Mortality alerts, actions taken and declining mortality: true effect or regression to the mean?

Perla J Marang-van de Mheen,[1] Gary A Abel,[2] Kaveh G Shojania[3]

[1]Department of Biomedical Data Sciences, Medical Decision Making, Leiden University Medical Centre, Leiden, Netherlands
[2]University of Exeter Medical School, Exeter, UK
[3]Department of Medicine, Sunnybrook Health Sciences Centre and the University of Toronto, Toronto, Canada

Correspondence to
Dr Perla J Marang-van de Mheen, Department of Biomedical Data Sciences, Medical Decision Making, Leiden University Medical Centre, Leiden, RC 2300, The Netherlands;
p.j.marang@lumc.nl

Linked

► http://dx.doi.org/10.1136/bmjqs-2017-007495
► http://dx.doi.org/10.1136/bmjqs-2018-008364

Check for updates

Alerts have become a routine part of our daily lives—from the apps on our phones to an increasing number of 'wearables' (eg, fitness trackers) and household devices. Within healthcare, frontline clinicians have become all too familiar with a barrage of alerts and alarms from electronic medical records and medical devices.

Somewhat less familiar to most clinicians, however, are the alerts received by institutions from regulators and other regional or national bodies monitoring healthcare performance. After the Bristol inquiry in 2001 in the UK,[1] research showed that given the available data Bristol could have been detected as an outlier and that it was not simply a matter of the low volume of cases.[2 3] Had the cumulative excess mortality been monitored using these routinely collected data, then an alarm could have given for Bristol after the publication of the 1991 Cardiac Surgical Register and could have saved children's lives.[4] Similar assertions have been made about detecting problems at Mid Staffordshire National Health Service Foundation Trust—that excessively high hospital standardised mortality ratios (SMRs) pre-dated the eventual recognition of exceptionally substandard care subsequently confirmed by other means.[5 6]

Following the Bristol inquiry, the UK implemented a national mortality surveillance system. This system alerts hospital trusts when they have higher than expected in-hospital mortality for at least one of 122 diagnosis/procedure groups, using cumulative sum (CUSUM) charts. In a CUSUM chart, the difference between the actual and expected outcome is plotted cumulatively so that a series of acceptable outcomes makes the chart vary randomly around the average or baseline, but a series of poor outcomes will make the chart moving away from the average (usually upwards). CUSUM charts were recently shown to be particularly useful, in comparison with other types of control charts, for faster detection of increases in adverse events.[7]

In the UK mortality surveillance system, the CUSUM charts are designed to detect twice or over the national average mortality, at which point an alert is triggered.[8] If a patient in a hospital had died but had a low expected mortality, the chart will move upward and closer towards the threshold of triggering an alert. By contrast, deaths among patients with higher expected mortality do little or nothing to move the hospital towards the threshold. Triggering an alert for a hospital could thus mean either that mortality has somewhat exceeded the predicted rate (an SMR above 1) over a longer period of time or that mortality has substantially exceeded expectation over a short period of time.

In *BMJ Quality & Safety*, two papers by Cecil *et al* report on the impacts of these mortality alerts.[9 10] Between 2007 and 2016, 860 alerts were generated across the roughly 135 trusts monitored by the programme. The authors focus on a subset of 204 alerts sent to 96 hospital trusts between 2011 and 2013. The Care Quality Commission (CQC), which regulates health and social care in England, pursued 75% (154) of these alerts. As Cecil and colleagues report, the CQC found areas of care that could be improved for 106 (69%) of the alerts and considered that failings in care could have affected patient outcomes for 38 (25%) of the pursued alerts.[9] Interestingly, hospitals receiving multiple alerts were less likely to find areas of care that could be improved compared with hospitals

1

receiving a single alert (52% vs 75%). This may seem surprising arguing that multiple alerts may reflect a clearer signal of consistent problems with care in such a hospital. On the other hand, as explained above, it may also simply be the result of a slightly higher than expected mortality over a longer period of time, generating multiple alerts but no real deficiencies in care. This explanation seems supported by the fact that these hospitals receiving multiple alerts seemed more likely to report case-mix or no improvement needed following these alerts, and the fact that the CQC was less likely to pursue cases in hospitals with multiple alerts (44% vs 79%).

The second paper investigates the more tantalising question of the possible impact these alerts have on subsequent mortality rates. The authors report that among hospitals receiving an alert, the relative risk of death—representing observed versus expected mortality—on average was 1.5 in the year preceding the alert.[10] After the alert had been generated, risk-adjusted mortality decreased by 61% in the 9 months after the alert—with a 38% decrease immediately in the month after the alert—and then levelled to reach the expected mortality rate.

Cecil *et al* give two potential explanations for this rapid decrease in risk-adjusted mortality following an alert. Hospitals might already monitor their performance and take action before receiving the alert. Alternatively, the rapid reduction in mortality following alerts might reflect the role of chance—specifically regression to the mean. Both of these explanations are consistent with the observation that the majority of the decline occurred directly after the alert.

For anyone interested in quality improvement, the first explanation is tempting. Yet, it does raise the question why hospitals monitoring their performance would let it come to receiving an alert to implement changes, rather than taking action long before that. Consider a hospital which receives an alert for mortality related to sepsis, the most common cause of mortality alerts.[9] It seems unlikely that a hospital would recognise that it has serious deficiencies in the management of sepsis but then wait until it receives a mortality alert to address any of these problems. Even if receiving the alert provided some sort of extra push, galvanising the hospital into action, it seems implausible that hospitals would so often succeed so quickly.

The particular example of sepsis—again, the most common cause of the mortality alerts—raises the issue of what hospitals can even do to improve mortality. Despite the various prominent campaigns focused on sepsis, a recent systematic review in a high-impact general medical journal concluded that "No high- or moderate-level evidence shows that SEP-1 (The Severe Sepsis and Septic Shock Early Management Bundle) or its haemodynamic interventions improve survival in adults with sepsis".[11] Interestingly, one thing hospitals can do—and, in fact are encouraged to do in campaigns

focused on sepsis—is to recognise sepsis earlier, which of course makes sense. The problem from a mortality monitoring point of view is that earlier recognition often translates into labelling more patients at lower risk of death with this diagnosis, increasing the incidence of sepsis while lowering the apparent mortality. This change does not represent deliberate changes in coding practices or gaming. But, it does make it hard to interpret reductions in mortality since many of the patients now labelled as septic have lower risks of death compared with those in earlier years.

Cecil and colleagues investigated changes in coding for sepsis and acute myocardial infarction.[10] They report that changes did occur, but generally too small to account for the magnitude of changes in mortality seen after the alerts. The problem with this investigation of coding practices is that, considering the case of sepsis, administrative data have serious problems. One fairly recent study estimated the incidence of sepsis over time in over 400 academic and community hospitals in the USA over a 5-year period (2009–2014) based on clinical criteria from electronic health records (EHRs) and claims-based data, much like the Hospital Episode Statistics used to generate mortality alerts in the UK.[12] Analysis of claims-based data for the more than 7 million adults in the sample indicated a significant increase in the incidence of sepsis over time as well as a marked decrease in sepsis mortality and death or discharge to hospice. By contrast, the incidence of sepsis based on clinical criteria obtained from EHRs remained stable. Inpatient mortality due to sepsis showed a small decline, but no reduction occurred in the combined outcome of death or discharge to hospice. In other words, the slight reduction observed for hospital mortality likely reflected more patients dying in hospice. So the reduction in mortality for patients with sepsis following an alert could be the result of more patients being included as septic rather than an effective intervention resulting in reduction of mortality. This may also constitute improved care and we can imagine that hospitals have also reported this as such or as coding changes to the CQC.

These arguments all involved sepsis, which was the most common cause of alerts, but still only accounted for 11.5% of all alerts. What about other conditions? Coronary artery bypass surgery (CABG) constituted the second most common trigger for mortality alerts. But, what changes in care can a hospital implement to successfully reduce mortality after CABG in under 9 months? Even for conditions with well-established, evidence-based processes of care, such as chronic obstructive pulmonary disease (COPD), better adherence to recommended care improves outcomes over the long term, not in-hospital mortality.

But, if actions taken by hospitals seem unlikely to explain the substantial and rapid reductions in mortality following alerts, how might they occur? An analogy can be drawn with a classic teaching example

used to illustrate regression to the mean. Suppose an exceptionally high number of accidents have occurred at a particular intersection during the past year. There is a natural tendency to take action to prevent further injuries and deaths—installation of speed cameras perhaps, or changes to signals and traffic lights. In the subsequent year, the number of accidents or collisions goes down. While it is tempting to attribute this reduction to the actions taken, we do not know what would have happened without the speed cameras or traffic lights with cameras. The high number of accidents in a short period of time may well have been bad luck. And, with the large number of roads and crossings, such runs of bad luck will occur regardless of whether or not a given road has serious safety problems. Of course, some intersections do truly pose greater risks for collisions than others, and some interventions might truly improve road safety. The problem is that chance can improve safety, too. Of all the intersections in a city with particularly high numbers of accidents during the past year, some have that outlier status on the basis of chance. And, most of those chance outliers will not have such a high number of accidents again in the following year. Thus, any reduction in harm after implementing changes at an extremely dangerous intersection might have happened even without taking any action. Failure to take into account this phenomenon of regression to the mean will overestimate the true effectiveness.[13]

Of course, we cannot prove that regression to the mean explains much of the changes in mortality observed by Cecil *et al* following mortality alerts. But, a mortality alert system based on extreme outliers comes as close as one can imagine to a textbook example of when regression to the mean will pose a problem. And, as we have pointed out, the actual improvements hospitals could make to achieve real improvements in mortality remain unclear, even for common conditions triggering these alerts, such as sepsis and CABG.

What might constitute the way forward? Some improvements could occur over the next few years as extraction of key clinical information from EHRs becomes feasible on a widespread basis. The study mentioned previously,[12] in which trends in the incidence of and mortality from sepsis were analysed using claims-based data and clinical criteria drawn from EHRs, illustrates the advantages of this approach.

Even with better data, however, improving requires serious effort, expertise and time. The National Surgical Quality Improvement Programme (NSQIP) involves data more robust than clinical criteria extracted from EHRs: trained personnel collect key data in quasi-real time. And, the outcomes use robust, validated risk adjustment. Yet, two independent studies showed that hospitals using NSQIP achieved only small improvements in outcomes, and these improvements did not differ from those seen in non-NSQIP hospitals.[14 15] What does this mean? Not that there is anything wrong with the NSQIP system. Just that improvements do not come easily. Many improvement efforts do not work, do not adhere to basic principles of improvement science[16] and often do not even have a clear rationale for the intervention.[17 18] Until the capacity to develop and execute effective improvements becomes more widespread, even the most accurate mortality alerts system possibly will struggle to show real reductions in mortality.

REFERENCES

1. The Bristol Royal Infirmary Inquiry. *The report of the public inquiry into children's heart surgery at the Bristol Royal Infirmary 1984–1995: learning from Bristol*. UK Government: The Bristol Royal Infirmary Inquiry, 2001.
2. Aylin P, Alves B, Best N, *et al*. Comparison of UK paediatric cardiac surgical performance by analysis of routinely collected data 1984–96: was Bristol an outlier? *Lancet* 2001;358:181–7.
3. Spiegelhalter DJ. Mortality and volume of cases in paediatric cardiac surgery: retrospective study based on routinely collected data. *BMJ* 2002;324:261–3.
4. Spiegelhalter D, Grigg O, Kinsman R, *et al*. Risk-adjusted sequential probability ratio tests: applications to Bristol, Shipman and adult cardiac surgery. *Int J Qual Health Care* 2003;15:7–13.
5. Hawkes N. How the message from mortality figures was missed at Mid Staffs. *BMJ* 2013;346:f562.
6. Macrae C. Early warnings, weak signals and learning from healthcare disasters. *BMJ Qual Saf* 2014;23:440–5.
7. Neuburger J, Walker K, Sherlaw-Johnson C, *et al*. Comparison of control charts for monitoring clinical performance using binary data. *BMJ Qual Saf* 2017;26:919–28.
8. Bottle A, Aylin P. Intelligent information: a national system for monitoring clinical performance. *Health Serv Res* 2008;43(1 Pt 1):10–31.
9. Cecil E, Wilkinson S, Bottle A, *et al*. A national hospital mortality surveillance system: a descriptive analysis. *BMJ Qual Saf* 2018.
10. Cecil E, Bottle A, Esmail A, *et al*. Investigating the association of alerts from a national mortality surveillance system with subsequent hospital mortality in England: an interrupted time series analysis. *BMJ Qual Saf* 2018:bmjqs-2017-007495.
11. Pepper DJ, Jaswal D, Sun J, *et al*. Evidence underpinning the centers for Medicare & Medicaid services' severe sepsis and septic shock management bundle (sep-1): a systematic review. *Ann Intern Med* 2018;168:558–68.

12 Rhee C, Dantes R, Epstein L, *et al*. Incidence and trends of sepsis in US hospitals using clinical vs claims data, 2009–2014. *JAMA* 2017;318:1241–9.

13 Retting RA, Ferguson SA, Hakkert AS. Effects of red light cameras on violations and crashes: a review of the international literature. *Traffic Inj Prev* 2003;4:17–23.

14 Etzioni DA, Wasif N, Dueck AC, *et al*. Association of hospital participation in a surgical outcomes monitoring program with inpatient complications and mortality. *JAMA* 2015;313:505–11.

15 Osborne NH, Nicholas LH, Ryan AM, *et al*. Association of hospital participation in a quality reporting program with surgical outcomes and expenditures for Medicare beneficiaries. *JAMA* 2015;313:496–504.

16 Taylor MJ, McNicholas C, Nicolay C, *et al*. Systematic review of the application of the plan-do-study-act method to improve quality in healthcare. *BMJ Qual Saf* 2014;23:290–8.

17 Davidoff F, Dixon-Woods M, Leviton L, *et al*. Demystifying theory and its use in improvement. *BMJ Qual Saf* 2015;24:228–38.

18 Shojania KG, Grimshaw JM. Evidence-based quality improvement: the state of the science. *Health Aff* 2005;24:138–50.