures.[55] Other programmes have also included participation in a quality database and availability of specific resources[85] as measures; however, to our knowledge, the validity of these has not been specifically evaluated.

Composite measures and single metrics

The STS uses a combination of structure, process, and outcome measures to provide a three-tier rating system.[16,70,86,87] A variety of measures were chosen, rescaled according to variability, and combined such that all domains contributed towards the overall score. In this way, they were able to provide a single, easy to understand metric that is capable of rating providers as above, meeting, or below the national average score. They were also able to improve on risk-adjusted mortality in discrimination of outlier units.

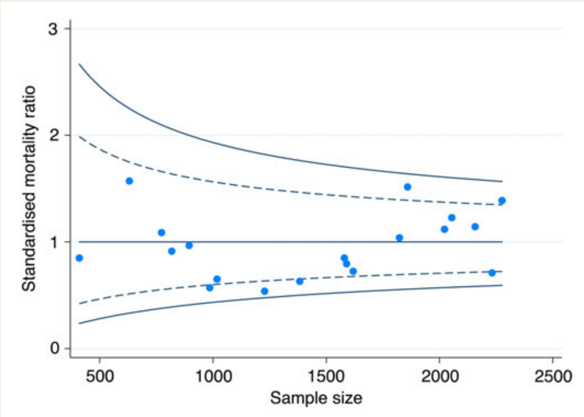

Figure 1 Funnel plot showing SMR for 19 different providers (hospitals in this case). *X*-axis position is determined by the number of cases. Six providers lie outside of 95% confidence limits, but none are outside of 99.8% limits.

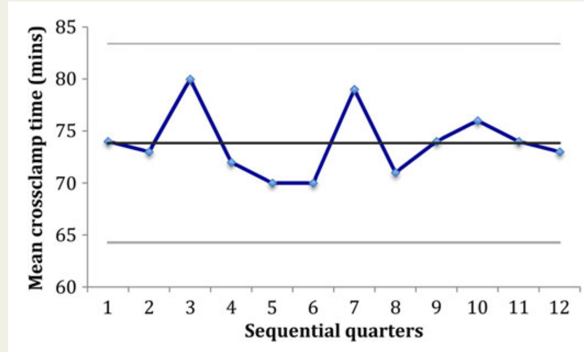

Figure 2 The Shewhart control chart showing mean aortic cross-clamp time for sequential quarters at one hospital. The middle line represents the overall mean. Outer lines represent three standard deviations. No point lies outside the line, and the process remains in control for the duration of the monitoring period.

The authors of this review formulated and investigated another measure, the ARC.[67] This metric assesses the change in probability of death from preoperative to postoperative care, using 'AusScore',[15] and recalibrated APACHE III scores. This score was able to discriminate between providers where risk-adjusted mortality could not and was strongly associated with morbidity measures and adverse intraoperative events.[88] This score may be useful in providing an early warning signal for poor-performing units and a perioperative focus point for quality improvement.

Statistical methods

Cross-sectional analysis

Funnel plots have traditionally been used to assess for publication bias in meta-analysis but are also used for institutional comparison. The ratio of observed to predicted events [commonly deaths represented via the standardized mortality ratio (SMR)] is plotted against either the total sample size or against the predicted number of events. Outcomes for units or providers may than be compared against the mean or a defined standard. Alert and alarm limits are then generated based on 95 and 99.8% confidence limits, respectively (*Figure 1*). Funnel plots avoid unrealistic ranking of units, as seen in league tables, and provide an effective visual method of assessing for outlier providers. They also allow easy examination of the case volume in relation to outcome.[89] However, to be useful, appropriate techniques for calculation of exact confidence intervals are required, particularly for rare binary events such as mortality in cardiothoracic surgical patients. Both over- and under-dispersion may occur, with the latter being most commonly observed when the prediction model is inaccurate. Risk modelling may be improved by using more complex methodologies that account for clustering of units or employ multilevel modelling techniques.

Longitudinal analysis

The measures described so far monitor a single or set of rates at a defined point in time retrospectively. While this may be useful, there is significant delay between the time of surgery, data acquisition, and analysis and quality improvement intervention. Techniques for continuous outcome monitoring may be helpful for local and national quality improvement programmes.

They may also be automated to minimize resource usage.

The control chart is a method of monitoring outcomes. It aims to differentiate between 'common-cause variation' (predictable variation as a result of phenomena within the system) and 'special-cause variation' (new, unpredictable variation as a result of a change within the system). Basic Shewhart control charts are plotted with three lines, the mean and two control limits either side, usually representing three standard deviations. *Figure 2* shows a basic control chart using mean cross-clamp times for a single hospital. The chart can be constructed by calculating the mean and standard deviation of all points. Points in this instance represent groups but could also represent single patients. Points outside the lines suggest special-cause variation.[90] While this technique may be used retrospectively, the control limits lines could also be used to prospectively monitor performance (with an appreciable false-positive rate). It has been suggested that this method could have identified poor-quality care in high-profile historical cases.[13,90,91]

A variety of other formulations of the control chart exist. Although they are more complex, they may be more suitable for use in the medical context. The CUSUM chart, for example, may be more sensitive to small changes in the mean than the Shewhart charts and therefore may more rapidly detect changes in performance.[92] In a CUSUM chart, sequential cases may be plotted on the *x*-axis. The change in *y*-axis value is then dependent on a binary outcome, for example survival or death. Where patient risk varies, risk adjustment is important to prevent false alarms and/or detect changes in performance.[93,94] Risk adjustment may be undertaken for a variety of control charts, including CUSUM charts[95] (*Figure 3A*); visual life-adjusted display (VLAD), also described as observed–expected plot (*Figure 3B*); and resetting sequential probability ratio test (RSPRT) charts (*Figure 3C*).[94,96–99] Risk adjustment alters the change in *y*-value in response to an outcome. The easiest of these to understand is the observed–expected plots. Here, a patient with a preoperative risk of death of 20% would result in a 0.8 unit increase in the y-axis if they died, and a 0.2 unit reduction in the y-axis value if they survived. Thus, a rising y-axis value may be the cause for alarm. The other risk-adjustment techniques operate by similar means, albeit with more complex statistical processing. The choice of CUSUM monitoring technique will vary according to the process being monitored,[98] for example, where a reduction in performance is as important as an increase in performance, the simultaneous risk-adjusted CUSUM chart may be most suitable (*Figure 3A*).

Exponentially weighted moving average (EWMA) charts are another technique of control chart monitoring and also include risk adjustment. Here, a weighted average of sequential observations is plotted using expected values to calculate control limits (*Figure 3D*). EWMA charts have the advantage that they are easier to interpret than the other forms of control chart.[35,100]

Control charts have some limitations that are important to recognize. It is possible to build up 'credit' such that sequential successes place a process well below the expected line. There may then be a reduction in performance, resulting in a rise in y-value, which is not detected due to the process remaining within control limits. As such, it is important to monitor trends as well as control limits. In addition, risk adjustment is required. Where the risk-adjustment model is inaccurate, out-of-control events may occur. The setting of control limits remains an important consideration, and it may be appropriate to set narrow control limits, allowing a higher false-positive rate.

Continuous assessment has been increasing in popularity. Control charts have the advantage that they can be monitored in real time. As a trend in outcomes is perceived (for example an increase in deaths), it may be rapidly recognized and investigated. A number of groups have reported the use of CUSUM, VLAD, and EWMA charts in local monitoring of outcomes.[35,39,47,101–103] Multiple different outcomes have been monitored including mortality, morbidity, length of stay, readmission, blood product use, and reoperation. In many cases, they have been able to detect an area of concern, investigate the cause, and rectify the problem. Visually, they may be interpreted relatively easily so that little statistical knowledge is required to participate in the quality improvement process. As such, continuous assessment represents a useful technique to maintain and improve clinical services.

Investigation of outlier units

Outlier institutions or practitioners may be identified. The response should use the pyramid of investigation model.[104] At the base of the pyramid is the most common reason for outlier identification, the data quality. Investigators then progress from the base to the top of the pyramid as follows, only moving on when each level has been explored and explained:

(1) Are there data quality issues that explain the variation?
(2) Are there case-mix factors that have resulted in false identification of an outlier?
(3) Are there structure or resource factors that may have resulted in higher mortality or morbidity?
(4) Are there processes of care that may result in higher mortality or morbidity?
(5) Is there a specific carer or team responsible for higher mortality or morbidity?

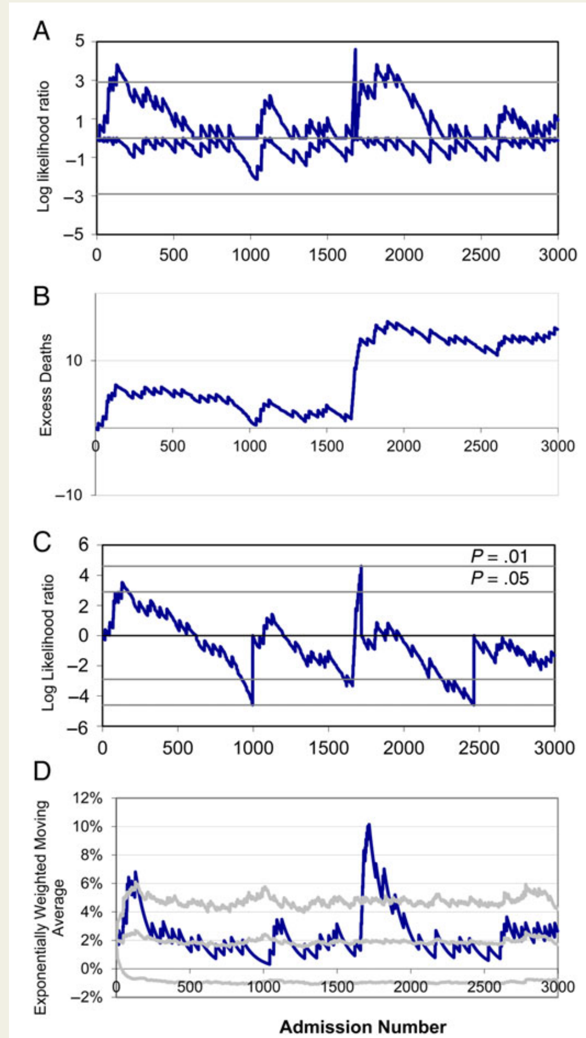

Figure 3 Risk-adjusted control charts showing the same data for 3000 sequential cardiothoracic procedures. Note: Data are simulated for demonstration purposes, and such large fluctuations in mortality rate are rarely seen. (*A*) Risk-adjusted CUSUM chart adjusting for predicted risk of death. Crossing of the upper boundary represents a doubling of the risk of death. This happens twice during the time period. Crossing the lower line represents a halving of the risk of death (with a 1% chance finding rate). (*B*) Observed–expected plot for deaths using AusScore score to estimate risk of death. For each death, there is a change in y-axis value of 1-(risk of death), and for each survival, there is a change in y-axis value of 0-(risk of death). Thresholds for alarm and investigation may be set individual institutions. There are two periods of increased 'excess deaths' during the period. (*C*) RSPRT adjusting for predicted risk of death. Crossing the uppermost boundary represents a doubling of the odds of death ($P = 0.01$), and crossing the lowermost boundary represents a halving of the odds of death. (*D*) EWMA chart showing EWMA of observed mortality (blue line) vs. EWMA AusScore-predicted risk of death (grey lines—upper and lower lines represent 95% confidence limits, and the middle line represents mean). Note that mortality exceeds predicted mortality twice during the period.

Where structures or processes are identified that may be resulting in worse outcomes, investigators consider how and whether they may be modified. Ongoing monitoring during and after the investigation should identify whether the outlier has returned to normal limits of variation. The responsibility for investigation of the outlier processes and teams varies according to health system. The outlier may be initially investigated internally, for example by case-note review for avoidable mortality. Where no improvement is made, the investigation does not identify areas of improvement, or where there are significant and multi-sourced concerns, a more comprehensive external review may occur.

Limitations

Limitations should be addressed. It is possible that some papers were not identified by the literature search, and some methodologies have been missed. However, this should have been avoided by the use of a thorough literature search plan. It was not possible to complete a quantitative systematic review to comply with published standards, such as the Meta-Analysis of Observation Studies in Epidemiology (MOOSE) guidelines.[105] This was in part due to the lack of common outcomes in assessment of quality measures. In addition, few papers empirically compared different quality measures, except in comparison with mortality. There is no gold standard of quality measurement; therefore, we had no reference to measure against. This meant that making quantitative judgements on utility was rarely possible.

Conclusion

Cardiac surgery is arguably the most developed specialty in terms of measuring and improving quality of care. In this review, we have described and evaluated the methods of measuring quality of care in cardiac surgery. Despite the focus on cardiac surgery, many of these concepts can be applied to other types of surgery or medical care. The structure, process, and outcome method proposed by Donebedian so many years ago continues to provide a useful framework to measure quality of care (and it may also aid in efforts to improve care).

Of the methods that remain unchanged, such as mortality, new and improved ways of combining measures and adjusting for risk have been developed and have proliferated. More recently, innovative ways of assessing these outcomes have been instituted, allowing contemporaneous evaluation of quality. In the context of a quality culture, and wider limits of certainty, tools such as control charts provide a real opportunity for rapid quality assessment and improvement. Other novel methods, such as failure to rescue and ARC, may also provide valuable insights into timeframes and contexts for improving quality.

We should use a multi-faceted approach to measuring quality. We should monitor morbidity and mortality measures contemporaneously using control chart techniques, with defined and robust processes in place for investigating out-of-control units. Simultaneously, we should monitor long-term outcomes at defined intervals, including patient-centred outcomes, such as quality of life and disability-free survival. Finally, we should continue to investigate new and innovative ways to measure and identify variation in care quality and pinpoint the processes that may have contributed to this variation. In this way, we may not only be able to maintain and improve the effectiveness of patient care but also the ability of individuals and populations to access that care.

Supplementary material

Supplementary material is available at *European Heart Journal – Quality of Care and Clinical Outcomes* online.

Funding

T.C. was funded by an Alfred-Monash Intensive Care Research Fellowship.

Conflict of interest: none declared.

References

1. Institute of Medicine (US) Committee to Design a Strategy for Quality Review and Assurance in Medicare. Medicare: A Strategy for Quality Assurance: Volume 1. Washington, DC: National Academies Press (US), 1990.
2. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: Institute of Medicine (IOM), National Academy Press, 2000.
3. Campbell SM, Roland MO, Buetow SA. Defining quality of care. *Soc Sci Med* 2000;**51**:1611–1625.
4. Pronovost PJ, Nolan T, Zeger S, Miller M, Rubin H. How can clinicians measure safety and quality in acute care? *Lancet* 2004;**363**:1061–1067.
5. Donabedian A. The quality of care. How can it be assessed? *JAMA* 1988;**260**:1743–1748.
6. Nelson EC, Mohr JJ, Batalden PB, Plume SK. Improving health care, Part 1: the clinical value compass. *Jt Comm J Qual Improv* 1996;**22**:243–258.
7. Brook RH, McGlynn EA, Cleary PD. Quality of health care. Part 2: measuring quality of care. *N Engl J Med* 1996;**335**:966–970.
8. Shroyer ALW, McDonald GO, Wagner BD, Johnson R, Schade LM, Bell MR, Grover FL. Improving Quality of Care in Cardiac Surgery: evaluating risk factors, processes of care, structures of care, and outcomes. *Semin Cardiothorac Vasc Anesth* 2008;**12**:140–152.
9. Society of Thoracic Surgeons. Quality measures. http://www.sts.org/quality-research-patient-safety/quality/quality-performance-measures (24 March 2015).
10. National Institute for Cardiovascular Outcomes Research. National Adult Cardiac Surgery Audit. Annual report, London, 2010–2011.
11. Kinsman R, Systems DC. The Society for Cardiothoracic Surgery in Great Britain & Ireland Sixth National Adult Cardiac Surgical Database Report 2008. 2009;1–516.
12. Tran L, Chand V, Newcombe B, Billah B, Shardey GC, Reid C. Cardiac surgery in Victorian public hospitals, 2010–11. Department of Health, 2012;1–28.
13. Arrowsmith JE, Powell SJ, Nashef SAM. Local clinical quality monitoring for detection of excess operative deaths. *Anaesthesia* 2006;**61**:423–426.
14. Nashef SAM, Roques F, Sharples LD, Nilsson J, Smith C, Goldstone AR, Lockowandt U. EuroSCORE II. *Eur J Cardiothorac Surg* 2012;**41**:734–744; discussion 744–745.
15. Billah B, Reid CM, Shardey GC, Smith JA. A preoperative risk prediction model for 30-day mortality following cardiac surgery in an Australian cohort. *Eur J Cardiothorac Surg* 2010;**37**:1086–1092.
16. Shahian DM, Grover FL, Anderson RP, Edwards FH. Quality measurement in adult cardiac surgery: introduction. *Ann Thorac Surg* 2007;**83**:S1–S2.
17. Glance LG, Dick A, Osler TM, Li Y, Mukamel DB. Impact of changing the statistical methodology on hospital and surgeon ranking: the case of the New York State cardiac surgery report card. *Med Care* 2006;**44**:311–319.
18. Hannan EL, Racz MJ, Jollis JG, Peterson ED. Using medicare claims data to assess provider quality for CABG surgery: does it work well enough? *Health Serv Res* 1997;**31**:659–678.
19. Osswald BR, Gegouskov V, Badowski-Zyla D, Tochtermann U, Thomas G, Hagl S, Blackstone EH. Overestimation of aortic valve replacement risk by EuroSCORE: implications for percutaneous valve replacement. *Eur Heart J* 2008;**30**:74–80.
20. Siregar S, Groenwold RHH, Versteegh MIM, Noyez L, ter Burg WJPP, Bots ML, van der Graaf Y, van Herwerden LA. Gaming in risk-adjusted mortality rates: effect of misclassification of risk factors in the benchmarking of cardiac surgery risk-adjusted mortality rates. *J Thorac Cardiovasc Surg* 2013;**145**:781–789.
21. Arenz C, Asfour B, Hraska V, Photiadis J, Haun C, Schindler E, Sinzobahamvya N. Congenital heart surgery: surgical performance according to the Aristotle complexity score. *Eur J Cardiothorac Surg* 2011;**39**:e33–e37.

22. Lacour-Gayet F, Jacobs JP, Clarke DR, Maruszewski B, Jacobs ML, O'Brien SM, Mavroudis C. Evaluation of the quality of care in congenital heart surgery: contribution of the Aristotle complexity score. *Adv Pediatr* 2007;**54**:67–83.
23. Lacour-Gayet F, Clarke DR. The Aristotle method: a new concept to evaluate quality of care based on complexity. *Curr Opin Pediatr* 2005;**17**:412–417.
24. Dimick JB, Welch HG, Birkmeyer JD. Surgical mortality as an indicator of hospital quality: the problem with small sample size. *JAMA* 2004;**292**:847–851.
25. Welke KF, Karamlou T, Ungerleider RM, Diggs BS. Mortality rate is not a valid indicator of quality differences between pediatric cardiac surgical programs. *Ann Thorac Surg* 2010;**89**:139–144; discussion 145–146.
26. Jacobs JP, Mavroudis C, Jacobs ML, Maruszewski B, Tchervenkov CI, Lacour-Gayet FG, Clarke DR, Yeh T, Walters HL, Kurosawa H, Stellin G, Ebels T, Elliott MJ, STS Congenital Database Taskforce, Joint EACTS-STS Congenital Database Committee. What is operative mortality? Defining death in a surgical registry database: a report of the STS Congenital Database Taskforce and the Joint EACTS-STS Congenital Database Committee. *Ann Thorac Surg* 2006;**81**:1937–1941.
27. Johnson ML, Gordon HS, Petersen NJ, Wray NP, Shroyer AL, Grover FL, Geraci JM. Effect of definition of mortality on hospital profiles. *Med Care* 2002;**40**:7–16.
28. Noyez L. Organized follow-up is essential for the evaluation of cardiac surgery. *J Cardiovasc Surg (Torino)* 2013;**54**:545–552.
29. Glance LG, Kellermann AL, Hannan EL, Fleisher LA, Eaton MP, Dutton RP, Lustik SJ, Li Y, Dick AW. The impact of anesthesiologists on coronary artery bypass graft surgery outcomes. *Anesth Analg* 2015;**120**:526–533.
30. Papachristofi O, Sharples LD, Mackay JH, Nashef SAM, Fletcher SN, Klein AA, Association of Cardiothoracic Anaesthetists (ACTA), and Association of Cardiothoracic Anaesthetists (ACTA). The contribution of the anaesthetist to risk-adjusted mortality after cardiac surgery. *Anaesthesia* 2016;**71**:138–146.
31. Guru V, Tu JV, Etchells E, Anderson GM, Naylor CD, Novick RJ, Feindel CM, Rubens FD, Teoh K, Mathur A, Hamilton A, Bonneau D, Cutrara C, Austin PC, Fremes SE. Relationship between preventability of death after coronary artery bypass graft surgery and all-cause risk-adjusted mortality rates. *Circulation* 2008;**117**:2969–2976.
32. Khuri SF, Henderson WG, DePalma RG, Mosca C, Healey NA, Kumbhani DJ. Determinants of long-term survival after major surgery and the adverse effect of postoperative complications. *Trans Meet Am Surg Assoc Am Surg Assoc* 2005;123:32–48.
33. Pappalardo F, Franco A, Landoni G, Cardano P, Zangrillo A, Alfieri O. Long-term outcome and quality of life of patients requiring prolonged mechanical ventilation after cardiac surgery. *Eur J Cardiothorac Surg* 2004;**25**:548–552.
34. Jidéus L, Liss A, Ståhle E. Patients with sternal wound infection after cardiac surgery do not improve their quality of life. *Scand Cardiovasc J* 2009;**43**:194–200.
35. Smith IR, Gardner MA, Garlick B, Brighouse RD, Cameron J, Lavercombe PS, Mengersen K, Foster KA, Rivers JT. Performance monitoring in cardiac surgery: application of statistical process control to a single-site database. *Heart Lung Circ* 2013;**22**:634–641.
36. Silber JH, Rosenbaum PR. A spurious correlation between hospital mortality and complication rates: the importance of severity adjustment. *Med Care* 1997;**35**(10 Suppl):OS77–OS92.
37. Moulton MJ, Creswell LL, Mackey ME, Cox JL, Rosenbloom M. Reexploration for bleeding is a risk factor for adverse outcomes after cardiac operations. *J Thorac Cardiovasc Surg* 1996;**111**:1037–1046.
38. Karthik S, Grayson AD, McCarron EE, Pullan DM, Desmond MJ. Reexploration for bleeding after coronary artery bypass surgery: risk factors, outcomes, and the effect of time delay. *Ann Thorac Surg* 2004;**78**:527–534; discussion 534.
39. Wolfe R, Bolsin S, Colson M, Stow P. Monitoring the rate of re-exploration for excessive bleeding after cardiac surgery in adults. *Qual Saf Health Care* 2007;**16**:192–196.
40. Paone G, Brewer R, Likosky DS, Theurer PF, Bell GF, Cogan CM, Prager RL, Membership of the Michigan Society of Thoracic and Cardiovascular Surgeons. Transfusion rate as a quality metric: is blood conservation a learnable skill? *Ann Thorac Surg* 2013;**96**:1279–1286.
41. Beaty CA, Haggerty KA, Moser MG, George TJ, Robinson CW, Arnaoutakis GJ, Whitman GJ. Disclosure of physician-specific behavior improves blood utilization protocol adherence in cardiac surgery. *Ann Thorac Surg* 2013;**96**:2168–2174.
42. Murphy GJ, Pike K, Rogers CA, Wordsworth S, Stokes EA, Angelini GD, Reeves BC, TITRe2 Investigators. Liberal or restrictive transfusion after cardiac surgery. *N Engl J Med* 2015;**372**:997–1008.
43. Carson SS. Outcomes of prolonged mechanical ventilation. *Curr Opin Crit Care* 2006;**12**:405–411.
44. Toumpoulis IK, Anagnostopoulos CE, Derose JJ, Swistel DG. The impact of deep sternal wound infection on long-term survival after coronary artery bypass grafting. *Chest* 2005;**127**:464–471.
45. Landoni G, Zangrillo A, Franco A, Aletti G, Roberti A, Calabrò MG, Slaviero G, Bignami E, Marino G. Long-term outcome of patients who require renal replacement therapy after cardiac surgery. *Eur J Anaesthesiol* 2006;**23**:17–22.
46. Roach GW, Kanchuger M, Mangano CM, Newman M, Nussmeier N, Wolman R, Aggarwal A, Marschall K, Graham SH, Ley C. Adverse cerebral outcomes after coronary bypass surgery. Multicenter Study of Perioperative Ischemia Research Group and the Ischemia Research and Education Foundation Investigators. *N Engl J Med* 1996;**335**:1857–1863.
47. Borracci RA, Rubio M, Ahuad Guerrero RA, Barrero C, Mauro V, Fairman E. Continuous monitoring of cardiac surgery quality indicators. *Arch Cardiol Mex* 2007;**77**:275–283.
48. Shroyer ALW, Coombs LP, Peterson ED, Eiken MC, DeLong ER, Chen A, Ferguson TB Jr, Grover FL, Edwards FH, Baskett RJF, Chen JC. The Society of Thoracic Surgeons: 30-day operative mortality and morbidity risk models. *Ann Thorac Surg* 2003;**75**:1856–1864; discussion 1864–1865.
49. Peterson ED, Coombs LP, Ferguson TB, Shroyer AL, DeLong ER, Grover FL, Edwards FH. Hospital variability in length of stay after coronary artery bypass surgery: results from the Society of Thoracic Surgeon's National Cardiac Database. *Ann Thorac Surg* 2002;**74**:464–473.
50. Liu M, Druschel CM, Hannan EL. Risk-adjusted prolonged length of stay as an alternative outcome measure for pediatric congenital cardiac surgery. *Ann Thorac Surg* 2014;**97**:2154–2159.
51. Tsai TC, Joynt KE, Orav EJ, Gawande AA, Jha AK. Variation in surgical-readmission rates and quality of hospital care. *N Engl J Med* 2013;**369**:1134–1142.
52. Shih T, Dimick JB. Reliability of readmission rates as a hospital quality measure in cardiac surgery. *Ann Thorac Surg* 2014;**97**:1214–1218.
53. Davies SM, Saynina O, McDonald KM, Baker LC. Limitations of using same-hospital readmission metrics. *Int J Qual Health Care* 2013;**25**:633–639.
54. Khandelwal N, Dale CR, Benkeser DC, Joffe AM, Yanez ND, Treggiari MM. Variation in tracheal reintubations among patients undergoing cardiac surgery across Washington State Hospitals. *J Cardiothorac Vasc Anesth* 2001;**29**:551–559.
55. Diya L, Van den Heede K, Sermeus W, Lesaffre E. The relationship between in-hospital mortality, readmission into the intensive care nursing unit and/or operating theatre and nurse staffing levels. *J Adv Nurs* 2012;**68**:1073–1081.
56. Auerbach AD, Hilton JF, Maselli J, Pekow PS, Rothberg MB, Lindenauer PK. Case volume, quality of care, and care efficiency in coronary artery bypass surgery. *Arch Intern Med* 2010;**170**:1202–1208.
57. Guru V, Anderson GM, Fremes SE, O'Connor GT, Grover FL, Tu JV, Canadian CABG Surgery Quality Indicator Consensus Panel. The identification and development of Canadian coronary artery bypass graft surgery quality indicators. *J Thorac Cardiovasc Surg* 2005;**130**:1257–1264.
58. Noyez L, Kievit P, Verkroost A, van Wetten H, Verhagen AFTM, van Swieten H. Evaluation of quality in adult cardiac surgery: let us speak the same language. *Neth Heart J* 2010;**18**:365–369.
59. Myles PS. Meaningful outcome measures in cardiac surgery. *J Extra Corpor Technol* 2014;**46**:23–27.
60. Gorman Koch C, Khandwala F, Blackstone EH. Health-related quality of life after cardiac surgery. *Semin Cardiothorac Vasc Anesth* 2008;**12**:203–217.
61. Silber JH, Williams SV, Krakauer H, Schwartz JS. Hospital and patient characteristics associated with death after surgery. A study of adverse occurrence and failure to rescue. *Med Care* 1992;**30**:615–629.
62. Pasquali SK, He X, Jacobs JP, Jacobs ML, O'Brien SM, Gaynor JW. Evaluation of failure to rescue as a quality metric in pediatric heart surgery: an analysis of the STS congenital heart surgery database. *Ann Thorac Surg* 2012;**94**:573–580.
63. Reddy HG, Shih T, Englesbe MJ, Shannon FL, Theurer PF, Herbert MA, Paone G, Bell GF, Prager RL. Failure to rescue. *Ann Thorac Surg* 2013;**95**:1976–1981.
64. Ahmed EO, Butler R, Novick RJ. Failure-to-rescue rate as a measure of quality of care in a cardiac surgery recovery unit: a five-year study. *Ann Thorac Surg* 2014;**97**:147–152.
65. Westaby S, De Silva R, Petrou M, Bond S, Taggart D. Surgeon-specific mortality data disguise wider failings in delivery of safe surgical services. *Eur J Cardiothorac Surg* 2015;**47**:341–345.
66. Shannon FL, Fazzalari FL, Theurer PF, Bell GF, Sutcliffe KM, Prager RL. A method to evaluate cardiac surgery mortality: phase of care mortality analysis. *Ann Thorac Surg* 2012;**93**:36–43.
67. Coulson TG, Bailey M, Reid CM, Tran L, Mullany DV, Smith JA, Pilcher D. Acute risk change for cardiothoracic admissions to intensive care: A new measure of quality in cardiac surgery. *J Thorac Cardiovasc Surg* 2014;**148**:3076–3081.
68. Shroyer AL, London MJ, VillaNueva CB, Sethi GK, Marshall G, Moritz TE, Henderson WG, McCarthy MJ Jr, Grover FL, Hammermeister KE. The processes, structures, and outcomes of care in cardiac surgery study protocol. *Med Care* 1995;**33**(10 Suppl):OS17–OS25.
69. National Quality Framework. National Voluntary Consensus Standards for Cardiac Surgery.

70. Shahian DM, Edwards FH, Ferraris VA, Haan CK, Rich JB, Normand SLT, DeLong ER, O'Brien SM, Shewan CM, Dokholyan RS, Peterson ED. Quality measurement in adult cardiac surgery: Part 1—conceptual framework and measure selection. *Ann Thorac Surg* 2007;**83**:S3–S12.
71. Karamichalis JM, Colan SD, Nathan M, Pigula FA, Baird C, Marx G, Emani SM, Geva T, Fynn-Thompson FE, Liu H, Mayer JE, del Nido PJ. Technical performance scores in congenital cardiac operations: a quality assessment initiative. *Ann Thorac Surg* 2012;**94**:1317–1323.
72. LaPar DJ, Isbell JM, Kern JA, Ailawadi G, Kron IL. Surgical Care Improvement Project measure for postoperative glucose control should not be used as a measure of quality after cardiac surgery. *J Thorac Cardiovasc Surg* 2014;**147**:1041–1048.
73. Sakata R, Kuwano H, Yokomise H. Hospital volume and outcomes of cardiothoracic surgery in Japan: 2005–2009 national survey. *Gen Thorac Cardiovasc Surg* 2012; **60**:625–638.
74. Spiegelhalter DJ. Mortality and volume of cases in paediatric cardiac surgery: retrospective study based on routinely collected data. *BMJ* 2002;**324**:261–263.
75. Kurlansky PA, Argenziano M, Dunton R, Lancey R, Nast E, Stewart A, Williams T, Zapolanski A, Chang H, Tingley J, Smith CR. Quality, not volume, determines outcome of coronary artery bypass surgery in a university-based community hospital network. *J Thorac Cardiovasc Surg* 2012;**143**:287–293.
76. Carey JS, Robertson JM, Misbach GA, Fisher AL. Relationship of hospital volume to outcome in cardiac surgery programs in California. *Am Surg* 2003;**69**:63–68.
77. Vinocur JM, Menk JS, Connett J, Moller JH, Kochilas LK. Surgical volume and center effects on early mortality after pediatric cardiac surgery: 25-year North American experience from a multi-institutional registry. *Pediatr Cardiol* 2013;**34**: 1226–1236.
78. Kalfa D, Chai P, Bacha E. Surgical volume-to-outcome relationship and monitoring of technical performance in pediatric cardiac surgery. *Pediatr Cardiol* 2014;**35**: 899–905.
79. Bazzani L, JP M. Case volume and mortality in pediatric cardiac surgery patients in California, 1998–2003. *Circulation* 2007;**115**:2652–2659.
80. Reames NB, Ghaferi AA, Birkmeyer JD, Dimick JB. Hospital volume and operative mortality in the modern era. *Ann Surg* 2014;**260**:244–251.
81. Chan T, Kim J, Minich LL, Pinto NM, Waitzman NJ. Surgical volume, hospital quality, and hospitalization cost in congenital heart surgery in the United States. *Pediatr Cardiol* 2015;**36**:205–213.
82. Patel HJ, Herbert MA, Drake DH, Hanson EC, Theurer PF, Bell GF, Prager RL. Aortic valve replacement: using a statewide cardiac surgical database identifies a procedural volume hinge point. *Ann Thorac Surg* 2013;**96**:1560–1565; discussion1565–6.
83. Gammie JS, O'Brien SM, Griffith BP, Ferguson TB, Peterson ED. Influence of hospital procedural volume on care process and mortality for patients undergoing elective surgery for mitral regurgitation. *Circulation* 2007;**115**:881–887.
84. Ch'ng SL, Cochrane AD, Wolfe R, Reid C, Smith CI, Smith JA. Procedure-specific cardiac surgeon volume associated with patient outcome following valve surgery, but not isolated CABG surgery. *Heart Lung Circ* 2015;**24**:583–589.
85. Jacobs JP, Jacobs ML, Austin EH, Mavroudis C, Pasquali SK, Lacour-Gayet FG, Tchervenkov CI, Walters H, Bacha EA, Nido PJD, Fraser CD, Gaynor JW, Hirsch JC, Morales DLS, Pourmoghadam KK, Tweddell JS, Prager RL, Mayer JE. Quality measures for congenital and pediatric cardiac surgery. *World J Pediatr Congenit Heart Surg* 2012;**3**:32–47.
86. O'Brien SM, Shahian DM, DeLong ER, Normand S-LT, Edwards FH, Ferraris VA, Haan CK, Rich JB, Shewan CM, Dokholyan RS, Anderson RP, Peterson ED. Quality measurement in adult cardiac surgery: part 2—statistical considerations in composite measure scoring and provider rating. *Ann Thorac Surg* 2007;**83**(4 Suppl):S13–S26.
87. Shahian DM, Xia H, Jacobs JP, Rankin JS, Welke KF, Edwards FH, Filardo G, Fazzalari FL, Furnary A, Kurlansky PA, Brennan JM, Badhwar V, O'Brien SM. The STS AVR+CABG composite score: a report of the STS Quality Measurement Task Force. *Ann Thorac Surg* 2014;**97**:1604–1609.
88. Coulson TG, Bailey M, Reid CM, Tran L, Mullany DV, Parker J *et al.* Acute risk change (ARC) identifies outlier institutions in perioperative cardiac surgical care when the standardized mortality ratio cannot. *BJA* [in press].
89. Spiegelhalter DJ. Funnel plots for comparing institutional performance. *Statist Med* 2005;**24**:1185–1202.
90. Mohammed MA, Cheng KK, Rouse A, Marshall T. Bristol, shipman, and clinical governance: Shewhart's forgotten lessons. *Lancet* 2001;**357**:463–467.
91. Mohammed MA, Worthington P, Woodall WH. Plotting basic control charts: tutorial notes for healthcare practitioners. *Qual Saf Health Care* 2008;**17**:137–145.
92. Huesch MD. False alarms in surgical quality control. *J Healthc Qual* 2008;**30**:4–10.
93. Novick RJ, Fox SA, Stitt LW, Forbes TL, Steiner S. Direct comparison of risk-adjusted and non–risk-adjusted CUSUM analyses of coronary artery bypass surgery outcomes. *J Thorac Cardiovasc Surg* 2006;**132**:386–391.
94. Grigg O, Farewell V. An overview of risk-adjusted charts. *J R Stat Soc Ser A Stat Soc* 2004;**167**:523–539.
95. Steiner SH, Cook RJ, Farewell VT, Treasure T. Monitoring surgical performance using risk-adjusted cumulative sum charts. *Biostatistics* 2000;**1**:441–452.
96. Lovegrove J, Valencia O, Treasure T, Sherlaw-Johnson C, Gallivan S. Monitoring the results of cardiac surgery by variable life-adjusted display. *Lancet* 1997;**350**: 1128–1130.
97. Cook DA, Coory M, Webster RA. Exponentially weighted moving average charts to compare observed and expected values for monitoring risk-adjusted hospital indicators. *BMJ Qual Saf* 2011;**20**:469–474.
98. Grigg OA, Farewell VT, Spiegelhalter DJ. Use of risk-adjusted CUSUM and RSPRT charts for monitoring in medical contexts. *Stat Methods Med Res* 2003;**12**: 147–170.
99. Cook DA, Duke G, Hart GK, Pilcher D, Mullany D. Review of the application of risk-adjusted charts to analyse mortality outcomes in critical care. *Crit Care Resusc* 2008;**10**:239–251.
100. Pilcher DV, Hoffman T, Thomas C, Ernest D, Hart GK. Risk-adjusted continuous outcome monitoring with an EWMA chart: could it have detected excess mortality among intensive care patients at Bundaberg Base Hospital? *Crit Care Resusc* 2010;**12**:36–41.
101. Kang N, Tsang VT, Gallivan S, Sherlaw-Johnson C, Cole TJ, Elliott MJ, de Leval MR. Quality assurance in congenital heart surgery. *Eur J Cardiothorac Surg* 2006;**29**: **693**–7;
discussion 697–698.
102. Pagel C, Utley M, Crowe S, Witter T, Anderson D, Samson R, McLean A, Banks V, Tsang V, Brown K. Real time monitoring of risk-adjusted paediatric cardiac surgery outcomes using variable life-adjusted display: implementation in three UK centres. *Heart* 2013;**99**:1445–1450.
103. Smith IR, Gardner MA, Garlick B, Brighouse RD, Cameron J, Lavercombe PS, Mengersen K, Foster KA, Rivers JT. Performance monitoring in cardiac surgery: application of statistical process control to a single-site database. *Heart Lung Circul* 2013;**22**:634–641.
104. Mohammed MA, Rathbone A, Myers P, Patel D, Onions H, Stevens A. An investigation into general practitioners associated with high patient mortality flagged up through the Shipman inquiry: retrospective analysis of routine data. *BMJ* 2004; **328**:1474–1477.
105. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000;**283**:2008–2012.