Identifying adverse events after outpatient surgery: improving measurement of patient safety

Amy K Rosen,1 Hillary J Mull1,2

Identification of adverse events (AEs) is critical for improving patient safety. However, accurate measurement continues to be challenging, and efforts to detect and track surgical AEs in the outpatient setting lag behind those of the inpatient setting. Although numerous methods have been utilised over the years to detect AEs (eg, voluntary reporting systems, chart review and patient interviews), these detection systems suffer from a variety of limitations including resource constraints.1 2 More recent development of automated surveillance systems to detect AEs using electronic medical record (EMR) data has greatly facilitated the identification of AEs, particularly among ambulatory patients.3–6 Menendez et al illustrate how EMR data and electronic triggers can contribute to better measurement of patient safety in outpatient surgery.7

Trigger methodology has substantially improved since the seminal work of the Institute for Healthcare Improvement (IHI) in the early 2000s that helped promulgate the use of chart-based trigger tools for retrospective detection of AEs.3–10 Although triggers are still evolving as informatics tools, and are likely in their ‘early stages’ of development, the trigger methodology represents a good compromise between two modalities: automated surveillance systems and manual chart review (ie, the ‘gold standard’). Triggers rely on both electronic and manual review processes to search for patterns in the data consistent with a possible AE. Triggers use surveillance rules or algorithms derived from clinical logic to flag patient medical records for the presence of an AE. Once a trigger is flagged in the data, then the patient’s medical record is reviewed to confirm the occurrence (yes/no) of the AE.11 Triggers facilitate more selective EMR review and also capitalise on the richness of EMR data, resulting in efficient and cost-effective tools that capture events missed by other methods (eg, voluntary reporting).12–15

The US Agency for Healthcare Research and Quality (AHRQ) sponsored a conference in 2008 that helped highlight the importance of triggers in identifying patient safety risks and hazards, while also endorsing future development of prospective triggers to enable timely interventions to prevent or reduce specific AEs.11 16 Triggers are now widely used to detect many types of AEs, including diagnostic errors, adverse drug events, hospital-acquired infections, delays in diagnoses and outpatient surgical AEs.17–26 Trigger algorithms are frequently applied to EMRs for automated surveillance, and increasingly to prospectively identify patients at risk of AEs.27 28

Menendez et al7 build off the EMR-based trigger methods in the field to develop and test 13 retrospective triggers used to identify 90-day postoperative events following orthopaedic outpatient surgery. The research has a number of limitations, some of which are noted by the authors; we point out a few that are not. While it is noteworthy that the authors produced a formula to predict the likelihood of identifying an AE in the medical record, limiting review of records to under 10%, the generalisability and applicability of this formula to other healthcare systems and AEs remains to be determined. Further, their methods for trigger development and AE definition were not well described, and it would be difficult for another group of researchers to replicate them. Their methods, as currently described, were generally not consistent with the considerations recommended in the AHRQ trigger conference calling for more rigorous development of such metrics through
the use of explicit criteria and techniques. While a few triggers had positive predictive values (PPVs) that were greater than 50%, most did not meet the criteria for ‘good’ performance, suggesting that as currently developed, they are probably not suitable for external monitoring or benchmarking purposes. However, they may very well be useful for quality improvement initiatives since they appear to be a more efficient way to screen for AEs than other methods. The AHRQ conference also recommended that researchers should estimate sensitivity as another method for evaluating trigger performance, since PPV is largely a function of event prevalence (which, in the case of AEs, is generally low). Regardless of which performance characteristic is selected, an important goal of trigger development should be to increase the accuracy of AE rate estimation in the population. We commend Menendez et al for developing a predictive formula that allows for more efficient identification of AEs; further refinement of this could potentially improve the accuracy of AE rate estimation.

As the authors also point out, “there is a need for better outpatient quality measures”, as well as further research “to understand the optimal use of electronic triggers as surgical quality indicators and as screening tools”. We agree that the field of outpatient surgery has largely been neglected until recently, with few studies focused on this area and where gaps in knowledge currently exist. There are, however, key considerations in developing such metrics. A well-designed quality measure should be reliable, valid and easy to use across care settings. Similarly, a trigger that performs well should have good face validity, and target those AEs that are likely to be preventable and that represent a relatively high proportion of patients within the population; it should also be able to assess harm. Trigger development should involve an iterative process of development and testing that includes guidance from clinical experts, review of the existing literature, discussion and testing of explicit criteria for each trigger, and incorporation of final selected criteria into trigger algorithms for empirical testing.

Thoughtful deliberations are also needed in order for researchers to identify what a ‘true’ AE is in the outpatient surgical setting. Identification can be difficult, in part because postoperative AEs are relatively infrequent, and those reported tend to include minor complaints and/or morbidities (eg, pain, nausea, dizziness). Thus, distinguishing whether an event is “an injury caused by medical management rather than the underlying condition of the patient” (ie, as defined by the Institute of Medicine), as well as assigning a level of harm to it, can be more subjective and ambiguous in the outpatient setting than in the inpatient setting. To improve identification of outpatient AEs, certain techniques have been used. For example, some authors have used standardised lists of AEs defined in the inpatient setting, expanding these to postoperative outcomes that are likely to occur in outpatient surgery. Additionally, AEs may be counted only if nurse reviewers agree on whether an AE was present or not and inter-rater reliability, as measured by a Cohen’s kappa >0.6, is achieved. An established severity classification scale and a harm scale may be employed to rate the severity of AEs and their associated harm, respectively. Finally, defining the time period from the index event (ie, outpatient surgery) to the postoperative AE is also an important component of establishing an accurate count of AEs in outpatient surgery. For example, a postoperative outcome of 90 days, as the authors use, may be too distant an event from the outpatient surgery to be meaningful for evaluating outpatient surgical quality. Thirty-day outcomes may be better for this purpose as they are likely to be more related to the index event.

In summary, Menendez et al make an important contribution to the patient safety measurement field in that it tackles a relatively understudied area, outpatient surgery (and specifically, orthopaedic surgery). The triggers showed some efficacy but could be enhanced with input from a wider group of clinical experts beyond one institution, additional performance criteria, and a more rigorous validation process for both the trigger algorithms and the AEs identified. Improvement of patient safety in outpatient surgery constitutes an important goal. We look forward to further work of the type carried out by Menendez et al providing efficient tools to identify targets for improvement, and, eventually metrics for assessing performance.

Contributors Both authors contributed to the development and writing of this manuscript.

Competing interests None declared.

Provenance and peer review Commissioned; internally peer reviewed.

REFERENCES


