

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

Babu Kunadian, Joel Dunning, Anthony P Roberts, Robert Morley, Darragh Twomey, James A Hall, Andrew G C Sutton, Robert A Wright, Douglas F Muir, Mark A de Belder

**EDITORIAL** by Holt and colleagues  
**RESEARCH** p 934

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

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

**BMJ** 2008;336:931-4  
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  $3\sigma$  upper control limit of 2.75% and  $2\sigma$  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  $3\sigma$  control limits to display and publish each operator's outcomes. The upper warning limit ( $2\sigma$  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

In 2004, the *Guardian* published mortality data from 244 named cardiac surgeons in the UK.<sup>1</sup> 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 but risk adjusted.<sup>2</sup>

Although comparative performance of UK cardiac surgeons has been published,<sup>3</sup> operator specific data for percutaneous coronary intervention are not yet available. We display operator specific outcomes after percutaneous coronary intervention using the north west quality improvement programme risk model and then cumulative funnels and funnel plots.

## METHODS

A detailed database of several variables has been maintained on all patients undergoing percutaneous coronary intervention at our unit since 1994, based on the British Cardiovascular Intervention Society national dataset.<sup>4</sup> 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. Our outcome of interest was major adverse cardiovascular and cerebrovascular events (see definition on bmj.com).

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 of 1786 patients and has been externally validated on our dataset.<sup>5,6</sup> The model calculates a patient's probability of in-hospital major adverse cardiovascular and cerebrovascular events on the basis of several risk factors (see bmj.com).

## Statistical analysis

We present 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

This article is an abridged version of a paper that was published on [bmj.com](http://bmj.com). Cite this article as: *BMJ* 2008, doi: 10.1136/bmj.39512.529120.BE

lower control limits calculated at  $3\sigma$  around the mean predicted events, using an exact binomial method.<sup>7</sup> 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 at  $2\sigma$  above the predicted mean. The control limits show how much variation to expect around the predicted event rate for a given volume of cases. If the observed rate varies more than this, a special cause is implied. Special cause variation is the fluctuation that is caused by unpredictable factors resulting in a non-random distribution of the data. Unlike common 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.

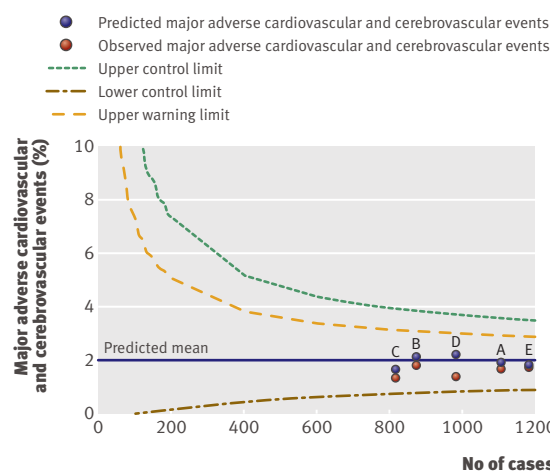
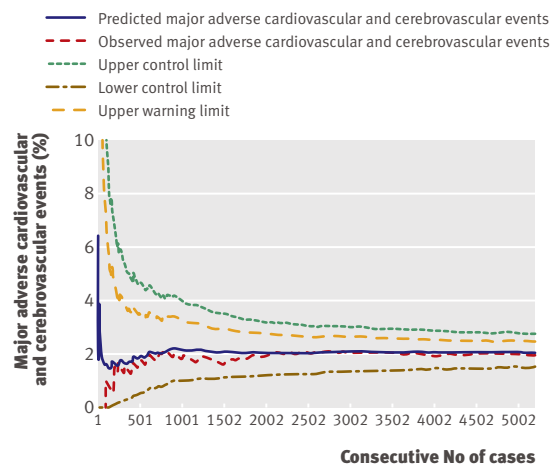
Cumulative funnel plots are produced to display the performance of each individual operator and the whole unit on a case series basis. In cumulative funnel plots the funnel shape is produced from the cumulative mean predicted major adverse cardiovascular and cerebrovascular events. The observed event 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 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. 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.<sup>8</sup>

**RESULTS**

Baseline characteristics of patients in the registry are on bmj.com. Among the 5198 patients undergoing percutaneous coronary intervention, 102 had procedural complications of interest (1.96%; see bmj.com).

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 (figure). For the individual operators the overall in-hospital events rates were within the upper control limit of 2.75% and the upper warning limit of 2.49% (figure).

The funnel plot for 2006 (the predicted means being recalculated for that year) is on bmj.com. It 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



(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 the 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

in that year; this operator has the lowest proportion of very low risk cases (see bmj.com). After risk adjustment of the same data using (observed outcome/expected outcome)<sup>8</sup> the results for all operators are clustered around the mean and are well below the upper control and warning limits (see bmj.com).

**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 using the north west quality improvement programme model.<sup>5</sup> Members of the public can see these outcomes. Quarterly internal monitoring using cumulative funnel plots, together with annual public reporting using

### 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.

funnel plots, would be a successful method for displaying operator specific data in the UK.

#### 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 feasible to collect. It is a valid model to use for predicting events, even with modern percutaneous coronary intervention practice. We propose 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.<sup>4</sup>

The British Cardiovascular Intervention Society is prospectively collecting data on which different analyses can be tested. Quality assessment can then be done at the level of each unit and at the level of individual doctors for all hospitals participating in the registry. One advantage is that the model does not depend on case volume of individual operators. 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 periodically published observed and risk adjusted patient mortality rates for cardiologists practising coronary angioplasty.<sup>9</sup> The data, possibly because of exhaustive validation, are about three years out of date when published. Public reporting of operator specific outcome data may influence doctors to withhold procedures from patients at higher risk, even when the procedure might be beneficial. The form of analysis depicted in our study should mitigate against such behaviour.

The expected low complication rate for percutaneous coronary interventions presents a statistical power

problem when comparing 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 events. Once the end point to be used has been determined, cumulative funnel plots are completely transparent. 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 high and low volumes of cases. The temporal display allows an analysis of when performance might deviate from an acceptable level.

#### Implications

This study allows internal and external monitoring of performance and will help inform the public about outcome data for individual cardiologists. We use the funnel plots on our website and use cumulative funnels internally to monitor observed major adverse cardiovascular and cerebrovascular events against predicted events for all our individual operators.

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. Risk prediction models can, however, overestimate or underestimate risk for specific patient groups.<sup>6</sup> Moreover, individual operators or institutions could become outliers by chance. One check would be to reassess the data using a different risk model. Yearly monitoring of three years' data may also provide reassurances.

If this model is universally adopted, it 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 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. Given that nearly all units in the UK collect data on their patients through the central cardiac audit database,<sup>4</sup> this service could be provided to all units through its central administration.

#### Limitations

Funnel charts produced using statistical process control may produce false negative or 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. Despite such false 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.

It is possible that outcomes after percutaneous coronary intervention improve and the risk model may therefore become less accurate over time. With a 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.

### Conclusion

In-hospital rates of major cardiovascular and cerebrovascular events after percutaneous coronary intervention can be monitored successfully using funnel plots and cumulative funnel plots. 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.

**Contributors:** See [bmj.com](http://bmj.com).

**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. NHS heart surgery. The data explained: 244 doctors and the problem of comparing mortality rates. *Guardian* 20 Oct;2007:10-1.
- 2 Healthcare Commission. <http://heartsurgery.healthcarecommission.org.uk>.
- 3 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.
- 4 Central cardiac audit database. Coronary heart disease audit supported by the central cardiac audit database. [www.ccad.org.uk](http://www.ccad.org.uk).
- 5 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.
- 6 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.
- 7 Spiegelhalter D. Funnel plots for comparing institutional performance. *Stat Med* 2005;24:1185-202.
- 8 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.
- 9 New York State Department of Health. Percutaneous coronary intervention (PCI) in New York State 2002-2004. [www.health.state.ny.us/nysdoh/heart/heart\\_disease.htm](http://www.health.state.ny.us/nysdoh/heart/heart_disease.htm).

**Accepted:** 24 February 2008

## Volume of procedures and risk of recurrence after repair of groin hernia: national register study

Pär Nordin, Willem van der Linden

**EDITORIAL** by Holt and colleagues  
**RESEARCH** p 931

Department of Surgery,  
Östersund Hospital, S 83183,  
Östersund, Sweden

Correspondence to: P Nordin  
[par.nordin@jll.se](mailto:par.nordin@jll.se)

**BMJ** 2008;336:934-7  
doi:10.1136/bmj.39525.514572.25

This article is an abridged version of a paper that was published on [bmj.com](http://bmj.com). Cite this article as: *BMJ* 2008, doi: 10.1136/bmj.39525.514572.25

### ABSTRACT

**Objective** To determine whether the association between volume and outcome found in major surgery also holds true for a minor operation.

**Design** Review of outcomes after hernia surgery in Sweden.

**Setting** Surgical units registered with the Swedish hernia register, which in 2004 covered about 95% of all hernia operations in Sweden.

**Participants** 86 409 patients over 15 years, who underwent 96 601 unilateral or bilateral groin hernia repairs (94 077 inguinal and 2524 femoral) in 1996-2004 at the participating surgical units.

**Main outcome measure** Re-operation for recurrence.

**Results** There was a significantly higher rate of re-operation in surgeons who carried out 1-5 repairs a year than in surgeons who carried out more repairs. There was no association between outcome and further increases in volume. Although about half of surgeons in Sweden who repair hernias are low volume operators, they performed only 8.4% of all repairs.

**Conclusions** Sweden's numerous low volume hernia surgeons perform such a small proportion of all operations that the impact of their inferior results on the nationwide re-operation rate is minimal. Volume indicates an approximate minimum value for the number of hernia repairs a surgeon should do each year but the outcome in surgeons who carry out more than that number disqualifies volume as an indicator of proficiency.

### INTRODUCTION

To determine whether or not the relation between volume and outcome might be extrapolated from major surgery to minor surgery, we used recurrence after groin hernia surgery as it can be reliably assessed and occurs often enough to serve as a measure of success. We studied the relation between the number of repairs carried out by surgeons and the incidence of re-operation in nearly 100 000 hernia repairs prospectively recorded in a hernia register in 1996-2004.