Systemic predictors of adverse events in a national surgical mortality audit

Analysis of peer-review data from ANZASM

Authors

Richard Clive Turner MBBS PhD; University of Tasmania, School of Medicine

Steve Simpson, Jr. MPH, PhD; The University of Melbourne, Melbourne School of Population and Global Health; University of Tasmania, Menzies Institute for Medical Research.

Mrunmayee Bhalerao MBBS; University of Tasmania, School of Medicine

This paper represents the elaboration of an oral communication presented by one of the authors (MB) at the 2018 Annual Scientific Congress of the Royal Australasian College of Surgeons: "Predictors of adverse events in 9 years of ANZASM data".

Number of Figures: 1

Number of Supplemental Figures: none

Number of Tables: 3

Number of Supplemental Tables: 4

Word count (abstract): 159

Word count (text): 3 250

Corresponding author

Professor Richard C Turner

richard.turner@utas.edu.au

Private Bag 96, Hobart. TAS. 7000

+61 417 736205

Field Code Changed

Abstract

Background
Peer review of surgical deaths can identify deficits in individual and systemic delivery of healthcare, ultimately informing quality improvement.

Methods
From 2008 to 2016, cases reported to the Australia and New Zealand Audit of Surgical Mortality were analyzed. Variables associated with peer-judged adverse events were sought.

Results
Of 21 045 cases evaluated, 24.8% incurred at least one adverse event judgement. The proportion of cases with reported adverse event significantly decreased over time. Following adjustment for demographic and clinical characteristics, significant negative patient-related associations were advanced age, greater ASA grade, and neurological and malignant comorbidities. Significant associations were also found with systemic or organisational factors, including State/Territory, surgical specialty, and hospital regionality.

Conclusion
Examination of this peer-reviewed database revealed systemic or organizational predictors of adverse events that may have implications for quality improvement at an institutional or jurisdictional level. The extent to which these associations are due to the peer-review process itself should be the focus of further research.

Background

Clinical peer review aims for quality improvement through self-reflection and value-added feedback.[1] Analysis of peer review data as a quality indicator often focuses on performance metrics of individual healthcare providers. However, aggregate data can also provide insight into variations of performance at an institutional or craft group level, potentially informing systemic quality improvements.

Some patient deaths are inevitable in surgical practice. Most are attributable to the nature of the presenting complaint, often with contribution from comorbidities. Whilst the primary focus of surgical mortality audit is individual performance, organisational factors may also be uncovered when reviewing events leading to an otherwise unexpected death.

The Australia and New Zealand Audit of Surgical Mortality (ANZASM) is a prospective database and peer review platform established in the early 2000s, whose principal objectives are informing, educating, facilitating change, and improving quality of surgical care.[2-4] It records patient, organizational and surgical team data regarding pre-, intra- and post-operative care for all deceased patients where a surgeon was involved. Since its inception, ANZASM has provided numerous insights into the determinants of surgical mortality and adverse events contributing to death.[4-8] We have sought to complement this work by examining the ANZASM database specifically for non-surgeon- or non-patient-related variables associated with adverse events flagged by peer review.

Methods

Data collection

To fulfil Continuing Professional Development requirements, surgeons in all Australian States and Territories (except New South Wales which had a separate registry at the time this data was acquired) should report in-hospital deaths of all patients in their care to ANZASM, which is administered by the Royal Australasian College of Surgeons (RACS). A Surgical Case Form (SCF) is completed by the relevant surgeon, including diagnosis, comorbidities, and management, as well as any Areas for Consideration, Areas of Concern, or Adverse Events. SCFs are sent to an assigned peer reviewer for First Line Assessment (FLA). From the information provided, the reviewer independently flags any Areas for Consideration, Areas of Concern, or Adverse Events. The FLA is subsequently fed back to the reporting surgeon.

For this study, data were extracted from all SCFs and their corresponding FLAs submitted between 1 Jan 2008 and 31 Dec 2016. The primary outcome measure was adverse events as judged by FLA. FLAEs were categorised into four non-mutually exclusive types: 1) pre-operative; 2) operation planning, including improving decision to operate, operation chosen, and operation timing; 3) intra-operative, including improving overall intra-operative, decisions made by surgeon, and the surgeon leading the operation; and 4) post-operative care.

Statistical analysis

Predictors of first-line adverse events (FLAEs) were evaluated by log-binomial regression, presented as prevalence ratios (95% CI). Multivariable models adjusted for age, emergency/elective status and surgical indication (categorised as functional derangement, neoplasia, infection, trauma, and other). New Zealand data was not included in the analysis, pending a bi-national data sharing agreement.

Ethics

Access to the ANZASM data was approved as a low-risk activity by the Royal Australasian College of Surgeons Ethics Committee.

Results

Between 1 January 2008 and 31 December 2016, 21 306 deaths were reported. Of these, 21 045 (98.8%) received an FLA. At least one FLAE was flagged in 5 225 (24.8%) patients, of which 2 900 (55.5%) were male and 4 002 (76.6%) over 65 years.

FLAE variability by patient characteristics

As per Supplemental Table 1, there were no overall significant differences in FLAE by sex. Patients in the 65-75-year increment were significantly more likely to incur FLAE, while the oldest patients (>85 years) were significantly less likely, both persisting on adjustment. The proportion of FLAE was significantly lower with increasing ASA grade. Patients with cardiovascular comorbidities had a significantly increased risk of FLAE, driven by pre-operative and post-operative complications (Supplemental Table 2). Patients with neurological and oncological comorbidities, however, had a significantly lower frequency of FLAEs. No differences in FLAE were seen by indigenous status.

FLAE variability by clinicopathological status

As per Table 1, elective procedures were roughly 2-fold more likely to have FLAEs compared to emergency procedures, persisting on adjustment; this increased risk was evident at all stages of surgery but particularly for intra-operative and post-operative care, where FLAE risk was 2.9 and 3.6-fold increased, respectively.

Compared to General Surgery, FLAEs were significantly more likely in Cardiothoracic and Other surgical specialties, but less likely in Neurosurgery and Paediatric Surgery. Cardiothoracic surgeons had a 75% increased risk of intra-operative complications, while Other specialty surgeons had roughly 2.5-fold increased frequencies of pre and intra-operative complications; the reduced FLAE frequency in Neurosurgery and Paediatric Surgery was across all aspects of care (Supplemental Table 3). Compared to admissions for functional or structural conditions, patients with neoplasia and trauma were less likely to have FLAE after adjustment. Compared to patients whose cause of death was a medical (e.g. respiratory or cardiac) condition, those dying from neoplasia or trauma had significantly lower likelihood of FLAE, though of these, only neoplasia persisted on adjustment. On the other hand, patients whose cause of death was infection or where death was sudden and/or of unknown cause were significantly more likely to have FLAE. Unsurprisingly, expectation of death was significantly associated with prevalence of FLAE, in that patients where death risk was minimal had greater likelihood of FLAE, while those with considerable or expected death risk had significantly lower likelihood.

<u>FLAEs variability by time and place</u>

As per Figure 1, there was an overall 25% reduction in FLAEs from 2008 to 2016, which remained significant after adjustment for age, surgical indication, and emergency/elective status (Table 2). This decline was largely driven by post-operative FLAEs (Supplemental Table 4). FLAEs were significantly less common on weekend days, though this difference was wholly abrogated on adjustment. Post-operative FLAEs were significantly less common after Monday, though there was no weekend/weekday difference for all FLAEs (Supplemental Table 4). A similar association was seen for time of surgery, with operations conducted on evening (1500-2300h) and night (2300-0700h) shifts having significantly less FLAEs than daytime (0700-1500h) ones; this persisted for evening shift operations, after adjustment.

Victoria had the highest proportion of FLAEs. Every other regional jurisdiction had significantly lower proportions, except the ACT and Northern Territory, following adjustment; these differences were broadly evident across all FLAE types (Supplemental Table 4). Capital city hospitals had the majority of cases and FLAEs, though the proportion of FLAEs was significantly greater in non-capital metropolitan and large rural centres, persisting on adjustment. These excesses were most evident for pre-operative and operation planning complications, though non-capital metropolitan centres had a significant excess of intra-operative complications, while post-operative complication were more common in large rural centres (Supplemental Table 4). Compared to public hospitals, private hospitals were significantly less likely to incur FLAEs after adjustment for age, indication and emergency/elective status, while private patients (including those treated in public hospitals) were not. This decreased frequency of FLAEs in private hospitals was only significant for intra-operative and post-operative FLAEs, while there was no difference in pre-operative FLAEs (Supplemental Table 4).

Discussion

A nine-year analysis of Australian surgical mortality data uncovered variation in FLAE reporting by time, health jurisdiction, hospital type and surgical specialty. Such systemic or organizational factors are distinct from surgeon- or patient-related factors associated with AEs. They have potential application in quality improvement of health services and of the peer review process itself.

Over 2008-2016, 24.8% of cases with an FLAE is not inconsiderable, although an AE is defined as something that *may* have *contributed* to death, rather than being the sole cause. FLAEs were significantly less common on weekend days and decreased in frequency over the audit period, though the weekend effect did not persist on adjustment for patient age, surgical indication, and emergency/elective admissions. The latter patient-related variable is evidenced as a confounder by the fact that 79.7% of weekday cases were emergencies, compared to 94.3% and 96.9% of Saturday and Sunday cases, respectively. Similarly, elective admissions were almost twice as likely to incur an FLAE compared to emergencies, possibly due to the former being conceived as planned procedures, not expected to have unforeseen complications. Conversely, greater leniency may be afforded to emergency procedures, where less time is available for reasoned decision-making.

The consistent downward temporal trend in the proportion of deaths with AEs may be attributed to systemic reforms, such as surgical safety checklists and evidence-based clinical care bundles. The former have been shown to result in a reduction in surgical mortality,[9] presumably mediated by increased vigilance for potential errors. Participation in a cross-institutional reporting system for surgical deaths may itself engender a culture of caution, reflection and best practice.[10-12] Indeed, the ANZASM process could be seen as an intentional application of the Hawthorne effect,[13] whereby an individual's behavior is altered simply by virtue of being observed.[14, 15]

Declining temporal trends in FLAE reporting may also partly reflect reviewer fatigue.[16] All RACS Fellows are encouraged to be first-line ANZASM reviewers through Continuing Professional Development rewards, and there is no limit on the length of time one can serve as a first-line reviewer. This could lead to increased tolerance to what may have once been considered the result of suboptimal management, or perhaps simply a desire to not have to provide as much justification for an adverse judgement. A reviewer's own accumulation of AEs over time may also make them more lenient when judging others' work. Constant recruitment of new reviewers may offset collective fatigue. As a quality indicator of the peer review process itself, ANZASM may therefore consider gathering reviewers' demographic data, e.g., years of practice, public/private work, rurality, etc. Although reviewers receive written instructions on how to complete an FLA and how to define an AE, robustness may be refreshed by additional education and periodic peer calibration exercises.

As well as changing over time, stringency of peer review may vary with the "culture" of certain health jurisdictions or specialty groups due to variable prescriptiveness of clinical practice guidelines, reviewers' desire to promote their own particular standard of care, or even the region- or specialty-specific nature of surgical training. Small subspecialties and state jurisdictions may be more sensitive to idiosyncratic judgments of relatively small numbers of reviewers; anonymisation of case reports may also be difficult in small practice communities. In the case of state jurisdictions, this effect has been moderated by pooling peer reviewers from a number of smaller jurisdictions, e.g., Western Australia, South Australia, and Tasmania. For Cardiothoracic Surgery, increased stringency may also be partly attributable to the historical factors that encourage and support self-criticism. One notable example is the Bristol Royal Infirmary Inquiry,[17] whose findings cited denial of adverse events and fear of punition as key contributors to excess postoperative mortality. For Neurosurgery, by contrast, relative leniency may be due to preconceived expectations regarding the high risk of surgery. Similarly, most deaths in Paediatric Surgery are likely to be only in those with an already poor outlook.

The RRMA (Rural, Remote and Metropolitan Areas) classification[18] was used to stratify treating hospitals by service accessibility. The two categories with significantly increased risk on adjustment of FLAE reporting were "large rural" (urban centre population 25,000–99,999) and "other metropolitan" (non-capital city population $\geq$100,000). Other rural and remote centres did not have excess FLAEs, concurring with Ferrah et al, who found no material difference over a similar period in the frequency of post-operative complications between urban and rural centres.[5] Evidence demonstrates an inverse relationship between surgical mortality and case volume,[19-21] with small rural centres usually transferring major cases to larger centres. In some instances, however, urgency or community expectations may dictate that definitive care be delivered at a secondary referral (regional or other metropolitan) hospital with a lower throughput of complex cases compared to tertiary referral (metropolitan) hospitals. This presumption is corroborated by the fact that the association was largely driven by Queensland, the second largest and most decentralised Australian state (data not shown). Furthermore, that FLAE reporting for these hospitals was largely driven by pre-operative AEs (data not shown) may reflect problems or delays in transferring patients to other centres. It may also be attributable to staffing issues at the point of entry, as might be expected where relatively junior staff have less senior supervision. In some cases, the hospital may have exceeded its capability for diagnosis or definitive management. All such issues warrant further exploration to determine what systemic reforms could be instituted.

This study's principal strength is the large FLA dataset, enhancing statistical power. Data are also highly representative, as compliance with reporting to ANZASM and subsequent FLA approach 100%. Another strength is that we used FLAEs as an outcome measure rather than self-reported AEs on SCFs, as the former are likely to be a more objective judgment. Indeed, Raju et al noted a three-fold increase in significant event reports by first-line assessors compared to treating surgeons.[4] In the univariable analysis, patient-related and clinical predictors of FLAEs largely behaved in an intuitively plausible manner. For example, advanced age, ASA and malignancy tend to accrue fewer AEs, presumably because they are afforded greater leniency in judgement of clinical management due to the already poor perceived prognosis of such patients. Thus, inclusion of these variables in the multivariable analysis serves to strengthen the validity of the findings with respect to system-related and organizational factors. Finally, ANZASM captures unwanted occurrences at all stages of the patient journey, including non-operative cases, rather than confining itself to post-operative complications only.

One limitation is that ANZASM does not provide a mortality *rate* to serve as a surrogate marker of surgical performance after adjustment for confounders.[22] However, the proportion of deaths associated with an adverse event is possibly more indicative of the quality of care, as it implies preventability, either through individual or systems improvement. Aggregate analysis of FLAs is subject to inter-rater variability. We sought to minimise this by reserving analysis for the highest level of seriousness rather than possibly more contentious Areas of Consideration and Areas of Concern. A Victorian study also showed good concordance between the more serious ANZASM second -line assessments (SLAs),[23] from which it could be inferred that there would be similarly low inter-rater variability in FLA judgments.

FLAEs rely on self-reported data in SCFs, which may limit our inferences. This is a potential confounder in all peer-on-peer reviews. We have attempted to indirectly adjust for it through the various patient-specific factors that may influence the rigor of self-reporting. It would otherwise be useful to adjust for individual surgeon demographics, were these data available.

Finally, the absence of New South Wales data limits the generalisability of our findings in the Australian setting. Moves are currently underway to merge the ANZASM and independently gathered New South Wales datasets. We therefore eventually wish to repeat our study with a full Australian dataset.

Conclusions

In the ANZASM database, we have demonstrated a consistent downward trend in peer-reported adverse events over several years. The increased risk of FLAEs in certain health jurisdictions, hospital categories, and surgical specialties may ultimately inform quality improvement at an institutional or craft group level, after more detailed analysis of the types of AEs driving these associations. The extent to which this is a function of the peer review process itself should also be subject to further enquiry, as well as ongoing education of reviewers regarding the ANZASM assessment process.

Acknowledgements

The authors thank Dr Mary Self for her feedback on the final draft of the manuscript.

References

1. Hill, T.E., P.F. Martelli, and J.H. Kuo, *A case for revisiting peer review: Implications for professional self-regulation and quality improvement.* PLOS ONE, 2018. **13**(6): p. e0199961.
2. Raju, R.S. and G.J. Maddern, *Lessons learned from national surgical audits.* BJS, 2014. **101**: p. 1485 – 1487.
3. Raju, R.S., et al., *The Australian and New Zealand Audit of Surgical Mortality-birth, deaths, and carriage.* Ann Surg, 2015. **261**(2): p. 304-8.
4. Raju, R.S., et al., *Australian and New Zealand Audit of Surgical Mortality: concordance between reported and audited clinical events and delays in management in surgical mortality patients.* ANZ J Surg, 2014. **84**(9): p. 618-23.
5. Ferrah, N., et al., *Rural centres do not have a higher prevalence of post-operative complications than urban centres: a retrospective analysis of a mortality audit.* ANZ J Surg, 2019.
6. Singla, A.A., et al., *No weak days? Impact of day in the week on surgical mortality.* ANZ Journal of Surgery, 2016. **86**: p. 15 - 20.
7. Jones, C.R., et al., *Deaths in incorrectly identified low-surgical-risk patients.* World Journal of Surgery, 2018. **42**(7): p. 1997 - 2000.
8. Gupta, A.K., et al., *Potentially avoidable issues in neurosurgical mortality cases in Australia: identification and improvements.* ANZ J Surg, 2017. **87**(1-2): p. 86 - 91.
9. Haynes, A.B., et al., *Mortality trends after a voluntary checklist-based surgical safety collaborative.* Ann Surg, 2017. **266**(6): p. 923 - 929.
10. Lui, C.-W., et al., *How participation in surgical mortality audit impacts surgical practice.* BMC Surgery 2017. **17**: p. 42 - 48.
11. Krahwinkel, W., et al., *The effect of peer review on mortality rates* International Journal for Quality in Health Care, 2016: p. 1 - 7.
12. Kiermeier, A., et al., *National surgical mortality audit may be associated with reduced mortality after emergency admission.* ANZ J Surg 2017. **87**: p. 830 - 836.
13. Roethlisberger, F.J. and W.J. Dickson, *Management and the Worker*. 1939, Cambridge, Mass.: Harvard University Press.
14. Mayo, E., *The human problems of an industrial civilization*. 2nd ed. Vol. 3. 1993, New York: MacMillan.
15. McCarney, R., et al., *The Hawthorne Effect: a randomised, controlled trial.* BMC Medical Research Methodology, 2007. **7**: p. 30.
16. Breuning, M., et al., *Reviewer fatigue? Why scholars decline to review their peers' work.* PS: Political Science & Politics, 2015. **48**(4): p. 595 - 600.
17. Dyer, C., *Bristol inquiry condemns hospital's "club culture".* BMJ, 2001. **323**: p. 181.
18. Welfare, A.I.o.H.a., *Rural, regional and remote health: a guide to remoteness classifications*. 2004: Canberra.
19. Birkmeyer, J.D., et al., *Hospital volume and surgical mortality in the United States.* N Engl J Med, 2002. **346**(15): p. 1128 - 1137.
20. Iwatsuki, M., et al., *Effect of hospital and surgeon volume on postoperative outcomes after distal gastrectomy for gastric cancer based on data from 145,523 Japanese patients collected from a nationwide web-based data entry system.* Gastric Cancer, 2018: p. 1 - 12.
21. Dudley, R.A., et al., *Selective referrral to high volume hospitals: Estimating potentially avoidable deaths.* JAMA, 2000. **283**(9): p. 1159 - 1166.
22. Heeney, A., et al., *Surgical mortality - an analysis of all deaths within a general surgical department.* Surgeon, 2014. **12**(3): p. 121 - 128.
23. Hansen, D., et al., *Validation of data submitted by the treating surgeon in the Victorian Audit of Surgical Mortality.* ANZ J Surg, 2019. **89**(1-2): p. 16 - 19.

Supplementary materials

Supplemental Table 1: Predictors of first line review adverse events: patient characteristics.

Supplemental Table 2: Predictors of first line review adverse events: patient characteristics, by FLAE type.

Supplemental Table 3: Predictors of first line review adverse events: clinicopathological status, by FLAE type

Supplemental Table 4: Predictors of first line review adverse events: time and place, by FLAE type