



Reflections on outcome analyses: Introducing the concept of near misses

Marc R. de Leval*

International Congenital Cardiac Centre,
Harley Street, London, UK

*Email:
Marc.deleval@hcahealthcare.co.uk

Medical audit is a form of acquisition of new knowledge related to outcomes of medical treatments. Philosophers and scientists remain divided with regard to the acquisition of new knowledge. Some maintain that the basic purpose of a scientific theory is to predict the outcomes of an experiment. This view is called 'instrumentalism' as it implies a theory is no more than an instrument for making predictions. Positivism, which was the prevailing theory of scientific knowledge in the 1930s and remains present in the views of many philosophers, is an extreme form of instrumentalism, it holds that all statements other than describing or predicting observations, are not only superfluous but meaningless. The basic purpose of a scientific theory is not to explain anything but to predict the outcomes of experiments.

The opposite view is that the explanatory power of a theory is paramount and its predictive power only supplementary. We can explain in hindsight what has happened, but we cannot predict what will happen in the future. This is expressed by the Danish philosopher Søren Kierkegaard in his famous phrase: "Life can only be understood backwards; but it must be lived forwards." Predictions cannot be justified by observational evidence, as the British philosopher Bertrand Russell illustrated in his story of the chicken. The chicken noticed that the farmer came every day to feed it. It predicted that the farmer would continue to bring food every day. Then one day the farmer came and wrung the chicken's neck. This illustrates the impossibility to extrapolate observations unless one has already placed them within an explanatory framework. Prediction, even perfect, is simply no substitute for explanation. Facts cannot be understood just by being summarised in a formula. They can be understood only by being explained. Furthermore, many unpredictable things can be explained. It is understanding, not mere knowing, that is important.

In reality the two approaches should be complementary, embodying deep explanations as well as accurately predicting outcomes.

These reflections are applied to my analysis of surgical outcomes.

There are also two distinct philosophical approaches to surgical audit: the statistical and the forensic. The former is based on prediction; the latter is based on explanation. The statistical approach has dominated surgical audit. Major advances have been made in statistical methods applied to outcome analysis, outcome prediction and performance monitoring since the pioneering work of Florence Nightingale [Fig. 1]. However, the problems she identified 140 years ago associated with data collection, risk stratification, data interpretation and the use of a single outcome measure such as mortality remain problematic to this day. A purely statistical approach has its pitfalls, of which I will mention just two.

An exclusive reliance on statistical methods to acquire knowledge in outcome analysis has led to the development of extremely complex data sets and to the increasing use of subjective probability in those analyses. This tendency has culminated recently, for example, in the proposition of the so-called Aristotle score in paediatric cardiac surgery as a complexity-adjusted method to evaluate surgical results. The data set includes 76 primary procedures, 85 associated procedures, 6 age groups, and 81 procedure-independent factors; altogether, 248 variables. A complexity score is allocated to each variable and is made up of the sum of 3 factors graded from 1 to 5: prediction of

DOI: <http://dx.doi.org/10.5339/ahcsps.2011.12>

Published: 29 December 2011
© 2011 de Leval, licensee
Bloomsbury Qatar Foundation
Journals. This is an open access
article distributed under the terms
of the Creative Commons
Attribution-NonCommercial license
CC BY-NC 3.0 which permits
unrestricted non-commercial use,
distribution and reproduction in
any medium, provided the original
work is properly cited.



Figure. 1 Florence Nightingale. Reproduced from Project Gutenberg eText 13103: “Great Britain and Her Queen”, by Anne E. Keeling <http://www.gutenberg.org/dirs/1/3/1/0/13103/13103-h/13103-h.htm>.

hospital mortality, length of ITU stay, and technical difficulty. The complexity score was based on subjective probability and established by a panel of over 50 experts from 23 countries.

Subjective probability, which is based on expert opinions, was very fashionable in the 1950s when up to 70% of North American PhD graduates in mathematics applied for a job at the ‘RAND Corporation’ (Research AND Development). This was the think tank par excellence, initially working mainly in the air force. It later diversified to offer services in other fields such as technology and economy forecasting. The pitfalls of expert opinion have been recognised. This applies to medical outcome analysis. As a personal example, at our weekly multidisciplinary medical conference we used to predict the risk of postoperative mechanical support for patients placed on the surgical waiting list, in order to rationalise the use of the available facilities such as intensive care unit beds. The outcome of this practice was reviewed for 157 patients. 141 were predicted to be low risk, 12 a moderate risk, and 4 a high risk. Seven of these patients required mechanical support, 6 were in the predicted low risk, 1 in the predicted moderate risk, and none in the predicted high risk. This gave a positive predictive value of 7%.

The use of expert opinion has been bereft of methodological guidance. The most important function of expert opinion in science is the representation of uncertainty. Opinion is by its very nature uncertain. Hence when expert opinion is used as input in a scientific inquiry, the question to be addressed is simply: is the uncertainty adequately represented? [1]

Statistics interrogate the phenotypes of outcomes (who, what, when, where, how many), a forensic analysis investigates their genotypes (how, why).

The forensic approach to surgical outcomes was pioneered by Ernest Amory Codman (1869–1940) [2].

Codman an extraordinary Boston surgeon who, in 1900, introduced the ‘End Result Card System’ on which were entered the presenting symptoms, the diagnosis, the treatment, the complications, and the result one year later. It also included why perfection had not been obtained, with a classification of errors and adverse events. In 1914 Codman resigned from the staff of the Massachusetts General Hospital (MGH) in protest at their promotion system not being based on performance. In 1915 he caused uproar at a public meeting by unveiling an 8 feet cartoon satirising the Boston medical world represented by humbugs grabbing the golden eggs produced by a large ostrich depicting the general public with its head buried in Beacon Hill [Fig. 2]. The Trustees of the hospital look on and say: “If we let her know the truth about our patients, do you suppose she would still be willing to lay?” The ostrich says: “If I only dare look and see, I might find a doctor who could cure my own ills.” Codman was dropped shortly afterwards from the position of Instructor in Surgery that he was still holding at Harvard. In 1917, speaking at the American College of Surgeons to promote standardisation of case records, he said: “The system is perfectly simple, the only difficulty with it being its revolutionary simplicity. It requires straightforward answers to these questions: What



Figure. 2 Codman cartoon from 1915. Reproduced from http://www.massgeneral.org/News/assets/images/102811codman_cartoon.jpg.

was the matter with the patient? What did the doctor do to him? What was the result? If the result was not good, what was the reason? Was it a fault of the doctor, the patient, the disease, or the hospital organisation or equipment?" Codman had recognised that explanation was indispensable in understanding and improving surgical outcomes and that negative outcomes could be as a result of human, organisational and equipment failures.

Progress in healthcare in general, and in cardiac surgery in particular, has been characterised by a progressive neutralisation and sometimes abolition of risks related to patients and treatments. In the arterial switch operation for transposition of the great arteries (TGA) as an example, coronary arterial anatomy is no longer a risk factor in many institutions. This is also the case for the size of the ascending aorta in the Norwood operation. The impact of aortic cross-clamping time on outcome has been greatly reduced by improvement in the techniques of myocardial protection. Consequently, outcomes depend increasingly on what I will call 'human factors' such as who did the operation, in which institution, under which circumstances, etc. This trend has long been recognised in high technology industries such as aviation, as illustrated by this statement by Foushee: "The only way to produce dramatic improvements in aviation safety is through increased emphasis on human factors" which are more amenable to forensic rather than statistical analyses.

A few years ago we embarked upon a research project aiming at analysing the impact of human factors on the outcome of the neonatal arterial switch operation (ASO) [3]. This was a multi-institutional study undertaken in collaboration with James Reason, 'father' of the organisational accident theory. All 243 neonatal ASO for TGA, with or without ventricular septal defect (VSD), performed by all 21 UK paediatric cardiac surgeons in 16 institutions over an 18-month period were entered into the study. The operations were observed by a human factors researcher who attended the operation from induction of anaesthesia to transfer of the patient to the intensive care unit. A detailed description of the operation was written down as the procedure was taking place. This included information on individual and team performance, information on communication with each team and between different teams, as well as situational and organisational data. From these reports we extracted errors or failures which were categorised as minor or major negative events. Minor events were failures that disrupted the surgical flow of the procedure but in isolation were not expected to have serious consequences for the safety of the patient. Major events, on the other hand, were failures that were likely to have serious consequences for the safety of the patient. Major and minor events were judged to be either compensated for or uncompensated. An event was deemed to have been compensated for when the patient recovered or when the consequences of the event were negated by appropriate actions. The clinical outcomes were divided into 4 categories: (1) extubation within 72 hours with no complications, which occurred in nearly half of the patients; (2) prolonged intubation or reversible morbidity, which occurred in just over a quarter of the patients; (3) survival with near-misses, which included major ischaemic or haemodynamic problems at the end of the operation, extracorporeal membrane oxygenation (ECMO) support, severe arrhythmias, deep-seated infections, and permanent neurological damage which occurred in 17.7%; (4) death, which occurred

in 6.6% in this population. With grouping of near-misses and death, the incidence of negative outcomes was 25%.

The impact of these events on outcomes over and above patient-related and procedural variables was then analysed. The results can be summarised as follows. Uncompensated major negative events have serious clinical consequences. However, because of their immediate potentially dramatic consequences, in most cases major events are quickly noticed and appropriate compensation can prevent their leading to a number of catastrophes. Minor negative events, on the other hand, are much more insidious. Because of their seemingly benign nature they sometimes go unnoticed and little effort is made to compensate. The main feature of minor events is their multiplicative effect. In isolation they have little impact, but their multiplication has a strong relationship to negative outcomes.

Human factors experts have proposed paradigms and theories to explain the dynamics of poor outcomes. The 'Swiss cheese' model by James Reason [4] is often used as a conceptual framework of organisational or system failure in healthcare. According to this theory, accidents in hazardous technologies often arise from the concatenation of latent and active failures and the breach of defence mechanisms. It is useful to depict complex systems such as healthcare as having a sharp and blunt end. At the sharp end, practitioners interact with the hazardous process in their roles as nurses, physicians, technicians, pharmacists and the like. At the blunt end of the system are regulators, administrators, economic policymakers, technology suppliers, etc. The blunt end of the system controls the resources and constraints that confront the practitioner at the sharp end. High-technology systems have many defence layers: some are engineered (alarms, physical barriers, automatic shutdowns, etc.), others rely on people (surgeons, anaesthetists, pilots), and others depend on procedures and administrative controls. Their function is to protect potential victims and threats from local hazards. In an ideal world each defence layer would be intact. In reality, however, they are more like slices of Swiss cheese, containing many holes. The presence of holes in any one slice does not normally result in a bad outcome. Failure occurs only if the holes in many layers are aligned so as to permit an error, a trajectory of accident opportunity, bringing hazards into damaging contact with the victims.

The holes in the defences arise for two reasons: active and latent conditions. Nearly all adverse events involve a combination of these two sets of factors. Active failures are the unsafe acts committed by people who are in direct contact with the patient or system. They take a variety of forms, such as errors and violations. Latent conditions are the inevitable resident pathogens within the system. They arise from decisions, from the designers, builders, procedure writers, and top-level management. Latent conditions have two kinds of adverse effect: they can translate into error-provoking conditions within the local workplace (for example time pressure, understaffing, inadequate equipment, fatigue, inexperience, etc.). Latent conditions may lie dormant within the system for many years before they combine with active failures and local triggers to create an accident opportunity. The following example is an illustration of an organisational accident that happened at Great Ormond Street Hospital, London, a few years ago.

A child who was an inpatient in a District General Hospital was due to receive chemotherapy under general anaesthesia at Great Ormond Street. He should have been fasted for six hours before the anaesthetic, but was allowed to eat and drink before leaving the local hospital. *Error number 1*: is a fasting error, a small negative event.

No beds were available for the patient on the oncology ward and he was admitted to a mixed specialty ward. *Error number 2*: lack of organisational resources.

The patient's notes were lost. *Error number 3*: loss of patient information.

The patient was due to receive intravenous vincristine to be administered by a specialist oncology nurse on the ward and intrathecal methotrexate to be administered in the operating theatre by an oncology specialist registrar. No oncology nurse specialist was available on the ward. *Error number 4*: communication failure between oncology and outlying ward.

Vincristine and methotrexate were transported together from pharmacy to the ward by a housekeeper instead of being kept separate at all times. *Error number 5*: drug delivery error due to non-compliance with hospital policy.

The housekeeper informed staff that both drugs were to go to theatre with the patient. *Error number 6*: incorrect information communicated by inexperienced staff.

The patient was consented only for intrathecal methotrexate and not for intravenous vincristine. *Error number 7*: consenting error.

A junior doctor abbreviated the route of administration to IV and IT, instead of using the full term in capital letters. *Error number 8*: poor prescribing practice.

When the fasting error was discovered, the chemotherapy procedure was postponed from the morning to the afternoon list. The doctor who was due to administer the intrathecal drug was on leave that afternoon and assumed that another doctor in charge of the ward would take over. No formal face-to-face handover was carried out between the two doctors. *Error number 9*: inappropriate task delegation.

The patient arrived in the anaesthetic room and the oncology senior registrar was called to administer the chemotherapy. However, that doctor was unable to leave his ward and assured the anaesthetist that he should go ahead and this was a straightforward procedure. *Error number 10*: goal conflict between ward and theatre.

The oncology senior registrar was not aware that both drugs had been delivered to theatre. *Error number 11*: situational awareness error.

The anaesthetist had the expertise to administer drugs intrathecally but had never administered chemotherapy. *Error number 12*: inappropriate task delegation.

He injected the methotrexate intravenously and the vincristine into the patient's spine and the patient died five days later. *Error number 13*: drug administration error.

It is the alignment of holes in each defence layer, each slice of the Swiss cheese which allowed the incident to occur.

Throughout this research we have established extensive contacts with high technology industries such as aviation and motor racing. A major step forward in safety initiatives in aviation was the introduction of crew resource management training in the 1980s. Crew resource management is about error management, which has three lines of defence: prevention, detection, and recovery. Current research places more and more emphasis on error prevention, on enhancing safety through anticipation and planning for unexpected events. Highly effective cockpit crews use one-third of their communications to discuss threats and errors in their environment, regardless of their workload, whereas poorly performing teams spend about 5% of their time doing so. Past successes should not be a reason for over-confidence. Safety is not a commodity but a value that requires continuous reinforcement and investment.

Error detection is the second line of defence. Our research has highlighted the impact of minor errors on outcomes. In Formula 1 motor-racing, 90% of the outcome of the race depends on the pitstop. Each pitstop is videotaped and observed by human factors experts, who score the errors. The highest score goes to the minor errors, those which are not detected by the performers. These are the minor events of our Switch study. Error recovery is the third line of defence. It includes tolerance, compensation and mitigation of adverse events.

Surgical outcomes are the result of complex interactions between an almost infinite number of variables related to patients, procedures, and care providers. As in any complex system, these interactions are non-linear. Mathematically, a linear relationship can be expressed as a simple equation in which the variables appear only to the power of one. In a non-linear system, the many variables in the equation vary with respect to one another by powers of two or greater. Non-linear systems do not obey the simple rules of addition. The whole is more than the sum of its components. In complex systems there is a high sensitivity to initial conditions. A tiny difference in the initial conditions can be amplified over time to enormous differences which, in the context of this discussion, could make the difference between life and death. Non-linearity produces feedbacks; the outcome of an effect goes on to trigger more changes, leading to unexpected behaviours. One of the most intriguing aspects of complex systems is that changes do not have to be related to external causes. We are capable of self-organisation. Non-linear interdependent dynamics can create patterns, coherence, networks, copying, synchronisation, synergy, etc.

Complexity theories have been applied to systems failures by Dietrich Dorner in his thought-provoking book entitled: "The Logic of Failure: Why things go wrong and what we can do to make them right" [5].

Tiny differences in the initial conditions of a complex system can change the global outcome. As an example a paralysed left diaphragm following a Norwood procedure for a hypoplastic left ventricle requires prolonged ventilatory support, prolonged indwelling catheters that can result in a thrombosis of the left superior vena cava leading to a chylothorax, leading to total parenteral nutrition and septicaemia and eventually precluding the Fontan operation.

Outcome analyses have been used increasingly to detect and discipline under-performance. Although the differences in detecting and disciplining are subtle, their implications are important. Detection of under-performance implies that any suspicion of under-performance should ring an alarm and be acted upon, whereas disciplining requires near scientific certainty before action is taken. Detection implies candour, self-criticism an openness to blame, and can lead to unjust sanctions, whereas genuinely divergent behaviours may remain undetected with disciplining. Sequential monitoring methods are commonly used to detect gradual trends, gradual deteriorations in outcomes. In a previous study [6] we examined the concept of near misses as a more sensitive method of outcome assessments which could serve as a warning sign of under-performance before the occurrence of failures to allow preventive methods to be taken.

Near misses are routinely reported to the Civil Aviation Authority and analysed for the purpose of furthering flight safety. The degree of risk inherent in each incident is assessed, trends are analysed and recommendations for remedial action are made. In our initial study of a cluster of deaths following the Arterial Switch operation we demonstrated that if the need to reinstitute cardiopulmonary bypass had been considered as a failure equivalent (near miss), action could have been taken to prevent the occurrence of failures.

When a performance problem has been suspected, how is it best to deal with it? This is a delicate issue that imposes serious soul searching. Defence in 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. It is the surgeon's duty not to be intimidated by those threats and to demonstrate the candour and the fortitude necessary to stand up to this pressure. There is a need for the medical profession to enter into a dialogue with the stakeholders represented by consumers (patients and their families), healthcare managers, healthcare organisations, media, litigation lawyers and alike to agree on a compromise between two opposite tendencies. A criminal court type of judgement based on proof beyond reasonable doubt should be used to initiate disciplinary measures. A civil court type of judgement based on proof beyond the balance of probability should initiate action recognising the limitations of outcome analysis methodology.

In summary, outcome analysis should rely on statistical and forensic analyses. Through medical progress, negative outcomes have become increasingly less predictable and therefore should always remain explainable. Progress in outcome analysis has been a shift from prediction to explanation, a shift from statistical to forensic analysis.

References

- [1] Cooke R. Expert in Uncertainty. Oxford University Press. 1991.
- [2] Millenson M. Demanding Medical Excellence. The University of Chicago Press. 1997.
- [3] de Leval MR, Carthey J, Wright D, Farewell V and Reason J. Human factors and cardiac surgery: a multicentre study. *Journal of Thoracic and Cardiovascular Surgery*. 2000;119:661.
- [4] Reason J. Human Error. Cambridge University Press. 1990.
- [5] Dorner D. The Logic of Failure. Metropolitan Books. 1989.
- [6] de Leval MR, Francois K, Bull C, Brawn W and Spiegelhalter D. Analyses of a cluster of surgical failures. *Journal of Thoracic and Cardiovascular Surgery*. 1994;107:914.