

Triangulating safety: applying social media analysis methods to revolutionize patient safety

Abstract

Too many people suffer from preventable medical errors. Medical error reporting systems have already been deployed in many countries to prevent similar errors from happening again. Through such systems, millions of errors have been reported and analyzed to yield information on safety incidents. Nevertheless, there is still huge room for improvement, especially in analyzing the free text description parts of those reports. Free text entries offer richer and more detailed information on incidents, but are too often ignored or wasted due to a lack of analytical resources. Recent developments in social media analysis can be utilized to effectively address this problem. Thanks to huge advancements in computing power and algorithm design, social media companies are effectively analyzing millions of free text postings every day to extract relevant information. Their automated analysis methods can be applied to analyze medical error reports. This includes collating detailed information on incident causation and even cultural issues like power gradients among health care organizations - all information that can dramatically improve safety. Many errors are reported each day, and many are waiting to be thoroughly analyzed. Leaving them unused is neither necessary nor justifiable. Currently available computing power and analytical methods can be utilized to activate their contribution to reducing medical errors.

Keywords: patient safety, medical error reporting system, automated content analysis, social media analysis methods

Volume 2 Issue 1 - 2015

Heon Jae Jeong,¹ Minji Kim²

¹Department of Health Policy and Management, Johns Hopkins Bloomberg School of Public Health, Johns Hopkins University, USA

²Annenberg School for Communication, University of Pennsylvania, USA

Correspondence: Heon-Jae Jeong, Department of Health Policy and Management, Johns Hopkins Bloomberg School of Public Health, Johns Hopkins University, 624 North Broadway, Rm. 455, Baltimore, 21205, USA, Tel 410-955-5315, Fax 410-955-6959, Email heong1@jhu.edu

Received: January 22, 2015 | **Published:** January 26, 2015

Introduction

A seminal report, 'To Err is Human,' stated that at least 44,000, and perhaps 98,000 patients may die in hospitals due to preventable medical errors in the US each year.¹ The US accounts for roughly 1/20 of the world's population. Therefore, even under the very generous assumption that every country has almost the same quality of care that the US has, at least approximately 880,000 patients around the world may die of preventable medical errors each year. Meanwhile, not every error causes harm; according to Heinrich's law that is well known in the field of safety, there are 300 non-harmful events (let us call them errors) behind each serious event.^{2,3} So how many medical errors occur in one single year? If we do the math, at least around 264,000,000 occur globally and 13,200,000 in the US occur each year.

Confronting these daunting numbers, we had two options. The seemingly easy option was just to feel helpless or angry, to blame the deaths on careless healthcare professionals, and do nothing. The other option was to analyze the problem and solve it. We, humans, though it appeared to be a thorny road, chose the latter, to improve patient safety. We decided to save lives. As patient safety takes center stage in health care, many changes took place. One of the most dramatic of these changes would be to bid farewell to the 'blame' culture, and to acknowledge the fact that everybody makes mistakes.⁴ Such a humble and honest mindset naturally led health care workers to admit their mistakes and take steps to ensure that they would not be repeated. Medical error reporting systems were developed to accomplish precisely this.

These systems encourage health care professionals to report all medical errors that they encounter. Because both harmful and non harmful errors (near misses) have the same mechanism,⁵ most systems welcome reports on both of them. Analyzing collected error reports can reveal systematic flaws, and more importantly, the lessons learned can be distributed widely to prevent the same or similar mistakes from

occurring again, which is why we call these systems 'reporting and learning systems' (RLS). The levels of RLS are hospital, regional, nation-wide and international systems.⁶ Health care workers have committed huge amounts of time to reporting such errors with the hope that those reports can save lives; millions of error reports have been collected. However, how well the collected data are being utilized is questionable. It is time to overhaul the systems and develop strategies to get the most out of the reporting systems.

The anatomy and pathology of medical error reports

Most RLS reporting forms have two major parts to describe an event: predetermined items with multiple options to choose (e.g., in which area did the event occurred? Ward, operation room, outpatient clinic and so forth) and space offered health care workers to describe the event in their own language. So far, the reports data from the 'quantitative' part has been thoroughly analyzed, providing information on the 'hot spots' of health care to which health care workers should pay extra attention, or where systematic changes are required. Indeed, these data served as a treasure map to save lives.

The latter part of RLS data, free text description, in contrast, has been poorly utilized, creating a vicious cycle, where such underutilization leads to lower reporting rates, leading to even lower utilization. It is obvious that free text description can paint a much more accurate picture of how an event occurred, including the information on interaction among multiple health care workers involved in the events, and temporal sequence of events. The more seasoned a safety manager of a hospital is, the more emphasis she would place on 'qualitative' data. Then, why are these data being underutilized? The lack of resources is probably an answer. Some of such data are composed of just a couple of sentences, but there are many reports with longer descriptions. The Johns Hopkins Hospital, using the

Patient Safety Net run by University Health System Consortium (UHC), allows health care workers to type up to 1,000 letters in an event description. No RLS, past or present, have the huge human resources to read all the reports and extract information from them. Therefore, safety researchers have usually compromised. We read only the free texts from the events that caused serious harm, in other words, unfortunately, many had to ignore the free texts description from less severe events. Though a seemingly logical compromise, it might be a dangerous choice. First, as mentioned above, less serious events or near misses have the same occurrence mechanism as serious events. Therefore, there is no reason to leave them unanalyzed; they are a huge missed opportunity to prevent future errors. In addition, health care workers are frustrated by their elaborate reports being ignored, and become less likely to report errors.

Triangulating safety: hints from social media analysis

What if we can analyze the apparently disorganized free text description of an event as much as the quantitative part with predefined variables of the RLS data? What if we can extract key factors of an incident and complex network of constituents of the event from the text? Combining this data with the structured data, we are obviously in a better position to see where and how the event occurred, leading us to predict future events in hospitals. This ‘corroboration’ of two major parts of error reports, in that they can locate the vulnerable sites, processes and picture detailed mechanism how events occurred, we call the ‘triangulation of safety.’ Indeed such mixing of quantitative and qualitative methods is called triangulation.⁷

Let us elaborate by using an illustrative example of submitted report about patient fall case. The quantitative part collected the following information, such as, type: patient fall, time: 11PM, location: corridor near a ward, harm: hip joint fracture. These gave us much information on the event, but let us take a look at the free text description that was filled out by a nurse: “A patient who was supposed to be discharged in two days fell while he went to bathroom and broke his left hip joint in a corridor on a ward at 11PM. Though no caregiver was around the patient, he did not call the nurse because he did not want to bother her. 9PM, two hours before the incident, a resident physician gave an order to a nurse to give the patient Lasix (diuretics). It was only to adjust fluid output to the fluid intake of the day. The nurse raised a concern in giving Lasix in the late night, but was ignored. The attending physician was very bossy to his trainees. For the resident, in and out of water was more important than the risk of patient fall.” This free text description gave us a completely different picture of the event.

In this incident, the hip fracture was the ultimate harm to the patient, as captured in the multiple-choice section of the report. However, the section was unable to capture the causal chain of events, including the attending physician’s bossiness, the power gradient among different Health care worker types and administration of Lasix at night, insufficient caregiver attention and patient’s reluctance to summon the nurse. A thorough examination of the free text description of the event can provide much richer information, although the quantitative part still helps us to conduct efficient exploratory analysis on a huge amount of RLS data. There is no doubt that triangulation would be much better in preventing errors.

There have been attempts to analyze the free text data from error reports,^{8,9} but they were all done manually by human coders. Considering the huge amount of the free text data, automated analysis

by computer or at least computer-assisted analysis is essential. Then, a couple of questions naturally arise: whether we have algorithms to analyze those free text data on medical errors, and if so, whether we have enough computing power to process such a huge flow of free text data. The answers would have been no and not only a few years ago, when most RLS were developed and started operation. But now, the answers might be yes and yes. Information Technology has evolved as fast as medical care has.

We can get some hints from social media. Every day 4.75 billion pieces of content items on Facebook and 500 million tweets are posted on Twitter,^{10,11} and they still manage to extract information to administer tailored advertisements and suggest new content that individual users might like. A professional social-networking site, LinkedIn, analyzes all the texts (though more categorized than Twitter or Facebook) you post as well as your relationships with other users, to suggest the best people to connect with or the most fitting job offers. Indeed, computing power can certainly satisfy the level to which all medical error reports collected with RLS can be analyzed in almost real-time, yielding information that had not been noted when an incident was reported. In addition, automated content analysis enables researchers to extract emotional status of a person or his or her feeling about the conversation or description.¹² The accuracy of automated analysis results was shown to match that of human-coded results.^{13,14} This can even be used to understand various cultural issues that were known to affect safety, such as teamwork climate, stress recognition, and job satisfaction of health care workers.¹⁵ The above-mentioned methods, despite their sure benefits, are yet to be applied to error reports data.

Conclusion

In sum, we can completely change the way of using RLS. Up to now, RLS has given us bird’s eye view of how many and what kind of errors and harmful events occur in the field of health care. Now we need a truffle-hunter’s view of each event to understand exactly what happened, who and how people were involved and in what context. Only then will we be able to effect a dramatic improvement in safety. Now, with all these new tools and techniques at our disposal, and more importantly with the huge medical error data waiting to be analyzed more thoroughly, we must ask ourselves, “Do we have the right not to do this work, the work that can save millions of lives?”

Acknowledgments

None.

Conflict of interest

The authors declare that there is no conflict of interest.

Funding

None.

References

1. Kohn LT, Corrigan JM, Donaldson MS. *To err is human: building a safer health care system*. National Academy Press: Washington, DC, USA; 2000.
2. Heinrich HW, Peterson DC, Roos RN. *Industrial accident prevention: A safety management approach*. 5th ed, McGraw-Hill: New York, USA; 1980. 468 p.
3. Taxis K, Gallivan S, Barber N, et al. *Can the Heinrich ratio be used to predict harm from medication errors?* London, UK; 2006.

4. Reason J. *Human error*. Cambridge University Press: London, UK; 1990. 320 p.
5. Myers JA, Dominici F, Morlock L. Learning from Near Misses in Medication Errors: A Bayesian Approach. Johns Hopkins University, Department of Biostatistics Working Papers. 2008.
6. Zikos D, Diomidous M, Mantas J. A framework for the development of patient safety education and training guidelines. *Stud Health Technol Inform*. 2010;155:189–195.
7. Jick TD. Mixing Qualitative and Quantitative Methods: Triangulation in Action. *Administrative Science Quarterly*. 1979;24(4):602–611.
8. Harris DM, Westfall JM, Fernald DH, et al. Mixed Methods Analysis of Medical Error Event Reports: A Report from the ASIPS Collaborative. In: Henriksen K, editors. *Advances in Patient Safety: From Research to Implementation (Volume 2: Concepts and Methodology)*. Rockville, MD: Agency for Healthcare Research and Quality (US); 2005.
9. Levinson DR. Office of Inspector General Hospital Incident Reports Systems Do Not Capture Most Patient Harm. Department of Health & Human Services, Office of the Inspector General, Washington: USA; 2012.
10. Facebook.com. Content items shared. 2013.
11. Twitter Inc. Unites States Securities and Exchange Commission: Form S-1 Registration statement. 2013.
12. Laver M, Benoit K, Garry J. Extracting policy positions from political texts using words as data. *American Political Science Review*. 2003;97(2):311–331.
13. Baek YM, Cappella JN, Bindman A. Automating Content Analysis of Open-Ended Responses: Wordscores and Affective Intonation. *Commun Methods Meas*. 2011;5(4):275–296.
14. Cardie C, Wilkerson J. Text annotation for political science research. *Journal of Information Technology & Politics*. 2008;5(1):1–6.
15. Sexton JB, Helmreich RL, Neilands TB, et al. The Safety Attitudes Questionnaire: psychometric properties, benchmarking data, and emerging research. *BMC Health Serv Res*. 2006;6:44.