Computer tools to assist the monitoring of outcomes in surgery

C. Sherlaw-Johnson\textsuperscript{a,}*, S. Gallivan\textsuperscript{a}, T. Treasure\textsuperscript{b}, S.A.M. Nashef\textsuperscript{c}

\textsuperscript{a}Clinical Operational Research Unit, University College London, Gower Street, London WC1E 6BT, UK
\textsuperscript{b}Guy’s and St Thomas’s Hospitals Medical School, London, UK
\textsuperscript{c}Papworth Hospital, Cambridge, UK

Received 7 April 2004; accepted 1 July 2004; Available online 13 August 2004

Abstract

Objective: In recent years, there has been increasing use of analytical and graphical methods to assist the monitoring of outcomes in adult cardiac surgery. In this paper, we present extensions to the basic VLAD methodology that add flexibility and assist in its interpretation. Methods: Using techniques from probability theory, we have devised graphical tools whereby deviations from expected outcomes can be monitored to see how likely they are to have occurred by chance. The methods are based upon pre-operative assessments of risk and use exact analytical techniques. Results: These tools allow deviations from expected outcomes to be readily assessed and compared with the distribution of chance outcomes. Appropriate colour coding allows interpretation in terms of a temperature gradient. Conclusions: Exact analysis methods based on the use of pre-operative risk assessment provide a useful means for assisting the interpretation of VLAD charts. Such analysis has the advantage that it is applicable even for relatively short series of operations. Also, it takes specific account of the heterogeneity of case mix when quantifying the variability that is expected. By displaying the overall history of outcomes in a visually intuitive manner, it complements the more formal tools for detecting isolated good and bad runs that are available.

Keywords: Audit; Performance monitoring; Cardiac surgery; Surgical outcomes

1. Introduction

In recent years, there has been increasing use of analytical and graphical methods to assist the monitoring of outcomes in adult cardiac surgery following ideas originally introduced in congenital heart surgery by de Leval [1]. Various modifications to the methods have been devised to allow case mix to be taken into account and to assist with the interpretation of the resulting charts [2–6]. One of the techniques that has emerged [2] goes by name VLAD, (Variable Life Adjusted Display) and is now in routine use in many hospitals.

In this paper, we discuss extensions to the basic VLAD methodology that add flexibility and assist in its interpretation. The most innovative of these extensions provides a facility whereby apparently disturbing trends can be monitored to see how likely they are to have occurred by chance.

* Corresponding author. Tel: +44-207-679-4507; fax: +44-207-813-2814.
E-mail address: c.sherlaw-johnson@ucl.ac.uk (C. Sherlaw-Johnson).
of exact methods and thus do not depend on the length of the series being examined.

There are other useful by-products of these methods. For example, it is possible to derive and compare mortality rates and associated prediction intervals for individual surgeons. These estimates take case-mix into account and, being exact, they do not rely on the validity of approximation formulae.

In addition to these new analytical methods, the software used to implement the new version of VLAD also includes the EuroSCORE for calculating preoperative risks. This software has been prepared using Microsoft Excel, and can be downloaded and used free of charge from the web site of the Clinical Operational Research Unit at University College London, www.ucl.ac.uk/operational-research/vladdownload.htm.

2. Background

So-called Cumulative Sum (CUSUM) methods are a popular means of examining overall surgical outcomes. At their simplest, they show a cumulative running total of total mortality plotted against total number of operations performed. This gives a jagged curve that climbs upwards (for example, the lower curve in Fig. 1 which uses fictitious data for illustration purposes). Deteriorating or improving performance can in principle be detected by observing how the slope of the curve changes. However, if the case-mix of a surgeon changes, which is likely for a surgeon in training, then apparent stability of the chart may shield the fact that the surgeon is dealing with more and more difficult cases. Equally, a surgeon under investigation, presented with a damning plot might rightly claim that his or her case-mix was particularly bad.

To compensate for variable case-mix, more sophisticated analyses are possible using pre-operative estimates for the risk of death based on the values of known risk factors. Such assessments might, for example, be derived from the Parsonnet score [8] or the EuroSCORE [9]. Given such information, one can take the surgeon’s case-mix into account and derive an estimate for the number of deaths that might be expected. This is shown in Fig. 1 where the difference between the curve expressing the expected number of deaths and the actual number gives an estimate for the ‘Net Life Gain’. Plotting this Net Life Gain against the number of operations that a surgeon has done gives an informative summary of the overall pattern of outcomes. This gives a jagged curve (Fig. 2) that moves up, for each surviving patient, and dips for each death. The amount of rise or fall of the chart depends on the pre-operative risk of the patient considered. The higher the risk, the less the surgeon is ‘penalised’ if there is a death; equally, the higher the ‘reward’ if the patient survives against the odds. Such charts are known as VLAD charts (Variable Life Adjusted Display) [2]. VLAD charts are now an accepted means for routine cardiac surgical audit [10] and have also been used in other contexts [11–13].

3. An exact method for assessing whether a given number of poor outcomes could have arisen by chance

If a surgeon has outcomes that are well-matched to the pre-operative risk scoring system being used, then for a

![Fig. 1. Calculation of net life gain from estimate of expected total mortality and actual cumulative mortality.](image)

![Fig. 2. VLAD chart plotting a running tally of net life gain derived from Fig. 1.](image)

![Fig. 3. A schematic representation of the different mortality states that are possible at different stages of a series of operations. Arrows indicate transitions that can occur between these states. The probability of the occurrence of a particular transition is assumed to occur according to the pre-operative risk scoring system used.](image)
given a series of cases, the associated VLAD chart would be expected to end at a point relatively close to the horizontal axis since this represents ‘par’. However as with the roulette wheel analogy, some random fluctuation is to be expected and it is important to be able to judge how concerned one should be about the final displacement from the horizontal axis. If the series is long enough, then approximation methods can be used to derive a normal distribution indicating the probability that a given displacement could have occurred by chance [4]. If the final displacement is large enough that it lies in the extreme lower tail of this bell-shaped distribution, then there may well be cause for concern since this suggests systematic differences from the pre-operative risks. Unfortunately, the approximation methods required to derive this bell-shaped curve are not valid for short case series.

We have now developed an alternative method of analysis that is based on exact methods and thus does not require any approximations. Full mathematical details of this are given in a technical appendix. Here, we shall simply indicate the way in which the method works.

Our analysis is based on a branch of probability theory that concerns the analysis of systems that evolve over time and which are influenced by chance [14]. In our analysis, we consider a series of operations and estimate the probabilities, however small, associated with all possible mortality outcomes that could have occurred by a particular stage in the series. A schematic diagram illustrating all such mortality outcomes is shown in Fig. 3, the arrows indicating how this system can evolve. The probability of movement along any particular branch of this figure is assumed to depend on the pre-operative risk assessment for each successive case, which may vary from patient to patient.

The central quantities of interest are the probabilities of occurrence of each of the mortality outcomes in Fig. 3. In view of the fact that pre-operative risks can vary from operation to operation, these probabilities are difficult to express in terms of explicit mathematical formulae. However, given the very structured nature of Fig. 3, it is possible to calculate these probabilities iteratively, basing the estimates at one stage in the series of operations on the values calculated for the previous operation. This is described in formal mathematical terms in Appendix. Once the series becomes long enough, the normal approximation can be used.

For a given series of operations, there may be many potential outcomes that could have occurred, each of which would give rise to a different VLAD values. Given exact estimates for the probabilities of occurrence of each of these, it is feasible to colour code all these potential values for the VLAD scores according to percentile ranges. Superimposing such a display on the VLAD chart provides

![VLAD chart for a series of 381 coronary artery bypass operations, with superimposed regions indicating the distribution of chance outcomes.](image1)

![Fig. 5. A comparison of mortality rates for five surgeons undertaking coronary artery bypass surgery with prediction intervals based on exact estimates derived from each surgeon’s case-mix. The black circles represent observed mortality.](image2)
4. Additional features

The use of the method shown in Fig. 3 for computing exact probabilities of all possible mortality outcomes makes it straightforward to calculate other statistical quantities that are useful within the audit process. A key statistic that will always be included in any audit summary is the overall mortality rate that any cardiac surgeon has experienced. Our exact probability calculations make it possible to combine the risk scores to derive the mortality rate that would be expected, given the surgeon’s case-mix. It is also possible to calculate prediction intervals that indicate the distribution of mortality rates that would occur if all variation were due to chance. Fig. 5 shows an example of how such information can be displayed graphically in order to compare the mortality rates recorded for a number of surgeons against what would be expected, given their case mix. Again this display makes use of colour coding to assist interpretation. The surgeons shown in Fig. 5 are all from the same centre, and the operations are all coronary artery bypass operations with risks calculated by the logistic EuroSCORE. Three of the five surgeons are within the lower 1% tail of the distribution of predicted death rates, (to the left of the dark blue region), another is within the lower 25% tail, and the fifth within the middle 50% of the distribution.

Similar displays can be produced showing prediction intervals for several surgeons’ final VLAD scores, although for the sake of brevity, illustrations of these are omitted.

5. Discussion

Exact analysis methods based on the use of a priori operative risk scores provide a useful means for assisting the interpretation of VLAD charts. Such analysis has the advantage that it is applicable even for relatively short series of operations. Also, it takes specific account of the heterogeneity of case-mix when quantifying the variability that is expected. Thus, it does not rely on assumptions concerning the validity of approximation formulae nor on assumptions that case-mix is homogeneous.

Again we should stress that the VLAD method is not a formal statistical test, but a method for exploring outcome data. The temperature gradient enhancements that we have introduced are purely a guide to assist interpretation and should not be used for formal hypothesis testing, particularly since no correction is made to account for sequential testing. However, by displaying the overall history of outcomes in a visually intuitive manner, it complements the more formal tools for detecting isolated good and bad runs that are available [5,6].

Appendix

This section describes how the exact distribution of outcomes were calculated for each point of the VLAD.

For a series of patients, let $D_n$ denote the total number of deaths observed after $n$ cases, and $y_i$ the risk associated with the $i$th patient, as derived from the risk score.

If $P_i(n)$ denotes the probability $D_n$ is equal to $i$, and we assume that the outcomes of successive patients are independent, then for each $i$ these probabilities can be calculated iteratively with the following equations:

$$P_0(0) = 1$$
$$P_0(n) = (1 - y_n)P_0(n - 1), \quad \text{for } n > 0$$
$$P_i(n) = (1 - y_n)P_i(n - 1) + y_nP_{i-1}(n - 1), \quad \text{for } 0 < i < n$$
$$P_n(n) = y_nP_{n-1}(n - 1)$$

If $V_n$ denotes the position of the VLAD after a series of $n$ patients, then $V_n = \sum_{i=1}^{n} y_i - D_n$, and $P(D_n \geq k) = P(V_n \leq \sum_{i=1}^{n} y_i - k)$, so the probability distribution for the VLAD can be derived directly from that for the number of deaths with the lower tail for one corresponding to an upper tail for the other.

References
