

Cumulative funnel plots for the early detection of interoperator variation: retrospective database analysis of observed versus predicted results of percutaneous coronary intervention

Babu Kunadian, research fellow, Joel Dunning, specialist registrar in cardiothoracic surgery, Anthony P Roberts, clinical effectiveness specialist adviser, Robert Morley, clinical audit lead, Darragh Twomey, clinical teaching fellow, James A Hall, consultant cardiologist, Andrew G C Sutton, consultant cardiologist, Robert A Wright, consultant cardiologist, Douglas F Muir, consultant cardiologist, Mark A de Belder, consultant cardiologist

Department of Cardiology, James Cook University Hospital, Middlesbrough TS4 3BW

Correspondence to: M A de Belder mark.debelder@stees.nhs.uk

doi:10.1136/bmj.39512.529120.BE

ABSTRACT

Objective To use funnel plots and cumulative funnel plots to compare in-hospital outcome data for operators undertaking percutaneous coronary interventions with predicted results derived from a validated risk score to allow for early detection of variation in performance.

Design Analysis of prospectively collected data.

Setting Tertiary centre NHS hospital in the north east of England.

Participants Five cardiologists carrying out percutaneous coronary interventions between January 2003 and December 2006.

Main outcome measures In-hospital major adverse cardiovascular and cerebrovascular events (in-hospital death, Q wave myocardial infarction, emergency coronary artery bypass graft surgery, and cerebrovascular accident) analysed against the logistic north west quality improvement programme predicted risk, for each operator. Results are displayed as funnel plots summarising overall performance for each operator and cumulative funnel plots for an individual operator's performance on a case series basis.

Results The funnel plots for 5198 patients undergoing percutaneous coronary interventions showed an average observed rate for major adverse cardiovascular and cerebrovascular events of 1.96% overall. This was below the predicted risk of 2.06% by the logistic north west quality improvement programme risk score. Rates of in-hospital major adverse cardiovascular and cerebrovascular events for all operators were within the 3 σ upper control limit of 2.75% and 2 σ upper warning limit of 2.49%.

Conclusion The overall in-hospital major adverse cardiovascular and cerebrovascular events rates were under the predicted event rate. In-hospital rates after percutaneous coronary intervention procedure can be monitored successfully using funnel and cumulative funnel plots with 3 σ control limits to display and publish

each operator's outcomes. The upper warning limit (2 σ control limit) could be used for internal monitoring. The main advantage of these charts is their transparency, as they show observed and predicted events separately. By this approach individual operators can monitor their own performance, using the predicted risk for their patients but in a way that is compatible with benchmarking to colleagues, encapsulated by the funnel plot. This methodology is applicable regardless of variations in individual operator case volume and case mix.

INTRODUCTION

Demand is growing for specialties to publish outcome data on their operative procedures. In the United Kingdom, collection of comparative data at all levels of health care has been prompted by incidents of failure of professional self regulation, particularly the Bristol Royal Infirmary¹ and Harold Shipman cases.² In 2004, under the Freedom of Information Act, the *Guardian* newspaper published mortality data from 244 named cardiac surgeons in the UK.³ The data were non-risk adjusted and from hospital episode statistics, which contain significant errors. Subsequently the Society for Cardiothoracic Surgery in Great Britain and Ireland produced its own outcome data in a risk adjusted fashion with the approval of its members.⁴ Such an approach is necessary if inappropriate conclusions and risk adverse behaviour are to be avoided.

Advancement in catheter technology and improved operator techniques have resulted in the continuous growth of percutaneous coronary interventions.^{5,6} Outcome analysis and quality control are important in interventional cardiology.⁷⁻¹¹ Benchmarking raw outcome data is difficult and is complicated by variation in case mix, referral patterns, procedural techniques, and adjunctive therapy. Need is pressing for workable risk models for patients undergoing percutaneous coronary interventions. The north west

quality improvement programme has provided a prediction model for major adverse cardiac events after percutaneous coronary intervention that has been subject to both internal and external validation.^{12,13}

The New York State department of health collects and reports the number of interventions and patient mortality statistics for all cardiologists who carry out percutaneous coronary interventions.¹⁴ This allows the public to make better informed decisions when choosing a physician or hospital, and provides physicians and healthcare organisations with comparative data that will serve to improve the quality of health care. Although comparative performance of UK cardiac surgeons has been published in the public arena,¹⁵ operator specific data for percutaneous coronary intervention are not yet available.

The task force of the American College of Cardiology and American Heart Association has recently published recommendations for standards to assess operator proficiency and institutional programme quality.¹⁶ We address these recommendations and provide a method to implement them in a UK setting. We used the north west quality improvement programme risk model and then used cumulative funnels and funnel plots to display the observed major adverse cardiovascular and cerebrovascular events against the predicted rate of these events. Comparative performance of UK cardiac surgeons has been disseminated using these plots.¹⁷ In cardiology, funnel plots have been used to interpret the dataset of the myocardial infarction national audit project (a UK cardiology dataset that provides specific performance tables).¹⁸ We aimed to show that operator specific outcomes after percutaneous coronary intervention can be monitored successfully using funnel plots and cumulative funnel plots.

METHODS

A detailed database of clinical, procedural, and angiographic variables has been maintained on all patients undergoing percutaneous coronary intervention in our unit since 1994. The dataset is based on the British Cardiovascular Intervention Society national dataset,¹⁹ with several additional data elements. The prospective acquisition of data is accomplished by immediate input from the operators after each procedure, with additional information and outcome data being collected and all the data being validated by a dedicated team of trained nurses. This data collection is part of a national quality assessment and quality improvement programme coordinated by the British Cardiovascular Intervention Society. We analysed data from 5198 consecutive percutaneous revascularisation procedures carried out between 1 January 2003 and 31 December 2006.

Study outcomes

Our outcome of interest was major adverse cardiovascular and cerebrovascular events, defined as one or more of: in-hospital death, Q wave myocardial infarction, emergency coronary artery bypass graft

surgery, and cerebrovascular accident. We defined Q wave myocardial infarction as a new pathological Q wave, with creatine kinase levels more than twice the laboratory upper limit of normal with increased creatine kinase MB fraction or troponin T. We did not include non-Q wave myocardial infarction because of difficulties in knowing whether minor changes in enzyme levels after procedures in the setting of acute coronary syndromes reflect the tail end of a preprocedural event or the procedure itself—this differentiation requires knowledge of preprocedural changes to enzyme levels but is not required in the national dataset. We considered Q wave myocardial infarction occurring in the context of angioplasty therapy for acute ST elevation myocardial infarction to be an outcome of the original coronary event and not a complication of percutaneous coronary intervention.

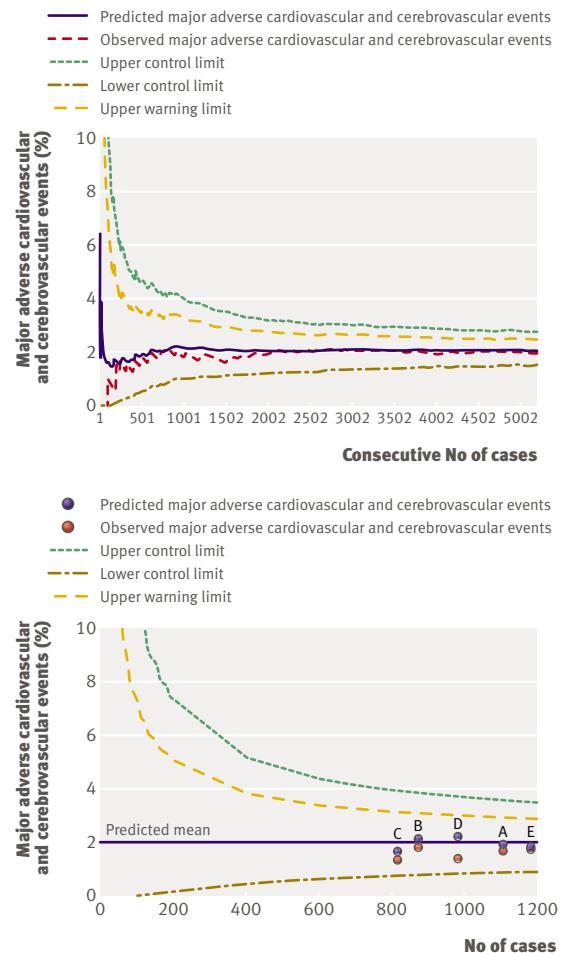


Fig 1 | (Top) Cumulative funnel plot displaying observed major adverse cardiovascular and cerebrovascular events for whole unit compared with mean predicted events (logistic north west quality improvement programme model) from 2003-6, with binomial funnel plot calculated around cumulative mean for unit's case series. **(Bottom)** Funnel plot displaying observed and predicted major adverse cardiovascular and cerebrovascular events, with denominator for that percentage (number of cases for each operator A-E) displayed as a scatter plot and compared with binomial funnel plot around mean predicted events for all cases

Calculation of expected major adverse cardiovascular and cerebrovascular events

The north west quality improvement programme model included 9914 consecutive patients undergoing percutaneous coronary intervention between 1 August 2001 and 31 December 2003 in the north west of England. The model was internally validated using a dataset consisting of 1786 patients and has been externally validated on our dataset.^{12,13} The model calculates a patient's probability of in-hospital major adverse cardiovascular and cerebrovascular events on the basis of several risk factors: age 70-79 years, age \geq 80 years, female sex, cerebrovascular disease, cardiogenic shock, urgent percutaneous coronary intervention, emergency percutaneous coronary intervention, left main stem lesion treated, and graft lesion treated. The calculation of predicted risk of major adverse cardiovascular and cerebrovascular events as a percentage is $P = \text{odds} / (1 + \text{odds}) \times 100$, where $\text{odds} = \exp(-5.4959 + [0.7048 \times \text{age} 70-79 \text{ years}] + [1.0106 \times \text{age} \geq 80 \text{ years}] + [0.4586 \times \text{female sex}] + [0.8618 \times \text{cerebrovascular disease}] + [3.2636 \times \text{cardiogenic shock}] + [0.4788 \times \text{urgent percutaneous coronary intervention}] + [1.3625 \times \text{emergency percutaneous coronary intervention}] + [1.6502 \times \text{left main stem lesion treated}] + [0.9101 \times \text{graft lesion treated}]$.

Statistical analysis

Funnel plots are based on statistical process control,²⁰⁻²⁴ a set of methods for ongoing improvement of systems, processes, and outcomes. We present two methods of display: funnel plots summarising overall performance and cumulative funnel plots of an individual operator's performance on a case series basis.

The mean predicted major adverse cardiovascular and cerebrovascular events for all cases is displayed as a percentage. We created a funnel plot using upper and lower control limits calculated at 3σ (calculated similarly to 99.8% confidence intervals, although control limits are prediction limits not precision limits) around the mean predicted major adverse cardiovascular and cerebrovascular events, using an exact binomial method.²¹ The mean of the predicted event rate is recalculated as each additional case is added to the series. Similarly, the observed event rate is recalculated as each case is added to the series, creating a cumulative mean. We also calculated an upper warning limit (calculated similarly to 95% confidence intervals) at 2σ above the predicted mean. Each operator is then displayed as a scatter point, showing the observed rate for major adverse cardiovascular and cerebrovascular events and the volume of cases for each operator, against the funnel. The control limits show how much variation to expect around the predicted event rate for a given volume of cases. If the observed event rate for any operator varies more than this, a special cause is implied. Special cause variation (also known as assignable cause variation) is the fluctuation that is caused by unpredictable factors resulting in a non-random distribution of the data. Unlike common cause variation (unassignable cause variation), special causes of variation can be eliminated by reacting to individual variations and have to be removed before a process can be improved by tackling sources of common cause variation. The more cases that are included, the more precisely the predicted event rate "constrains" the observed event rate. For this reason the control limits become narrower when the number of cases done by each operator increases.

Cumulative funnel plots are produced to display the performance of each individual operator and the whole unit on a case series basis. In cumulative funnel plots the funnel shape is produced from the cumulative mean predicted major adverse cardiovascular and cerebrovascular events—that is, it is calculated successively adding each new case. The observed events rate is also shown cumulatively, with each observed line for events beginning with 0%. The first event in the series produces a rise to the percentage of observed events and is followed by a falling line as the number of cases increases without an event, until the next event in the series. This produces a distinctive saw-tooth pattern, making the number of events and cases transparent while allowing comparison of rates. The predicted risk for patients in the series is also clearly displayed and makes it possible to calculate the number of events that would be needed to cross the control limits. The right hand end of the trace for an individual operator's

Table 1 | Preoperative characteristics of patients in percutaneous coronary intervention registry of James Cook University Hospital

Variables	No (%) in registry (n=5198)
Age (years):	
<50	733 (14.1)
50-59	1445 (27.8)
60-69	1689 (32.5)
70-79	1149 (22.1)
>79	135 (2.6)
Women	1487 (28.6)
Diabetes	738 (14.2)
Renal dysfunction	83 (1.6)
Peripheral vascular disease	379 (7.3)
Cerebrovascular disease	296 (5.7)
Previous coronary bypass grafts	338 (6.5)
Extent of vessel disease:	
1	3342 (64.3)
2	1487 (28.6)
3	369 (7.1)
Left main stem >50%	146 (2.8)
Indication for percutaneous coronary intervention:	
Stable angina	1778 (34.2)
Unstable angina	2521 (48.5)
ST elevation myocardial infarction	811 (15.6)
Cardiogenic shock	94 (1.8)
Priority of procedure:	
Elective	1778 (34.2)
Urgent	2459 (47.3)
Emergency	962 (18.5)

Variables defined according to central cardiac audit database definitions.¹⁹

cumulative observed event rate indicates where they will be positioned on the funnel plot.

The risk adjusted major adverse cardiovascular and cerebrovascular event rate is a comparison of observed to expected numbers of events as a ratio, with confidence intervals, to reflect errors due to random sampling. A risk adjusted rate is the event rate that would be expected for the operators had their patients been identical to the unit wide mix.²⁵

We also assessed whether the predicted risk of patients entering the unit is changing, by plotting the average logistic north west quality improvement programme score monthly and comparing this with the overall complication rates for these same periods.

RESULTS

Tables 1 and 2 list the baseline personal, clinical, and procedural characteristics of patients in the registry of the James Cook University Hospital, Middlesbrough. Among the 5198 patients undergoing percutaneous coronary intervention, 102 had procedural complications of interest (1.96%). Most patients experienced a single procedural event (72 deaths, 11 myocardial infarctions, two strokes, and eight emergency coronary artery bypass grafting operations) and nine patients experienced combined outcomes (three emergency coronary artery bypass grafting and death, four myocardial infarction and death, one myocardial infarction and coronary artery bypass grafting, and one stroke and death).

The cumulative funnel plot for all patients undergoing percutaneous coronary interventions shows that overall the observed major adverse cardiovascular and cerebrovascular events rate (1.96%) was less frequent than the predicted risk of 2.06% using the logistic north west quality improvement programme model (fig 1). For the individual operators the overall in-hospital events rates (fig 1) were within the 3 σ (equivalent to the 99.8% confidence interval) upper control limit of 2.75% and the 2 σ upper warning limit of 2.49%. For example, in an individual operator's series of 1178 percutaneous coronary interventions (operator E, fig 2) over three years, the mean predicted risk was 1.78%. Therefore in this case series of 1178 cases, 21 major adverse cardiovascular and cerebrovascular events

would have been expected to occur. The 3 σ control limits would be crossed if there were fewer than 10 events (0.81%, the lower control limit) or greater than 39 in-hospital events (3.32%, the upper control limit). The upper warning limit would be crossed if more than 32 in-hospital events occurred.

Figure 3 displays the funnel plot for 2006 (the predicted means being recalculated for that year). This shows observed and expected major adverse cardiovascular and cerebrovascular events, together with the denominator for the percentage (number of cases for each operator A-E), displayed as a scatter plot and compared with the binomial funnel plot calculated around the mean predicted for all cases reported that year. This funnel shows that all operators had complication rates similar to the predicted rates. Operator C, with the highest number of cases, had a slightly higher than predicted event rate; this operator has the lowest proportion of very low risk cases (table 3). After adjustment of the same data for further risk by the method (observed outcome/expected outcome) \times unit observed outcome²⁵ the results for all operators are clustered around the mean and are well below the upper control and warning limits (fig 3).

Case mix has not changed significantly in the unit in the three year study period, with a predicted risk (calculated using logistic north west quality improvement programme) of in-hospital major adverse cardiovascular and cerebrovascular events for all cases of 1.74% in 2003 and 1.9% in 2006 (fig 4). In addition, no evidence has been found of a drift in the observed rate of events over time, although variation around the mean seems to be less since this method was introduced.

DISCUSSION

We have presented the outcome data for all interventional cardiologists carrying out percutaneous coronary interventions in our centre between 2003 and 2006. We have risk adjusted the outcome data for percutaneous coronary intervention using the north west quality improvement programme model.¹² Members of the public can now see the outcomes for these individual doctors and can be reassured that all are performing to satisfactory standards. Quarterly internal monitoring using cumulative funnel plots, together with annual public reporting using funnel plots, would be a successful method for displaying operator specific data in the UK. This provides an opportunity for interventionalists, as well as referring doctors and patients, to review overall results on a regular basis.

Strengths and weaknesses

The dataset needed to calculate the predicted risk of major adverse cardiovascular and cerebrovascular events according to the north west quality improvement programme is small and feasible to collect. We have shown that it is a valid model to use for predicting events, even with modern percutaneous coronary intervention practice (for example, high rate of stenting, use of thienopyridines and glycoprotein IIb/

Table 2 | Procedural characteristics in percutaneous coronary intervention registry of James Cook University Hospital

Variables	No (%) of patients in registry
Restenotic lesions	192 (3.7)
American Heart Association type C lesions	1653 (31.8)
Left main stem lesions	88 (1.7)
Graft lesions	151 (2.9)
Thrombus containing lesion*	717 (13.8)
Multivessel percutaneous coronary intervention	1128 (21.7)
Chronic occlusion	348 (6.7)
Stent	4964 (95.5)
Glycoprotein IIb/IIIa use	3909 (75.2)

*Defined as visible filling defect in non-occluded vessel. All other vessels with acute occlusion in emergency setting are also assumed to contain thrombus.

IIIa inhibitors). It is proposed that the north west quality improvement programme model be used to calculate the risk adjusted outcome nationally using the UK national percutaneous coronary intervention audit dataset coordinated by the British Cardiovascular Intervention Society.¹⁹

The data collection exercise by the British Cardiovascular Intervention Society provides prospectively collected data on which different methods of analysis can be tested. Quality assessment can then be done both at the level of each unit and at the level of individual doctors across the country for all the hospitals participating in the society's registry. This will allow institutions to measure risk adjusted outcomes and to compare them with national benchmarks for improving quality of care. One advantage of the model is that it does not depend on the case volume of the individual operator. When a difference in case mix, referral patterns, procedural techniques, or adjunctive therapy is sufficiently dominant to create a special cause in the funnel plot for a unit, the individual operator's cumulative funnel will show whether the operator is within the limits for the predicted risk in their personal case series.

Comparison with other studies

Since 1994 the New York State department of health has collected and periodically published observed and risk adjusted patient mortality rates for all the interventional cardiologists practising coronary angioplasty in the state.¹⁴ The New York data, possibly because of exhaustive validation, are about three years out of date by the time they are published. Public reporting of operator specific outcome data may influence doctors to withhold procedures from patients at higher risk, even when a doctor believes that the procedure might be beneficial. The form of analysis depicted in our study should mitigate against such behaviour in that it provides risk adjusted data. Using such a risk adjusted approach, one study has shown how the results of cardiac surgery have improved over time despite a higher risk patient profile, providing evidence that publication of individual results does not necessarily lead to surgeons avoiding high risk patients and, moreover, participation in the quality programme has been associated with better outcomes.²⁶ We have presented our data in funnel plots and cumulative funnel plots summarising overall and individual performance on a case series basis. The routine publication of these funnel plots on performance attainments will encourage reflection, analysis of processes, and subsequent improvements in delivery of health care.

The expected low complication rate for percutaneous coronary interventions presents a statistical power problem when attempting to compare results of individual operators with different case volumes. For this reason we used a combined end point of clinically important outcomes. With a large enough database the methodology could be used for individual components of major adverse cardiovascular and cerebrovascular

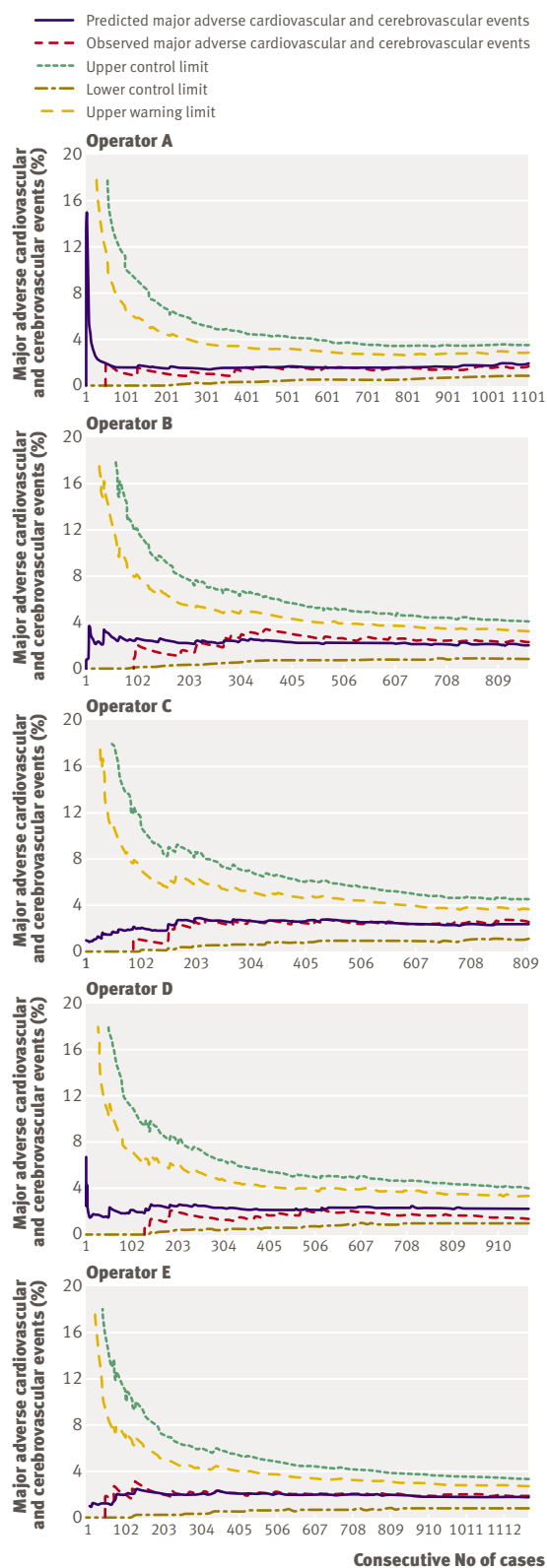


Fig 2 | Cumulative funnel plots displaying observed major adverse cardiovascular and cerebrovascular events for each operator (A-E) compared with mean predicted events (logistic north west quality improvement programme model) from 2003-6, with binomial funnel plot calculated around mean for this unit's case series

events. Once the end point to be used has been determined, cumulative funnel plots are completely transparent in showing the observed events and the number of cases and events included. They have acceptable sensitivity to deteriorating performance if the predicted events are not previously excessively above that observed. They can be applied to operators with both a high volume and a low volume of cases. The temporal display allows an analysis of when performance might deviate from an acceptable level, which itself allows for appropriate analysis and corrective action. A study using similar statistical monitoring techniques, has shown how an overall impression that there was a problem with excessive bleeding rates within a cardiothoracic unit was translated into a statistical demonstration that the problem was real and allowed for evaluation of possible reasons and corrective action.²⁷

Sequential probability ratio tests²⁸ have been suggested for monitoring mortality related to cardiac surgery. These plots (and variants of them, such as cumulative sum techniques) have the best sensitivity to changing performance; however, they have two main disadvantages. The first is their lack of a direct, transparent display of observed and expected mortality, combining the two as they do into a single trace. Secondly, they involve several arbitrary assumptions—specifically they are designed to detect a doubling (or other multiple) of the risk and require α and β (effectively the acceptable true and false positive rates) to be chosen in advance. The methodology described here is visually easy to understand, and its use (especially if fed back to the units by the national society) will encourage individual departments to continuously monitor their results and act accordingly.

Implications

This study will allow internal and external monitoring of performance and will help inform the public about outcome data for individual interventional cardiologists. The analysis for our unit shows good overall results, and predicted mean complication rates for individual operators in this study fell below the 3σ limit (upper control) and 2σ limit (upper warning). The results compare favourably with national and international data. We use the funnel plots on our hospital trust website and use cumulative funnels internally to monitor observed major adverse cardiovascular and cerebrovascular events against predicted events for all our individual operators carrying out percutaneous coronary interventions. Patients in north east England will now be able to scrutinise an individual operator's outcomes, and this should provide reassurance about overall quality. Although such plots could theoretically be used to aid in patient choice of doctors, our own data suggest that a patient could expect equivalent short term outcomes from any of the operators. This type of analysis should be updated regularly, to give ongoing reassurance to patients. This also has a potential role in professional revalidation and could be fed into systems of doctor appraisal. This may require that the process should be made even more robust by developing methods for external validation of data quality.

If such a monitoring system showed that the results for an individual operator (or institution) fell outside the warning limits, this should trigger a response. Results outside the upper warning could initiate an internal review and results outside the upper control limit could trigger an external audit to determine opportunities to improve quality of care. Such reviews should first determine the completeness and accuracy of data collection and establish the case mix in comparison with the case mix overall for the unit and in the national dataset. Risk prediction models are not perfect, however, and for specific groups of patients can either overestimate or underestimate risk.¹³ Moreover, individual operators or institutions could become outliers just by chance. One check would be to reassess the data using a different, but also validated, risk model.

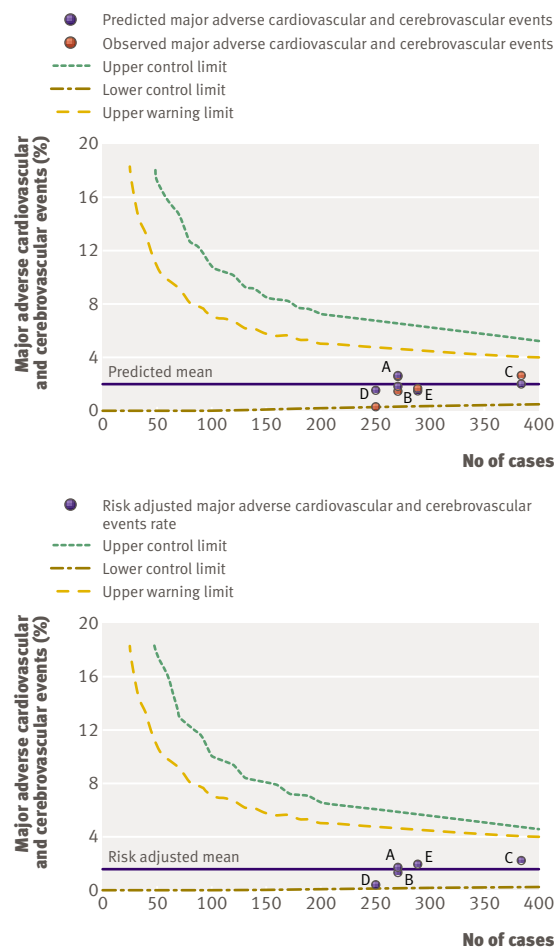


Fig 3 | (Top) Funnel plots displaying observed and predicted major adverse cardiovascular and cerebrovascular events, with denominator for that percentage (number of cases for each operator A-E) displayed as a scatter plot and compared with binomial funnel plot calculated around mean predicted (north west quality improvement programme model) for all cases reported in 2006. (Bottom) Additional risk adjusted major adverse cardiovascular and cerebrovascular events for each operator in 2006. Risk adjustment=(observed outcome/expected outcome)×unit observed outcome²⁵

Table 3 | Number (percentage) of cases done by each operator according to north west quality improvement programme model variables in 2006

Variable	Operator				
	A	B	C	D	E
Total No	271	271	384	250	289
Age (years):					
70-79	57 (21)	54 (19.9)	78 (20.3)	59 (23.6)	62 (21.5)
≥80	8 (3)	5 (1.8)	8 (2.1)	7 (2.8)	14 (4.8)
Female	79 (29.2)	75 (27.7)	105 (27.3)	70 (28)	95 (32.9)
Cerebrovascular disease	10 (3.7)	17 (6.3)	22 (5.7)	6 (2.4)	17 (5.9)
Cardiogenic shock	9 (3.3)	4 (1.5)	7 (1.8)	4 (1.6)	2 (0.7)
Type of percutaneous coronary intervention:					
Urgent	137 (50.6)	118 (43.5)	175 (45.6)	85 (34)	143 (49.5)
Emergency	60 (22.1)	53 (19.6)	77 (20.1)	56 (22.4)	64 (22.1)
Left main stem treated	4 (1.5)	2 (0.7)	19 (4.9)	4 (1.6)	4 (1.4)
Graft lesion treated	6 (2.2)	10 (3.7)	2 (0.5)	4 (1.6)	5 (1.7)
Risk level (score*):					
Very low (1-5)	223 (82.3)	230 (84.9)	300 (78.1)	208 (83.2)	230 (79.6)
Low (6-8)	34 (12.5)	34 (12.5)	68 (17.7)	32 (12.8)	50 (17.3)
Moderate (9-11)	5 (1.8)	3 (1.1)	8 (2.1)	6 (2.4)	7 (2.4)
High (12-14)	0	0	1 (0.3)	0	0

*Score assigned by integer coefficient to each independent variable based roughly on odds estimate in multivariable north west quality improvement programme model.¹²

Yearly monitoring of the previous three years' data may also provide reassurances. The funnel plots provide summary data for a predetermined period (or for a predetermined number of cases) whereas the cumulative funnel plots can be used continuously, thus reflecting contemporary activity. Should there still be a concern about results after evaluating data quality and case mix then changes aimed at improving outcomes could be proposed. Such changes might include a review of case selection or techniques, "buddying-up" with colleagues in specific cases, or retraining. It has been shown that participation in such audit processes improves the quality of data collection, but, more importantly, it can improve patient care without

resulting in risk averse behaviour on the part of individual operators or institutions.²⁹

If this model is universally adopted by all centres carrying out percutaneous coronary interventions, this would allow a well structured benchmarking system in which the public could have full confidence. A system that could activate internal and external reviews would provide an appropriate response to deal with variations in practice. Equally, for most clinicians, attainment of results within these confidence limits would provide them with a mark of excellence and a track record of results recognised as being equivalent to the high standards seen nationally. To be accepted by the profession, such a system must persuade those undertaking high risk cases that their case mix is catered for by the risk model and that the system is primarily about quality improvement rather than seeking just to identify poor operators. We recommend that this methodology, which accommodates these concerns, be considered by the national society. Given that nearly all units in the UK collect data on their patients through the central cardiac audit database,¹⁹ this service could be provided to all units through its central administration, thus avoiding replication of effort.

We believe that cumulative funnel plots have acceptable sensitivity to poor performance using 3σ (the widely accepted best practice in statistical process control) rather than 2σ or lower (for example, 90% limits). Using a tighter upper control limit than 3σ would increase the risk of a false positive and this is not justified if this method is to be applied to a large number of operators across the UK where an unjust examination of an individual operator would become too probable. These 2σ warning limits and 3σ upper

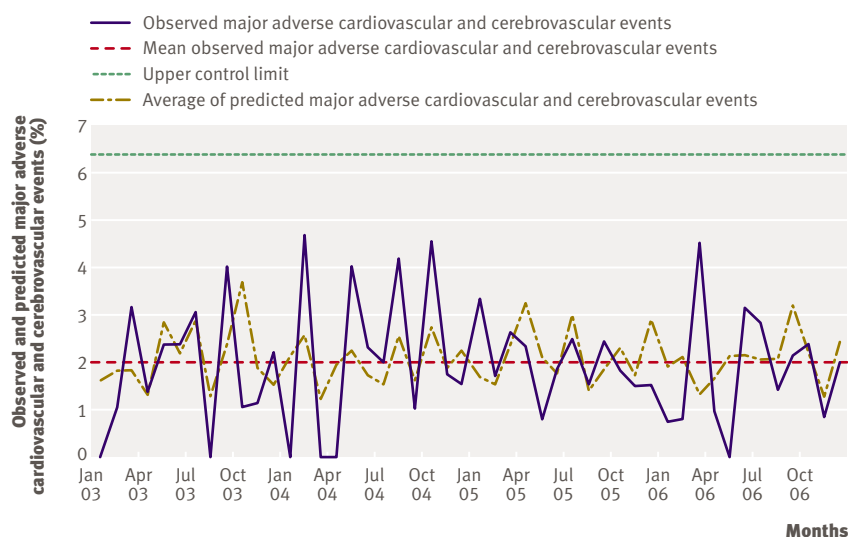


Fig 4 | Monthly observed and predicted major adverse cardiovascular and cerebrovascular events in James Cook University Hospital

WHAT IS ALREADY KNOWN ON THIS TOPIC

Comparative performance of UK cardiac surgeons has been published in the public arena; data on operator specific percutaneous coronary interventions (PCIs) are not yet available in the UK. The north west quality improvement programme (NWQIP) has provided a prediction model for major adverse cardiac events after PCI, and has been subject to internal and external validation.

WHAT THIS STUDY ADDS

The NWQIP model can be used to compare actual with predicted outcomes for operators doing PCIs; the data can be displayed using cumulative plots and funnel plots.

Monthly or quarterly internal monitoring using cumulative funnel plots, together with annual public reporting using funnel plots would be a useful method for displaying operator specific data on PCIs.

This study will allow internal and external monitoring of performance and will help inform the public about outcome data for individual interventional institutions or cardiologists.

control limits have been successfully implemented by the Society for Cardiothoracic Surgery, allowing widespread acceptance of the national publication of risk adjusted mortality in adult cardiac surgery.⁴

Limitations

As with all performance measurement and analysis methods, funnel charts produced using statistical process control may produce both false negative results and false positive results. Performing within limits does not guarantee that a unit may not be underperforming although this may be too slight to be detected or masked by other factors. Being outside a control limit is not always abnormal (special cause) either, even when set at 3σ , equivalent to 99.8% confidence intervals. Despite such false negative results and false positive results, detecting a “special cause” encourages further investigation at the level of the individual operator. Similarly, funnel plots help prevent investigation of an outcome resulting from common cause variation as if it were a special cause phenomenon. Such plots are more meaningful and useful than league tables.

It is possible that outcomes after percutaneous coronary intervention improve and the risk model may therefore become less accurate over time. With a large central database that combines data from a large number of units, the risk model could be periodically re-evaluated so that the concept of benchmarking is not lost.

It would be important to interpret a new operator’s early experience appropriately. The concept of learning curves should be acknowledged. This is implicit in the methodology described because with low numbers of procedures the confidence intervals are wide. Nevertheless, should the results fall outside the control limits, this should trigger some form of response, such as buddying-up with a more senior operator for a predefined number of cases.

Conclusion

Our operators’ in-hospital major adverse cardiovascular and cerebrovascular events rates were lower than the predicted event rate. In-hospital rates after

percutaneous coronary intervention can be monitored successfully using funnel plots and cumulative funnel plots with 3σ control limits to display and publish each operator’s outcomes. The upper warning limit (2σ control limit) could be used for internal monitoring. The main advantage of these charts is their transparency, as they show observed and predicted events separately. By this approach individual operators can monitor their own performance, using the predicted risk for their patients in a way that allows benchmarking against colleagues. The methodology allows scrutiny of outcomes regardless of variations in individual operator case volume.

Contributors: BK, JD, MdeB, and APR designed the study. JAH, AGCS, RAW, DFM, and MdeB carried out the procedures and collected the in-lab data. BK, DT, JD, and RM ensured completeness of the database and did the analysis with support from PR. BK, MdeB, and PR led on the writing of the manuscript but all authors contributed to the submitted versions of the manuscript. MdeB is the guarantor for the study.

Funding: None.

Competing interests: BK, JAH, AGCS, RAW, DFM, and MdeB have received travel grants from manufacturers of coronary stents and percutaneous coronary intervention related pharmaceutical companies. MdeB has sat on advisory boards for stent manufacturers and percutaneous coronary intervention related pharmaceutical companies and has received research grants from a few stent manufacturers. As members or fellows of the Royal College of Physicians we have an interest in methods of revalidation and may be involved in standard setting. As, respectively, president of the British Cardiovascular Intervention Society and council member of the British Cardiovascular Society, MdeB and JAH will have a role in informing the debate about these issues. The authors do not believe that any of these declarations constitute a conflict of interest as regards this study.

Ethical approval: Not required.

Provenance and peer review: Not commissioned; externally peer reviewed.

- 1 Anon. Learning from Bristol: the report of the public inquiry into children’s heart surgery at the Bristol Royal Infirmary 1984-1995. www.bristolinquiry.org.uk.
- 2 Anon. Shipman inquiry. Fifth report: safeguarding patients, lessons from the past—proposals for the future. 2004. www.the-shipman.org.uk.
- 3 Anon. NHS heart surgery. The data explained: 244 doctors and the problem of comparing mortality rates. *Guardian* 20 Oct;2007:10-1.
- 4 Healthcare Commission. <http://hearturgery.healthcarecommission.org.uk>.
- 5 Detre K, Yeh W, Kelsey S, Williams D, Svigne-Nickens P, Holmes D Jr, et al. Has improvement in PTCA intervention affected long-term prognosis? The NHLBI PTCA Registry experience. *Circulation* 1995;91:2868-75.
- 6 Williams DO, Holubkov R, Yeh W, Bourassa MG, Al-Bassam M, Block PC, et al. Percutaneous coronary intervention in the current era compared with 1985-1986: the National Heart, Lung, and Blood Institute registries. *Circulation* 2000;102:2945-51.
- 7 McGrath PD, Malenka DJ, Wennberg DE, Shubrooks SJ Jr, Bradley WA, Robb JF, et al. Changing outcomes in percutaneous coronary interventions: a study of 34,752 procedures in northern New England, 1990 to 1997. Northern New England Cardiovascular Disease Study Group. *J Am Coll Cardiol* 1999;34:674-80.
- 8 O’Connor GT, Malenka DJ, Quinton H, Robb JF, Kellett MA Jr, Shubrooks S, et al. Multivariate prediction of in-hospital mortality after percutaneous coronary interventions in 1994-1996. Northern New England Cardiovascular Disease Study Group. *J Am Coll Cardiol* 1999;34:681-91.
- 9 Kimmel SE, Berlin JA, Strom BL, Laskey WK. Development and validation of simplified predictive index for major complications in contemporary percutaneous transluminal coronary angioplasty practice. The Registry Committee of the Society for Cardiac Angiography and Interventions. *J Am Coll Cardiol* 1995;26:931-8.
- 10 Block PC, Peterson ED, Krone R, Kesler K, Hannan E, O’Connor GT, et al. Identification of variables needed to risk adjust outcomes of coronary interventions: evidence-based guidelines for efficient data collection. *J Am Coll Cardiol* 1998;32:275-82.
- 11 Moscucci M, O’Connor GT, Ellis SG, Malenka DJ, Sievers J, Bates ER, et al. Validation of risk adjustment models for in-hospital

- percutaneous transluminal coronary angioplasty mortality on an independent data set. *J Am Coll Cardiol* 1999;34:692-7.
- 12 Grayson AD, Moore RK, Jackson M, Rathore S, Sastry S, Gray TP, et al. Multivariate prediction of major adverse cardiac events after 9914 percutaneous coronary interventions in the north west of England. *Heart* 2006;92:658-63.
 - 13 Kunadian B, Dunning J, Das R, Roberts AP, Turley AJ, Twomey D, et al. External validation of established risk adjustment models for procedural complications after percutaneous coronary intervention on an independent data set in a UK population setting. *Heart* Nov 2007; doi:10.1136/hrt.2007.129197.
 - 14 New York State Department of Health. Percutaneous coronary intervention (PCI) in New York State 2002-2004. www.health.state.ny.us/nysdoh/heart/pdf/pci_2002-2004.pdf.
 - 15 Bridgewater B, Grayson AD, Jackson M, Brooks N, Grotte GJ, Keenan DJ, et al. Surgeon specific mortality in adult cardiac surgery: comparison between crude and risk stratified data. *BMJ* 2003;327:13-7.
 - 16 King SBI, Aversano TM, Ballard WLM, Beekman RHI, Cowley MJM, Ellis SGM, et al. ACCF/AHA/SCAI 2007 Update of the clinical competence statement on cardiac interventional procedures: a report of the American College of Cardiology Foundation/American Heart Association/American College of Physicians Task Force on Clinical Competence and Training (writing committee to update the 1998 clinical competence statement on recommendations for the assessment and maintenance of proficiency in coronary interventional procedures). *Circulation* 2007;116:98-124.
 - 17 Bridgewater B, Adult Cardiac Surgeons of North West England. Mortality data in adult cardiac surgery for named surgeons: retrospective examination of prospectively collected data on coronary artery surgery and aortic valve replacement. *BMJ* 2005;330:506-10.
 - 18 Gale CP, Roberts AP, Batin PD, Hall AS. Funnel plots, performance variation and the myocardial infarction national audit project 2003-2004. *BMC Cardiovasc Disord* 2006;6:34.
 - 19 Central cardiac audit database. Coronary heart disease audit supported by the central cardiac audit database. www.ccad.org.uk.
 - 20 Spiegelhalter D. Funnel plots for institutional comparison. *Qual Safety Health Care* 2002;11:390-1.
 - 21 Spiegelhalter D. Funnel plots for comparing institutional performance. *Stat Med* 2005;24:1185-202.
 - 22 Carey RG. *Improving healthcare with control charts: basic and advanced SPC methods and case studies*. Milwaukee: Quality Press, 2004.
 - 23 Bolsin S, Patrick A. Was Bristol an outlier? *Lancet* 2001;358:2084.
 - 24 Colson M, Bolsin S. The use of statistical process control methods in monitoring clinical performance. *Int J Qual Health Care* 2003;15:445.
 - 25 Hannan EL, Racz M, Ryan TJ, McCallister BD, Johnson LW, Arani DT, et al. Coronary angioplasty volume-outcome relationships for hospitals and cardiologists. *JAMA* 1997;277:892-8.
 - 26 Bridgewater B, Grayson AD, Brooks N, Grotte G, Fabri BM, Au J, et al. Has the publication of cardiac surgery outcome data been associated with changes in practice in northwest England: an analysis of 25,730 patients undergoing CABG surgery under 30 surgeons over eight years. *Heart* 2007;93:744-8.
 - 27 Wolfe R, Bolsin S, Colson M, Stow P. Monitoring the rate of re-exploration for excessive bleeding after cardiac surgery in adults. *Qual Safety Health Care* 2007;16:192-6.
 - 28 Spiegelhalter D, Grigg O, Kinsman R, Treasure T. Risk-adjusted sequential probability ratio tests: applications to Bristol, Shipman and adult cardiac surgery. *Int J Qual Health Care* 2003;15:7-13.
 - 29 Narins CR, Dozier AM, Ling FS, Zareba W. The influence of public reporting of outcome data on medical decision making by physicians. *Arch Intern Med* 2005;165:83-7.

Accepted: 24 February 2008