Analysis of a cluster of surgical failures

Application to a series of neonatal arterial switch operations

A pediatric cardiac surgeon performed 104 neonatal arterial switch operations for transposition of the great arteries with or without ventricular septal defect between June 1987 and February 1993. Initial euphoria on having only one death in the first 52 patients gave way to increasing concern when patients 53, 55, 59, 63, 64, 67, and 68 died. Sensing a problem, the surgeon visited a low-risk institution after patients 55 and 64 had died and then decided to retrain after patient 68 died. One death has occurred since. To find out whether the cluster of failures could have been related to chance alone, to variability of risk factors across time, or to suboptimal performance, we conducted the following analyses: First, identification of trends with the cumulative sum procedure was undertaken and actual mortality compared with the mortality predicted from an equation derived from a multiinstitutional study. Second, logistic regression analysis of risk factors was done. If a mechanism of continuous monitoring had been in place, unfavorable trends and a need for change in protocol would have been detected earlier. Retrospective risk factor analysis suggested an excessive risk for patients with origin of the circumflex or left anterior descending coronary arteries from sinus 2 and a protective effect of phenoxybenzamine. However, about half of the risk associated with the cluster of failures was not accounted for by the variables analyzed. There was therefore an indication of suboptimal performance that appears to have been neutralized by retraining. (J Thorac Cardiovasc Surg 1994;107:914-24)

Marc R. de Leval, MD, FRCS, Katrien François, MD (by invitation), Catherine Bull, MRCP (by invitation), William Brawn, FRCS (by invitation), and David Spiegelhalter, PhD (by invitation), *London, England*

For many surgical procedures, mortality rates decrease with experience of the surgeon. Under normal circumstances failures occur as rare, isolated events; should they tend to cluster, the surgeon will wonder if the "bad run" is simply a matter of chance or the result of a deterioration in performance requiring remedial measures.

This paper analyzes a cluster of failures in the experience of one surgeon with the neonatal arterial switch operation for transposition of the great arteries (TGA).

The following issues were addressed: (1) Could the cluster of failures be due to chance alone? (2) Could procedural risk factors and their variation across time explain the mortality? (3) Could human error account for the cluster of failures? (4) Could appropriate monitoring techniques allow early detection of trends in surgical outcomes? (5) Could outcome measures other than death provide a refined way of monitoring surgical performance? (6) If the surgeon's performance varies across time how is it best expressed? (7) How can the failure rate be reset to its nominal low level or below after a period of suboptimal performance?

Patients, methods, and mortality

Overall experience. The first author of this paper prepared himself to perform the arterial switch operation for TGA by a series of on-site tutorials. Between June 1987 and February 1993 he then performed 104 consecutive neonatal switch operations for simple TGA (89 patients) or TGA with ventricular septal defect (VSD) (15 patients). The age of the patient at the time of the operation varied from 2 to 31 days (mean 12.5) and the weight from 2.3 to 5 kg (mean 3.4). Fifty-three percent received an infusion of prostaglandin. Balloon atrial septostomy was usually done and the prostaglandin infusion was discontinued thereafter.

Initial euphoria on having only one death among the first 52

From the Cardiothoracic Unit, Hospital for Sick Children, Great Ormond Street, London, England.

Read at the Seventy-third Annual Meeting of The American Association for Thoracic Surgery, Chicago, Ill., April 25-28, 1993.

Address for reprints: M. R. de Leval, MD, Cardiothoracic Unit, Hospital for Sick Children, Great Ormond Street, London WC1N 3JH, England.

0022-5223/94 $3.00 + 0 **12/6/52313**

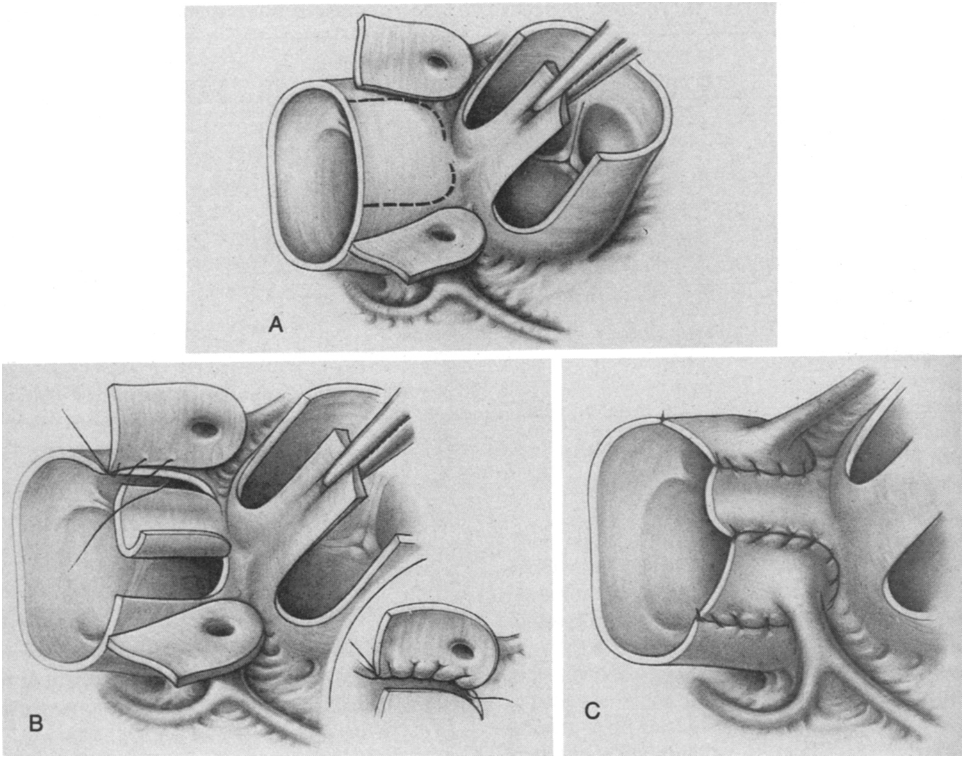


Fig. 1. Illustrations of coronary arterial transfer removing large button of aorta and tucking it into centrally hinged trap door with minimal mobilization.

patients gave way to increasing concern when patients 53, 55, 59, 63, 64, 67, and 68 died. The surgeon instinctively sensed problems and visited another institution known for its low mortality after patient 55 and again after patient 64. He then elected to retrain at a third institution after patient 68, and one death has occurred since.

Surgical techniques

Before patient 68. Operations were done with cardiopulmonary bypass and deep hypothermia with reduced flow to improve exposure when necessary. Direct bicaval cannulation was used in 13 patients and a single atrial cannula in 55 patients. The atrial septostomy and the VSD (when present) were closed during a period of total circulatory arrest when a single venous cannula was used. A single dose of cold crystalloid cardioplegic solution was infused at low pressure in 36 patients, whereas multiple doses were infused in 32 patients. By and large, the aorta was transected distally. The space between the aorta and the pulmonary artery was freed of all connective tissue, so as to obtain a complete mobilization of both great arteries and to expose the root of the pulmonary artery. The coronary arteries were then excised from the aortic sinuses with a button of aortic wall and their proximal portions were dissected and mobilized to facilitate their transfer to the pulmonary artery. Up to patient 55, the coronary buttons were implanted below the transection site of the neo-aorta into circular defects created in the corresponding sinuses. The defects created in the neo-pulmonary artery by the button excision of the coronary ostia were filled by suturing in place two pieces of glutaraldehyde-treated autologous pericardium. The pulmonary arterial bifurcation was brought in front of the aorta (Lecompte maneuver) in all but seven patients.[1] A number of maneuvers were inconsistently applied during the initial period. They include (1) the administration of phenoxybenzamine (α-blocker) immediately before the start of cardiopulmonary bypass, (2) the use of aprotinin to reduce postoperative bleeding, and (3) the application of fibrin glue on the suture lines.

After the visits to another institution after patients 55 and 64 died, a number of modifications were introduced: less general use of phenoxybenzamine, use of a lesser degree of hypothermia and multiple doses of cardioplegic solution, more extensive mobilization of the proximal coronary arteries, and the use of a single scalloped patch of autologous pericardium for the reconstruction of the aortic sinuses.

After patient 68. As a method of retraining the surgeon elected to copy and consistently reproduce to their smallest details all the surgical procedures and support systems used by another surgeon known for his very low mortality with the arterial switch operation.

All operations were done with a single venous cannula in the right atrium (right-angled Rygg-Kyvsgaard cannula; Cardiovascular Specialties Ltd., Scarborough, Ontario, Canada) at 18° C nasopharyngeal temperature with a single dose of cardioplegic solution and a period of total circulatory arrest to close the atrial septal defect and the VSD when present. The aorta was transected distally and the coronary arteries were excised with most of the intercommissural sinus wall, which was later used to lengthen the coronary artery. The periadventitial tissues surrounding the aorta, the pulmonary artery, and the

Table I. *Distribution of patient-specific factors, procedural factors, and outcome measures among the three eras*

	Patients 1-35 (35 patients)	*Patients 36-70 (35 patients)*	*Patients 71-104 (34 patients)*	*Overall (%)*
Patient-specific factors				
LAD and/or circumflex artery from sinus 2	11	11	13	34
VSD present	4	4	5	13
Aorta and PA side-by-side	3	0	2	5
Procedural factors				
Phenoxybenzamine used	33	23	34	87
Aprotinin used	7	34	3	42
Fibrin glue	17	33	7	55
Two atrial cannulas used	3	11	0	13
Less than one cardioplegic infusion	18	14	1	32
Less than one patch for PA repair	33	23	0	54
Extensive coronary mobilization	35	28	0	61
Lecompte maneuver	29	34	35	94
Outcome				
Died	1	7	1	9
Back on bypass (died)	3 (1)	8 (4)	4 (0)	16
Renal failure	9	6	1	18

LAD, Left anterior descending; *PA*, pulmonary artery.

Table II. *Logistic regression for outcome of early death*

	Univariate		*Multivariate*					
	p *Value*	*Relative risk*	p *Value*	*Relative risk*	p *Value*	*Relative risk*	p *Value*	*Relative risk*
Patients 1-35		1		1		1		1
36-70	0.052	8.50	0.043	9.88	0.17	4.8	0.14	5.77
71-104	0.982	1.03	0.947	0.91	0.85	1.33	0.91	1.18
Coronaries								
Cx sinus 1		1		1				1
Cx sinus 2	0.041	4.55	0.025	5.96			0.044	5.37
No phenoxybenzamine		1				1		1
Phenoxybenzamine given	0.001	0.086			0.041	0.172	0.06	0.19

Cx, Circumflex.

coronary arteries were left in place so that the medial margin of the cuff remained loosely attached to the pulmonary artery. The incision in the pulmonary artery was made alongside this line of attachment and then extended centrally so as to make a trapdoor flap. This resulted in a reduction of the anteroposterior displacement of the coronary artery. The cuff was then tucked into the trapdoor taking wider bites on the cuff than on the neo-aorta thus producing a tubular extension to the coronary artery. Rather than moving the coronary arteries from the aorta and transferring them to the pulmonary artery, the aorta was taken away from the coronaries and the pulmonary artery was brought to them (Fig. 1). The neo-pulmonary artery was reconstructed with a generous trouser-shaped patch of autologous pericardium and the Lecompte maneuver was done in all cases.

Table III. *Logistic regression for outcome of death or near miss*

	Univariate			*Multivariate*		
	Relative risk	*95% confidence*	*limits*	*Relative risk*	*95% confidence*	*Limits*
Patients						
1-35	1			1		
36-70	4.89	1.23	19.47	1.07	0.17	6.88
71-104	1.84	0.40	8.38	3.41	0.52	22.18
Coronaries						
Cx sinus 1	1			1		
Cx sinus 2	3.49	1.25	9.75	4.10	1.30	12.92
No phenoxybenzamine	1			1		
Phenoxybenzamine given	0.31	0.09	1.08	0.55	0.12	2.55
No aprotinin	1			1		
Aprotinin given	4.93	1.60	15.19	9.06	1.23	66.65

Cx, Circumflex.

Phenoxybenzamine was used in all cases, whereas the administration of aprotinin and the application of fibrin glue were abandoned.

Methods of analysis

Outcome measures. We explored the suggestion that outcome events other than death could be considered as failures and used in the same way as early death to monitor the quality of performance and the risk factor analysis. In this context we adopted the concept of "near misses" familiar from the aviation industry. The criterion chosen as having near-miss analogies was the need to reinstitute cardiopulmonary bypass after a trial of weaning. Other outcome measures reflecting postoperative morbidity were also reviewed. The need for postoperative peritoneal dialysis was compared across the experience in an attempt to define safety margins under different surgical protocols.

Identification of trends

CUMULATIVE SUM PROCEDURE (CUSUM). We explored the use of the CUSUM procedure developed for monitoring quality on a production line and occasionally used in medical practice, which avoids the well-known problem associated with repeated significance testing.[2-4] First, lower and upper targets must be specified; these were taken as 2% and 5% for mortality and 5% and 10% for death or near miss. Boundaries can then be constructed such that if the cumulative number of failures crosses the upper boundary then a significant departure from the lower target value is concluded, whereas crossing the lower boundary indicates a significant departure from the upper target value. The position of a boundary line corresponds to a chosen level of statistical significance: 0.2 and 0.05 were selected to correspond to an "alert" and an "alarm." Thus when the sequence crosses above the "alert" line one can be 80% sure and above the "alarm" line one can be 95% sure that the lower failure rate had been transgressed. Similarly, there are two boundary lines that provide reassurance (see appendix).

When an upper boundary is crossed, action is taken aimed at rectifying the unacceptable failure rate. The boundaries are then redrawn with the start point at the protocol change.

COMPARISON WITH MULTICENTER DATA. The large multiinstitutional prospective study of patient, support, procedural, and institutional risk factors for death after the arterial switch operation reported in 1992 by Kirklin and associates[5] from the Congenital Heart Surgeons Society (CHSS) provides an ideal basis for the surgeon to monitor his or her performance. The findings of the study can be summarized in an equation that provides a probability of death for a specific patient with simple transposition when the coronary artery anatomy and the experience of the institution where the operation is done are known. With our hospital categorized as a "low-risk" institution, the actual number of deaths encountered was compared sequentially with the number predicted from the CHSS equation.

Risk factor analysis. The series was arbitrarily divided into thirds: patients 1 through 35, patients 36 through 70, and patients 71 through 104. The prevalence of several anatomic and procedural variables in each of the three eras was documented. Data about anesthetic and assisting personnel were also collected. The influence of factors such as fatigue (surgical performance after sleepness night for transplantation work) was also examined. None of these were found to have any association with the deaths and none are shown. Associations between early death and era of operation, coronary arterial anatomy (left anterior descending or circumflex coronary artery from sinus 1 or sinus 2), use of phenoxybenzamine, and other variables were estimated for each possible risk factor separately. Relevant variables were then explored jointly with multiple logistic regression. The distribution of potential risk factors across the experience was examined to try to establish whether the cluster of failures could be accounted for by a cluster of high-risk cases or by an observable procedural factor (Tables I through III). Associations were expressed as relative risks with 95% confidence intervals.

Expression of "current mortality." A running estimate of the current failure rate was calculated conventionally and also with the use of an "exponentially weighted moving average," in which previous observations are systematically downweighted. It is as if there is a memory loss as each new case is observed. After some exploration, the rate of "memory loss" at each case was taken as 7% so that, for example, the experiences with the tenth previous patient carried about half the weight of the last patient seen. A standard error for the estimate is calculated under the same weighting.

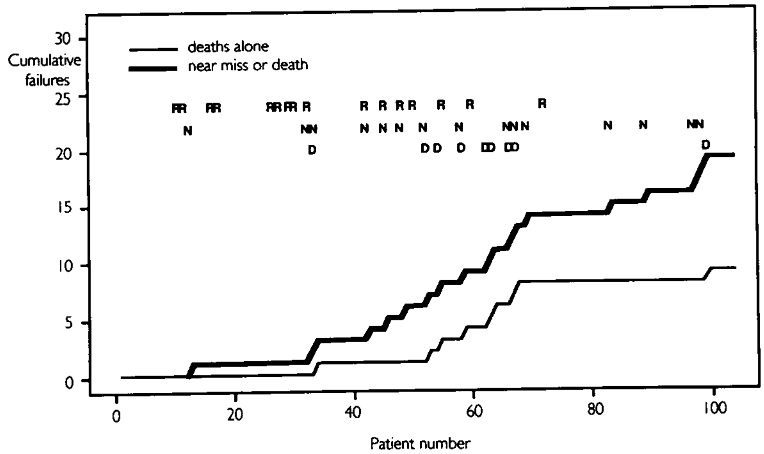


Fig. 2. Cumulative failure. "Failure" as death or death/near miss. *R*, Peritoneal dialysis; *N*, near miss; *D*, death.

Results

Outcome measures. There were nine deaths in the series of 104 patients. Of the 15 patients who needed secondary support with cardiopulmonary bypass, five died within 24 hours and two died several months later with severe left ventricular impairment. Peritoneal dialysis was required in 16 of 95 patients who survived at least 24 hours; none subsequently died. The whole experience viewed retrospectively as a sequence of successes and failures is summarized in Fig. 2. Early death and near-miss events are plotted cumulatively against patient number. On examination of the deaths there appear to be three distinct phases in the experience. There is an initial phase (patients 1 to 52) in which mortality is low (one death in 52 patients). Subsequently a cluster of deaths appears between patients 53 and 68. The chance of seven deaths occurring in 16 patients if the mortality rate truly was 2% is 0.00000001, for 5% is 0.000006, for 10% is 0.0004, and for 20% is 0.02 (binomial). In the third phase, mortality has fallen again toward the original low level (one death in 36 patients). The rise in cumulative failure rate for the category "death or near miss" occurs before the category "death." Near misses (taken as a need for reinstituting cardiopulmonary bypass) continue to be a feature in the third, low-mortality phase. In the third phase, the need for postoperative peritoneal dialysis has decreased below the level of that of the previous two phases.

Detection of trends. The CUSUM method of sequential monitoring of outcomes is shown in Fig. 3, *A* and *B*. After the third death in the cluster (patient 59), the surgeon could have been 80% sure that performance had deteriorated and that the lower limit of mortality had risen above 2%, and 95% sure after the death of patient 64. Fig. 3, *B*, reviews "failures" defined as deaths or near misses, expecting a 5% to 10% "failure" rate. With this configuration, the surgeon would have been "alerted" at patient 49 and "alarmed" at patient 59 by which time he would have been 95% sure that a deterioration in performance had occurred.

After the surgeon has been retrained, the CUSUM boundaries are reset and the series is currently approaching the first of the lower boundaries. If no further deaths occur, at patient 130 the surgeon will be 80% sure that his long-run mortality rate is below 5%.

Comparison with the CHSS database. Performance early in the experience justified classification with low-risk institutions in the CHSS study. At patient 50 only 1 patient had died, compared with the 3.5 predicted by the CHSS equation. However by the time of the death of patient 59, the observed and predicted numbers of deaths were equal, and by patient 70 there were 3.7 more deaths than predicted. For the whole experience, observed and expected deaths are again approaching each other. Viewed since surgeon retraining, mortality is less than predicted by the CHSS equation, as pertained before the cluster of failures (Fig. 4).

Risk factor analysis. Table I documents the distribution of patient-specific, procedural factors and outcome measures within the experience divided into three equal eras. All the anatomic and procedural factors were explored as predictive of death and death or near miss.

Table II summarizes the logistic regression model for the whole experience; variables with no predictive value on univariate analysis are not shown. The emerging model suggests an excessive risk for patients with the origin of the circumflex coronary artery from sinus 2 and a protective effect of phenoxybenzamine. After patient 59, when

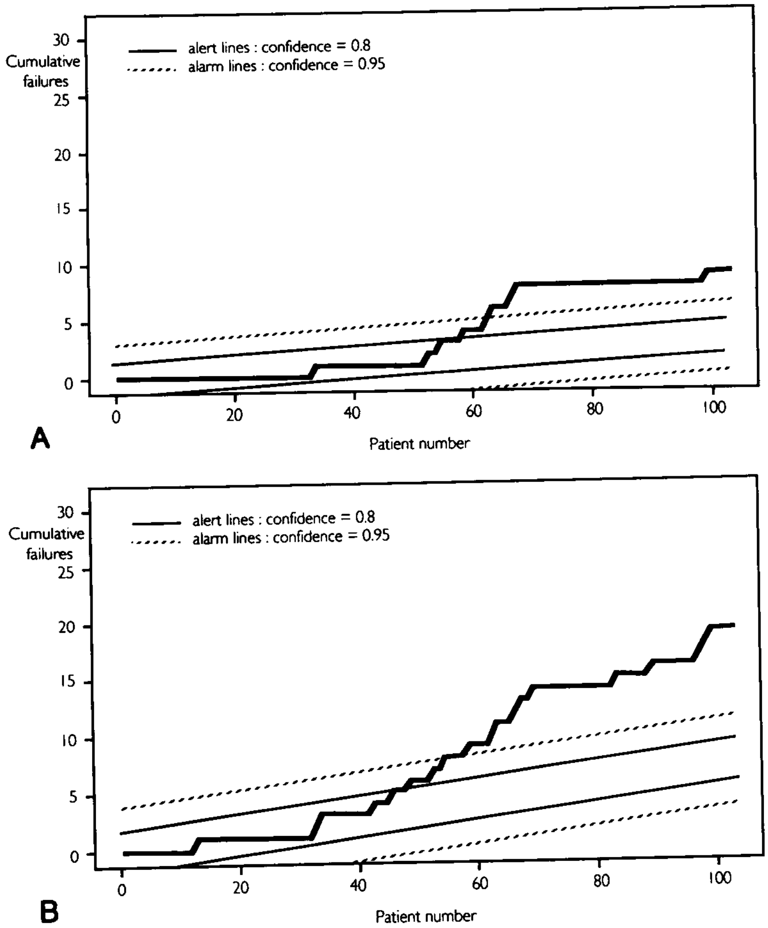


Fig. 3. Cumulative failure rate with CUSUM boundaries. Alert lines correspond to 80% confidence; alarm lines to 95% confidence (see text). **A,** Death: targets 2% and 5%. **B,** Death or near miss: targets 5% and 10%.

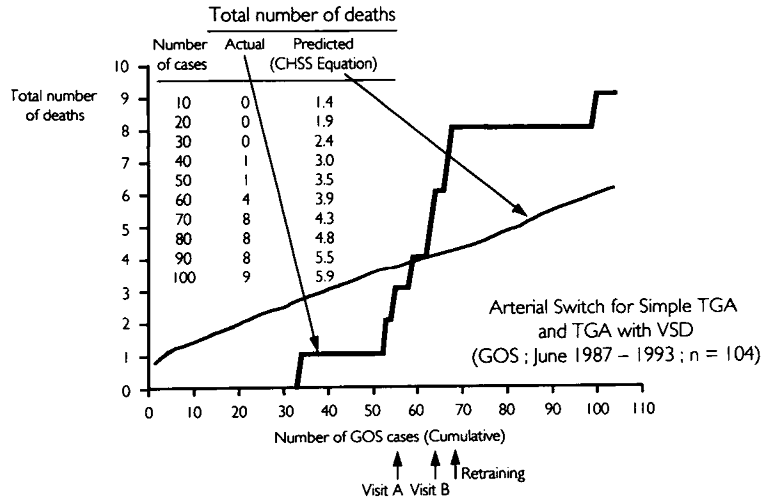
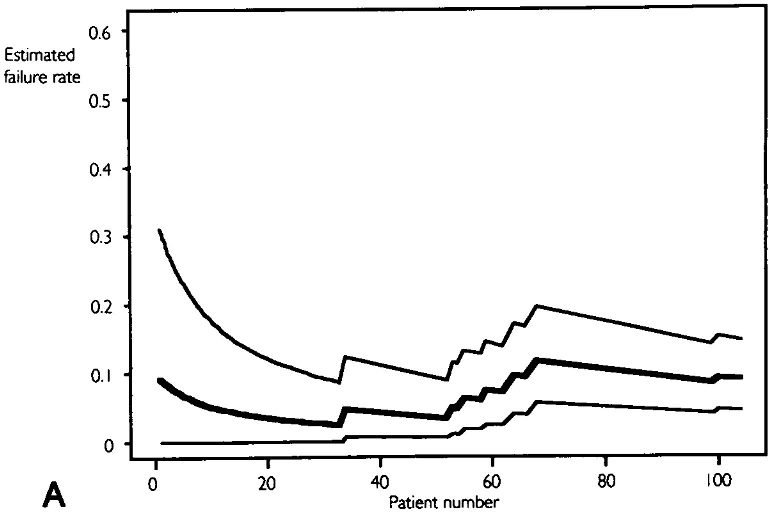


Fig. 4. Cumulative failure rate: comparison of actual and predicted (from CHSS equation) number of deaths. *GOS,* Great Ormond Street.



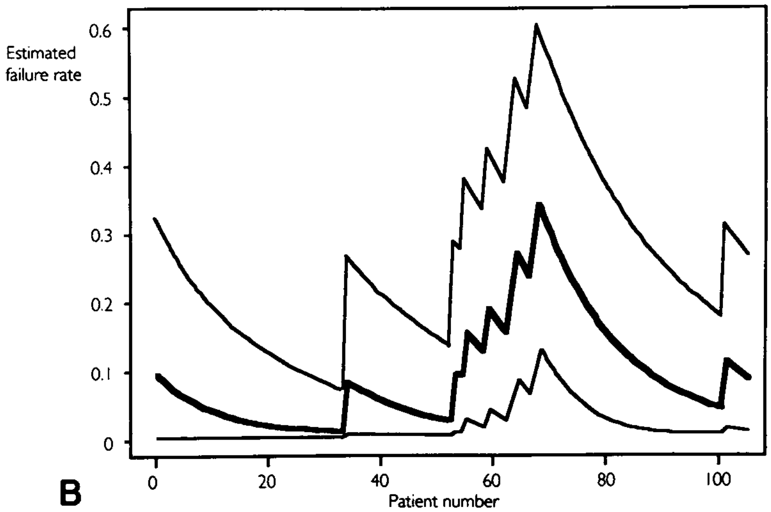


Fig. 5. Serial estimates of failure rate with 95% confidence limits. **A,** With no "memory loss." **B,** With down-weighting at 7% per case.

the surgeon could have been almost certain that a deterioration in performance had occurred, none of the possible risk factors explored (Table I) would have emerged as predictive of early death. However, by patient 68 (the time of retraining), the influence of coronary anatomy and the protective effect of phenoxybenzamine would have been demonstrable.

Table II also explores how far the distribution of the risk factors identified explains the large number of deaths in the middle third of the experience. When the patient-specific risk factor (coronary anatomy) and the procedural risk factor (failure to use phenoxybenzamine) are added individually to the era information, the scatter of higher-risk patients through the experience does not account for the excessive risk of the middle era whereas the failure to use phenoxybenzamine is associated in time with the cluster of failures and appears to explain about half the relative risk associated with the middle era. The remainder of the risk associated with the middle era is not accounted for by variables included in the analysis.

Table III shows the logistic regression model with the outcome of death or near miss. Coronary anatomy and era are again relevant in predicting this outcome; the use of phenoxybenzamine is not a powerful predictor and the use of aprotinin emerges as an important risk factor.

Mortality estimates along the experience. The point estimates of mortality reviewed along the experience are shown in Fig. 5. The estimated failure rate using the whole series (Fig. 5, *A*) and with "memory fading" at 7% per case (Fig. 5, *B*) is illustrated. The running estimate, which apportions equal weight to every case in the experience, is relatively insensitive to a cluster of deaths. According to the estimate that downweights the relevance of previous cases, the "current estimate" of mortality at patient 33 would have been 1% (0% to 7%), at 68 would have been 34% (12% to 57%), and currently is 8% (2% to 25%).

Discussion

The learning curve describing a decline in risks as a function of cumulative experience has been widely recognized and well documented in surgical apprenticeship. The occurrence of a cluster of failures beyond the learning curve appears to be uncommon or at least much less publicized.

A cluster of failures can be simply a matter of chance: a run of good luck can be followed by a run of bad luck. Alternatively, something that is not necessarily immediately apparent may account for the shift in failure rate and require investigation and remedial measures.

Unfavorable trends in outcomes after cardiac operations can be related to some specific procedural risk factors, to some less specified part of the total surgical protocol (e.g., anesthetic protocol, support system protocol, postoperative protocol), to unmonitored elements of the overall procedure (introduction of a new piece of equipment or a new drug), or to a deterioration of the surgeon's own performance (human error).

Audit aimed at monitoring performance is best organized prospectively, if any deterioration is to be recognized promptly. Sequential monitoring methods chosen in most situations will be designed to detect gradual trends; a gradual deterioration in performance is more difficult to perceive than a cluster of failures. Two of many possible methods of sequential monitoring are illustrated in this paper. Both use the concept of a target value or an agreed standard. Comparison of the observed and expected deaths predicted from the CHSS database has the advantage of incorporating knowledge available about the influence of recognized risk factors for a procedure. The CUSUM method has the advantage of avoiding the need to address a computer or calculator after each case and offers a statistical statement as a boundary is crossed. More sophisticated methods of sequential monitoring are available and may be more suitable for detection of more abrupt changes in failure rate: these include on-line detection of change points in failure rates and using estimate of the rate of change of failure rate as patients accumulate.[6]

Near misses: a refinement in monitoring surgical performance? The concept of near miss was examined to find out whether the early favorable results were misleadingly optimistic and whether a more sensitive method of outcome assessment could have alarmed the surgeon earlier. Although commercial jet transport aircraft are lost at a rate of no more than 1/1,000,000 flying hours, errors in aviation have been investigated more thoroughly than errors in any other sort of endeavor.[7] Near misses are routinely reported to the Civil Aviation Authority and analyzed for the purpose of furthering flight safety. The degree of risk inherent in each incident is assessed, trends are analyzed, and recommendations for remedial action are made.

Taking the need to reinstitute cardiopulmonary bypass as a near miss (failure equivalent), the surgeon would have crossed the alarm line earlier. Although deaths have reverted to the original low level after retraining, near misses continue to occur, so there appears to be some remaining hazard associated with the origin of the circumflex or the left anterior descending coronary artery from sinus 2 that has not been annulled.

It would be misleading to incriminate surgical techniques alone in the need to reinstitute cardiopulmonary bypass. One cannot lose sight, for instance, of the particularly vulnerable position of an unprepared left ventricle on its pressure-volume function curve at the end of the operation. However, none of the patients in whom cardiopulmonary bypass had to be reinstituted has died since surgeon retraining.

Introducing the need for postoperative peritoneal dialysis as an outcome measure discloses that the early experience was less straightforward than it first appeared and that in this respect the recent experience is more uneventful. The concept of near miss could perhaps be used to define a safety index for a surgical procedure, a surgeon, or an institution.

Procedural risk factors. Exploration of risk factors for early death suggested an excessive risk for patients with the origin of the left anterior descending or the circumflex coronary artery from sinus 2. This concurs with results of the CHSS study in which that coronary artery pattern was the only risk factor related to the malformation of TGA.[5] The potential difficulties in transferring that subset of coronary arteries was described very early in the arterial switch experience.[8, 9] The translocation of the coronary arteries imposes an anteroposterior displacement and a rotation of the vessel. The latter is along the horizontal axis when the coronaries arise from the lateral aspect of the aorta and along the sagittal axis when the coronaries arise from the posterior aspect of the aorta. We believe that the technique recently used reduces the risk of coronary artery distortion. The excision of a smaller aortic wall cuff, a more extensive dissection of the proximal coronary arteries, and the creation of a circular defect or a slitlike incision in the neo-aorta impose more displacement to the coronary artery with potential risks of elongation and torsion.

The apparent protective effect of phenoxybenzamine is not clearly understood. Could phenoxybenzamine prevent coronary spasm induced by the manipulation of the coronary arteries? What is the effect of phenoxybenzamine on a heart that is essentially denervated by the transection of the great arteries? Does phenoxybenzamine improve cardiac performance by reducing afterload? One can only postulate that the specific procedural alterations introduced after retraining contributed to the improvement of the results. Furthermore, non-use of phenoxybenzamine, multiple doses of cardioplegic solution, and transfer of the coronary arteries into circular openings of the pulmonary artery do not preclude excellent results.[9, 10]

Estimation of mortality during a cluster of failures. It is common practice for a surgeon to provide his or her patients with an estimate of the risks of a particular procedure. How should this risk be calculated in the middle of a cluster of failures? The surgeon addressing the parents of, say, patient 65 could have explained that 4 of the last 10 (40%) or that 6 of the last 64 (9.4%) of similar patients had died, both statements being true. To what extent estimates should be influenced by local rather than remote experience is obviously a matter of memory and judgment. Estimation of failure rate, allowing for "memory fading" for each case, discounts remote results and emphasizes current experience. The rate of discounting greatly influences the mortality estimates: at patient 65 the estimated mortality with no "memory loss" would be 11%, with memory loss of 5% per case 21%, with 10% per case 30%, and with 20% per case 48%.

Dealing with suboptimal performance. There is a wide acceptance of the hypothesis that, other things being equal, the quality of care improves with the experience of those providing it. This has been confirmed for the arterial switch operation in a recently published multiinstitutional report. Increasing institutional experience was associated with improvement in results. The surgeon in the present study would have quickly qualified for being included in the low-risk group and he had operated on enough patients to feel confident that he was mastering the procedure. In the absence of an immediately apparent explanation, it was therefore appropriate that he should question his own performance and incriminate human error during a cluster of failures. Retrospectively, this was supported by the risk factor analysis that did not account for half of the risk associated with the era of deaths.

Surgical performance is dependent on integrated cognitive, physical, and affective skills. As in any other skilled handicraft, increased experience leads to better technical performance. Akin to the musician, the surgeon needs practice to maintain peak performance and the more critical the procedure the more important it is that this practice is provided. In the present study, however, performance paradoxically deteriorated despite constant practice.

Retrospectively, it is likely that if the surgeon had developed a process of assessing his performance with outcome criteria more sensitive than death he could have noticed that his early results were falsely optimistic and he would have been motivated to take steps to improve them.

When a performance problem has been identified, how is it best to deal with it? This is a delicate issue that imposes serious soul searching. Defensive thinking dominates many aspects of our social and professional lives. Incompetence is tolerated at many levels because dealing effectively with it may lead to nightmares of legal entanglements with employers, families, colleagues, unions, grievance committees, threats of suits, and so forth. It is the surgeon's duty not to be intimidated by these threats and to demonstrate the candor and the fortitude necessary to stand up to this pressure. Unless specific aspects of a surgical procedure have been identified as the likely causes of the failures, it can be hazardous and counterproductive to introduce elements of management protocols borrowed from other institutions. After his visits to a low-risk institution after patients 55 and 64 died, the surgeon in this series had a tendency to abandon the use of phenoxybenzamine, which, retrospectively, proved to be a protective factor in his own experience. In the absence of identified procedural risk factors, retraining implies therefore the adoption of another management protocol in its smallest details.

Could subtle alterations in cognitive, psychomotor, and/or affective functions be incriminated in a cluster of failures? By analogy with the learning curve one could postulate the existence of a "falling curve" that could, for example, be related to age. There is general literature on the effects of aging on skilled performance in which slips and lapses become more likely as skilled behavior and motor subroutines are less actively monitored. However, it is often considered that increased experience counterbalances any technical decline.

The medical literature is sparse on failing surgical skill in the aging surgeon. Harvey Cushing, before going to the Brigham Hospital, wrote to Henry Christian on November 20, 1911:

> Why not put the surgical age of retirement for the attending surgeon at 60 and the physician at 63 or 65, as you think best? I have an idea that the surgeon's fingers are apt to get a little stiff and thus make him less competent before

the physician's cerebral vessels do. However, as I told you, I would like to see the day when somebody would be appointed surgeon somewhere who had no hands, for the operative part is the least part of the work."*

Going back to the analogy with the aviation industry, in 1960 the Federal Aviation Administration fixed the retirement age of airline pilots at 60 years to prevent risks of physical incapacitation and hazards caused by decreased cognitive abilities. Both issues are currently strongly challenged.[11] Numerous recent data suggest that with practice and experience continuous psychomotor performance can be maintained at a high level well into old age.[12] Those data showed that practice eliminates age differences, the higher liability being associated with inexperience. The accident rate for pilots with extensive practice decreases with age. More important, differences in cognitive performance associated with aging are small relative to the total range of individual differences. Older experienced persons are known to use effective compensatory mechanisms to offset age-related deficiencies and increasing pilot age and experience is correlated with increasing safety.

Inferences and conclusions

1. A system of continuous monitoring of surgical performance can usefully be instituted to allow early detection of unfavorable trends. An external standard can be chosen from recognized leaders in the field or taken from published data of multicenter studies.
2. Outcome measures other than death can be used as a refinement of quality-control methods.
3. Once unfavorable trends, unlikely to be due to chance alone, have been detected, it is suggested that an analysis of possible risk factors be made.
4. Should the risk factor analysis be inconclusive, the need for change of protocol may imply retraining.
5. The present experience demonstrates that retraining can reset the failure rate to its initial low level.
6. From the retrospective risk factor analysis of the whole experience, the following recommendations can be made for the arterial switch operation: use of phenoxybenzamine, removal of as large an aortic button with the coronary arteries as is possible, minimal mobilization of the coronary arteries that are implanted into a medially hinged trapdoor flap, and non-use of aprotinin.

We appreciate the provision of Fig. 4 by Drs. John W. Kirklin and E. H. Blackstone, University of Alabama at Birmingham, Birmingham, Alabama.

*From Fulton JF. Harvey Cushing. A biography. The classics of medicine library. New York: Gryphon Editions, 1991:351-2. Reprinted with permission.

REFERENCES

1. Lecompte Y, Neveux JY, Zannini L, Tu TV, Duboys Y, Jarreau MM. Reconstruction of the pulmonary outflow tract without prosthetic conduit. J THORAC CARDIOVASC SURG 1982;84:727-33.
2. McPherson K. Statistics: the problem of examining accumulating data more than once. N Engl J Med 1974; 290: 501-2.
3. Kenett R, Pollak M. On sequential detection of a shift in the probability of a rare event. J Am Stat Assoc 1983;78:389-95.
4. Williams SM, Perry BR, Schlup MMT. Quality control: an application of the CUSUM. BMJ 1992;304:1359-61.
5. Kirklin JW, Blackstone EH, Tchervenkov CI, Castaneda AR, and the Congenital Heart Surgeons Society. Clinical outcomes after the arterial switch operation for transposition: patient, support, procedural, and institutional risk factors. Circulation 1992;86:1501-15.
6. Pollak M. Optimal detection of a change in distribution. Ann Stat 1985;13:206-27.
7. Green R. Human error on the flight deck. Philos Trans R Soc Lond 1990;B-327:503-12.
8. Yacoub MH, Radley-Smith R. Anatomy of the coronary arteries in transposition of the great arteries and methods for their transfer in anatomical correction. Thorax 1978; 33:418-24.
9. Quaegebeur JM, Rohmer J, Ottenkamp J, et al. The arterial switch operation: an eight-year experience. J THORAC CARDIOVASC SURG 1986;92:361-84.
10. Planché C, Bruniaux J, Lacour-Gayet F, et al. Switch operation for transposition of the great arteries in neonates: a study of 120 patients. J THORAC CARDIOVASC SURG 1988; 96:354-63.
11. Moser R. Is the "age 60 rule" still appropriate? Med Aeronautique Spatiale 1991;30:43-6.
12. Tsang PS. A reappraisal of aging and pilot performance. Int J Aviation Psychol 1992;2:193-212.

Discussion

Dr. Serafin Y. DeLeon *(Maywood, Ill.).* I noticed in your Fig. 1 that the coronary arteries end up behind the pulmonary arteries. There are two ways of deciding where the coronary arteries should be implanted. Before bypass the site can be marked on the pulmonary artery or during bypass the site can be marked after transection of the pulmonary artery and aorta.

We noticed early in our experience that with the technique of eyeballing where the coronary arteries should go after the aorta and pulmonary artery have been transected we had some coronary arteries ending up behind the main pulmonary artery, which made these vulnerable to any episode of pulmonary hypertension.

Currently we mark where to implant the arteries before cardiopulmonary bypass is instituted. With this technique the coronary arteries consistently end up on the lateral portion of the neo-aorta and therefore are not sandwiched between the main pulmonary artery and the aorta. The coronary arteries will be less vulnerable to any episode of pulmonary hypertension. Additionally, older children may be less susceptible to myocardial ischemia with exercise, similar to what happens in cases of aberrant left coronary artery located between the aorta and pulmonary artery.

It is interesting that when we reviewed the literature and we looked at various illustrations of arterial switches, a lot of them showed the coronary arteries located behind the main pulmonary artery. Early in our experience we also used to reconstruct the main pulmonary artery with synthetic materials, which probably added to coronary artery compression. We had some unexplained deaths early in our experience.

Dr. William G. Williams *(Toronto, Ontario, Canada).* Mr. Marc de Leval is no stranger to setting precedents and I think he set another precedent with this paper, which presents a means of monitoring surgical results.

Having analyzed his own results, the obvious implication of this work is to extend the same analysis to an institution. The reality is that none of us can do this kind of work forever, and as we phase in new people we must assure ourselves of consistent performance by the institution as a whole. I wonder if Mr. de Leval could discuss how these analytic techniques could be applied to the entire institution in which he works and to other institutions?

Dr. Davis C. Drinkwater, Jr. *(Los Angeles, Calif.).* I want to add my congratulations to Mr. de Leval for being so honest. Not many people would do this. I think the paper points out the importance of visiting programs that are doing these cases so that one can learn by direct communication.

Looking back at the science and art aspect, was there something technical that was done differently in the first 50 cases? You had such a good success rate that I doubt skill as a surgeon per se would be a large factor. Could it have been a technical thing that was improved on or that was returned to through direct communication that has resulted in such excellent results in the last 50 cases?

Dr. Hillel Laks *(Los Angeles, Calif.).* It is interesting that there was no mention of myocardial protection. Could you comment on that?

Mr. de Leval. Dr. DeLeon, I do agree with you that there is a risk of coronary obstruction as a result of compression by the pulmonary artery. However, I do not believe that the surgeon can impose any position to the coronary arteries without taking into consideration their site of origin. If they arise, for example, between the great arteries on either side of the posterior commissure of the aorta, they must be translocated to a corresponding position behind the neo-pulmonary artery.

Dr. Williams, I did not look at the overall experience of our institution because I thought that I was dealing with a problem of individual performance. I do think, however, that using refined ways to look at surgical results, such as near misses, could help introduce the concept of safety index, not only for a surgeon but also for an institution or a particular operation. Going one step further, one could apply the near-miss concept to cost and define cost-efficiency indexes.

Dr. Drinkwater, the present experience has convinced me of the value of retraining and I would encourage surgeons to observe each other more often. This could become, for example, an extended function of the CHSS as a form of continuing education and assessment of surgical performance.

As far as the myocardial protection is concerned, Dr. Laks, we have looked into this in great detail and it did not come up as a risk factor. What we do now is use hypothermic cardiopulmonary bypass (18° C) with a single dose of cold crystalloid cardioplegia for up to 80 minutes of ischemic time. We also use topical cooling and do not use reperfusion solutions.

Appendix

Repeatedly doing significance tests as data accumulate can lead to uninterpretable conclusions.[2] "Sequential analysis" was invented during World War II for monitoring quality on production lines using schemes that guaranteed prespecified error rates despite testing after each observation, and one of the most popular techniques has been based on the cumulative sum (CUSUM) of failures.

Suppose we have a target failure rate of p_0 and would like to be alerted if the failure rate rose to an "alternative" p_1. We set a type I (false alarm) error rate α as the probability of incorrectly rejecting the target p_0 in favor of the alternative p_1, and a type II (false reassurance) error rate β as the probability of incorrectly rejecting the alternative p_1 in favor of the target p_0. Let $a = \ln\{(1-\beta)/\alpha$; $b = \ln\{(1-\alpha)/\beta$; $P = \ln\{p_1/p_0\}$; $Q = \ln\{(1-p_0)/(1-p_1)\}$; and $s = Q/(P+Q)$. We accumulate the total number of failures X after n operations. Then the upper "alert" boundary line is given by the formula $X_0 = ns - b/(P+Q)$, and the lower "reassurance" boundary line is $X_1 = ns + a/(P+Q)$.